



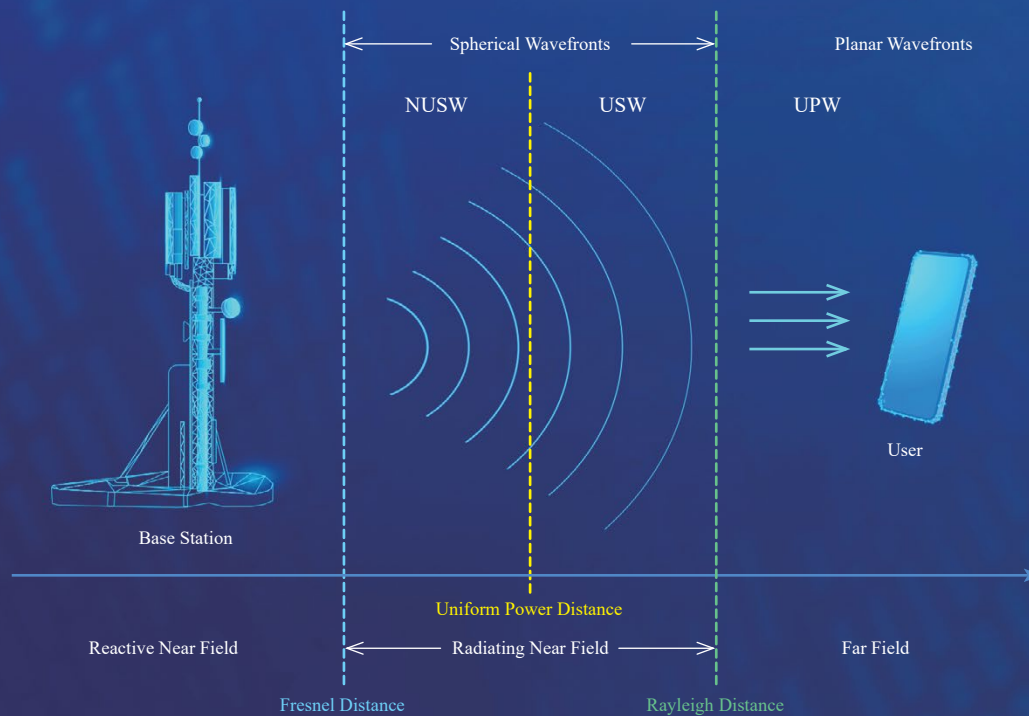
ZTE COMMUNICATIONS

中兴通讯技术(英文版)

<http://zte.magtechjournal.com>

March 2024, Vol. 22 No. 1

Special Topic: Near-Field Communication and Sensing Towards 6G



ISSN 1673-5188



9 771673 518246



(See Fig. 1 on P. 4)

The 9th Editorial Board of ZTE Communications

Chairman

GAO Wen, Peking University (China)

Vice Chairmen

XU Ziyang, ZTE Corporation (China) | **XU Chengzhong**, University of Macau (China)

Members (Surname in Alphabetical Order)

AI Bo	Beijing Jiaotong University (China)
CAO Jiannong	The Hong Kong Polytechnic University (China)
CHEN Chang Wen	The Hong Kong Polytechnic University (China)
CHEN Yan	Northwestern University (USA)
CHI Nan	Fudan University (China)
CUI Shuguang	UC Davis (USA) and The Chinese University of Hong Kong, Shenzhen (China)
GAO Wen	Peking University (China)
GAO Yang	Nanjing University (China)
GE Xiaohu	Huazhong University of Science and Technology (China)
HE Yejun	Shenzhen University (China)
HWANG Jenq-Neng	University of Washington (USA)
Victor C. M. LEUNG	The University of British Columbia (Canada)
LI Xiangyang	University of Science and Technology of China (China)
LI Zixue	ZTE Corporation (China)
LIAO Yong	Chongqing University (China)
LIN Xiaodong	ZTE Corporation (China)
LIU Chi	Beijing Institute of Technology (China)
LIU Jian	ZTE Corporation (China)
LIU Yue	Beijing Institute of Technology (China)
MA Jianhua	Hosei University (Japan)
MA Zheng	Southwest Jiaotong University (China)
PAN Yi	Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences (China)
PENG Mugen	Beijing University of Posts and Telecommunications (China)
REN Fuji	Tokushima University (Japan)
REN Kui	Zhejiang University (China)
SHENG Min	Xidian University (China)
SU Zhou	Xi'an Jiaotong University (China)
SUN Huifang	Pengcheng Laboratory (China)
SUN Zhili	University of Surrey (UK)
TAO Meixia	Shanghai Jiao Tong University (China)
WANG Chengxiang	Southeast University (China)
WANG Haiming	Southeast University (China)
WANG Xiang	ZTE Corporation (China)
WANG Xiaodong	Columbia University (USA)
WANG Xiyu	ZTE Corporation (China)
WANG Yongjin	Nanjing University of Posts and Telecommunications (China)
XU Chengzhong	University of Macau (China)
XU Ziyang	ZTE Corporation (China)
YANG Kun	University of Essex (UK)
YUAN Jinhong	University of New South Wales (Australia)
ZENG Wenjun	EIT Institute for Advanced Study (China)
ZHANG Honggang	Zhejiang Lab (China)
ZHANG Jianhua	Beijing University of Posts and Telecommunications (China)
ZHANG Yueping	Nanyang Technological University (Singapore)
ZHOU Wanlei	City University of Macau (China)
ZHUANG Weihua	University of Waterloo (Canada)

Special Topic ►

Near-Field Communication and Sensing Towards 6G

- 01 Near-Field Communication and Sensing Towards 6G WEI Guo, ZHAO Yajun, CHEN Li
- 03 Towards Near-Field Communications for 6G: Challenges and Opportunities
..... LIU Mengyu, ZHANG Yang, JIN Yasheng, ZHI Kangda, PAN Cunhua
- 16 Link Budget and Enhanced Communication Distance for Ambient Internet of Things
..... YANG Yibing, LIU Ming, XU Rongtao, WANG Gongpu, GONG Wei
- 24 Impacts of Model Mismatch and Array Scale on Channel Estimation for XL-HRIS-Aided Systems
..... LU Zhizheng, HAN Yu, JIN Shi
- 34 Degree of Freedom Analysis for Holographic MIMO Based on a Mutual-Coupling-Compliant Channel Model SUN Yunqi, JIAN Mengnan, YANG Jun, ZHAO Yajun, CHEN Yijian
- 41 Near-Field Beam Training for Holographic MIMO Communications: Typical Methods, Challenges and Future Directions
..... SHEN Jiayu, YANG Jun, ZHU Chen, DENG Zhiji, HUANG Chongwen
- 53 Near-Field Wireless Power Transfer, Sensing and Communication with Bessel Beams
..... CAO Xinghan, YIN Huarui, YOU Changsheng

Review ►

- 62 Recent Advances in Video Coding for Machines Standard and Technologies
..... ZHANG Qiang, MEI Junjun, GUAN Tao, SUN Zhewen, ZHANG Zixiang, YU Li
- 77 RIS-Assisted Cell-Free MIMO: A Survey
..... ZHAO Yaqiong, KE Hongqin, XU Wei, YE Xinquan, CHEN Yijian
- 87 Research on Multi-Core Processor Analysis for WCET Estimation
..... LUO Haoran, HU Shuisong, WANG Wenyong, TANG Yuke, ZHOU Junwei

Research Papers ►

- 95 Filter Design of Wireless Base Station Power Supply LI Wei, GUO Wei, WANG Zhida
- 106 Real-Time 4-Mode MDM Transmission Using Commercial 400G OTN Transceivers and All-Fiber Mode Multiplexers REN Fang, LI Yidan, YE Bing, LIU Jianguo, CHEN Weizhang

Serial parameters: CN 34-1294/TN*2003*q*16*110*en*P*¥20.00*2200*12*2024-03

Submission of a manuscript implies that the submitted work has not been published before (except as part of a thesis or lecture note or report or in the form of an abstract); that it is not under consideration for publication elsewhere; that its publication has been approved by all co-authors as well as by the authorities at the institute where the work has been carried out; that, if and when the manuscript is accepted for publication, the authors hand over the transferable copyrights of the accepted manuscript to *ZTE Communications*; and that the manuscript or parts thereof will not be published elsewhere in any language without the consent of the copyright holder. Copyrights include, without spatial or timely limitation, the mechanical, electronic and visual reproduction and distribution; electronic storage and retrieval; and all other forms of electronic publication or any other types of publication including all subsidiary rights.

Responsibility for content rests on authors of signed articles and not on the editorial board of *ZTE Communications* or its sponsors.
All rights reserved.

Statement

This magazine is a free publication for you. If you do not want to receive it in the future, you can send the "TD unsubscribe" mail to magazine@zte.com.cn. We will not send you this magazine again after receiving your email. Thank you for your support.

ZTE Communications Guidelines for Authors

Remit of Journal

ZTE Communications publishes original theoretical papers, research findings, and surveys on a broad range of communications topics, including communications and information system design, optical fiber and electro-optical engineering, microwave technology, radio wave propagation, antenna engineering, electromagnetics, signal and image processing, and power engineering. The journal is designed to be an integrated forum for university academics and industry researchers from around the world.

Manuscript Preparation

Manuscripts must be typed in English and submitted electronically in MS Word (or compatible) format. The word length is approximately 3 000 to 8 000, and no more than 8 figures or tables should be included. Authors are requested to submit mathematical material and graphics in an editable format.

Abstract and Keywords

Each manuscript must include an abstract of approximately 150 words written as a single paragraph. The abstract should not include mathematics or references and should not be repeated verbatim in the introduction. The abstract should be a self-contained overview of the aims, methods, experimental results, and significance of research outlined in the paper. Five carefully chosen keywords must be provided with the abstract.

References

Manuscripts must be referenced at a level that conforms to international academic standards. All references must be numbered sequentially in-text and listed in corresponding order at the end of the paper. References that are not cited in-text should not be included in the reference list. References must be complete and formatted according to *ZTE Communications* Editorial Style. A minimum of 10 references should be provided. Footnotes should be avoided or kept to a minimum.

Copyright and Declaration

Authors are responsible for obtaining permission to reproduce any material for which they do not hold copyright. Permission to reproduce any part of this publication for commercial use must be obtained in advance from the editorial office of *ZTE Communications*. Authors agree that a) the manuscript is a product of research conducted by themselves and the stated co-authors; b) the manuscript has not been published elsewhere in its submitted form; c) the manuscript is not currently being considered for publication elsewhere. If the paper is an adaptation of a speech or presentation, acknowledgement of this is required within the paper. The number of co-authors should not exceed five.

Content and Structure

ZTE Communications seeks to publish original content that may build on existing literature in any field of communications. Authors should not dedicate a disproportionate amount of a paper to fundamental background, historical overviews, or chronologies that may be sufficiently dealt with by references. Authors are also requested to avoid the overuse of bullet points when structuring papers. The conclusion should include a commentary on the significance/future implications of the research as well as an overview of the material presented.

Peer Review and Editing

All manuscripts will be subject to a two-stage anonymous peer review as well as copyediting, and formatting. Authors may be asked to revise parts of a manuscript prior to publication.

Biographical Information

All authors are requested to provide a brief biography (approx. 100 words) that includes email address, educational background, career experience, research interests, awards, and publications.

Acknowledgements and Funding

A manuscript based on funded research must clearly state the program name, funding body, and grant number. Individuals who contributed to the manuscript should be acknowledged in a brief statement.

Address for Submission

<http://mc03.manuscriptcentral.com/ztecom>



Special Topic on Near-Field Communication and Sensing Towards 6G

Guest Editors



 WEI Guo



 ZHAO Yajun



 CHEN Li

6G wireless technologies involve dense device deployment, utilize large-scale antenna arrays, and operate in the millimeter-wave and terahertz bands. This will shift the challenges of communication and sensing from the far field to the radiative near field, where traditional models and algorithms may incur errors or even become obsolete due to mismatch. Therefore, there is an urgent need to discuss localization and communication issues in the near-field region. The near-field assumption allows us to capture more information in electromagnetic signals, unlocking new possibilities for improving communication quality and localization accuracy. It can be anticipated that research on the near-field will play an increasingly important role in 6G and future wireless networks.

In this special issue on near-field communication and sensing towards 6G, a series of articles are presented to summarize the opportunities and challenges of communication sensing in the near-field region and offer innovative solutions to some key issues. A comprehensive review of near-field communications for 6G is provided, and other research articles encompass a wide range of topics, including ambient Internet of Things (ambient IoT), channel estimation, degree of freedom (DoF) analysis, beam training, and Bessel beams. The call for papers for this special issue has attracted a series of high-quality submissions, indicating a high level of interest among researchers in the field of near-field region. After two rounds of rigorous review, the following six papers are presented in this special issue organized as follows.

The first paper titled “Towards Near-Field Communications for 6G: Challenges and Opportunities” conducts a comprehensive survey on the challenges and opportunities in near-field communication. The key technologies in 6G mobile networks bring about a completely new near-field assumption. This review article formulates a general model for near-field channels and discusses the challenges in beam training, localization, and transmission scheme design within the near-field region. Finally, a series of promising research directions for near-field communication are proposed and summarized.

The second paper titled “Link Budget and Enhanced Communication Distance for Ambient Internet of Things” proposes a low-noise amplifier (LNA) module to increase the communication range of backscatter communication. Backscatter communication is a key technology for ambient IoT and plays a crucial role in connecting everything in B5G and 6G systems. By utilizing the LNA module before the envelope detection at the tag to enhance the strength of the incident signal, the downlink communication range is increased by nearly 20 m, which holds the potential to address the challenges of high losses and limited communication distance in backscatter communication.

The third paper titled “Impacts of Model Mismatch and Array Scale on Channel Estimation for XL-HRIS-Aided Systems” investigates the lower bound of the estimated parameters for the extremely large-scale hybrid reconfigurable intelligent surface (XL-HRIS) and discusses the impact of its near-field effects on estimation accuracy. XL-HRIS is an improved version of the reconfigurable intelligent surface (RIS), with the potential to further enhance communication performance. However, the increase in the size of RIS introduces near-field

DOI: 10.12142/ZTECOM.202401001

Citation: G. Wei, Y. J. Zhao, and L. Chen, “Editorial: near-field communication and sensing towards 6G,” *ZTE Communications*, vol. 22, no. 1, pp. 1–2, Mar. 2024. doi: 10.12142/ZTECOM.202401001.

effects, where the phase variation of the signal becomes significant, and mismatched channel models may reduce the accuracy of the estimation. In this case, the paper meticulously studies the impact of the array scale, distance, and signal-to-noise ratio on the lower bound of parameter estimation for XL-HRIS in the near-field region.

The fourth paper titled “Degree of Freedom Analysis for Holographic MIMO Based on a Mutual-Coupling-Compliant Channel Model” utilizes a mutual-coupling-compliant circuit multiple-input multiple-output (MIMO) model to study the DoF of the holographic MIMO (HMIMO) channel. The DoF is a key indicator of the spatial multiplexing layers of a wireless channel. Recent studies indicate that in near-field communication, when considering strong mutual coupling, the effective DoF may be less than the number of antennas. This paper investigates the DoF in HMIMO systems, provides an upper bound on DOF under strong coupling conditions, and validates the results through numerical simulations.

The fifth paper titled “Near-Field Beam Training for Holographic MIMO Communications: Typical Methods, Challenges and Future Directions” conducts a comprehensive survey of near-field beam training methods in HMIMO systems. HMIMO is a promising technology in future wireless systems for achieving ultra-high frequency spectral efficiency and spatial resolution. Due to the increase in antenna aperture, the spherical wavefront effect in near-field communication becomes more pronounced in HMIMO systems, necessitating the use of new near-field beam training methods. The paper introduces, analyzes, and compares the performance of several typical near-field beam training methods, validating them on a hardware platform. Additionally, it addresses the challenges and outlines future research directions in near-field beam training.

The sixth paper titled “Near-Field Wireless Power Transfer, Sensing and Communication with Bessel Beams” conducts a comprehensive survey on the generation and application of Bessel beams in the near-field region. Bessel beams are a type of beam with limited propagation distance. Unlike traditional phased-array beamforming, Bessel beams exhibit unique non-diffracting characteristics, maintaining a stable beam intensity along the propagation direction without spreading. The paper introduces the concept and basic theory of Bessel beams, and categorizes and discusses research on their generation and applications, with particular emphasis on their application in near-field wireless power transfer and communi-

cation. Finally, the research challenges and opportunities associated with Bessel beams are discussed.

To conclude, it is hoped that this special issue serves as a solid foundation, helping researchers understand and address the upcoming near-field challenges in wireless networks. Additionally, we hope that this special issue sparks more research interest in near-field communication and sensing in 6G networks. Finally, we sincerely express our gratitude to all the authors and reviewers whose outstanding work contributes to the success of this special issue. We sincerely hope that the papers in this special issue provide both clarity and insight for all readers in the field.

Biographies

WEI Guo received his BS degree in electronic engineering from the University of Science and Technology of China (USTC) in 1983 and MS and PhD degrees in electronic engineering from the Chinese Academy of Sciences, China in 1986 and 1991, respectively. He is currently a professor with the School of Information Science and Technology, USTC. He has won the second prize of the National Science and Technology Progress Award and published more than 100 papers. He holds dozens of national invention patents. His current research interests include wireless and mobile communications, wireless multimedia communications, and wireless information networks.

ZHAO Yajun received his BE, MS, and PhD degrees. Since 2010, he has assumed the role of Chief Engineer at the Wireless and Computing Product R&D Institute, ZTE Corporation. Prior to this, he contributed to wireless technology research at the Wireless Research Department, Huawei. Currently, his primary focus centers on 5G standardization technology and the advancement of future mobile communication technology, particularly 6G. His research pursuits encompass a broad spectrum, including reconfigurable intelligent surfaces (RIS), spectrum sharing, flexible duplex, CoMP, and interference mitigation. He has played an instrumental role in founding the RIS Tech Alliance (RISTA) and currently holds the position of Deputy Secretary General within the organization. Additionally, he is a founding member of the RIS task group under the purview of the China IMT-2030 (6G) Promotion Group, where he serves as the deputy leader.

CHEN Li received his BE degree in electrical and information engineering from Harbin Institute of Technology, China in 2009 and PhD degree in electrical engineering from the University of Science and Technology of China (USTC) in 2014. He joined The University of Hong Kong, China as a postdoctor from 2016 to 2017. He is currently an associate professor with the Department of Electronic Engineering and Information Science, USTC. He has hosted a number of national projects and won the second prize of the Anhui Provincial Natural Science Award and the Anhui Provincial Outstanding Youth Fund. His research interests include integrated communication and computation and integrated communication and sensing.



Towards Near-Field Communications for 6G: Challenges and Opportunities

LIU Mengyu¹, ZHANG Yang¹, JIN Yasheng¹,
ZHI Kangda², PAN Cunhua¹

(1. Southeast University, Nanjing 210096, China;
2. Queen Mary University of London, London E1 4NS, UK)

DOI: 10.12142/ZTECOM.202401002

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240229.1113.002.html>,
published online March 1, 2024

Manuscript received: 2023-11-30

Abstract: Extremely large-scale multiple-input multiple-output (XL-MIMO) and terahertz (THz) communications are pivotal candidate technologies for supporting the development of 6G mobile networks. However, these techniques invalidate the common assumptions of far-field plane waves and introduce many new properties. To accurately understand the performance of these new techniques, spherical wave modeling of near-field communications needs to be applied for future research. Hence, the investigation of near-field communication holds significant importance for the advancement of 6G, which brings many new and open research challenges in contrast to conventional far-field communication. In this paper, we first formulate a general model of the near-field channel and discuss the influence of spatial nonstationary properties on the near-field channel modeling. Subsequently, we discuss the challenges encountered in the near field in terms of beam training, localization, and transmission scheme design, respectively. Finally, we point out some promising research directions for near-field communications.

Keywords: near-field communications; extremely large-scale antenna arrays; spatial non-stationarity; beam training; localization

Citation (Format 1): LIU M Y, ZHANG Y, JIN Y S, et al. Towards near-field communications for 6G: challenges and opportunities [J]. *ZTE Communications*, 2024, 22(1): 3 - 15. DOI: 10.12142/ZTECOM.202401002

Citation (Format 2): M. Y. Liu, Y. Zhang, Y. S. Jin, et al., "Towards near-field communications for 6G: challenges and opportunities," *ZTE Communications*, vol. 22, no. 1, pp. 3 - 15, Mar. 2024. doi: 10.12142/ZTECOM.202401002.

1 Introduction

The advent of the 5G wireless network marks the inception of an era characterized by high speed, low latency, and robust connectivity^[1-4]. Millimeter waves (mmWave) and various key technologies like massive multiple-input multiple-output (MIMO) have contributed significantly to the success of 5G, particularly in areas such as smart cities and the Internet of Things (IoT)^[5-7]. Although 5G has achieved remarkable results, it still faces limitations in addressing diverse and complex business scenarios, including virtual reality, driverless vehicles and metaverse^[8]. These emerging scenarios impose greater demands on the network capacity, prompting researchers to explore 6G wireless networks, which will further surpass the limits of the current technologies to meet much more complex communication needs^[9].

In contrast to 5G, 6G has the potential to deliver higher communication rates, ultra-reliable and low latency communi-

cations (URLLC), and ubiquitous coverage^[10-11]. In order to achieve these goals for 6G, some technologies considered promising for 6G have attracted significant attention, such as extremely large-scale MIMO (XL-MIMO) and Terahertz (THz) communications^[12]. Specifically, XL-MIMO can realize ultra-high network throughput and support a large number of users by further increasing the number of antennas on massive MIMO. Meanwhile, emerging technologies like reconfigurable intelligent surface (RIS) and artificial intelligence (AI) techniques empower XL-MIMO to attain increased spectral efficiency, enhanced positioning accuracy, and broader network coverage across a more diverse frequency range. Also, THz communications, with its capacity to exploit richer spectrum resources for enhanced data transmission rates, is a key candidate spectrum technology with significant potential.

However, the new technologies in 6G, e.g., THz communications and XL-MIMO, also lead to operating frequency band escalation and antenna array aperture expansion. As a result, Rayleigh distance, the crucial parameter, that distinguishes between the near-field and far-field boundaries of electromagnetic (EM) wave propagation changes significantly^[12]. For example, Ref. [13] gave the Rayleigh distances corresponding to different array apertures, where the near-field range corresponding to an array of 0.5 m aperture was already as high as

This work was supported in part by National Key Research and Development Young Scientist Project 2023YFB2905100, the National Natural Science Foundation of China under Grant Nos. 62201137 and 62331023, the Fundamental Research Funds for the Central Universities under Grant No. 2242022k60001, and the Research Fund of National Mobile Communications Research Laboratory, Southeast University, China under Grant No. 2023A03.

47 m for millimeter waves at 28 GHz. As the Rayleigh distance increases, the user will be easily located in the near field rather than the far field of the base station (BS), meaning that the consideration of the near-field effect is crucial in 6G wireless communications. Within the near field, far field models that approximate EM waves as plane waves are no longer accurate and the spherical shaping of the wavefront cannot be neglected^[14]. This new property renders current wireless communication models and findings based on far field assumptions inapplicable in the near-field cases. Therefore, it is necessary to re-investigate the challenges, potential benefits, and solutions introduced by considering near-field communications to advance 6G communication capabilities.

Nowadays, there are relatively few articles about near-field communication in 6G^[15-18]. The author of Ref. [15] provided an overview of 6G wireless systems including challenges, insights, and related opportunities. Although the article elaborated on the challenges brought by the near field in future 6G wireless communication, there is a lack of explanation of its specific issues and opportunities. The authors of Ref. [16] presented an overview of near-field communications, contrasted it with far-field communications, and analyzed the key challenges. However, the description of the channel modeling process is neglected. The authors of Ref. [17] studied the basic channel models, antenna structures, and analytical foundations. Furthermore, the authors of Ref. [18] studied the new opportunities brought by the near-field beam focusing caused by the near-field communications, which is different from that in the far field. Different from the above articles, we focus on more specific research problems in near-field communications, especially the combination of near-field communication and new technologies, including the latest research results in the fields of deep learning and near-field beam training, RIS-aided near-field localization, and near-field transmission scheme design with visibility regions (VRs).

In this paper, we systematically present the fundamental models, recent advancements, and solutions in 6G near-field communication. To begin with, we introduce the overall spatial channel model, where the communication region is divided into three parts given different characteristics of array propagation signal phases and power. Subsequently, we introduce the formulation of the channel model encompassing the MIMO systems, accounting for both the far field and the near field. Then, we present recent advances in near-field communication for 6G, including near-field beam training, near-field localization, and near-field

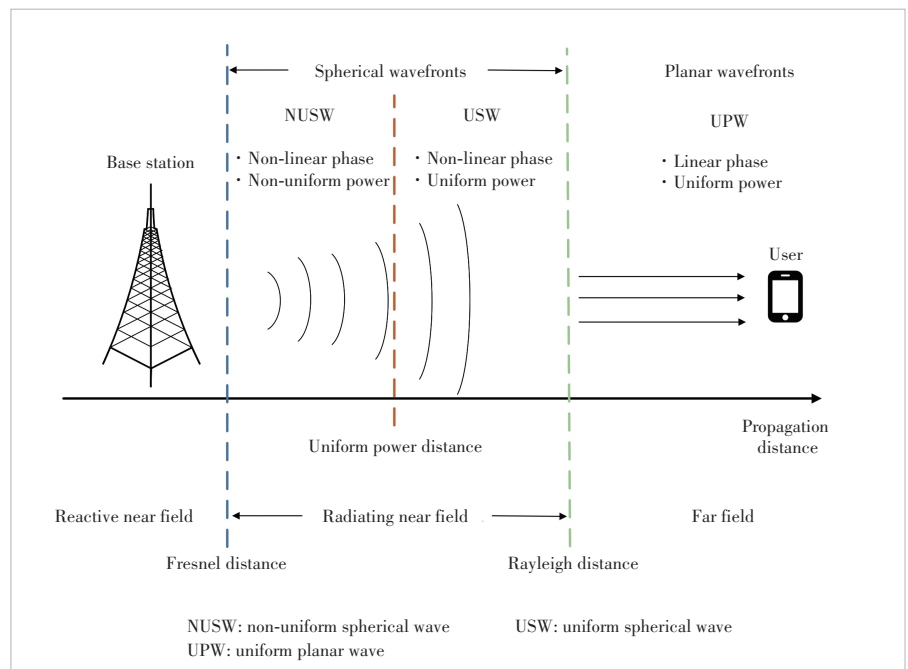
transmission scheme design. We elaborate on the challenges encountered in near-field communication within these use cases and provide corresponding solutions. Finally, we share some promising research directions for near-field communication, such as investigating the effectiveness of near-field beam training in complex environments, proposing near-field localization algorithms with low complexity and overhead, and carrying out channel measurements to verify and modify the near-field EM channel model.

2 Fundamentals of Near-Field Communications

In this section, we first introduce the overall space channel model. Subsequently, we elaborate on the modeling approach to the near-field channel. Additionally, the spatial nonstationary property of the channel in the near field is investigated.

2.1 Analysis of the Overall Space Channel Model

As illustrated in Fig. 1, the space from BS to the user can be divided into three regions: far field, radiating near field, and reactive near field. These regions are divided according to the Rayleigh distance and Fresnel distance^[19-21], respectively. The far field refers to the region where the distance between the transmitter and the receiver exceeds the Rayleigh distance, allowing the propagating signal to be safely approximated as the plane wave. In the radiating near field, the distance between the transmitter and the receiver is typically less than the Rayleigh distance but greater than the Fresnel distance, where the propagating signal can no longer be modeled as the plane wave; instead, the accurate spherical wave model



▲ Figure 1. Overall channel model

should be utilized. In the reactive near field where the EM fields are reactive, signals are not propagated as EM waves but stored or released as energy. Since the distance of the reactive near field is very small, the near field in the following discussion mainly refers to the radiating near field.

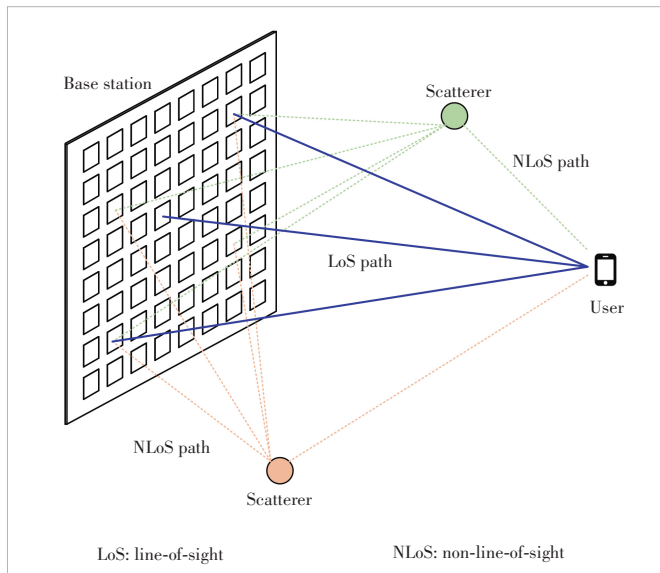
Additionally, as depicted in Fig. 1, we categorize the far and near fields into three models^[17]. 1) Uniform planar wave (UPW): in the far field, signal propagation typically conforms to this model where the phase difference in signal transmission is linear and the transmitted power is uniform for different antenna elements. 2) Uniform spherical wave (USW): in the near field, when the distance between the transmitter and receiver is less than the Rayleigh distance but greater than the uniform power distance^[17, 22], the phase difference between the signals propagated by different antennas is nonlinear while the transmitted power is uniform. 3) Non-uniform spherical wave (NUSW): in the near field, when the distance between the transmitter and the receiver is less than the uniform power distance, the phase difference between the signals propagated by different antennas is nonlinear and the transmitted power is not uniform.

2.2 Generic Channel Modeling

In this section, we first derive a general channel model considering MIMO systems. After that, we analyze a simple scenario of a multiple-input single-output (MISO) system with a uniform linear array (ULA) array at the base station to gain more insights.

2.2.1 MIMO Channel Model

As shown in Fig. 2, we first analyze the channel model of the MIMO system. Let us consider a MIMO system that includes a transmitter with N_t antennas and user equipment (UE) with N_r antennas. We first set the center of the BS antenna as the origin,



▲ Figure 2. Near-field channel model

i.e., $\mathbf{s}_0 = [0, 0, 0]^T$. Then, the locations of the m -th element of BS and the n -th antenna of UE can be expressed by $\mathbf{s}_m = [s_x^m, s_y^m, s_z^m]^T$, $\forall m = -(N_t - 1)/2, \dots, (N_t - 1)/2$ and $\mathbf{u}_n = [u_x^n, u_y^n, u_z^n]^T$, $\forall n = -(N_r - 1)/2, \dots, (N_r - 1)/2$, respectively. In addition, the distance between the m -th antenna of UE and the n -th element of BS is given by $r_{m,n} = \|\mathbf{u}_n - \mathbf{s}_m\|$. Then, the line-of-sight (LoS) channel coefficient can be expressed as:

$$[\mathbf{H}_{\text{LoS}}]_{m,n} = \alpha_{m,n} e^{-j\frac{2\pi}{\lambda}r_{m,n}}, \quad (1)$$

where $\alpha_{m,n}$ represents the amplitude for the channel link between the m -th antenna of BS and the n -th element of UE.

1) Far-field channel

Firstly, we can define the propagation direction vector from the transmitter to UE as follows:

$$\mathbf{v}(\theta, \phi) = [\cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi]^T. \quad (2)$$

Based on the plane-wave assumption^[23] for the far-field channel, the propagation distance can be approximated by $r_{m,n} \approx r_0 - \mathbf{v}^T(\theta, \phi)\mathbf{s}_m - \mathbf{v}^T(\theta, \phi)(\mathbf{u}_n - \mathbf{u}_0)$, where $r_0 = \|\mathbf{u}_0 - \mathbf{s}_0\|$ denotes the distance between the central elements of the receiving and the transmitting antenna arrays. Moreover, we assume $\alpha_{m,n} \approx \alpha$, where α denotes the amplitude for the channel link between the central antenna elements of the transmitter and the UE. Then, the LoS far-field channel coefficient can be rewritten as:

$$[\mathbf{H}_{\text{LoS}}^{\text{far}}]_{m,n} = \alpha_{m,n} e^{-j\frac{2\pi}{\lambda}r_{m,n}} \approx \alpha e^{-j\frac{2\pi}{\lambda}r_0} e^{j\frac{2\pi}{\lambda}\mathbf{v}^T(\theta, \phi)\mathbf{s}_m} e^{j\frac{2\pi}{\lambda}\mathbf{v}^T(\theta, \phi)(\mathbf{u}_n - \mathbf{u}_0)}. \quad (3)$$

From the channel coefficients in Eq. (3), we can decouple the MIMO channel into the product of two array response vectors. Therefore, the far-field LoS MIMO channel can be expressed as:

$$\mathbf{H}_{\text{LoS}}^{\text{far}} = \alpha \mathbf{a}_{\text{far}}^{\text{UE}}(\theta, \phi) (\mathbf{a}_{\text{far}}^{\text{TR}}(\theta, \phi))^T, \quad (4)$$

where

$$\mathbf{a}_{\text{far}}^{\text{UE}}(\theta, \phi) = \left[e^{j\frac{2\pi}{\lambda}\mathbf{v}^T(\theta, \phi)(\mathbf{u}_{-(N_r-1)/2} - \mathbf{u}_0)}, \dots, e^{j\frac{2\pi}{\lambda}\mathbf{v}^T(\theta, \phi)(\mathbf{u}_{(N_r-1)/2} - \mathbf{u}_0)} \right]^T, \quad (5)$$

$$\mathbf{a}_{\text{far}}^{\text{TR}}(\theta, \phi) = \left[e^{j\frac{2\pi}{\lambda}\mathbf{v}^T(\theta, \phi)\mathbf{s}_{-(N_t-1)/2}}, \dots, e^{j\frac{2\pi}{\lambda}\mathbf{v}^T(\theta, \phi)\mathbf{s}_{(N_t-1)/2}} \right]^T. \quad (6)$$

Furthermore, the form of the far-field non-line-of-sight (NLoS) channel is similar to the far-field LoS channel. Therefore, the overall far-field MIMO channel model with multipath components is given by:

$$\mathbf{H}^{\text{far}} = \alpha \mathbf{a}_{\text{far}}^{\text{UE}}(\theta, \varphi) (\mathbf{a}_{\text{far}}^{\text{TR}}(\theta, \varphi))^T + \sum_{l=1}^L \beta_l \mathbf{a}_{\text{far}}^{\text{UE}}(\theta_l, \varphi_l) (\mathbf{a}_{\text{far}}^{\text{TR}}(\theta_l, \varphi_l))^T. \quad (7)$$

2) Near-field channel

As shown in Fig. 2, the NLoS channels in the near field exist between the transmitter and the UE due to the presence of scatterers. Clearly, the NLoS channel can be divided into the combination of two MISO channels, i.e., the channel from the UE to the scatterer and the channel from the scatterer to the transmitter. Therefore, the NLoS channel can be decoupled as the product of two array response vectors:

$$\mathbf{H}_{\text{NLoS}}^{\text{near}} = \sum_{l=1}^L \beta_l \mathbf{a}_{\text{near}}^{\text{UE}}(\mathbf{r}_l) (\mathbf{a}_{\text{near}}^{\text{TR}}(\mathbf{r}_l))^T, \quad (8)$$

where β_l and \mathbf{r}_l denote the channel coefficients for the l -th NLoS path and the location of the l -th scatterer, respectively. And the near-field array response vector $\mathbf{a}_{\text{near}}^{\text{UE}}(\mathbf{r}_l)$ and $\mathbf{a}_{\text{near}}^{\text{TR}}(\mathbf{r}_l)$ can be given by:

$$\mathbf{a}_{\text{near}}^{\text{UE}}(\mathbf{r}_l) = \left[e^{-j\frac{2\pi}{\lambda} \|\mathbf{u}_{-(N_t-1)/2} - \mathbf{r}_l\|}, \dots, e^{-j\frac{2\pi}{\lambda} \|\mathbf{u}_{(N_t-1)/2} - \mathbf{r}_l\|} \right]^T, \quad (9)$$

$$\mathbf{a}_{\text{near}}^{\text{TR}}(\mathbf{r}_l) = \left[e^{-j\frac{2\pi}{\lambda} \|\mathbf{s}_{-(N_t-1)/2} - \mathbf{r}_l\|}, \dots, e^{-j\frac{2\pi}{\lambda} \|\mathbf{s}_{(N_t-1)/2} - \mathbf{r}_l\|} \right]^T. \quad (10)$$

However, the LoS channel in the near field cannot be decoupled as the product of two array response vectors. Specifically, the LoS channel coefficient can be expressed as:

$$\left[\mathbf{H}_{\text{LoS}}^{\text{near}} \right]_{m,n} = \alpha_{m,n} e^{-j\frac{2\pi}{\lambda} r_{m,n}} = \alpha_{m,n} e^{-j\frac{2\pi}{\lambda} \|\mathbf{u}_n - \mathbf{s}_m\|}. \quad (11)$$

Therefore, the overall near-field channel with multi-path components can be modeled as

$$\mathbf{H}^{\text{near}} = \mathbf{H}_{\text{LoS}}^{\text{near}} + \sum_{l=1}^L \beta_l \mathbf{a}_{\text{near}}^{\text{UE}}(\mathbf{r}_l) (\mathbf{a}_{\text{near}}^{\text{TR}}(\mathbf{r}_l))^T. \quad (12)$$

2.2.2 MISO Channel Model Based on ULA Base Station

To more intuitively understand the properties of the near-field channel, we next simplify it to the uniform linear array (ULA) model.

We assume that the base station is equipped with N_t -ULA antennas and the user is equipped with a single antenna. In the case of ULA, we can ignore the z -axis and assume the locations of the m -th antenna element of BS and UE as $\mathbf{u} = [r_0 \cos \theta, r_0 \sin \theta]^T$ and $\mathbf{s}_m = [md, 0]^T, \forall m = -(N_t - 1)/2, \dots, (N_t - 1)/2$, respectively, where d denotes spacing between two adjacent antenna elements. Therefore,

the distance between the m -th antenna element of the transmitter and UE can be expressed as:

$$r_m = \|\mathbf{u} - \mathbf{s}_m\| = \sqrt{r_0^2 - 2r_0md \cos \theta + m^2d^2} \stackrel{(a)}{\approx} r_0 - md \cos \theta + \frac{m^2d^2 \sin^2 \theta}{2r_0}. \quad (13)$$

where (a) denotes the Taylor expansion approximation, which is also called the Fresnel approximation^[24]. Then, the near-field array response vector for the ULA channel is given by

$$\mathbf{a}_{\text{ULA}}^{\text{near}}(\theta) = \left[e^{j\frac{2\pi}{\lambda} \left(-(N_t-1)/2 d \cos \theta - \frac{-(N_t-1)^2 d^2 \sin^2 \theta}{2r} \right)}, \dots, e^{j\frac{2\pi}{\lambda} \left((N_t-1)/2 d \cos \theta - \frac{(N_t-1)^2 d^2 \sin^2 \theta}{2r} \right)} \right]^T. \quad (14)$$

For the far-field scenario, the distance r_m can be approximated as $r_m = r_0 - md \cos \theta$. Based on this, the far-field array response vector for the ULA channel can be expressed by:

$$\mathbf{a}_{\text{ULA}}^{\text{far}}(\theta) = \left[e^{-j\frac{2\pi}{\lambda} (-(N_t-1)/2 d \cos \theta)}, \dots, e^{-j\frac{2\pi}{\lambda} (N_t-1)/2 d \cos \theta} \right]^T. \quad (15)$$

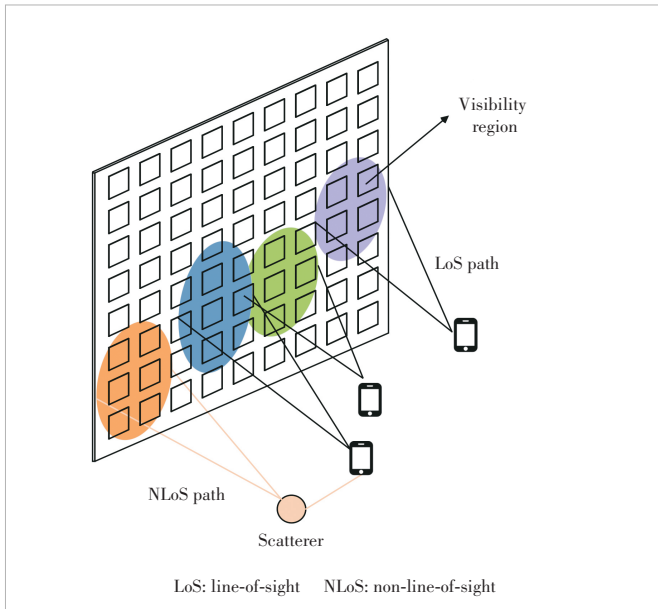
Clearly, it can be observed from Eq. (14) that the phase of the near-field array response vector is not a linear function of m . However, for the far-field array response vector in Eq. (15), the phase is a linear function of m .

From the above derivation, it is evident that the far-field plane wave assumption will result in an obvious error in the practical near-field scenario. Moreover, the array response vectors in the near-field channel are more complex, posing challenges to conventional transmission scheme design and channel estimation.

2.3 Near-Field Channel with Spatial Non-Stationarity Property

In the near-field scenario, since the size of the transmitter's antenna array is significantly increased, spatial non-stationary properties^[25-26] start to appear. Consequently, we introduce the concept of spatial non-stationarities and VRs based on the modeling of the near-field channel.

As illustrated in Fig. 3, spatial non-stationarities mean that due to the large array aperture, different parts of the antenna array could have different views of the propagation environment. This property restricts the user to receiving only a portion of the signal transmitted by the antenna array, which is referred to as VR of the user. The VRs are formed primarily due to two reasons as follows^[12]. 1) Unequal path loss: Based on the near-field channel, the increase in antenna array size makes it impossible to ignore the difference in path loss from different array elements to the user. Some of the array elements that are farther away from the user suf-



▲ Figure 3. Illustration of spatial non-stationary property

fer greater path loss. As a result, most of the power of the signal received by the user comes from those array elements with lower path loss. This phenomenon leads to the appearance of VRs. 2) Blockage: The large antenna size increases the possibility of obstacles between the user and some antennas. This obstruction prevents the user from receiving signals from some sub-arrays which are blocked by obstacles, bringing the emergence of VRs.

Then, we can obtain the near-field channel considering the VRs^[27-28]:

$$\mathbf{h}^{\text{near}} = \alpha e^{-j\frac{2\pi}{\lambda}r} \mathbf{a}^{\text{near}}(\mathbf{r}) \odot \mathbf{f}(\Phi) + \sum_{l=1}^L \beta_l \mathbf{a}^{\text{near}}(\mathbf{r}_l) \odot \mathbf{f}(\Phi_l), \quad (16)$$

where Φ and Φ_l represent the index of the array that is visible to the user and the l -th scatterer, respectively, and $\mathbf{f}(\Phi)$ denotes the vector of VRs, which is given by:

$$[\mathbf{f}(\Phi)]_n = \begin{cases} 1, & n \in \Phi, \\ 0, & n \notin \Phi. \end{cases} \quad (17)$$

Moreover, when the prior information about the actual environment is unknown, the user's VR information can be modeled as a Markov process or birth and death process^[29-30].

3 Challenges and Research Progress in Near-Field Communications

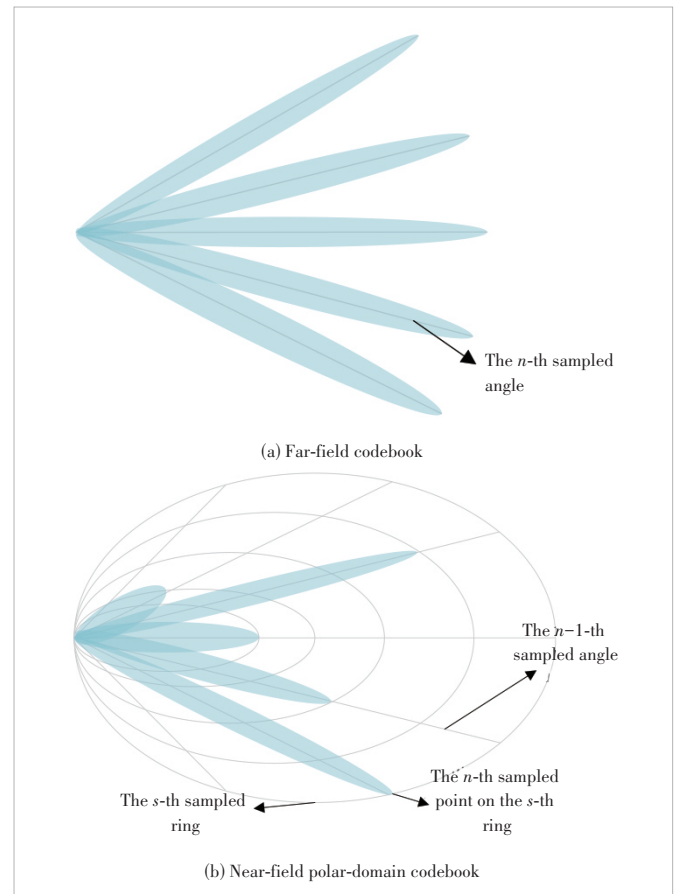
In this section, we will present some challenging directions in near-field communication, including beam training, localization and transmission scheme design. For each challenge, we then provide a detailed literature review and the latest research progress.

3.1 Beam Training for Near-Field Communications

3.1.1 Challenges

Codebook-based beam training intends to identify the optimal combination of transmitting and receiving beams for subsequent data transmission. Most of the existing research on beam training has been conducted under the assumption of far field^[31-34]. However, for next-generation communication systems, the use of XL-MIMO techniques invalidates existing far-field assumptions, making it inevitable to consider the near-field model^[14]. The change from far-field to near-field presents opportunities and challenges to beam training.

In the far-field domain, the codebook design only considers the angle domain information as shown in Fig. 4(a), where the angles are uniformly sampled. However, the Rayleigh distance that distinguishes the near field from the far field expands with the growth of the antenna array, necessitating the consideration of near-field effects^[20]. Unlike the far field, the distance information is added in the near-field domain codebook design, making the codebook dimension dramatically higher^[17]. Therefore, it becomes crucial to investigate codebook-based beam training methods under the near-field domain to reduce overhead.



▲ Figure 4. Codebook for the far field and near field

3.1.2 Existing Works

Similar to far-field beam training, near-field beam training can be accomplished using both traditional and machine learning-based methods, with beam sweeping and hierarchical beam training^[35].

The most straightforward way for codebook-based near-field beam training is beam sweeping; however, this approach introduces unacceptable pilot overhead due to the expansion of the codebook dimension. To tackle this problem, researchers designed a hierarchical beam training scheme based on a hierarchical near-field codebook, which reduces the pilot overhead^[36]. Furthermore, the authors of Ref. [24] designed a polar-domain near-field codebook in which the angular domain was uniformly sampled while the distance was sampled as inhomogeneous distance rings as shown in Fig. 4(b). By utilizing the polar-domain near-field codebook, the authors in Ref. [37] proposed a new two-phase beam training method. Specifically, beam training is accomplished by angular domain sweeping based on a far-field codebook and distance sweeping based on a polar-domain near-field codebook, respectively, which significantly reduces the training overheads. Nevertheless, the necessary pilot overhead remains excessive, and the impact of noise on the hierarchical beam training scheme cannot be neglected. Deep learning constitutes a subdivision of machine learning that relies on artificial neural networks to emulate and acquire intricate data representations and features through multi-layer neural network architectures. It gradually abstracts high-level features and representations in the input data through multi-layer nonlinear transformations to enable learning and reasoning about complex tasks. Recently, deep learning has been widely used in wireless communication to reduce the cost of beam training^[34, 38 - 41]. The authors of Ref. [34] took the received signals corresponding to a few beams as the input for deep neural networks (DNN) to estimate the beam that best matches the channel. Ref. [34] introduced a DNN-based method with location information to reduce beam training overhead. Inspired by the above work, the authors of Refs. [38 - 41] have extended the deep learning method to near-field beam training. The authors of Ref. [38] used a part of the near-field beam as the input of DNN for prediction. By contrast, Ref. [40] utilized the information of far-field wide beam as the input of the deep learning network to predict the optimal beam to reduce the overhead. In the following part, we present more details of the research work in Ref. [40].

3.1.3 Near-Field Beam Training Based on Deep Learning

A well-trained deep learning approach is used in Ref. [40] to uncover the angle

and distance information carried by the far-field wide beam to find the optimal near-field codeword. Specifically, the authors constructed the neural network structure shown in Fig. 5 for the angle and distance domains, respectively. The deep learning-based near-field beam training problem can be formulated as:

$$n^* = f_1(\mathbf{y}^w), n^* \in \{1, 2, \dots, N\}, \quad (18)$$

$$s^* = f_2(\mathbf{y}^w), s^* \in \{1, 2, \dots, S\}, \quad (19)$$

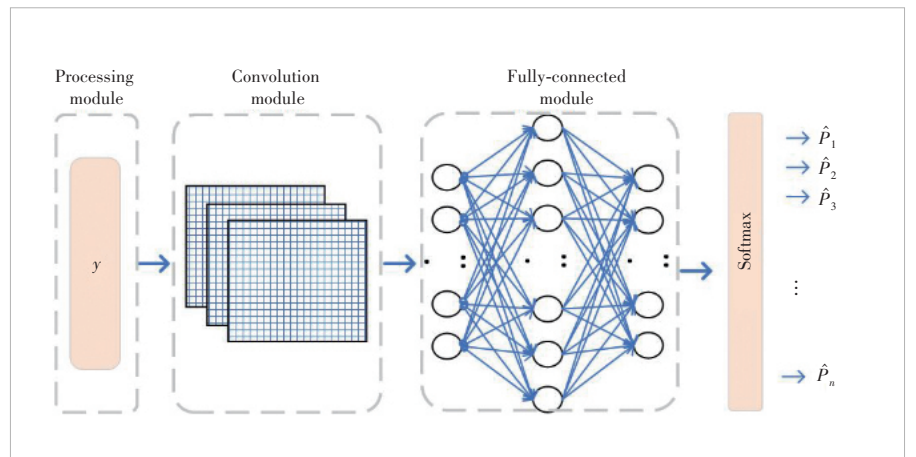
where \mathbf{y}^w is received signals of the far-field wide beams. n^* and s^* denote the optimal angle and ring indices in the polar domain, respectively.

The input of the neural network is the real and imaginary part of \mathbf{y}^w , which is transformed into a matrix form, respectively. Then with carefully designed convolutional and fully connected layers, the final output can be written as:

$$\hat{\mathbf{P}}^a = [\hat{p}_1^a, \hat{p}_2^a, \dots, \hat{p}_N^a]^T, \hat{\mathbf{P}}^r = [\hat{p}_1^r, \hat{p}_2^r, \dots, \hat{p}_S^r]^T, \quad (20)$$

where \hat{p}_n^a and \hat{p}_s^r denote the estimated probability of the optimal angle and ring indices, respectively. The authors determine the index corresponding to the angle and distance under the polar-domain codebook by finding the maximum estimated probability in $\hat{\mathbf{P}}^a$ and $\hat{\mathbf{P}}^r$ to achieve the optimal beam. In addition, the authors offer an improved scheme, in which it performs additional tests by obtaining K maximum possible angle indices and L maximum possible distance indices in $\hat{\mathbf{P}}^a$ and $\hat{\mathbf{P}}^r$. It can be formulated as:

$$\begin{aligned} \{\hat{p}_{\sigma_1}^a, \hat{p}_{\sigma_2}^a, \dots, \hat{p}_{\sigma_n}^a\} &= \langle \langle \{\hat{p}_1^a, \hat{p}_2^a, \dots, \hat{p}_N^a\} \rangle \rangle, \{\hat{p}_{\gamma_1}^r, \hat{p}_{\gamma_2}^r, \dots, \hat{p}_{\gamma_s}^r\} = \\ &\langle \langle \{\hat{p}_1^r, \hat{p}_2^r, \dots, \hat{p}_S^r\} \rangle \rangle, \end{aligned} \quad (21)$$



▲ Figure 5. Proposed neural network structure for beam training

$$\mathcal{L}_a = \{\sigma_1, \sigma_2, \dots, \sigma_K\}, \mathcal{L}_d = \{\gamma_1, \gamma_2, \dots, \gamma_L\}, \quad (22)$$

where $\langle \cdot \rangle$ denotes the order operation. By performing KL additional near-field beam training tests, we can get the optimal near-field beam.

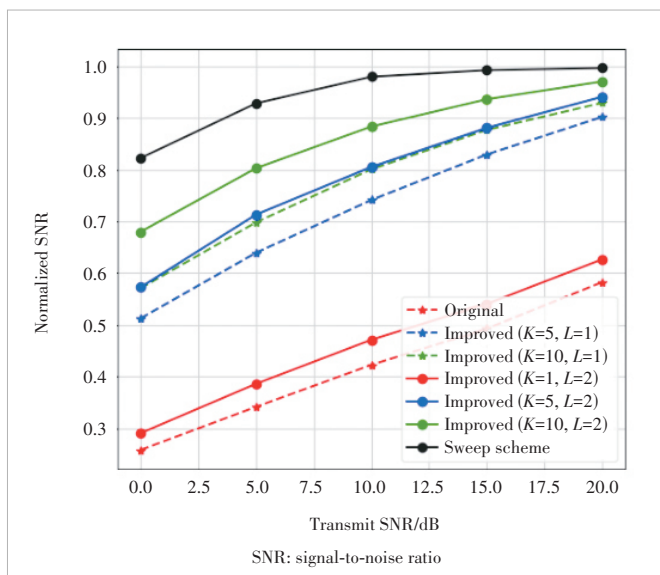
Simulation results validate the advantages of original and improved schemes compared with the sweep scheme as shown in Fig. 6. The improved scheme can achieve over 90% of the data rate and reduce overhead by 95% compared with the sweep scheme. This stems from the remarkable data processing capabilities in neural networks and the full leveraging of the output in the improved scheme.

3.2 Localization for Near-Field Communications

3.2.1 Challenges

In THz systems, the performance of positioning can be improved since all the geometric information is included in the channel state information (CSI) measurements, which can be estimated more accurately with narrow beams^[42]. In addition, RIS can provide reliable and high-precision estimation performance with low energy consumption, and the large size of the RIS panel enables high-accuracy positioning parameter estimation^[43]. Therefore, research on RIS-aided localization has attracted increasing attention, while most of them adopted the planar wavefront based on the far-field channel model^[44-52].

However, in some envisioned scenarios with the XL-RIS panel, the typical indoor communication distances may not guarantee the validity of far-field conditions. In addition, UE is usually located in the near-field area of the XL-RIS in the THz localization system owing to the short wavelength^[53]. Therefore, the spherical wavefront of near-field transmission should be considered in XL-RIS-based THz localization, where the transmitted wavefront will arrive at each element of



▲ Figure 6. Normalized SNR for the original and improved schemes

the RIS panel with different Angles of Arrival (AoAs). The assumption that all the RIS elements share a common AoA of the same path is no longer valid, making the channel estimation and localization more challenging.

3.2.2 Existing Works

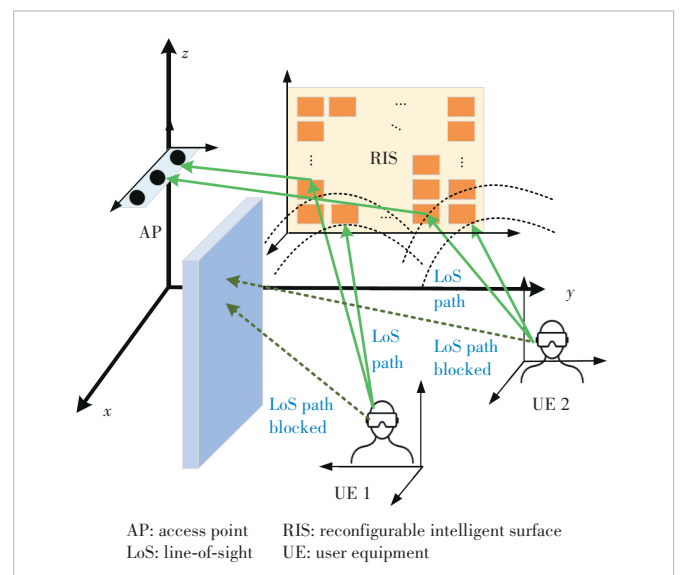
Although near-field RIS behavior has attracted increasing attention recently, the study on RIS-aided near-field localization is still in its infancy. Most researchers considering the spherical wavefront mainly analyzed the Fisher information matrix (FIM) and the Cramer-Rao lower bound (CRLB) to illustrate the position error bound (PEB) and orientation error bound (OEB), which could be derived as performance benchmarks for the practical algorithms^[48, 54-55]. In addition, the authors of Ref. [56] proposed an algorithm of RIS-aided near-field joint channel estimation and localization (NF-JCEL) in THz systems, which addressed the specific issue of channel estimation and UE positioning in the near-field scenario.

3.2.3 RIS-Aided Near-Field Localization Algorithm

The authors of Ref. [56] considered the uplink transmission of a THz localization system as shown in Fig. 7, where the direct links between the AP and the UE are assumed to be obstructed. The receiving antenna arrays of the AP and the RIS are assumed to be ULA and UPA, respectively.

Assuming that the UE is located at the near field of the RIS panel, the authors derived the two-dimensional Fresnel approximation of the distance between the m -th element and UE u , and approximated the near-field array response as:

$$\mathbf{a}_{R,u}^{N,i}(\omega_u, \varphi_u, d_u^0)[m] = \exp\left(-j \frac{2\pi}{\lambda_i} \left(J_m(\omega_u, \varphi_u) + Q_m(\omega_u, \varphi_u, d_u^0) \right)\right), \quad (23)$$



▲ Figure 7. System model and communication scenario

where $J_m(\omega_u, \varphi_u) = m_z \Delta_R \omega_u - m_y \Delta_R \varphi_u$, and $Q_m(\omega_u, \varphi_u, d_u^0) = \left((m_z \Delta_R)^2 + (m_y \Delta_R)^2 - J_m^2(\omega_u, \varphi_u) \right) / 2d_u^0$.

Then, the authors carried out the NF-JCEL algorithm by estimating the AoAs $\omega_u = \sin \phi_{u,0}$, $\varphi_u = \sin \theta_{u,0} \cos \phi_{u,0}$, the distance d_u^0 , and the cascaded channel attenuation $g_{u,i}$, and finally obtained the position of UE u by utilizing the geometric relationship. Furthermore, because the steering vectors were frequency-dependent, the authors processed the channel estimation on each sub-band.

1) RIS training phase shifts and pilot design

The authors utilized different RIS phase shift vectors to obtain a unique estimation of the channel matrix. The pseudo-random (PN) sequences were utilized as the transmitting pilot at different time slots, which are statistically orthogonal with each other, and the RIS phase shift vector changed for S times in a pilot data duration. Then the composited signal vector which collected the received signals in the S different phase shifts could be obtained, and we had a least square (LS) estimator of $\mathbf{A}\mathbf{x}_t$ as $\overline{\mathbf{A}\mathbf{x}_t}$, where \mathbf{A} is the channel matrix to be estimated, and \mathbf{x}_t is the pilot sequence in the t -th time slot.

Note that the rank of the cascaded channel moves up by employing diverse phase shifts of the RIS to avoid noise amplification when utilizing the LS estimation. Since the channel coefficients are unknown, the existing RIS phase designs such as the random phase shift and the Discrete Fourier Transform (DFT) matrix are considered. By collecting the received signal in the τ pilot durations, the covariance matrix of the received signal is estimated as the mean covariance matrix of $\overline{\mathbf{A}\mathbf{x}_t}$.

2) Estimation of AoAs

The authors provided a novel method to separate the distance and AoAs in the covariance matrix. With complicated mathematical manipulations, they constructed a down-sampled Toeplitz matrix \mathbf{T} to decouple the distance and the AoAs. Therefore, the far-field angle estimation methods could be obtained with the sampled correlation matrix.

Specifically, the authors applied a computationally efficient subspace-based method to estimate the angles. By leveraging the key idea of the MUSIC algorithm, the AoAs could be estimated by minimizing the following cost function:

$$f(\omega, \varphi) = \mathbf{b}^H(\omega, \varphi) \mathbf{I}_m \mathbf{b}(\omega, \varphi), \quad (24)$$

where $\mathbf{b}(\omega, \varphi) = \mathbf{v}(\omega) \otimes \mathbf{s}(\varphi)$, $\mathbf{v}(\omega)_{[z]} = \exp\left(-j \frac{4\pi}{\lambda} z \Delta_R \omega\right)$, $\mathbf{s}(\varphi)_{[y]} = \exp\left(-j \frac{4\pi}{\lambda} y \Delta_R \varphi\right)$, and \mathbf{I}_m is designed by utilizing the property of the sampled correlation matrix \mathbf{T} .

By applying the Lagrangian multiplier method, the relationship between each optimal ω_u and φ_u is obtained. Then, substitute the equation into the objective function and the op-

timal value of ω can be reached by solving a one-dimensional problem. Since $\omega_u = \sin \phi_u$, we can search in the interval of $\omega_u \in [-1, 1]$ and obtain the U largest peaks as the estimator of ω_u . Finally, by utilizing the relationship between ω_u and φ_u , we have U vectors $\hat{\mathbf{s}}(\varphi_u)$ corresponding to $\hat{\omega}_u$, and the estimator of φ_u is obtained by applying the LS estimation.

3) Estimation of distances, channel gains and locations

Similar to the method of tackling ω and φ , the authors transform the estimation of distance into the following problem.

$$d_u^{0*} = \arg \min_{d_u^0} \left(\left(\text{diag}\{\mathbf{p}_u\} \mathbf{q}_u \right)^H \mathbf{I}_n \left(\text{diag}\{\mathbf{p}_u\} \mathbf{q}_u \right) \right), \quad (25)$$

where $\mathbf{p}_u = \left[\exp\left(-j \frac{2\pi}{\lambda} J_{-N_0}\right), \dots, \exp\left(-j \frac{2\pi}{\lambda} J_{N_0}\right) \right]^T$, $\mathbf{q}_u = \left[\exp\left(-j \frac{\pi Q_{-N_0}}{\lambda d_u^0}\right), \dots, \exp\left(-j \frac{\pi Q_{N_0}}{\lambda d_u^0}\right) \right]$, and \mathbf{I}_n is designed in the similar way as \mathbf{I}_m . By conducting one-dimensional search, the U largest peaks of the searching results are obtained as the estimated distances.

Finally, the LS method can be applied to estimate the channel gain of sub-band i , and the localization of UE u can be obtained according to the geometric relationship with the estimation over different frequency bands.

4) Simulation results

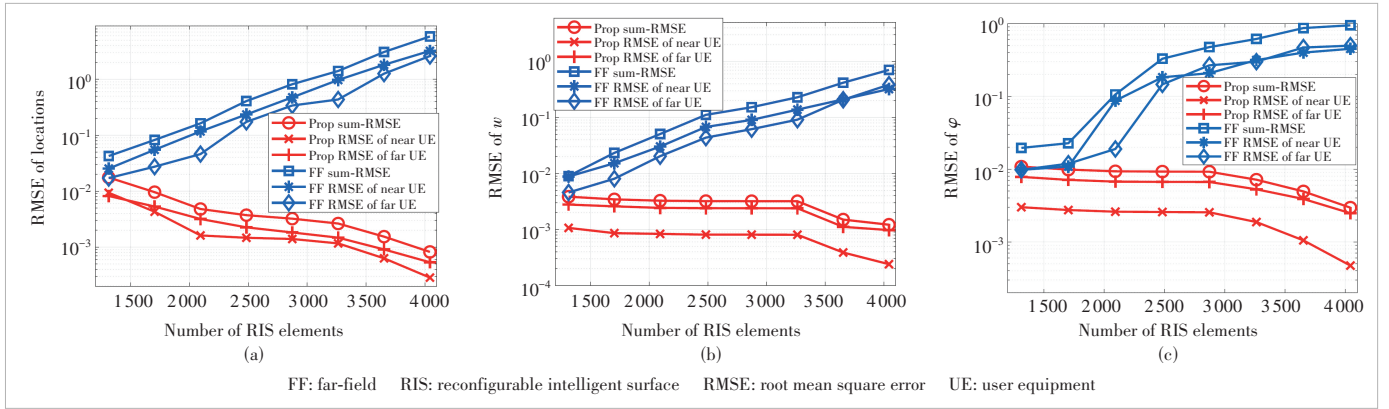
Simulation results validate the advantages of the NF-JCEL algorithm over its far-field counterpart in the near-field scenario. As shown in Fig. 8(a), the NF-JCEL algorithm outperforms the corresponding far-field cases in the root mean square error (RMSE) of UE locations. In addition, along with the increase of elements, the NF-JCEL algorithm obtains better performance while the conventional far-field model suffers severe degradation. When it comes to the RMSE performance of the estimated channel coefficients, similar trends in the sum RMSE can be observed in Figs. 8 (b) and 8(c).

This phenomenon is owing to the fact that more elements can provide more angular information, thus leading to higher angular resolution when estimating the AoAs. In contrast, the conventional far-field model neglects the distinction of AoAs at the RIS, so that the approximation error of the planar wavefront becomes more severe though the correlation matrices of larger dimensions are exploited.

3.3 Transmission Scheme Design for Near-Field Communications

3.3.1 Challenges

As the antenna array size increases, the Rayleigh distance significantly extends from a few meters to tens or hundreds of meters^[17]. Consequently, the plane wave assumption in the far field becomes invalid. Based on the spherical wave character-



▲ Figure 8. Element number versus RMSE of estimated parameters

istics, distance information needs to be considered in the near-field channel modeling. Furthermore, with the emergence of large-scale antenna arrays, the dimension of the channel becomes large and the computational complexity for transmission scheme design also greatly increases, which becomes a key design challenge in near-field scenarios. Moreover, the unique spatial non-stationary characteristics in the near field make the performance characteristics and algorithm design different from those in the far field with conventional stationary channels. This necessitates the exploration of more accurate near-field channel models and raises a new challenge: how to design new transmission schemes based on the characteristics of the near-field channel.

3.3.2 Existing Works

The near-field channel introduces several new channel characteristics: the spherical wave property of the propagating signal and the spatial non-stationary property. Recently, some researchers have confirmed the existence of spatial non-stationary properties in near-field channels^[26, 57], i.e., the user's transmitted signal may only be received by a small part of the array. Based on this characteristic, the authors of Ref. [26] proposed a channel model based on VRs. To reduce the computational complexity and by exploiting the nature of VRs, the authors of Ref. [58] proposed a graph-based linear receiving algorithm inspired by random access. Moreover, an antenna selection algorithm was proposed in Ref. [59] to improve the total energy efficiency of the system. However, the above work assumes that the VR information is known and the arrays other than the VR receive signal amplitude with values of zero, which is impractical. To fill these gaps, the authors of Ref. [60] first derived a near-field channel model considering antenna polarization based on the EM field theory. Then, a detection algorithm for VRs was proposed, which can obtain VR information of different users. Finally, the authors proposed a linear detection algorithm based on VRs, exploiting VRs information for different users to reduce the computational complexity. In the following part, we present details of the research work in Ref. [60].

3.3.3 Transmission Scheme Design for Near-Field with Spatial Non-Stationaries

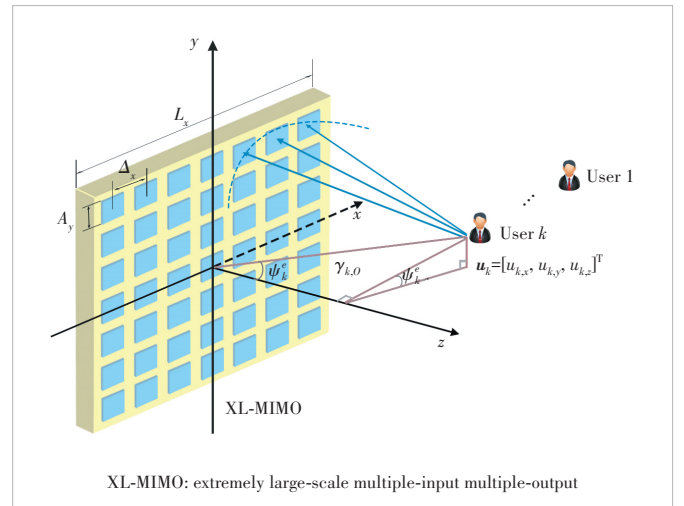
The authors of Ref. [60] considered an uplink transmission system with multiple single-antenna users and an XL-MIMO base station, as shown in Fig. 9.

The channel between the (m_x, m_y) -th antenna element of the base station and user k can be expressed as

$$h_{k,m_x,m_y} = \sqrt{\xi_{k,m_x,m_y}} e^{-j\chi_{k,m_x,m_y}}, \quad (26)$$

where ξ_{k,m_x,m_y} and χ_{k,m_x,m_y} denote the channel coefficient and phase, respectively.

Then, the authors derive a near-field channel model that takes polarization mismatch into account by using the inhomogeneous Helmholtz wave equation and Green's function. We first define the Δ_x and Δ_y as the antenna spacing along the x -axis and y -axis, respectively. After that, by assuming the location of user k and the (m_x, m_y) -th antenna element as $\mathbf{u}_k = [u_{k,x}, u_{k,y}, u_{k,z}]^T$ and $\mathbf{p}_{m_x,m_y} = [m\Delta_x, m\Delta_y, 0]^T$, respectively, the channel coefficient and phase are given by:



▲ Figure 9. System model in Ref. [60]

$$\xi_{k,m_x,m_y} \approx \frac{A}{4\pi} \frac{u_{k,z} \left((m_x \Delta_x - u_{k,x})^2 + u_{k,z}^2 \right)}{\left\{ (m_x \Delta_x - u_{k,x})^2 + (m_y \Delta_y - u_{k,y})^2 + u_{k,z}^2 \right\}^{\frac{5}{2}}}, \quad (27)$$

$$\chi_{k,m_x,m_y} \approx \frac{2\pi}{\lambda} \left\| \mathbf{p}_{m_x,m_y} - \mathbf{u}_k \right\|. \quad (28)$$

The authors derived an explicit expression for single-user SNR based on the maximum ratio combining (MRC) detector. As shown in Fig. 10(a), the SNR gradually increases to saturation with the increase of antenna numbers. Furthermore, the near-field channel model considering polarization mismatch satisfies the introduced theoretical upper bound.

Next, based on the proposed near-field channel model, the authors proposed a VR-based low-complexity transmission scheme by exploiting the spatial non-stationarity of XL-MIMO. The dimensionality of the channel matrix in the transmission design can be greatly reduced by utilizing this method, thus reducing the computational complexity accordingly. Furthermore, the authors proposed a graph-theory based user grouping algorithm to categorize users with high VRs overlap ratios into one group. The authors then utilized the partial zero-forcing (PZF) detection algorithm to eliminate the mutual interference within each user group, which could further reduce the computational complexity.

As shown in Fig. 10(b), as the number of antenna elements increases, the computational complexity of the whole array (WA)-based design rises considerably. For the VR-based zero-forcing (ZF) algorithm and the graph-theory based PZF algorithm, low computational complexity can be achieved.

4 Future Directions

In this section, several promising directions of near-field communications are discussed according to the contributions and deficiencies of the aforementioned appealing works.

4.1 AI-Aided Near-Field Beam Training

The AI-based near-field beam training method proposed in Ref. [40] effectively reduces the pilot overhead, revealing the nonlinear signal processing capacity of AI. Therefore, the AI-based method is perceived as a promising solution to the complex challenges in near-field communications. Specifically, it is reasonable to employ AI to deal with the massive parameters in near-field problems, such as beamforming and channel estimation.

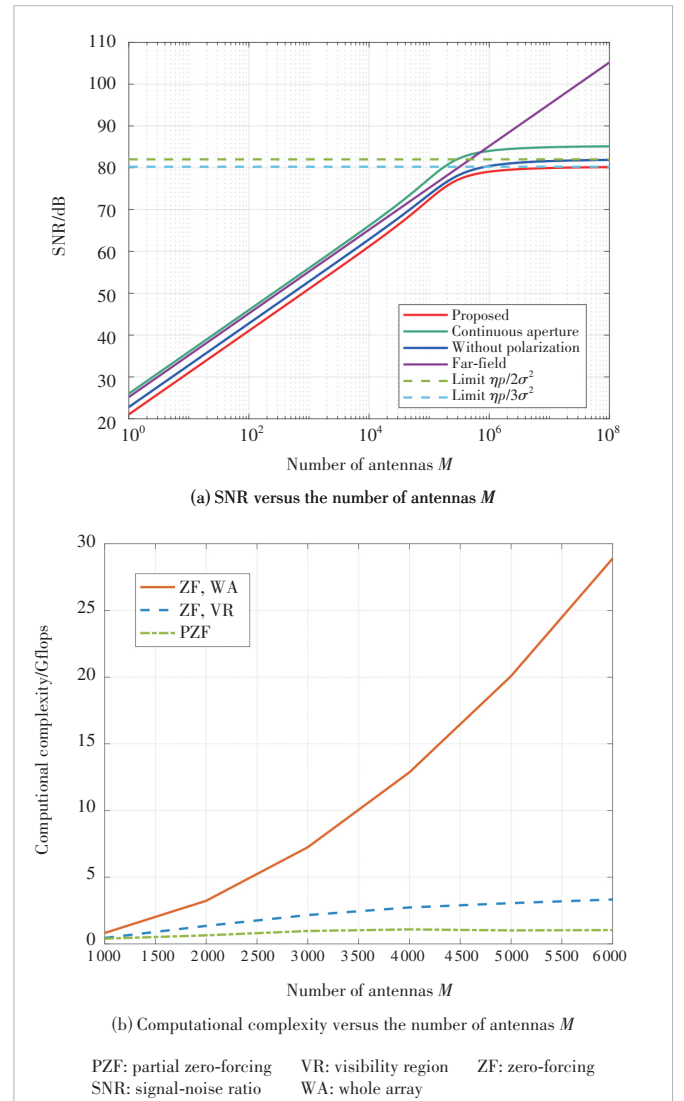
In addition, most research mainly focuses on reducing the pilot overhead in the design of near-field beam training, while the effectiveness in complex environments is not fully explored. Hence, it is crucial to study the trade-off between low pilot overhead and high accuracy and propose an effective

near-field beam training method that can be applied in complicated propagation environment with multiple mobile users and scatters.

4.2 XL-RIS-Aided Near-Field Localization

As derived in Ref. [56], the proposed algorithm with XL-RIS can achieve higher positioning resolution with proper design to utilize the richer AoA information. It is also possible for a single-RIS-aided system to provide a 3D localization function by exploiting the near-field channel model. Therefore, near-field XL-RIS-aided localization can be considered as a promising innovation, and it is valuable to carry out more studies in terms of theoretical error distribution analysis as well as practical positioning algorithms.

Although Ref. [56] has achieved great performance in XL-RIS-aided joint channel estimation and localization in a near-



▲ Figure 10. Effects of the number of antennas on system performance and computational complexity

field scenario, the computational complexity is relatively high due to the massive elements. Therefore, it is valuable to design a near-field localization algorithm with low complexity and pilot overhead.

4.3 Near-Field Behavior of XL-Array

As one of the main innovations of 6G wireless communications, the successful application of XL-MIMO depends on the investigation and understanding of the near-field propagation. In addition, it is valuable to study the physically accurate near-field propagation model of XL-arrays.

Specifically, the near-field EM channel model proposed in Ref. [60] accurately depicts the physical EM wave transmission in the near-field scenario, which can be utilized to analyze the impact of discrete array aperture and polarization mismatch. Nevertheless, the physical channel model in Ref. [60] is totally based on theoretical analysis with Maxwell's EM field theory, so it is also meaningful to carry out channel measurements to verify and modify the near-field EM channel model.

Moreover, the authors in Ref. [58 – 60] utilized the information of VRs to remove the elements of the antenna array that contributed less to the signal transmission, and constructed equivalent dimensionality reduction channels to reduce the computational complexity. However, how to accurately measure VR using channel estimation methods is still an open question.

5 Conclusions

The far-field assumptions tend to be invalid owing to the large array size and short wavelength in 6G wireless communications. Hence, the near-field effect and spherical wavefront are no longer dispensable in some envisioned 6G scenarios. In this paper, we discussed the challenges and opportunities of near-field communication in terms of the fundamentals and applications in various schemes. We first introduced the three regions of free-space propagation and the corresponding wavefronts, with an emphasis on the radiating near-field and spherical wavefront. Then, we derived the near-field MIMO channel model and discussed the impact of spatial non-stationarity property on near-field propagation. In addition, we introduced several appealing works with respect to beam training and localization in the near-field scenarios and a recent work exploiting the spatial non-stationarity in the near field to achieve low-complexity transmission design. Finally, some insights into the future improvements of near-field communication were discussed.

References

- [1] CHETTRI L, BERA R. A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems [J]. *IEEE Internet of Things journal*, 2020, 7(1): 16 – 32. DOI: 10.1109/JIOT.2019.2948888

- [2] GARCIA M H C, MOLINA-GALAN A, BOBAN M, et al. A tutorial on 5G NR V2X communications [J]. *IEEE communications surveys & tutorials*, 2021, 23(3): 1972 – 2026. DOI: 10.1109/COMST.2021.3057017
- [3] LI S C, LI D X, ZHAO S S. 5G Internet of Things: a survey [J]. *Journal of industrial information integration*, 2018, 10: 1 – 9. DOI: 10.1016/j.jii.2018.01.005
- [4] SHAFI M, MOLISCH A F, SMITH P J, et al. 5G: A tutorial overview of standards, trials, challenges, deployment, and practice [J]. *IEEE journal on selected areas in communications*, 2017, 35(6): 1201 – 1221. DOI: 10.1109/JSAC.2017.2692307
- [5] SHAFIQUE K, KHAWAJA B A, SABIR F, et al. Internet of Things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios [J]. *IEEE access*, 2020, 8: 23022 – 23040. DOI: 10.1109/ACCESS.2020.2970118
- [6] WANG C X, HAIDER F, GAO X Q, et al. Cellular architecture and key technologies for 5G wireless communication networks [J]. *IEEE communications magazine*, 2014, 52(2): 122 – 130. DOI: 10.1109/MCOM.2014.6736752
- [7] AGI WAL M, SAXENA N, ROY A. Towards connected living: 5G enabled Internet of Things (IoT) [J]. *IETE technical review*, 2019, 36(2): 190 – 202. DOI: 10.1080/02564602.2018.1444516
- [8] YOU X H, WANG C X, HUANG J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts [J]. *Science China information sciences*, 2020, 64(1): 110301. DOI: 10.1007/s11432-020-2955-6
- [9] ZHANG Z Q, XIAO Y, MA Z, et al. 6G wireless networks: vision, requirements, architecture, and key technologies [J]. *IEEE vehicular technology magazine*, 2019, 14(3): 28 – 41. DOI: 10.1109/MVT.2019.2921208
- [10] YANG P, XIAO Y, XIAO M, et al. 6G wireless communications: vision and potential techniques [J]. *IEEE network*, 2019, 33(4): 70 – 75. DOI: 10.1109/MNET.2019.1800418
- [11] ZHANG L, LIANG Y C, NIYATO D. 6G Visions: Mobile ultra-broadband, super internet-of-things, and artificial intelligence [J]. *China communications*, 2019, 16(8): 1 – 14. DOI: 10.23919/JCC.2019.08.001
- [12] LU H Q, ZENG Y, YOU C S, et al. A tutorial on near-field XL-MIMO communications towards 6G [EB/OL]. (2023-10-17) [2023-11-10]. <http://arxiv.org/abs/2310.11044>
- [13] PIZZO A, SANGUINETTI L, MARZETTA T L. Fourier plane-wave series expansion for holographic MIMO communications [J]. *IEEE transactions on wireless communications*, 2022, 21(9): 6890 – 6905. DOI: 10.1109/TWC.2022.3152965
- [14] CUI M Y, WU Z D, LU Y, et al. Near-field MIMO communications for 6G: Fundamentals, challenges, potentials, and future directions [J]. *IEEE communications magazine*, 2023, 61(1): 40 – 46. DOI: 10.1109/MCOM.004.2200136
- [15] TATARIA H, SHAFI M, MOLISCH A F, et al. 6G wireless systems: Vision, requirements, challenges, insights, and opportunities [J]. *Proceedings of the IEEE*, 2021, 109(7): 1166 – 1199. DOI: 10.1109/JPROC.2021.3061701
- [16] AN J C, YUEN C, DAI L L, et al. Toward beam focusing-aided near-field communications: research advances, potential, and challenges [EB/OL]. (2023-09-17) [2023-11-10]. <http://arxiv.org/abs/2309.09242>
- [17] LIU Y W, WANG Z L, XU J Q, et al. Near-field communications: a tutorial review [EB/OL]. [2023-11-10]. <http://arxiv.org/abs/2305.17751.pdf>
- [18] ZHANG H Y, SHLEZINGER N, GUIDI F, et al. 6G wireless communications: from far-field beam steering to near-field beam focusing [J]. *IEEE communications magazine*, 2023, 61(4): 72 – 77. DOI: 10.1109/MCOM.001.2200259
- [19] SHERMAN J. Properties of focused apertures in the Fresnel region [J]. *IRE transactions on antennas and propagation*, 1962, 10(4): 399 – 408.

- DOI: 10.1109/TAP.1962.1137900
- [20] SELVAN K T, JANASWAMY R. Fraunhofer and Fresnel Distances: unified derivation for aperture antennas [J]. *IEEE antennas and propagation magazine*, 2017, 59(4): 12 – 15. DOI: 10.1109/MAP.2017.2706648
- [21] KRAUS J D, MARHEFKA R J. *Antennas for all applications* [M]. New York, USA: McGraw-Hill School, 2001
- [22] LU H Q, ZENG Y. Communicating with extremely large-scale array/surface: unified modeling and performance analysis [J]. *IEEE transactions on wireless communications*, 2022, 21(6): 4039 – 4053. DOI: 10.1109/TWC.2021.3126384
- [23] HEATH R W, GONZÁLEZ-PRELCIC N, RANGAN S, et al. An overview of signal processing techniques for millimeter wave MIMO systems [J]. *IEEE journal of selected topics in signal processing*, 2016, 10(3): 436 – 453. DOI: 10.1109/JSTSP.2016.2523924
- [24] CUI M Y, DAI L L. Channel estimation for extremely large-scale MIMO: far-field or near-field? [J]. *IEEE transactions on communications*, 2022, 70(4): 2663 – 2677. DOI: 10.1109/TCOMM.2022.3146400
- [25] MARTÍNEZ À O, DE CARVALHO E, NIELSEN J Ø. Towards very large aperture massive MIMO: a measurement based study [C]/*IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2014: 281 – 286. DOI: 10.1109/GLOCOMW.2014.7063445
- [26] YUAN Z Q, ZHANG J H, JI Y L, et al. Spatial non-stationary near-field channel modeling and validation for massive MIMO systems [J]. *IEEE transactions on antennas and propagation*, 2023, 71(1): 921 – 933. DOI: 10.1109/TAP.2022.3218759
- [27] HAN Y, JIN S, WEN C K, et al. Channel estimation for extremely large-scale massive MIMO systems [J]. *IEEE wireless communications letters*, 2020, 9(5): 633 – 637. DOI: 10.1109/LWC.2019.2963877
- [28] TIAN J C, HAN Y, JIN S, et al. Low-overhead localization and VR identification for subarray-based ELAA systems [J]. *IEEE wireless communications letters*, 2023, 12(5): 784 – 788. DOI: 10.1109/LWC.2023.3244000
- [29] LÓPEZ C F, WANG C X. Novel 3-D non-stationary wideband models for massive MIMO channels [J]. *IEEE transactions on wireless communications*, 2018, 17(5): 2893 – 2905. DOI: 10.1109/TWC.2018.2804385
- [30] WANG J, WANG C X, HUANG J, et al. A general 3D space-time-frequency non-stationary THz channel model for 6G ultra-massive MIMO wireless communication systems [J]. *IEEE journal on selected areas in communications*, 2021, 39(6): 1576 – 1589. DOI: 10.1109/JSAC.2021.3071850
- [31] BAI T Y, HEATH R W. Analysis of beam sweep channel estimation in MmWave massive MIMO networks [C]/*IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2016: 615 – 619. DOI: 10.1109/GlobalSIP.2016.7905915
- [32] CHEN K J, QI C H, LI G Y. Two-step codeword design for millimeter wave massive MIMO systems with quantized phase shifters [J]. *IEEE transactions on signal processing*, 2021, 68: 170 – 180. DOI: 10.1109/TSP.2019.2959250
- [33] MA K, HE D X, SUN H C, et al. Deep learning assisted mmWave beam prediction with prior low-frequency information [C]/*IEEE International Conference on Communications*. IEEE, 2021: 1 – 6. DOI: 10.1109/ICC42927.2021.9500788
- [34] QI C H, WANG Y J, LI G Y. Deep learning for beam training in millimeter wave massive MIMO systems [J]. *IEEE transactions on wireless communications*, 2020. DOI: 10.1109/TWC.2020.3024279
- [35] QI C H, DONG P H, MA W Y, et al. Acquisition of channel state information for mmWave massive MIMO: traditional and machine learning-based approaches [J]. *Science China information sciences*, 2021, 64(8): 181301. DOI: 10.1007/s11432-021-3247-2
- [36] WEI X H, DAI L L, ZHAO Y J, et al. Codebook design and beam training for extremely large-scale RIS: far-field or near-field? [J]. *China communications*, 2022, 19(6): 193 – 204. DOI: 10.23919/JCC.2022.06.015
- [37] ZHANG Y P, WU X, YOU C S. Fast near-field beam training for extremely large-scale array [J]. *IEEE wireless communications letters*, 2022, 11(12): 2625 – 2629. DOI: 10.1109/LWC.2022.3212344
- [38] REZAIE S, MANCHÓN C N, DE CARVALHO E. Location- and orientation-aided millimeter wave beam selection using deep learning [C]/*IEEE International Conference on Communications (ICC)*. IEEE, 2020: 1 – 6. DOI: 10.1109/ICC40277.2020.9149272
- [39] JIANG G L, QI C H. Near-field beam training based on deep learning for extremely large-scale MIMO [J]. *IEEE communications letters*, 2023, 27(8): 2063 – 2067. DOI: 10.1109/LCOMM.2023.3289513
- [40] LIU W, REN H, PAN C H, et al. Deep learning based beam training for extremely large-scale massive MIMO in near-field domain [J]. *IEEE communications letters*, 2023, 27(1): 170 – 174. DOI: 10.1109/LCOMM.2022.3210042
- [41] LIU W, PAN C H, REN H, et al. Low-overhead beam training scheme for extremely large-scale RIS in near field [J]. *IEEE transactions on communications*, 2023, 71(8): 4924 – 4940. DOI: 10.1109/TCOMM.2023.3278728
- [42] FAN D, GAO F F, LIU Y W, et al. Angle domain channel estimation in hybrid millimeter wave massive MIMO systems [J]. *IEEE transactions on wireless communications*, 2018, 17(12): 8165 – 8179. DOI: 10.1109/TWC.2018.2874640
- [43] ZHOU G, PAN C H, REN H, et al. Channel estimation for RIS-aided multi-user millimeter-wave systems [J]. *IEEE transactions on signal processing*, 2021, 70: 1478 – 1492. DOI: 10.1109/TSP.2022.3158024
- [44] HE J G, WYMEERSCH H, KONG L, et al. Large intelligent surface for positioning in millimeter wave MIMO systems [C]/*The 91st Vehicular Technology Conference (VTC2020-Spring)*. IEEE, 2020: 1 – 5. DOI: 10.1109/VTC2020-Spring48590.2020.9129075
- [45] LIU Y, HONG S, PAN C H, et al. Optimization of RIS configurations for multiple-RIS-aided mmWave positioning systems based on CRLB analysis [EB/OL]. (2021-11-28) [2023-12-12]. <http://arxiv.org/abs/2111.14023>
- [46] WANG R, XING Z, LIU E W. Joint location and communication study for intelligent reflecting surface aided wireless communication system [EB/OL]. (2021-05-01) [2023-12-12]. <http://arxiv.org/abs/2103.01063>
- [47] FENG Z Y, WANG B, ZHAO Y P, et al. Power optimization for target localization with reconfigurable intelligent surfaces [J]. *Signal processing*, 2021, 189: 108252. DOI: 10.1016/j.sigpro.2021.108252
- [48] ELZANATY A, GUERRA A, GUIDI F, et al. Reconfigurable intelligent surfaces for localization: position and orientation error bounds [J]. *IEEE transactions on signal processing*, 2021, 69: 5386 – 5402. DOI: 10.1109/TSP.2021.3101644
- [49] ZHANG H B, ZHANG H L, DI B Y, et al. Towards ubiquitous positioning by leveraging reconfigurable intelligent surface [J]. *IEEE communications letters*, 2021, 25(1): 284 – 288. DOI: 10.1109/LCOMM.2020.3023130
- [50] HE J G, WYMEERSCH H, SANGUANPUAK T, et al. Adaptive beam-forming design for mmWave RIS-aided joint localization and communication [C]/*IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 2020: 1 – 6. DOI: 10.1109/WCNCW48565.2020.9124848
- [51] FASCISTA A, COLUCCIA A, WYMEERSCH H, et al. RIS-aided joint localization and synchronization with a single-antenna mmwave receiver

- [C]/IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021: 4455 - 4459. DOI: 10.1109/ICASSP39728.2021.9413515
- [52] WU T, PAN C H, PAN Y J, et al. Fingerprint-based mmWave positioning system aided by reconfigurable intelligent surface [J]. IEEE wireless communications letters, 2023, 12(8): 1379 - 1383. DOI: 10.1109/LWC.2023.3275204
- [53] CHEN J C, HUDSON R E, YAO K. Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field [J]. IEEE transactions on signal processing, 2002, 50(8): 1843 - 1854. DOI: 10.1109/TSP.2002.800420
- [54] ABU-SHABAN Z, KEYKHOSRAVI K, KESKIN M F, et al. Near-field localization with a reconfigurable intelligent surface acting as lens [C]/IEEE International Conference on Communications. IEEE, 2021: 1 - 6. DOI: 10.1109/ICC42927.2021.9500663
- [55] RAHAL M, DENIS B, KEYKHOSRAVI K, et al. RIS-enabled localization continuity under near-field conditions [C]/The 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2021: 436 - 440. DOI: 10.1109/SPAWC51858.2021.9593200
- [56] PAN Y J, PAN C H, JIN S, et al. RIS-aided near-field localization and channel estimation for the terahertz system [J]. IEEE journal of selected topics in signal processing, 2023, 17(4): 878 - 892. DOI: 10.1109/JSTSP.2023.3285431
- [57] PAYAMI S, TUFVESSON F. Channel measurements and analysis for very large array systems at 2.6 GHz [C]/The 6th European Conference on Antennas and Propagation (EUCAP). IEEE, 2012: 433 - 437. DOI: 10.1109/EuCAP.2012.6206345
- [58] AMIRI A, ANGJELICHINOSKI M, DE CARVALHO E, et al. Extremely large aperture massive MIMO: low complexity receiver architectures [C]/IEEE Globecom Workshops (GC Wkshps). IEEE, 2018: 1 - 6. DOI: 10.1109/GLOCOMW.2018.8644126
- [59] MARINELLO J C, ABRÃO T, AMIRI A, et al. Antenna selection for improving energy efficiency in XL-MIMO systems [J]. IEEE transactions on vehicular technology, 2020, 69(11): 13305 - 13318. DOI: 10.1109/TVT.2020.3022708
- [60] ZHI K D, PAN C H, REN H, et al. Performance analysis and low-complexity design for XL-MIMO with near-field spatial non-stationarities [EB/OL]. (2023-03-21) [2023-12-12]. <http://arxiv.org/abs/2304.00172>

Biographies

LIU Mengyu received his BE degree from the School of Electronic and Information Engineering, Beijing Jiaotong University, China in 2023. He is currently pursuing his ME degree with the School of Information Science and Engineering, Southeast University, China. His research interests include massive MIMO, reconfigurable intelligent surfaces (RIS), and near-field communications.

ZHANG Yang received his BE degree from the Harbin Institute of Technology, China in 2023. He is currently pursuing his ME degree with the School of Information Science and Engineering, Southeast University, China. His research interests include reconfigurable intelligent surfaces (RIS), array signal processing, near-field communications, and localization.

JIN Yasheng received his BE degree from the Bell Honors School, Nanjing University of Posts and Telecommunications, China in 2023. He is currently pursuing his ME degree with the School of Information Science and Engineering, Southeast University, China. His research interests include massive MIMO, reconfigurable intelligent surfaces (RIS), machine learning and near-field communications.

ZHI Kangda received his BE degree from the School of Communication and Information Engineering, Shanghai University, China in 2017, ME degree from School of Information Science and Technology, University of Science and Technology of China in 2020, and PhD degree from the School of Electronic Engineering and Computer Science, Queen Mary University of London, UK in 2023. His research interests include reconfigurable intelligent surface (RIS), massive MIMO, and near-field communications. He received the Exemplary Reviewer Certificate of the IEEE Wireless Communications Letters in 2021 and 2022.

PAN Cunhua (cpan@seu.edu.cn) received his BS and PhD degrees from the School of Information Science and Engineering, Southeast University, China in 2010 and 2015, respectively. From 2015 to 2016, he was a research associate at the University of Kent, UK. He held a post-doctoral researcher position at the Queen Mary University of London, UK from 2016 to 2019, where he was a lecturer from 2019 to 2021. Since 2021, he has been a full professor with Southeast University. His research interests include reconfigurable intelligent surfaces (RIS), intelligent reflection surface (IRS), ultra-reliable low latency communication (URLLC), machine learning, UAV, the Internet of Things, and mobile edge computing. He has published over 120 IEEE journal articles. He received the IEEE ComSoc Leonard G. Abraham Prize and IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award both in 2022.

Link Budget and Enhanced Communication Distance for Ambient Internet of Things



YANG Yibing¹, LIU Ming¹, XU Rongtao²,

WANG Gongpu¹, GONG Wei³

(1. School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China;

2. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;

3. School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026, China)

DOI: 10.12142/ZTECOM.202401003

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240223.1839.002.html>,
published online February 26, 2024

Manuscript received: 2023-12-02

Abstract: Backscatter communications will play an important role in connecting everything for beyond 5G (B5G) and 6G systems. One open challenge for backscatter communications is that the signals suffer a round-trip path loss so that the communication distance is short. In this paper, we first calculate the communication distance upper bounds for both uplink and downlink by measuring the tag sensitivity and reflection coefficient. It is found that the activation voltage of the envelope detection diode of the downlink tag is the main factor limiting the backscatter communication distance. Based on this analysis, we then propose to implement a low-noise amplifier (LNA) module before the envelope detection at the tag to enhance the incident signal strength. Our experimental results on the hardware platform show that our method can increase the downlink communication range by nearly 20 m.

Keywords: ambient IoT (AIoT); B5G; backscatter communication; link budget; low-noise amplifier (LNA); Release 19; tag chip sensitivity; upper bounds

Citation (Format 1): YANG Y B, LIU M, XU R T, et al. Link budget and enhanced communication distance for ambient Internet of Things [J]. *ZTE Communications*, 2024, 22(1): 16 - 23. DOI: 10.12142/ZTECOM.202401003

Citation (Format 2): Y. B. Yang, M. Liu, R. T. Xu, et al., "Link budget and enhanced communication distance for ambient Internet of Things," *ZTE Communications*, vol. 22, no. 1, pp. 16 - 23, Mar. 2024. doi: 10.12142/ZTECOM.202401003.

1 Introduction

Backscatter communications (BackCom), a key technology for ambient Internet of Things (AIoT), have the advantages of low hardware cost, low maintenance, easy deployment and flexible extension, and will play an important role in connecting everything for B5G and 6G systems. It was first introduced in the late 1940s by STOCKMAN^[1]. Afterward, the radio frequency identification (RFID)^[2] system enabled by BackCom has been widely used for commodity identification, item tracking, logistic management, etc. Recently, the third-generation partnership project (3GPP) started investigating the scenarios, use cases, services and targets of the potential use of BackCom in future wireless networks, and the corresponding network is referred to as ambient IoT^[3] in Release 19.

Different from traditional wireless communications, the

backscatter transmitters (e.g., tags, devices, and sensors) do not need to generate the carrier signal by itself. BackCom transmits information by reflecting the received RF signal in a modulated way. So, it leaves out the RF units of traditional communication equipment such as a local oscillator, a mixer, and a duplexer, which greatly simplifies the RF frontend and reduces the power consumption dedicated to RF signal transmission. As a consequence, backscatter devices do not need batteries but can accomplish wireless signal reception and transmission by wireless energy harvesting (EH) and BackCom, which significantly improves its sustainability^[4-6]. Therefore, BackCom has been regarded as a promising solution to passive and green IoT.

One open challenge for BackCom is that the signals suffer a round-trip path loss so that the communication distance is short, which is largely different from traditional point-to-point communications. Specifically, the signals in traditional point-to-point communication systems experience a one-way path loss, and both the transmitter and receiver are equipped with power amplifiers to enhance signal quality. As a result, it is relatively easier to achieve long-range communication. In con-

This work was supported in part by National Natural Science Foundation of China under Grant Nos. 61971029 and U22B2004, and in part by Beijing Municipal Natural Science Foundation under Grant No. L222002.

trast, in BackCom systems, the information is delivered through RF signal reflection. Therefore, the backscatter signals undergo a round-trip path loss, which requires a larger link budget. For example, the signal coverage of the ultra-high frequency (UHF) radio frequency identification (RFID) system, namely Electronic Product Code Global Class1 Generation 2 (EPC C1G2)^[7], is at the level of roughly 10 m in practical deployment, in comparison with coverage up to several hundred meters for typical cellular communication systems at the same frequency band with similar radiation power level. The communication range of BackCom could be further reduced if higher communication rates are required.

The short coverage inevitably limits the application of BackCom. Therefore, it is necessary to identify the major factors that restrict the communication range of BackCom systems and find corresponding effective solutions to boost its transmission capability. The major factors that constrain the communication distance may come from uplink or downlink. Accordingly, it is necessary to perform the link budget to calculate and compare the upper bound downlink and uplink communication distance.

Existing works have studied the link budget of BackCom mostly through theoretical analysis. Friis free-space formula was used to determine the theoretical upper bound communication distance in Refs. [8 – 9]. The authors in Ref. [10] incorporated more practical factors such as multipath propagation to derive more practical path losses of passive UHF tags. Four types of propagation models of BackCom were discussed in Ref. [11] and some insights into the link quality of BackCom transmission were provided.

Using the link budget, we can determine the power received at the tag and the reader. When the received power decreases to the sensitivity of the tag and the reader, we can calculate the upper bounds for the downlink and uplink communication distance. Next, comparing the two bounds, we can find the bottleneck of the BackCom, uplink or downlink. And the result is that BackCom is mainly limited by the downlink. Therefore, we need to improve the downlink sensitivity. That is, enhancing the tag sensitivity can increase the communication range of BackCom systems.

Generally speaking, the overall sensitivity of the tag is determined by the envelope detection circuit and the comparator. The sensitivity of the envelope detection circuit is related to the forward voltage drop of the diode. In order to improve the sensitivity, the authors in Ref. [12] suggested that the envelope detection circuit used diode-connected low positive channel Metal Oxide Semiconductor (PMOS) transistors for low-voltage operation and less energy loss. Ref. [13] introduced a detector-first wake-up receiver (WuRx) that used boundary condition-based envelope detector optimization, improving sensitivity and performance while conserving power. The envelope detector described in Ref. [14] employed the threshold compensation technology to effectively mitigate the

impact of the transistor's threshold voltage on demodulation sensitivity. A voltage doubler rectifier was proposed to improve the passive envelope detection of the tag, thereby enhancing the tag sensitivity^[15–16]. Unfortunately, when a multiple-stage voltage doubler rectifier is used, it is a challenge to design and implement the RF circuit. Besides, using a multiple-stage voltage doubler rectifier will cause annoying parasitic parameters that may lead to a relatively slower rectification rate.

Actually, the factors that can affect the tag sensitivity may be the forward voltage drop of the diode, the sensitivity of the comparator, and the algorithm limit imposed by the adopted decoding scheme. In summary, the bottleneck that restricts the communication range remains unclear and the results have not been validated over practical hardware implementations, which motivates our present work.

In this paper, we first derive the link budget of the BackCom transmission by taking into account practical factors including the impedance modulation mode, impedance matching, and polarization mismatch. Using the more comprehensive and accurate link budget model, the uplink and downlink communication ranges can be calculated with the nominal value of reader sensitivity and the measured tag sensitivity. It is found through the analysis that the key factor that sets the upper bound of our practical BackCom system is the capability of the envelope detection circuit of the downlink tags. To address this issue, we propose to incorporate a low-noise amplifier (LNA) module to enhance the signal strength before envelope detection, which consequently improves tag sensitivity and increases the communication distance of the downlink.

The remainder of this paper is organized as follows. Section 2 introduces two communication systems based on backscatter, then describes the downlink signal receiving and the uplink reflection signal process of the tag, and derives the reflection coefficient and radar cross section. In Section 3, firstly, the tag reflection coefficient and other parameters are measured to calculate the sensitivity. Secondly, through the received power and sensitivity, the upper bound distance between the upper and lower lines of the system is analyzed, and the upper bound of the system is obtained as the starting voltage of the diode of the label envelope detection module. Finally, a low-noise power amplifier module is implemented in the receiving circuit of the tag. In Section 4, the performance improvement of the communication system is verified through theoretical and practical measurements. Section 5 concludes the paper.

2 Backscatter Communication Model

Backscatter communication systems can be classified into two categories: monostatic backscatter communication systems (MBCS) and bistatic backscatter communication systems (BBCS)^[17], as shown in Figs. 1(a) and 1(b), respectively.

The MBCS in Fig. 1(a) illustrates that the RF carrier signal

transmitter and the backscatter signal receiver are collocated in one device commonly referred to as a reader. A typical reader may communicate with multiple backscatter tags in the surrounding area. The popular EPC C1G2 RFID system adopts this MBCS configuration.

In the downlink of MBCS, the reader broadcasts the RF carrier signal and signaling information simultaneously. The RF carrier signal can be leveraged by the tag to transmit its information back to the reader. The signaling information is used to activate specific tags and deliver associated commands. More specifically, the tag will be awakened and decode the signaling. Depending on signaling from the reader, the tag transmits its own information by reflecting and modulating the RF carrier signals^[18]. Consequently, the reader performs signal processing to the received backscatter signal to retrieve the tag information. It is worth noting that the backscattered signals that carry the tag information suffer from the power attenuation of both downlink and uplink.

As shown in Fig 1(b), the BBCS consists of three components: a backscatter reader, a tag, and an RF carrier source. In comparison with MBCS, BBCS separates the RF carrier source functionality from the backscatter signaling and information transceiver. This can not only offer flexibility to the practical deployment but improve the signal strength at the reader. The backscatter signal received by the reader also undergoes the path loss of both uplink and downlink. Whereas, since the RF carrier source and the reader are not necessarily collocated and the former can be close to the tag, the incident signal power at the tag antenna in BBCS can be higher than that in MBCS.

Instead of generating an active RF signal in traditional wireless communication systems, the tag in MBCS and BBCS controls and changes the load impedance of the antenna to modulate the incident carrier signals. This yields the RF carrier reflected from the tag antenna in an alternating manner. The tag can encode its information by properly controlling the changing pattern of the reflected signal. The reader, on the other hand, can retrieve the information from the tag by observing the variation of the backscatter signal.

Let Z_a denote the complex tag antenna impedance of the tag

antenna. The reflection coefficient of the tag antenna can be expressed as:

$$\Gamma = \frac{Z_c - Z_a^*}{Z_c + Z_a}, \quad (1)$$

where Z_c is the complex load impedance. When $\Gamma = 0$ ($Z_c = Z_a^*$), it is called the impedance matching state or the impedance absorption state, and the tag does not generate any reflected signals. When $\Gamma \neq 0$, the tag modulates the impedance of load to alter the reflection coefficient, thereby causing changes in the amplitude and phase of the reflected signals. This process is known as the impedance modulation, and the alteration of the reflected signal is detectable and decodable into data by the reader.

The process of impedance modulation is analogous to target identification in radar systems, where the radar cross section (RCS) serves as a metric to quantify the target's ability to reflect radar waves. By considering the tag as the target for assessing its reflective capability, we can express the tag antenna RCS^[19] as follows:

$$\sigma_i = \frac{\lambda^2 G_{\text{tag}}^2}{4\pi} |\Gamma_i - A_s|^2, \quad (2)$$

where λ is the wavelength and G_{tag} is the tag antenna gain. Tag RCS affects the field backscattered from a loaded antenna RCS. It can be divided into load-dependent and load-independent components. The subscript i indicates that the load Z_{ci} is connected to the tag antenna, corresponding to the reflection coefficient Γ_i . For example, when there are only two load states, $i = 1$ is for the tag information bit "0" and $i = 2$ is for the tag information bit "1". A_s corresponds to the structural mode of the tag antenna, which is a load-independent antenna characteristic. Its value is complex, in general, and depends on the geometry and materials used for the antenna construction.

The tag modulates its information by changing the RCS, and the modulated backscatter signal is proportional to the difference of tag RCS ($\Delta\sigma$). Furthermore, the magnitude of $\Delta\sigma$ is shown as below.

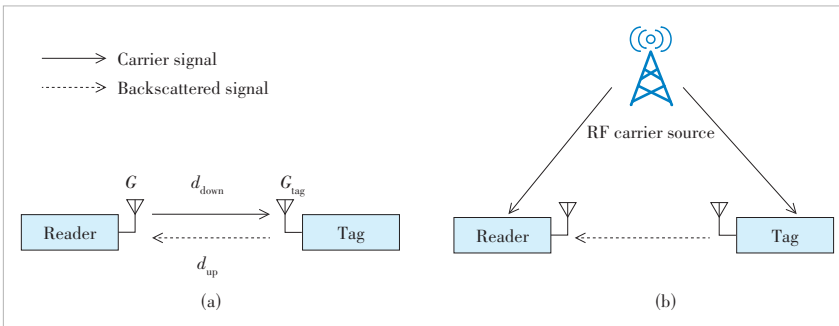
$$\Delta\sigma = \frac{\lambda^2 G_{\text{tag}}^2}{4\pi} \alpha |\Gamma_1 - \Gamma_2|^2, \quad (3)$$

where the coefficient α depends on the tag's modulation system. When the duty cycle of the tag's square-wave modulated signal is 50%, we can have $\alpha=1/4$ ^[20].

3 Link Budget

3.1 Downlink Budget

In the downlink communication, the reader



▲ Figure 1. Two backscatter communication systems: (a) monostatic backscatter communication system and (b) bistatic backscatter communication system

generates and broadcasts signaling that includes reflection indication information to activate the tag. Note that the signal strength is a critical parameter to evaluate the reliability of data transmission. Therefore we use effective isotropic radiated power (EIRP), a value that measures the coverage range and communication capability of a wireless system. The transmitted EIRP of the reader can be expressed as:

$$\text{EIRP}_t = P_t G, \quad (4)$$

where P_t represents the transmission power used by the reader, which indicates the actual power output from the transmitter to the antenna, and G is the gain of the reader antenna, which is the efficiency of the antenna relative to an ideal isotropic antenna. The higher the EIRP, the stronger the transmission capability of the system, allowing it to propagate over longer distances.

In a multipath environment, the path loss between the reader and the tag antennas can be written as^[10]

$$L_p(d) = \left(\frac{\lambda}{4\pi d} \right)^2 \left| 1 + \sum_{n=1}^N \Gamma_n \frac{d}{d_n} e^{-jk(d_n - d)} \right|^2, \quad (5)$$

where d is the length of the direct path, and d_n is the length of the n -th path with a reflection coefficient of Γ_n . N is the total number of reflections.

The signal strength at the tag antenna location can be obtained from the transmitted EIRP and path loss as follows:

$$P_{\text{tag}} = \text{EIRP}_t G_{\text{tag}} \beta L_p(d_{\text{down}}) = P_t G G_{\text{tag}} \beta L_p(d_{\text{down}}), \quad (6)$$

where β is the polarization mismatch, also known as polarization isolation, which measures the degree of signal loss caused by polarization mismatch in signal transmission from the reader antenna to the tag antenna. G_{tag} is the gain of the tag antenna, and d_{down} is the downlink communication distance.

As shown in Fig. 2, the signal received from the tag antenna will go through the impedance matching circuit. Since the reflection coefficient is usually not equal to 0, a small part of the signal will be reflected by the tag antenna. The signal that is

eventually captured and processed by the signal reception circuit of the tag is:

$$P_{\text{tag, chip}} = P_{\text{tag}} (1 - |\Gamma|^2), \quad (7)$$

where Γ is the reflection coefficient of impedance matching. The power reflection coefficient $|\Gamma|^2$ shows the fraction of the maximum power available from the generator that is not delivered to the load^[8]. A well-designed impedance matching is of vital importance to the system performance. The impedance matching circuit is normally designed to achieve $\Gamma = 0$ to maximize the signal reception.

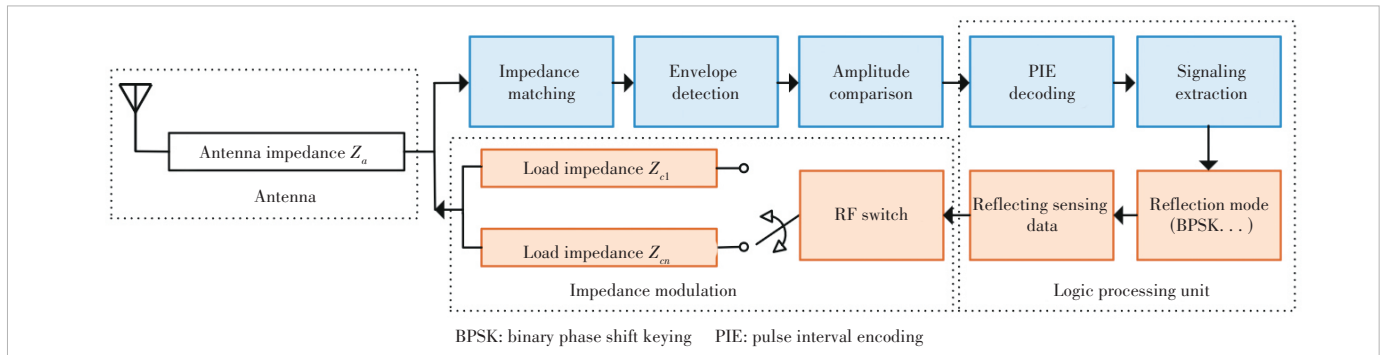
After that, the received signal is processed by the envelope detection module for demodulation. It converts the received RF signal into a low-frequency envelope signal with varying amplitudes. The following comparison module calculates a threshold based on the long-term average value of the varying envelope, and determines the information carried on the envelope by comparing the instantaneous envelope with the threshold. Subsequently, the bi-level output is processed by the logic processing unit. Consequently, the tag extracts control information such as the reflection rate and the modulation scheme from the decoded sequence. With this information, the tag specifies its working mode and prepares the signal transmission that will be carried out later.

3.2 Uplink Budget

In uplink communications, the tag modulates and reflects its information through impedance modulation. The reader processes the received reflected signal. The tag modulates and reflects a modulated backscatter signal with differential EIRP (ΔEIRP). The backscattered power of the tag can be calculated as $\Delta\text{EIRP} = S\Delta\sigma$. The following equation exhibits a proportional relationship with the received power P_{tag} ^[21]:

$$\frac{\Delta\text{EIRP}}{P_{\text{tag}}} = \frac{S\Delta\sigma}{SA} = G_{\text{tag}} K, \quad (8)$$

where $K = \alpha |\Gamma_1 - \Gamma_2|^2$ represents the tag modulation loss factor. S is the power density of an electromagnetic (EM) wave



▲ Figure 2. Typical functional diagram inside a tag

incident on the tag. $A = \frac{\lambda^2 G_{\text{tag}}}{4\pi}$ is the effective area of the tag antenna.

The power of the modulated tag signal received by reader antennas can be written in terms of differential EIRP as

$$P_r = \Delta \text{EIRP} G \beta L_p(d_{\text{up}}) = P_{\text{tag}} G_{\text{tag}} K G \beta L_p(d_{\text{up}}) = P_t G^2 G_{\text{tag}}^2 \beta^2 L_p(d_{\text{down}}) L_p(d_{\text{up}}) \alpha |\Gamma_1 - \Gamma_2|^2, \quad (9)$$

where d_{down} is the downlink communication distance. It can be observed that the signal received by the reader undergoes path loss in both the uplink and downlink. Based on the analysis of the sensitivity and received power at readers and tags, respectively, which is conducted in the next section, it can be seen that the limiting bottleneck for the communication distance of MBCS is the downlink communication distance.

4 Downlink Improvement of Tag

4.1 Tag Chip Sensitivity

In downlink communications, the tag receives and extracts the signaling indication information of RF signals sent from the reader, and sets the corresponding reflection mode. However, signal decoding and information extraction from RF signals can be implemented by an analog-to-digital converter (ADC), as in traditional wireless systems. However, it consumes a significant amount of power and therefore is not suitable for ultra-low power systems. In a signal reception module of the tag, a threshold comparison circuit with lower power consumption is used to achieve frequency down-conversion.

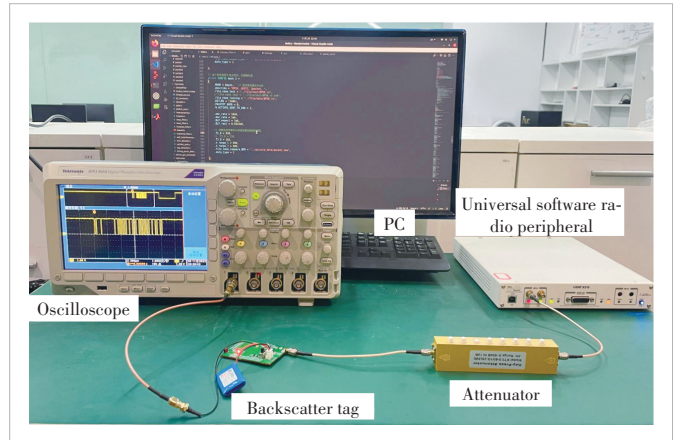
The tag chip sensitivity is also known as the activation threshold, indicating the minimum power of signaling that can be recognized by the tag's logic processing unit and used to perform impedance modulation. Depending on the tag's reception processing, factors affecting sensitivity may be the forward voltage drop of the diode, the sensitivity of the comparator, and the algorithm limit imposed by the adopted decoding scheme.

More specifically, if the input signal is lower than the level of forward voltage drop, the received signal cannot be demodulated and no output will be generated. Similarly, when the input signal is too weak, the signal amplitudes that represent different information states are very close to each other. Especially, when the difference is lower than the minimum difference that can be detected by the comparator (i.e. the sensitivity of the comparator), the comparator cannot output reliable results for the following processing. In addition, the decoding algorithm employed by the tag may also have certain performance upper bound, beyond which the decoding algorithm cannot constantly provide correct output information. This performance upper bound is usually associated with the signal quality.

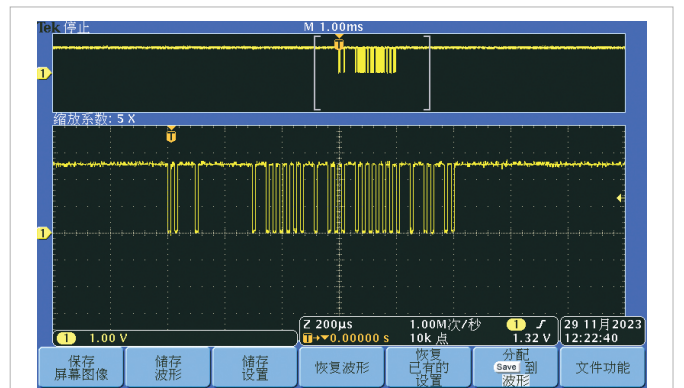
We measure the chip sensitivity of the tag with the following experiments. As shown in Fig. 3, we use universal software radio peripheral (USRP) to transmit a carrier signal and signaling with certain power, through an attenuator connected to the tag, and connect the comparator output of the tag to an oscilloscope. The operating frequency is 925 MHz, and the measurement steps are as follows. Firstly, we measure the threshold power P_{th} at the tag antenna when the comparator has a stable waveform output corresponding to the signaling. That is, the tag correctly decodes the signaling and initiates reflection. Fig. 4 shows the comparator's output waveform measured by the oscilloscope at $P_{\text{th}} = -19.90$ dBm, matching the signaling from the reader. Due to an increase in the attenuation value of the attenuator, the comparator waveforms no longer match. The threshold power P_{th} is called the tag sensitivity. Secondly, the reflection coefficient $|\Gamma| = 0.1384$ of impedance matching is measured by using a vector network analyzer (VNA) at the P_{th} power level. The tag chip sensitivity $P_{\text{chip, sen}} = -19.98$ dBm.

4.2 Communication Distance Upper Bound

In MBCS, when $L_p(d_M) = \left(\frac{\lambda}{4\pi d_M}\right)^2$, the upper bound of the downlink distance can be determined by:



▲ Figure 3. Sensitivity measurement experiment



▲ Figure 4. Output waveform of the comparator

$$d_{\text{down}} = \frac{\lambda}{4\pi} \sqrt{\frac{P_t G G_{\text{tag}} \beta (1 - |\Gamma|^2)}{P_{\text{tag,chip}}}} \quad (10)$$

Similarly, the upper bound of the uplink distance is

$$d_{\text{up}} = \frac{\lambda}{4\pi} \sqrt[4]{\frac{P_t G^2 G_{\text{tag}}^2 \beta^2 \alpha |\Gamma_1 - \Gamma_2|^2}{P_r}} \quad (11)$$

We set the reader sensitivity $P_{\text{reader, sen}} \approx -86$ dBm, the reader antenna gain $G = 5$ dBi, the tag antenna gain $G_{\text{tag}} = 2$ dBi, polarization mismatch $\beta = 0.5$, and coefficient $\alpha = 1$ ^[22].

Suppose that the tag adopts the binary phase shift keying (BPSK) modulation mode. Due to the influence of factors such as RF switch and transmission line impedance, the actual reflection coefficient is often less than 1. When the tag adopts the BPSK modulation, the values of the measured reflection coefficients are $\Gamma_1 = -0.69 + 0.32j$ and $\Gamma_2 = 0.67 - 0.23j$, respectively. Figs. 5 and 6 are respectively signal power received by the tag chip and the reader when transmission power of the reader $P_t = 10, 15$ and 20 dBm, respectively.

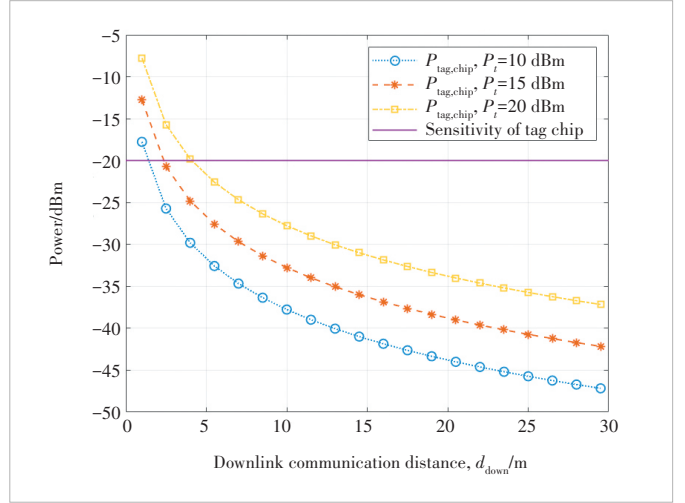
When the tag chip received power equals to the envelope detection threshold power ($P_{\text{tag,chip}} = P_{\text{chip, sen}}$), the downlink upper bound communication distance $d_{\text{down}} \approx 2.30$ m. Similarly, when $P_r = P_{\text{reader, sen}}$ the uplink upper bound communication distance $d_{\text{up}} \approx 16.76$ m. Final upper bound communication distance $d_M = \min\{d_{\text{down}}, d_{\text{up}}\} = 2.30$ m. In the real-world test, the restriction of the software program on the signaling preamble is reduced, and the tag cannot receive signaling information when approaching the upper bound distance. At this time there is no comparison voltage output from the viewing comparator and envelope signal output from the envelope detection. Accordingly, we conclude that the downlink distance is mainly limited by the envelope detection threshold.

Based on the previous measurements and analysis, we propose to implement an LNA module after the tag antenna reception. As shown in Fig. 7, the added LNA amplifies the strength of the signal received by the tag while minimizing the noise in the signal to ensure that the signal-to-noise ratio (SNR) can be maintained at a high level when the signal is transmitted and processed in the system. Since the LNA amplifies the voltage input to the envelope detection, the tag is easier to detect the downlink signaling. The gain of LNA

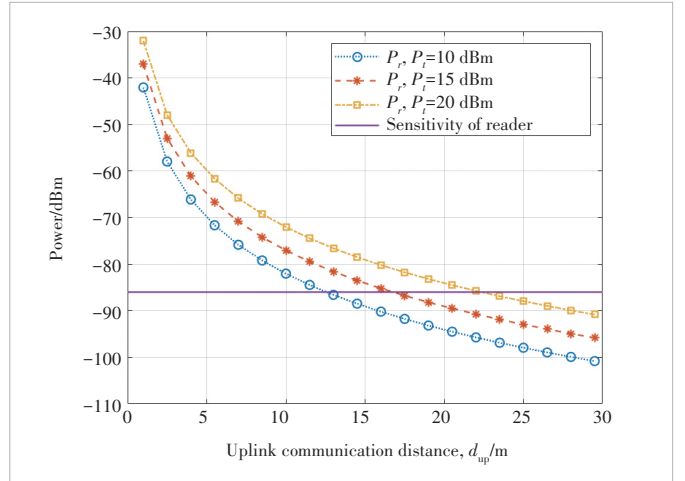
$G_{\text{LNA}} = 21$ dB. The tag chip sensitivity remains unchanged, but the threshold power $P_{\text{th, LNA}}$ at the arrival of the tag antenna decreases to $P_{\text{th}} - G_{\text{LNA}}$. Eventually, the sensitivity of the labels is roughly improved by G_{LNA} .

5 Performance Analysis

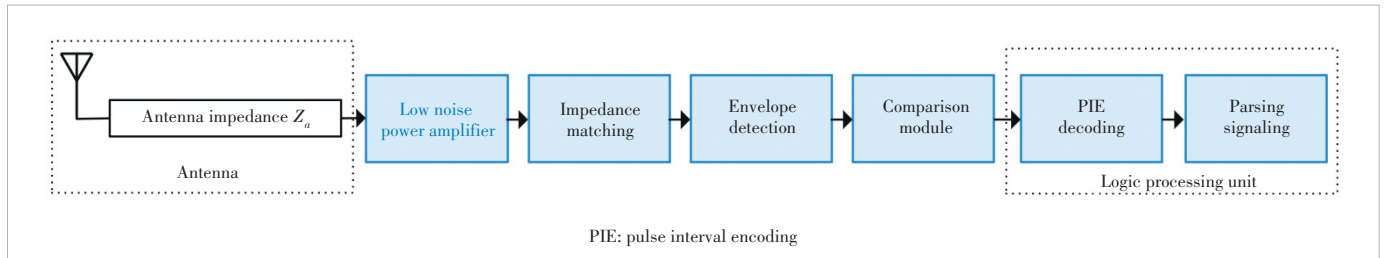
The signal power $P_{\text{tag,chip}}$ and $P_{\text{tag,chip, LNA}}$ received by tag chips versus the distance are plotted in Fig. 8, where the transmission power of the reader is set as $P_t = 15$ dBm. We use the



▲ Figure 5. Signal power received at the tag



▲ Figure 6. Signal power received by the reader

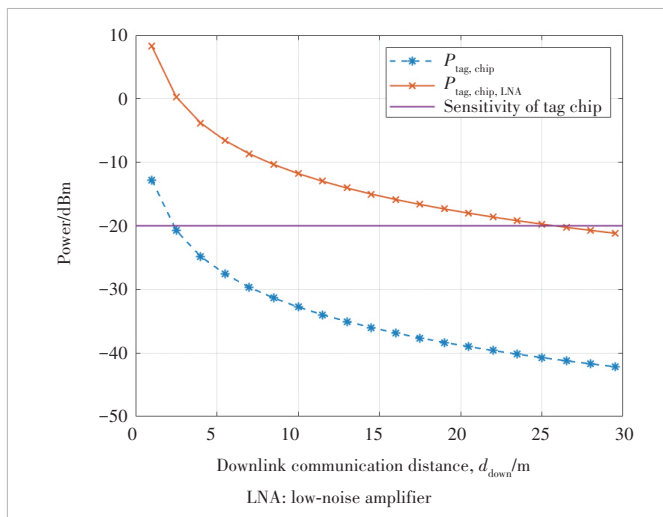


▲ Figure 7. Functional diagram of the backscatter tag with enhanced downlink

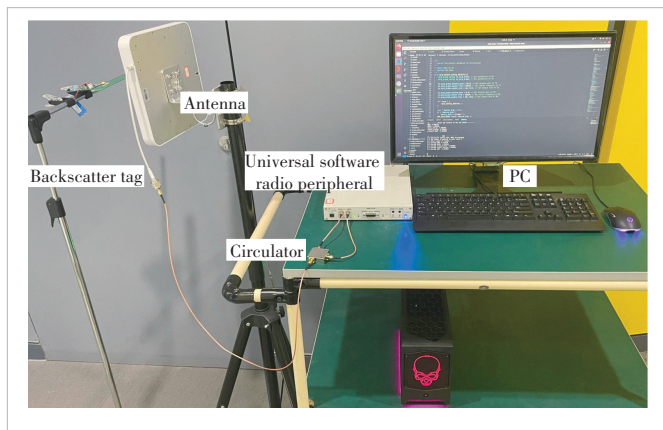
same configuration as those in Fig. 5.

When $P_{\text{tag,chip,LNA}} = P_{\text{chip,sen}}$, we can obtain that the upper bound of downlink communication distance can be 25.75 m, i. e., $d_{\text{down,LNA}} \approx 25.75$ m. After implementing the LNA, the downlink upper bound communication range is improved by 23.45 m.

The test environment is depicted in Fig. 9, where a USRP is in conjunction with a PC as an integrated system. The USRP is typically used as the radio front-end hardware, responsible for receiving and transmitting radio signals. The PC is used for processing and analyzing these received signals. The transmission and reception of radio frequency signals are realized by connecting the antenna with a circulator. The operating frequency is 925 MHz, and the transmission power $P_t = 15$ dBm. We measure the upper bound distance as the maximum distance at which the tag can correctly recognize the signaling. Our final measurement results indicate that, after implementing the LNA module, the communication range is increased by approximately 20 m.



▲ Figure 8. Signal power received at the tag



▲ Figure 9. Downlink upper bound communication distance measurement experiment

In MBCS, the tag can identify smaller signals after implementing the LNA module. According to our analysis, the downlink sensitivity is the bottleneck that sets the upper bounds of the overall communication range. The improvement in the downlink communication distance consequently leads to an improved uplink distance, lower reader power consumption and enhanced communication performance. It is worth noting that these benefits can also be achieved in some BBCS configurations. Improving tag sensitivity will result in an extended maximum effective communication distance between the RF carrier source and the tag.

6 Conclusions

In this paper, we analyze the uplink and downlink budgets and measure the chip sensitivity of the tag for ambient IoT. According to the relationship between sensitivity and the received power, we derive the upper bounds of both uplink and downlink distances. It is found that the backscatter communications are mainly limited by the downlink distance, instead of the uplink distance. It is also found that the tag sensitivity results from its hardware constraints, especially the activation voltage of the envelope detection diode. Accordingly, we propose to incorporate an LNA in the tag to amplify the incident signals before the envelope detection, and the corresponding experimental results show that the downlink communication range could be extended over 20 m.

References

- [1] STOCKMAN H. Communication by means of reflected power [J]. Proceedings of the IRE, 1948, 36(10): 1196 - 1204. DOI: 10.1109/JRPROC.1948.226245
- [2] ROY S, JANDHYALA V, SMITH J R, et al. RFID: from supply chains to sensor nets [J]. Proceedings of the IEEE, 2010, 98(9): 1583 - 1592. DOI: 10.1109/jproc.2010.2053690
- [3] 3GPP. Study on Ambient IoT (Internet of Things) in RAN: TR 38.848 [S]. 2023
- [4] LIU X L, ANSARI N. Toward green IoT: energy solutions and key challenges [J]. IEEE communications magazine, 2019, 57(3): 104 - 110. DOI: 10.1109/MCOM.2019.1800175
- [5] TORO U S, WU K S, LEUNG V C M. Backscatter wireless communications and sensing in green Internet of Things [J]. IEEE transactions on green communications and networking, 2022, 6(1): 37 - 55. DOI: 10.1109/TGCN.2021.3095792
- [6] SONG C Y, DING Y, EID A, et al. Advances in wirelessly powered backscatter communications: from antenna/RF circuitry design to printed flexible electronics [J]. Proceedings of the IEEE, 2022, 110(1): 171 - 192. DOI: 10.1109/JPROC.2021.3125285
- [7] EPCglobal. EPC radio-frequency identity protocols Class-1 Generation-2 UHF RFID protocol for communications at 860 MHz-960 MHz: GB3100-3102 [S]. 2015
- [8] NIKITIN P V, RAO K V S, LAM S F, et al. Power reflection coefficient analysis for complex impedances in RFID tag design [J]. IEEE transactions on microwave theory and techniques, 2005, 53(9): 2721 - 2725. DOI: 10.1109/TMTT.2005.854191
- [9] DE VITA G, IANNACCONE G. Design criteria for the RF section of UHF

- and microwave passive RFID transponders [J]. *IEEE transactions on microwave theory and techniques*, 2005, 53(9): 2978 – 2990. DOI: 10.1109/TMTT.2005.854229
- [10] LAZARO A, GIRBAU D, SALINAS D. Radio link budgets for UHF RFID on multipath environments [J]. *IEEE transactions on antennas and propagation*, 2009, 57(4): 1241 – 1251. DOI: 10.1109/TAP.2009.2015818
- [11] GRIFFIN J D, DURGIN G D. Complete link budgets for backscatter-radio and RFID systems [J]. *IEEE antennas and propagation magazine*, 2009, 51(2): 11 – 25. DOI: 10.1109/MAP.2009.5162013
- [12] WEI P, CHE W Y, BI Z Y, et al. High-efficiency differential RF front-end for a Gen2 RFID tag [J]. *IEEE transactions on circuits and systems II: Express briefs*, 2011, 58(4): 189 – 194. DOI: 10.1109/TCSII.2011.2124530
- [13] MOODY J, BASSIRIAN P, ROY A, et al. Interference robust detector-first near-zero power wake-up receiver [J]. *IEEE journal of solid-state circuits*, 2019, 54(8): 2149 – 2162. DOI: 10.1109/JSSC.2019.2912710
- [14] TIAN Y, YU N, LI J, et al. A high sensitivity and high dynamic range ASK demodulator for passive UHF RFID with over-voltage protection [C]/*International conference on electronics technology*. IEEE, 2023: 1012 – 1016. DOI: 10.1109/ICET58434.2023.10211808
- [15] MOODY J, BASSIRIAN P, ROY A, et al. A -76 dBm 7.4 nW wakeup radio with automatic offset compensation [C]/*IEEE International Solid-State Circuits Conference*. IEEE, 2018: 452 – 454. DOI: 10.1109/ISSCC.2018.8310379
- [16] ALAJI I, AOUIMEUR W, GHANEM H, et al. Design of zero bias power detectors towards power consumption optimization in 5G devices [J]. *Microelectronics journal*, 2021, 111: 105035. DOI: 10.1016/j.mejo.2021.105035
- [17] KIMIONIS J, BLETSAS A, SAHALOS J N. Increased range bistatic scatter radio [J]. *IEEE transactions on communications*, 2014, 62(3): 1091 – 1104. DOI: 10.1109/TCOMM.2014.020314.130559
- [18] DOBKIN D M. *The RF in RFID* [M]. Amsterdam, Netherlands: Elsevier, 2008. DOI: 10.1016/B978-0-7506-8209-1.X5001-3
- [19] GREEN R B. *The general theory of antenna scattering* [D]. Columbus, USA: The Ohio State University, 1963
- [20] NIKITIN P V, RAO K V S. Antennas and propagation in UHF RFID systems [C]/*IEEE International Conference on RFID*. IEEE, 2008: 277 – 288. DOI: 10.1109/RFID.2008.4519368
- [21] NIKITIN P V, RAO K V S. Theory and measurement of backscattering from RFID tags [J]. *IEEE antennas and propagation magazine*, 2006, 48(6): 212 – 218. DOI: 10.1109/MAP.2006.323323
- [22] NIKITIN P V, RAO K V S, MARTINEZ R D. Differential RCS of RFID tag [J]. *Electronics letters*, 2007, 43(8): 431. DOI: 10.1049/el: 20070253

Biographies

YANG Yibing received her BS degree from Anhui University, China in 2022. She is currently working toward her PhD degree with the School of Computer Science and Technology, Beijing Jiaotong University, China. Her research interests include mobile and Internet networks

LIU Ming (mingliu@bjtu.edu.cn) received his BE and ME degrees in electrical engineering from Xi'an Jiaotong University, China in 2004 and 2007, respectively, and PhD degree in electrical engineering from the National Institute of Applied Sciences, France in 2011. He was with the Institute of Electronics and Telecommunications of Rennes, France, as a postdoctoral researcher from 2011 to 2015. He is currently with Beijing Jiaotong University, China, as an associate professor. His main research interests include beyond 5G/6G, PHY security, and AI for wireless communications.

XU Rongtao received his BS degree in radio technology from Xi'an Jiaotong University, China in 1997, MS degree in communication and information system from the Beijing University of Posts and Telecommunications, China in 2000, and PhD degree in electronic and information engineering from The Hong Kong Polytechnic University, China in 2007. From 2000 to 2003, he was a system engineer at Siemens Ltd., China. In 2007, he joined Beijing Jiaotong University, China, where he is currently an associate professor with the State Key Laboratory of Rail Traffic Control and Safety. His research interests include wideband mobile communications, railway communications, and wireless sensor networks.

WANG Gongpu received his BE degree in communication engineering from Anhui University, China in 2001, MS degree from the Beijing University of Posts and Telecommunications, China in 2004, and the PhD degree from University of Alberta, Canada in 2011. From 2004 to 2007, he was an assistant professor at the School of Network Education, Beijing University of Posts and Telecommunications. He is currently a full professor with the School of Computer and Information Technology, Beijing Jiaotong University, China. His research interests include wireless communication theory, signal processing technologies, and the Internet of Things.

GONG Wei received his BE degree in computer science from Huazhong University of Science and Technology, China, and ME degree in software engineering and PhD degree in computer science from Tsinghua University, China. He is a professor with the School of Computer Science and Technology, University of Science and Technology of China. He had also conducted research with Simon Fraser University, Canada, and the University of Ottawa, Canada. His current research interests include wireless networks, Internet of Things, and distributed computing.

Impacts of Model Mismatch and Array Scale on Channel Estimation for XL-HRIS-Aided Systems



LU Zhizheng, HAN Yu, JIN Shi

(National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China)

DOI: 10.12142/ZTECOM.202401004

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240308.1541.002>,
published online March 11, 2024

Manuscript received: 2024-01-25

Abstract: Extremely large-scale hybrid reconfigurable intelligence surface (XL-HRIS), an improved version of the RIS, can receive the incident signal and enhance communication performance. However, as the RIS size increases, the phase variations of the received signal across the whole array are nonnegligible in the near-field region, and the channel model mismatch, which will decrease the estimation accuracy, must be considered. In this paper, the lower bound (LB) of the estimated parameter is studied and the impacts of the distance and signal-to-noise ratio (SNR) on LB are then evaluated. Moreover, the impacts of the array scale on LB and spectral efficiency (SE) are also studied. Simulation results verify that even in extremely large-scale array systems with infinite SNR, channel model mismatch can still limit estimation accuracy. However, this impact decreases with increasing distance.

Keywords: XL-HRIS; near-field; LB; model mismatch; parameter estimation

Citation (Format 1): LU Z Z, HAN Y, JIN S. Impacts of model mismatch and array scale on channel estimation for XL-HRIS-aided systems [J]. *ZTE Communications*, 2024, 22(1): 24 – 33. DOI: 10.12142/ZTECOM.202401004

Citation (Format 2): Z. Z. Lu, Y. Han, and S. Jin, “Impacts of model mismatch and array scale on channel estimation for XL-HRIS-aided systems,” *ZTE Communications*, vol. 22, no. 1, pp. 24 – 33, Mar. 2024. doi: 10.12142/ZTECOM.202401004.

1 Introduction

Reconfigurable intelligence surface (RIS) is a candidate enabling technology for future 6G wireless communication systems^[1]. It can increase the spectral efficiency (SE) through reconfiguring the communication environment with low hardware cost and little consumption^[2-4]. To increase the SE of the communication systems, the phase shift of each element at the RIS should be devised specifically, which requires the acquisition of channel state information (CSI) in advance. However, for the passive RIS^[5-13] that is composed of nearly passive elements, only the cascade channel between the transmitter and the receiver can be estimated, leading to high computational complexity and high pilot overhead that is equal to the number of elements at the RIS. CSI of the two separate channels (i.e., the channel between the user equipment (UE) and the RIS and that between the RIS and the BS) cannot be utilized directly to enhance communication performance. Hence, the hybrid RIS (HRIS), which is composed of both pas-

sive and active elements, has been proposed to estimate the CSI in separate channels. The passive elements can only reflect the signal with a programmable phase shift, while the active elements can also process the receive signal^[14-18]. Many studies have been conducted on channel estimation for HRIS-aided systems^[19-21]. The authors of Ref. [19] proposed a two-stage channel estimation scheme, in which the HRIS estimated the channel between the UE and the HRIS firstly, followed by the BS estimating the channel between the HRIS and the BS. In Ref. [20], the UE and the BS transmitted pilots to the HRIS, respectively, and then the HRIS estimated the two separate channels. In Ref. [21], the HRIS was equipped with only one single RF chain, which enabled the estimation of the two separate channels with low hardware cost.

However, as more elements are employed in the HRIS, it becomes necessary to consider extremely large-scale HRIS (XL-HRIS) aided systems for further 6G wireless communication systems. When the aperture of the array is large, the phase variations of the receive signal among the whole array cannot be ignored, and the channel should be modeled by the spherical-wave model. This will lead to high computational complexity for channel estimation, because the high-dimension channel vector and the emerging distance parameter need to be considered in the near-field region^[22-29]. Espe-

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant Nos. 62301148, 62341107, and 62261160576, in part by the Natural Science Foundation of Jiangsu Province under Grant No. BK20230824, and in part by the Key Technologies R&D Program of Jiangsu (Prospective and Key Technologies for Industry) under Grant Nos. BE2023022 and BE2023022-1.

cially, in order to decrease the hardware cost, only a few active elements are employed in the XL-HRIS, which will also bring challenges to full dimensional channel estimation. Many studies have been conducted on near-field channel estimation^[30–32]. Ref. [30] considered a near-field channel scenario, where the channel was modeled by a spherical wave model. In Ref. [31], a large lens antenna array-aided system was considered, which can estimate the position parameters and locate UE precisely. Ref. [32] considered a spherical wave channel model and proposed an iteration-based algorithm based on the Newtonized orthogonal matching pursuit (NOMP) algorithm^[33], which can detect the channel parameters and reconstruct the near-field channel accurately. Nevertheless, most studies on channel estimation in RIS-aided systems still estimate the channel parameters based on the planar wave channel model and ignore the differences between the realistic spherical wave channel model and the assumed planar wave channel model. This can lead to the estimation error lower bound (LB) analysis being lower than the true results. There are few studies on the impacts of the array scale and the variation between the assumed channel model and realistic channel model on channel estimation accuracy.

In this paper, to study the impacts of the large array scale on channel estimation accuracy and communication performance, an estimation error LB of channel parameters is derived, which considers the accuracy degradation caused by the channel model mismatch between the realistic spherical channel wave model and the assumed planar wave channel model. A true LB result is obtained and the impacts of system parameters on LB, such as the distance, the number of elements and transmitted power, are analyzed. Moreover, the impacts of the large array scale and channel model mismatch on the SE performance are evaluated, and the simulation results are given. To reduce the computational complexity and pilot overhead in the XL-RIS-aided system, a unique XL-HRIS architecture that uses active and passive elements simultaneously is introduced in this paper. A decoupled channel estimation algorithm, which estimates the angle and distance parameters separately, is also considered.

The rest of the paper is organized as follows. In Section 2, the system model, signal model, and near-field channel estimation algorithm under the proposed XL-HRIS architecture are introduced. In Section 3, an LB that considers the channel model mismatch is defined and derived, and the impacts of the large array scale on LB and SE are then studied. Section 4 shows the simulation results regarding the impacts of the array scale on parameter estimation accuracy and communication performance. Finally, the conclusions are presented in Section 5.

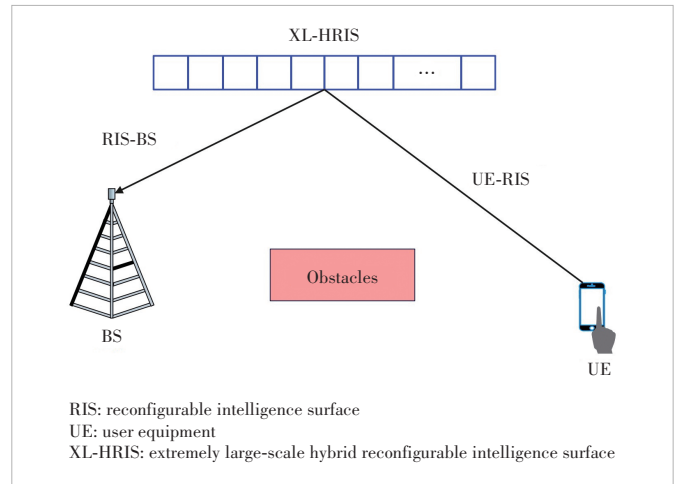
2 Proposed XL-HRIS-Aided System

2.1 System Model

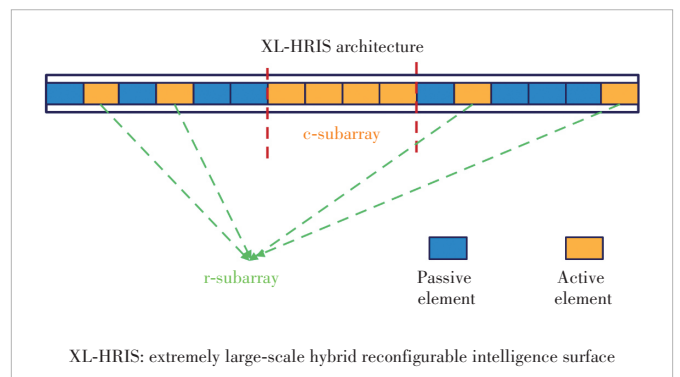
An XL-HRIS-aided time division duplexing (TDD) system

is proposed in this paper, as shown in Fig. 1. In this system, the BS is equipped with an M -element uniform linear array (ULA), while the UE only has one single antenna. The carrier frequency is denoted as f_c , and the wave length is $\lambda_c = c/f_c$, where c denotes the speed of light. The spacing between adjacent antennas is the half of the wave length. However, the channel between the UE and the BS is blocked by obstacles, which can cause a serious damage on the communication performance. Therefore, an XL-HRIS is introduced to provide an additional link to enhance the communication quality.

This paper proposes a unique XL-HRIS architecture, as shown in Fig. 2, to decrease the computational complexity and pilot overhead in the decoupled separate channel estimation scheme, and to improve system communication performance. The architecture is employed with an N -elements ULA and the spacing between adjacent elements is $d = \lambda_c/2$. $N_{\text{act}} \ll N$ active elements exist in the XL-RIS, which can be divided into two parts. One part is an N_{sub} -element subarray ULA (sULA) at the center of the whole array, which is called the core subarray (c-subarray). The other part is composed of N_{ran} active elements, which are randomly distributed in the ULA, forming a random subarray (r-subarray). We have $N_{\text{act}} = N_{\text{sub}} +$



▲ Figure 1. Proposed XL-HRIS-aided system model



▲ Figure 2. Proposed XL-HRIS architecture with active c-subarray and active r-subarray. Other elements are passive

N_{ran} . Each active element has two modes: active and passive. In the active mode, the active elements can receive and process the signal, and the XL-HRIS can estimate the channel between the transmitter and the active elements at the ULA. In the passive mode, each element can only reflect the signal with a programmable phase shift. All other $N - N_{\text{act}}$ elements in the ULA are passive elements, which can only operate in the passive mode.

2.2 Signal Model and Channel Model

The channel $\mathbf{h}_R \in \mathbb{C}^{N \times 1}$ between the transmitter and the ULA can be expressed as

$$\mathbf{h}_R = g\mathbf{w}_R(\theta, r), \quad (1)$$

where g denotes the complex gain; θ is the azimuth angle of the receive signal at the center of the XL-HRIS; r denotes the distance between the transmitter and the center of the XL-HRIS. The steering vector can be expressed as

$$\mathbf{w}_R(\theta, r) = \left[e^{jk_c(r_1 - r)}, \dots, e^{jk_c(r_N - r)} \right]^T, \quad (2)$$

where $k_c = 2\pi/\lambda_c$ is the wave number; r_n denotes the distance from the transmitter to the n -th element at the ULA, $n = 1, \dots, N$, which can be expressed as

$$r_n = \sqrt{r^2 - 2r\chi_n d\theta + \chi_n^2 d^2}, \quad (3)$$

where $\chi_n = n - (N + 1)/2$. Assume that all the pilots are transmitted by the transmitter, and when the XL-HRIS is in the active mode, the receive signal $\mathbf{y}_{\text{sub}} \in \mathbb{C}^{N_{\text{sub}} \times 1}$ at the c -subarray can be denoted as

$$\mathbf{y}_{\text{sub}} = \sqrt{P}\mathbf{h}_{\text{sub}} + \mathbf{z}_{\text{sub}}, \quad (4)$$

where \mathbf{z}_{sub} is the additional Gaussian noise with 0 zero and σ^2 variance, and the channel between the UE and the c -subarray can be expressed as

$$\mathbf{h}_{\text{sub}} = g\mathbf{w}_{\text{sub}}(\theta, r), \quad (5)$$

where

$$\mathbf{w}_{\text{sub}}(\theta, r) = \left[e^{jk_c(r_{\text{sub},1} - r)}, \dots, e^{jk_c(r_{\text{sub},N_{\text{sub}}} - r)} \right]^T, \quad (6)$$

where $r_{\text{sub},n}$ denotes the distance from the transmitter to the n -th element at the c -subarray, which can be denoted as

$$r_{\text{sub},n} = \sqrt{r^2 - 2r\delta_n d\theta + \delta_n^2 d^2}, \quad (7)$$

where $\delta_n = n - (N_{\text{sub}} + 1)/2$, for $n = 1, \dots, N_{\text{sub}}$. Similarly, the channel between the UE and the r -subarray can be denoted as \mathbf{h}_{ran} .

Considering that the aperture of the c -subarray is small

compared to the whole ULA, the phase variations of the receive signal among the c -subarray can be ignored. As a result, the channel can be simplified as the planar wave model, which can be denoted as

$$\mathbf{h}_{\text{sim}} = g\mathbf{a}_{\text{sim}}(\theta), \quad (8)$$

where the simplified steering vector can be expressed as

$$\mathbf{a}_{\text{sim}}(\theta) = \left[e^{-j\delta_1 k_c d\theta}, \dots, e^{-j\delta_{N_{\text{sub}}} k_c d\theta} \right]^T. \quad (9)$$

Due to the large number of elements at the XL-HRIS, estimation of the cascade channel between the UE and the BS has high computational complexity. Hence, a two-stage decoupled separate channel estimation scheme is proposed in this paper, in which the UE transmits pilots to the RIS firstly, while the RIS estimates the channel between the UE and the RIS through the two active subarrays. Then, the RIS reflects the signal to the BS, and the BS estimates the RIS-BS channel. Since the channel between the RIS and the BS is stable, it only needs to be estimated once at the beginning of the channel correlation time. However, the channel between the UE and the RIS needs to be estimated frequently. This paper only focuses on the estimation of the UE-RIS channel.

As only part of elements at the XL-HRIS are active, the full dimensional UE-RIS channel estimation faces many challenges. Due to the spatial sparsity of the UE-RIS channel, we only estimate the channel parameters through the active elements in this paper, and then reconstruct the full dimensional channel. In the XL-HRIS-aided system, the computational complexity remains high due to the need to estimate more parameters under the spherical wave model. Fortunately, the receive signal of the c -subarray and that of the r -subarray in the proposed XL-HRIS architecture can be processed respectively. A decoupled estimation strategy is adopted to reduce the computational complexity. Firstly, the angle of arrival (AoA) parameter is estimated from the receive signal of the c -subarray, and then the distance parameter is detected from the receive signal of the r -subarray. However, since the c -subarray detects the AoA parameter based on the planar-wave model, which differs from the practical spherical-wave model, channel model mismatch can occur, which affects the estimation accuracy. In the next section, the impacts of the c -subarray scale on the AoA parameter estimation accuracy will be analyzed, and an estimation error lower bound will be given. The influence of the estimation error on the system communication performance will be also studied.

3 LB and SE Analysis Under Model Mismatch

3.1 Problem Description

When the XL-HRIS is in the active mode, the receive signal at the c -subarray can be denoted as

$$\mathbf{y}_{\text{sub}} = \sqrt{P} \mathbf{g} \mathbf{w}_{\text{sub}}(\theta, r) + \mathbf{z}_{\text{sub}}, \quad (10)$$

where the realistic likelihood function under the realistic channel model can be expressed as

$$p(\mathbf{y}_{\text{sub}}|\{\theta, r\}) = \frac{1}{(2\pi\sigma^2)^{N_{\text{sub}}}} e^{-\frac{1}{\sigma^2} \|\mathbf{y}_{\text{sub}} - g_{\text{equ}} \mathbf{w}_{\text{sub}}(\theta, r)\|^2}, \quad (11)$$

where $g_{\text{equ}} = \sqrt{P} g$ is thought as constant. However, in the proposed channel estimation algorithm, the channel between the UE and the c-subarray is simplified as Eq. (8) for estimating the AoA parameter from the receive signal of the c-subarray, which can be expressed as

$$\mathbf{y}_{\text{sub}} \approx \sqrt{P} \mathbf{g} \mathbf{a}_{\text{sim}}(\theta) + \mathbf{z}_{\text{sub}}, \quad (12)$$

where the simplified likelihood function under the planar channel model can be expressed as

$$\tilde{p}(\mathbf{y}_{\text{sub}}|\theta) = \frac{1}{(2\pi\sigma^2)^{N_{\text{sub}}}} e^{-\frac{1}{\sigma^2} \|\mathbf{y}_{\text{sub}} - g_{\text{sim}} \mathbf{a}_{\text{sim}}(\theta)\|^2}. \quad (13)$$

When the distance between the UE and XL-HRIS is far or the array scale of the c-subarray is small, the variation among the realistic and simplified channel models can be ignored, and the Cramer-Rao lower bound (CRLB) of the estimated AoA parameter can be directly derived based on Eq. (13). Nevertheless, with the increase of the scale of the c-subarray in the proposed XL-HRIS-aided systems, the phase variations among the whole array become large. Then, the estimation of θ based on the simplified model in Eq. (13) will face accuracy degradation due to the non-negligible variation between the simplified model and the realistic model. Hence, the LB of the estimated AoA parameter, which considers the channel model mismatch, should be studied to get a more accurate estimation error lower bound.

3.2 LB of Estimated AoA Parameter

To obtain the LB of the estimated AoA parameter, the channel model mismatch must be considered. Firstly, we recall the estimation lower bound derivation based on the pseudo-true parameter η_{pse} when the model mismatch happens^[34]. The pseudo-true parameter can minimize the Kullback-Leibler (KL) divergence between the true probability density function with the true parameter $\bar{\eta}$ and the mismatched probability density function with the pseudo-true parameter η_{pse} . Then, the Preliminary 1 can be obtained:

Preliminary 1: When model mismatch happens, a pseudo-true parameter η_{pse} can be estimated firstly, and the LB of the estimated parameter $\bar{\eta}$ can be expressed as

$$\text{LB}(\bar{\eta}) = \text{CRLB}(\eta_{\text{pse}}) + (\bar{\eta} - \eta_{\text{pse}})^{\text{H}} (\bar{\eta} - \eta_{\text{pse}}), \quad (14)$$

where $\text{CRLB}(\eta_{\text{pse}})$ denotes the CRLB of the estimated pseudo-true parameter η_{pse} under the mismatched model.

Based on Preliminary 1, the LB of the estimated AoA parameter under the channel model mismatch can be given by Theorem 1.

Theorem 1: When estimating the AoA parameter from the realistic receive signal based on the planar wave model, the LB of the estimated AoA parameter can be expressed as

$$\text{LB}(\theta) = \frac{6\sigma^2}{k_c^2 d^2 P |g|^2 N_{\text{sub}} (N_{\text{sub}} - 1)} + \left(1 + \frac{3N_{\text{sub}} + 4}{N_{\text{sub}}^2 - 1}\right)^2 \frac{9d^2(1 - 2\theta^2 + \theta^4)}{16r^2}. \quad (15)$$

Proof: Let $\{\bar{\theta}, \bar{r}\}$ denote the true channel parameters, and the pseudo-true parameter is defined as

$$\theta_{\text{pse}} = \arg \min_{\theta \in [0, 2\pi]} D\left(p(\mathbf{y}_{\text{sub}}|\{\bar{\theta}, \bar{r}\}) \parallel \tilde{p}(\mathbf{y}_{\text{sub}}|\theta)\right), \quad (16)$$

where $D\left(p(\mathbf{y}_{\text{sub}}|\{\bar{\theta}, \bar{r}\}) \parallel \tilde{p}(\mathbf{y}_{\text{sub}}|\theta)\right)$ denotes the KL divergence between the distributions $p(\mathbf{y}_{\text{sub}}|\{\bar{\theta}, \bar{r}\})$ and $\tilde{p}(\mathbf{y}_{\text{sub}}|\theta)$. Since $\hat{\theta}$ is an unbiased estimate of θ_{pse} , i.e., $\mathbb{E}\{\hat{\theta}\} = \theta_{\text{pse}}$, the CRLB of the unbiased estimate can be denoted as

$$\mathbb{E}_p\left\{\left(\hat{\theta} - \theta_{\text{pse}}\right)^2\right\} \geq \text{CRLB}(\theta_{\text{pse}}), \quad (17)$$

where $\mathbb{E}_p\{\cdot\}$ denotes taking the expectation operation, and $\text{CRLB}(\theta_{\text{pse}})$ is CRLB of the pseudo-true parameter θ_{pse} , which can be calculated as

$$\text{CRLB}(\theta_{\text{pse}}) = \sqrt{\mathbf{FIM}^{-1}(\theta_{\text{pse}})}. \quad (18)$$

The Fisher Information Matrix (FIM) $\mathbf{FIM}(\theta_{\text{pse}})$ can be calculated as

$$\mathbf{FIM}(\theta_{\text{pse}}) = \frac{2}{\sigma^2} \mathbb{E} \left\{ \frac{\partial \mathbf{y}_{\text{sub}}^{\text{H}}}{\partial \theta_{\text{pse}}} \frac{\partial \mathbf{y}_{\text{sub}}}{\partial \theta_{\text{pse}}} \right\}. \quad (19)$$

Substituting Eq. (12) into Eq. (19), the FIM matrix can be rewritten as

$$\mathbf{FIM}(\theta_{\text{pse}}) = \frac{2P|g|^2}{\sigma^2} \mathbb{E} \left\{ \frac{\partial \mathbf{a}^{\text{H}}}{\partial \theta_{\text{pse}}} \frac{\partial \mathbf{a}}{\partial \theta_{\text{pse}}} \right\} = \frac{2k_c^2 d^2 P |g|^2}{\sigma^2} \sum_{n=1}^{N_{\text{sub}}} \delta_n^2 = \frac{k_c^2 d^2 P |g|^2 N_{\text{sub}} (N_{\text{sub}} - 1)}{6\sigma^2}. \quad (20)$$

Hence, the CRLB of the pseudo-true parameter θ_{pse} can be

denoted as

$$\text{CRLB}(\theta_{\text{pse}}) = \frac{6\sigma^2}{k_c^2 d^2 P |g|^2 N_{\text{sub}} (N_{\text{sub}}^2 - 1)}. \quad (21)$$

Due to the non-negligible variation between the pseudo-true parameter θ_{pse} and the true parameter $\bar{\theta}$, the parameter estimation error lower bound of the true AoA $\bar{\theta}$ can be denoted as

$$E_p \left\{ (\hat{\theta} - \bar{\theta}) \right\} \geq \text{LB}(\bar{\theta}), \quad (22)$$

where $\text{LB}(\bar{\theta})$ is defined as an LB of the true AoA parameter $\bar{\theta}$, which can be rewritten as

$$\text{LB}(\bar{\theta}) = \text{CRLB}(\theta_{\text{pse}}) + (\theta_{\text{pse}} - \bar{\theta})^2, \quad (23)$$

where $(\hat{\theta} - \bar{\theta})^2 = (\hat{\theta} - \theta_{\text{pse}})^2 + (\theta_{\text{pse}} - \bar{\theta})^2 + 2(\hat{\theta} - \theta_{\text{pse}})(\theta_{\text{pse}} - \bar{\theta})$ and $E_p \left\{ (\hat{\theta} - \theta_{\text{pse}})(\theta_{\text{pse}} - \bar{\theta}) \right\} = E_p \left\{ (\hat{\theta} - \theta_{\text{pse}}) \right\} (\theta_{\text{pse}} - \bar{\theta}) = 0$ due to the unbiased estimate. Hence, in order to calculate the LB of the true AoA parameter precisely, the variation between the pseudo-true parameter and the true AoA parameter $\vartheta = \theta_{\text{pse}} - \bar{\theta}$ should be detected firstly.

To obtain the accurate value of ϑ , we need to find a pseudo-true parameter θ_{pse} that can minimize the KL divergence between the distributions under the realistic channel model and the simplified channel model, as shown in Eq. (16). However, this step has a high computational complexity due to the non-convex function expression. Fortunately, in order to minimize the KL divergence between $p(\mathbf{y}_{\text{sub}}, \{\bar{\theta}, \bar{r}\})$ and $\tilde{p}(\mathbf{y}_{\text{sub}}, \theta_{\text{pse}})$, the value of ϑ can also be obtained through minimizing the KL divergence between the realistic steering vector $\mathbf{w}_{\text{sub}}\{\bar{\theta}, \bar{r}\}$ and the simplified steering vector $\mathbf{a}_{\text{sim}}(\bar{\theta} + \vartheta)^{[34]}$, which can be expressed as

$$\vartheta = \arg \min_{\vartheta \in [0, 2\pi]} \left\| \mathbf{a}_{\text{sim}}(\bar{\theta} + \vartheta) - \mathbf{w}_{\text{sub}}(\bar{\theta}, \bar{r}) \right\|^2. \quad (24)$$

Let $\varpi(\vartheta) = \left\| \mathbf{a}_{\text{sim}}(\bar{\theta} + \vartheta) - \mathbf{w}_{\text{sub}}(\bar{\theta}, \bar{r}) \right\|^2$, then

$$\begin{aligned} \varpi(\vartheta) &= 2N_{\text{sub}} - 2 \sum_{n=1}^{N_{\text{sub}}} \cos \left(k_c (\delta_n d (\theta + \vartheta) + r_n - r) \right) \stackrel{(a)}{\approx} \\ &= 2N_{\text{sub}} - 2 \sum_{n=1}^{N_{\text{sub}}} \cos \left(k_c \delta_n d \left(\vartheta + \frac{\delta_n d (1 - \theta^2)}{2r} \right) \right) \stackrel{(b)}{\approx} \\ &= \sum_{n=1}^{N_{\text{sub}}} k_c^2 \delta_n^2 d^2 \left(\vartheta + \frac{\delta_n d (1 - \theta^2)}{2r} \right)^2 = \\ &= k_c^2 d^2 \sum_{n=1}^{N_{\text{sub}}} \delta_n^2 \left(\vartheta^2 + \frac{\delta_n^2 d^2 (1 - \theta^2)^2}{4r^2} + \vartheta \frac{\delta_n d (1 - \theta^2)}{r} \right) = \\ &= Q + k_c^2 d^2 \left(\sum_{n=1}^{N_{\text{sub}}} \delta_n^2 \vartheta^2 + \frac{d(1 - \theta^2)}{r} \sum_{n=1}^{N_{\text{sub}}} \delta_n^3 \vartheta \right), \end{aligned} \quad (25)$$

where (a) is derived by $\sqrt{1+x} \approx 1 + \frac{1}{2}x - \frac{1}{8}x^2$, i. e., $r_n \approx r - \delta_n d \theta + \frac{\delta_n^2 d^2 (1 - \theta^2)}{2r}$, and (b) is derived by $\cos x \approx 1 - \frac{1}{2}x^2$, where $x = k \delta_n d \left(\vartheta + \frac{\delta_n d (1 - \theta^2)}{2r} \right)$. Meanwhile, the constant Q can be denoted as $Q = \frac{k^2 d^4 (1 - 2\theta^2 + \theta^4)}{4r^2} \sum_{n=1}^{N_{\text{sub}}} \delta_n^4$. Eq. (25) can be rewritten as

$$\varpi(\vartheta) = A\vartheta^2 + 2B\vartheta + Q' = A \left(\vartheta + \frac{B}{A} \right)^2 + Q' - \frac{B^2}{A}, \quad (26)$$

where

$$A = k_c^2 d^2 \sum_{n=1}^{N_{\text{sub}}} \delta_n^2 = \frac{N_{\text{sub}} (N_{\text{sub}}^2 - 1) k_c^2 d^2}{12} > 0, \quad (27)$$

and

$$\begin{aligned} B &= \frac{k_c^2 d^3 (1 - \theta^2)}{2r} \sum_{n=1}^{N_{\text{sub}}} \delta_n^3 = \\ &= \frac{N_{\text{sub}} (N_{\text{sub}}^2 + 3N_{\text{sub}} + 3) k_c^2 d^3 (1 - \theta^2)}{16r}. \end{aligned} \quad (28)$$

Hence, in order to minimize $\varpi(\vartheta)$, the value of ϑ can be taken as the symmetry axis of the quadratic convex function in Eq. (26), i. e., $\vartheta = -B/A$, which can be calculated as

$$\vartheta = \frac{3(N_{\text{sub}}^2 + 3N_{\text{sub}} + 3)d(1 - \theta^2)}{4(N_{\text{sub}}^2 - 1)r}. \quad (29)$$

Finally, the LB of the true AoA parameter $\bar{\theta}$ can be calculated as

$$\text{LB}(\bar{\theta}) = \text{CRLB}(\theta_{\text{pse}}) + \vartheta^2. \quad (30)$$

By substituting Eqs. (21) and (29) into Eq. (30), Theorem 1 has been proved.

The LB of the AoA parameter is composed of two parts. The first part is the CRLB of the pseudo-true parameter estimate, which is inversely proportional to the signal-to-noise ratio (SNR) and N_{sub}^3 . The SNR is defined as

$$\text{SNR} = \frac{P |g|^2}{\sigma^2}. \quad (31)$$

The second part is the variance caused by the channel model mismatch, which is inversely proportional to r^2 . We can find that the AoA's LB mainly depends on the scale of the c-subarray, the SNR and the distance from the transmitter to the center of the XL-HRIS. The following lemmas can be obtained

via Eq. (15).

Lemma 1: When the scale of the c -subarray is large and tends towards infinity, but the distance between the transmitter and the XL-HRIS is finite, the CRLB of the pseudo-true parameter estimate will tend towards 0, while the value of LB mainly depends on the channel model mismatch and tends towards a non-zero constant value, which can be expressed as

$$\text{LB}_{N_{\text{sub}} \rightarrow \infty}(\theta) = \frac{9d^2(1 - 2\theta^2 + \theta^4)}{2r^2}. \quad (32)$$

Lemma 2: If the distance between the UE and the XL-HRIS is infinite, the variation caused by the channel model mismatch will tend towards 0, and the value of LB tends towards the CRLB of the pseudo-true parameter estimate, i.e., $\text{LB}_{r \rightarrow \infty}(\theta) = \text{CRLB}(\theta_{\text{pse}})$.

Lemma 3: With the increase of the transmit power, the LB tends towards a constant value, which can be expressed as

$$\text{LB}_P(\theta) = \left(1 + \frac{3N_{\text{sub}} + 4}{N_{\text{sub}} - 1}\right)^2 \frac{9d^2(1 - 2\theta^2 + \theta^4)}{2r^2}. \quad (33)$$

Note that the CRLB of the estimated pseudo-true parameter is inversely proportional to the SNR, while the variation caused by the channel model mismatch is independent of the SNR. Hence, with the transmit power increasing to infinity, the CRLB of the estimated pseudo-true parameter will tend towards 0, and the AoA parameter estimation accuracy is mainly limited by the channel model mismatch.

3.3 Impacts of Channel Model Mismatch and Array Scale on SE

To reduce computational complexity in the proposed XL-HRIS-aided system, the AoA parameter is estimated by the c -subarray, and then the distance parameter is detected through the receive signal of the r -subarray. Finally, the complex gain can be calculated through the least square (LS) method. By estimating the two separate channels between the UE/BS and the XL-HRIS, the cascade channel from the UE to the BS can be reconstructed accurately. The impacts of the c -subarray's scale on the SE are studied as follows.

When the XL-HRIS is in the passive mode, the BS receives the signal reflected by the XL-HRIS, which can be denoted as

$$\mathbf{y}_{\text{BS}} = \sqrt{P} \mathbf{w}_{\text{rec}}^H \mathbf{H}_{\text{BS}} \mathbf{\Omega} \mathbf{h}_{\text{R}} + \mathbf{n}_{\text{BS}}, \quad (34)$$

where $\mathbf{H}_{\text{BS}} \in \mathbb{C}^{M \times N}$ denotes the channel between the XL-HRIS and the BS; $\mathbf{n}_{\text{BS}} \in \mathbb{C}^{M \times 1}$ is the additional Gaussian white noise with 0 mean and σ^2 variance; $\mathbf{\Omega} = \text{diag}(e^{j\varphi_1}, \dots, e^{j\varphi_N})$ is the phase shift matrix of the XL-HRIS; φ_n follows the random distribution within $[0, 2\pi)$, for $n = 1, \dots, N$; $\mathbf{w}_{\text{rec}} \in \mathbb{C}^{M \times 1}$ denotes the beamforming vector. If the perfect CSI has been ob-

tained^[35-36], and the maximal ratio combining (MRC) receiver is adopted, i.e., $\mathbf{w}_{\text{rec}} = \frac{\mathbf{H}_{\text{BS}} \mathbf{\Omega} \mathbf{h}_{\text{R}}}{\|\mathbf{H}_{\text{BS}} \mathbf{\Omega} \mathbf{h}_{\text{R}}\|}$. Hence, the SE with the perfect CSI can be calculated as

$$C = \text{E}\{\log_2(1 + \text{SNR}_{\text{rec}})\}, \quad (35)$$

where the received SNR at the BS is defined as

$$\text{SNR}_{\text{rec}} = \frac{P \|\mathbf{H}_{\text{BS}} \mathbf{\Omega} \mathbf{h}_{\text{R}}\|^2}{\sigma^2}. \quad (36)$$

Nevertheless, the perfect CSI cannot be obtained in advance in the real communication system, and the CSI should be estimated through a channel estimation algorithm, which can lead to errors in the estimated channel parameters and a decrease in the SE. Considering a stable environment between the XL-HRIS and the BS, we assume that the perfect CSI of \mathbf{H}_{BS} has been obtained. Hence, only the impacts of the estimated parameters' errors in the channel between the UE and the XL-HRIS on the SE performance should be analyzed. To calculate the maximum SE when the estimation error exists (which will decrease the SE) in the proposed system, the minimum errors of the estimates are assumed as $\{\Delta\theta_{\text{min}}, \Delta r_{\text{min}}\} = \{\text{LB}(\theta), \text{CRLB}(r)\}$, where $\text{CRLB}(r)$ can be calculated as

$$\text{CRLB}(r) = \frac{2P |g|^2}{\sigma^2} \text{E}\left\{ \frac{\partial \mathbf{w}_{\text{R}}^H}{\partial r} \frac{\partial \mathbf{w}_{\text{R}}}{\partial r} \right\}. \quad (37)$$

The detailed expression of $\text{CRLB}(r)$ can be seen in Ref. [32]. Typically, a special scenario is considered, i.e., $\theta = 1$, and the receive signal is rewritten as

$$\mathbf{y} = \tilde{g} [e^{jk_r r_1}, \dots, e^{jk_r r_N}]^T, \quad (38)$$

where $\tilde{g} = g r e^{-jk_r r}$. Then, we have $r_n = |r - \chi_n d|$, and the $\text{CRLB}(r)$ can be denoted as

$$\text{CRLB}(r) = \sum_{n \in \Theta} \frac{2P |g|^2}{\sigma^2} \frac{1}{(r - \chi_n d)^2} \left(\frac{r^2}{(r - \chi_n d)^2} + k_c^2 \right), \quad (39)$$

where Θ denotes the vector constructed by the index of the elements in the c -subarray and the r -subarray. Meanwhile, in the special scenario $\theta = 1$, the estimation accuracy lower bound of the realistic parameter $\text{LB}(\theta)$ is equal to that of the pseudo-true parameter $\text{CRLB}(\theta_{\text{pse}})$, which can be expressed as

$$\text{LB}(\theta) = \text{CRLB}(\theta_{\text{pse}}) = \frac{6\sigma^2}{k_c^2 d^2 P |g|^2 N_{\text{sub}} (N_{\text{sub}}^2 - 1)}. \quad (40)$$

Hence, in the scenario $\theta = 1$, the estimation error lower bound will tend towards 0 when the transmitted signal power or the number of active elements at the c -subarray is infinite, and different from Eq. (15), the LB is independent of the distance parameter.

Then, the received SNR at the BS, which considers the estimation errors, can be rewritten as

$$\text{SNR}_{\text{rec}} = \frac{P \|\hat{\mathbf{w}}_{\text{rec}}^H \mathbf{H}_{\text{BS}} \boldsymbol{\Omega} \mathbf{h}_{\text{R}}\|^2}{\sigma^2}, \quad (41)$$

where the receiver can be denoted as

$$\hat{\mathbf{w}}_{\text{rec}} = \frac{\hat{g} \mathbf{H}_{\text{BS}} \boldsymbol{\Omega} \mathbf{w}_{\text{R}}(\theta + \Delta\theta_{\text{min}}, r + \Delta r_{\text{min}})}{\|\hat{g} \|\mathbf{H}_{\text{BS}} \boldsymbol{\Omega} \mathbf{w}_{\text{R}}(\theta + \Delta\theta_{\text{min}}, r + \Delta r_{\text{min}})\|}, \quad (42)$$

in which the complex gain \hat{g} can be calculated through the LS method as

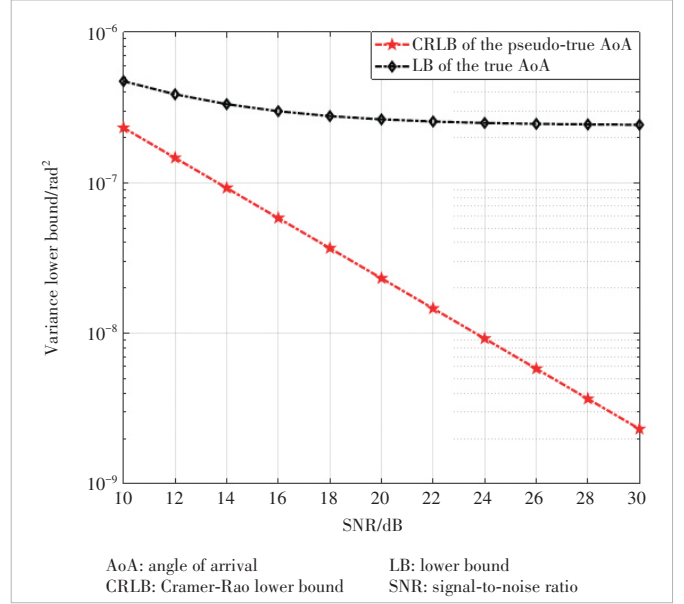
$$\hat{g} = \frac{[\hat{\mathbf{w}}_{\text{sub}}^H, \hat{\mathbf{w}}_{\text{ran}}^H][\mathbf{y}_{\text{sub}}^H, \mathbf{y}_{\text{ran}}^H]^H}{P \|\hat{\mathbf{w}}_{\text{sub}}^H, \hat{\mathbf{w}}_{\text{ran}}^H\|^2}, \quad (43)$$

where $\hat{\mathbf{w}}_{\text{sub}} = \mathbf{w}_{\text{sub}}(\theta + \Delta\theta_{\text{min}}, r + \Delta r_{\text{min}})$ and $\hat{\mathbf{w}}_{\text{ran}} = \mathbf{w}_{\text{ran}}(\theta + \Delta\theta_{\text{min}}, r + \Delta r_{\text{min}})$, respectively. Hence, the existence of channel model mismatch will increase the AoA parameter estimation error and thus damage the SE of the communication system, even in the scenario where the number of active elements at the c -subarray is infinite. Note that when the distance between the transmitter and the XL-HRIS is short, a lower CRLB of the estimated AoA parameter can be obtained, and the SE will further decrease.

4 Simulation Results

In this section, we evaluate the impacts of the channel model mismatch on the AoA parameter estimation error lower bound, as well as the impacts of the scale array on both the parameter estimation accuracy and communication performance of the proposed XL-HRIS-aided system. The parameters of the proposed system are set as follows: the carrier frequency is $f_c = 30$ GHz; the XL-HRIS has $N = 1024$ elements; the c -subarray has $N_{\text{sub}} = 64$ active elements while the r -subarray has $N_{\text{ran}} = 64$ active elements; the SNR is set as 10 dB, and the angle and distance are randomly generated within $\theta \in [0, 2\pi)$ and $r \in (10, 50)$ meters, respectively.

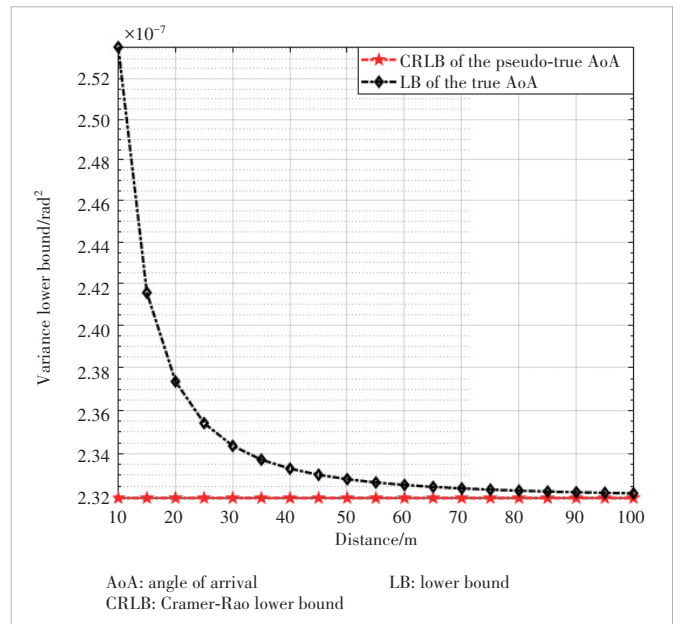
The LB of the AoA parameter is evaluated first as shown in Figs. 3, 4 and 5, where the distance between the transmitter and the center of the ULA is set as 3 m, and the azimuth angle is set as $\pi/4$. In Fig. 3, the LB of the AoA parameter and the CRLB of the pseudo-true AoA parameter with the SNR are highlighted. It is observed that the channel model mismatch leads to an increase in the parameter estimation LB, which fur-



▲ Figure 3. LB of the AoA parameter

ther decreases the estimation accuracy of channel estimation. With the increase of SNR, the CRLB of the pseudo-true parameter tends towards 0, but the LB approaches a non-zero constant, which demonstrates that the estimation error caused by the channel model mismatch will not disappear even the transmitted power is infinite. Hence, for the parameter estimation algorithm based on the planar wave channel model, further increased power will not bring significant performance improvement if the transmitted power exceeds a threshold.

In Fig. 4, the LB and CRLB are plotted against the distance between the transmitter and the XL-HRIS. With the increase



▲ Figure 4. LB versus the distance

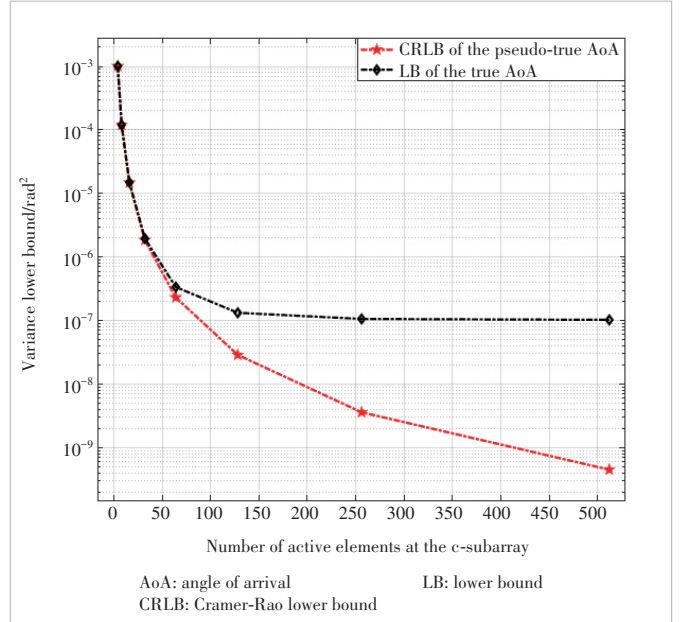
of distance, the error caused by the channel model mismatch tends towards 0, and the curves for LB and CRLB begin to coincide, because the variation of the realistic channel model and the simplified channel model can be ignored when the transmitter is far from the c -subarray, and only little difference exists between the spherical wave model and the planar wave model. However, if the distance is short, the channel model mismatch will significantly increase the LB. Therefore, ignoring the channel model mismatch will only result in a decreased estimation performance.

In Fig. 5, the LB and CRLB are plotted against the number of active elements at the c -subarray. In accordance with Eq. (15), while the CRLB of the pseudo-true AoA parameter tends towards 0 when the number of active elements at the c -subarray grows to infinite, the accuracy of the estimated AoA parameter is still limited by the channel model mismatch and tends towards a non-zero constant. This verifies that the accuracy of the estimated AoA parameter still has a non-zero lower bound even in communication systems with extremely large-scale arrays. Hence, as the number of elements at the whole array increases, the channel model mismatch must be considered to enhance the estimation accuracy and improve communication performance.

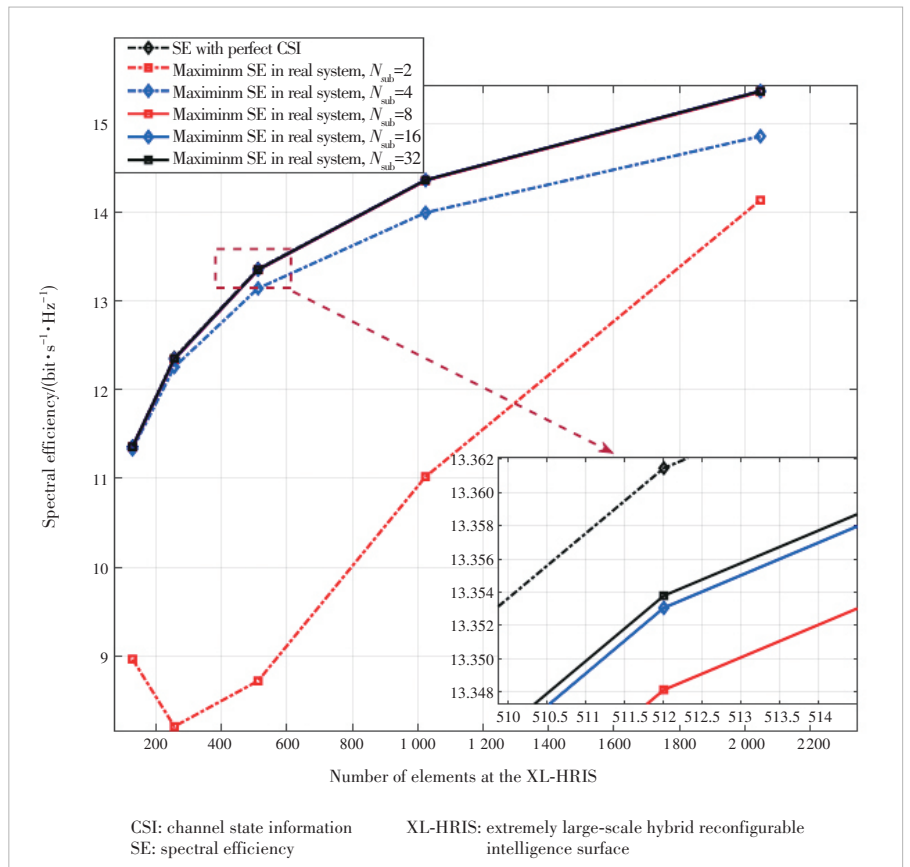
Fig. 6 evaluates the SE of the proposed system with various array scale settings. The distance between the transmitter and the XL-HRIS is $r = 5$ m, and the AoA is $\theta = 0.5$. We can find that the obtained maximum SE, which considers the estimation error in real communication systems, is far lower than that of the perfect CSI scenario when the number of active elements at the c -subarray is small, because the unavoidable error is caused by the channel model mismatch as Eq. (15). Fortunately, with the increase of the array scale of the c -subarray, a more accurate estimate of AoA can be obtained, as shown in Fig. 5, and thus the SE will be improved, which verifies the benefits of the extremely large-scale array for future 6G wireless communication systems. However, as the number of active elements at the c -subarray grows to infinity, the obtained maximum SE is always lower than the SE with the perfect CSI. That is because, when the scale of the c -subarray is infinite, the LB of the estimated AoA tends towards a non-zero constant inversely proportional to r^2 as Eq. (15), which can damage the SE of the proposed system when the distance between the transmitter and the XL-HRIS is short.

5 Conclusions

In this paper, an XL-HRIS-aided system is introduced with a proposed near-field channel estimation algorithm. With the



▲ Figure 5. LB versus the number of active elements at the c -subarray



▲ Figure 6. SE with the array scale

increasing number of antennas and the higher frequency band for 6G wireless communication system, the variations among the realistic spherical wave channel model and the simplified planar wave channel model cannot be ignored. This paper studies the accuracy degradation of estimated parameters caused by the channel model mismatch and provides a closed-form solution of the LB. The LB with system parameters is also studied, which verifies that the LB will tend towards a non-zero constant even though the number of active elements at the c -subarray or the SNR is infinite. Moreover, the impacts of the array scale of the XL-HRIS on the SE is evaluated, which proves that the extremely large array can indeed enhance the SE and further improves the system communication performance.

References

- [1] MATTHAIYOU M, YURDUSEVEN O, NGO H Q, et al. The road to 6G: ten physical layer challenges for communications engineers [J]. *IEEE communications magazine*, 2021, 59(1): 64 - 69. DOI: 10.1109/MCOM.001.2000208
- [2] YANG H H, CAO X Y, YANG F, et al. A programmable metasurface with dynamic polarization, scattering and focusing control [J]. *Scientific reports*, 2016, 6: 35692. DOI: 10.1038/srep35692
- [3] ZHANG S W, ZHANG R. Capacity characterization for intelligent reflecting surface aided MIMO communication [J]. *IEEE journal on selected areas in communications*, 2020, 38(8): 1823 - 1838. DOI: 10.1109/JSAC.2020.3000814
- [4] DI RENZO M, ZAPPONE A, DEBBAH M, et al. Smart radio environments empowered by reconfigurable intelligent surfaces: how it works, state of research, and the road ahead [J]. *IEEE journal on selected areas in communications*, 2020, 38(11): 2450 - 2525. DOI: 10.1109/JSAC.2020.3007211
- [5] HUANG C, ZAPPONE A, ALEXANDROPOULOS G C, et al. Reconfigurable intelligent surfaces for energy efficiency in wireless communication [EB/OL]. (2018-10-06) [2019-01-10]. <https://arxiv.org/abs/1810.06934>
- [6] WU Q Q, ZHANG R. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming [J]. *IEEE transactions on wireless communications*, 2019, 18(11): 5394 - 5409. DOI: 10.1109/TWC.2019.2936025
- [7] HAN Y, TANG W K, JIN S, et al. Large intelligent surface-assisted wireless communication exploiting statistical CSI [J]. *IEEE transactions on vehicular technology*, 2019, 68(8): 8238 - 8242. DOI: 10.1109/TVT.2019.2923997
- [8] TANG W K, CHEN M Z, CHEN X Y, et al. Wireless communications with reconfigurable intelligent surface: path loss modeling and experimental measurement [J]. *IEEE transactions on wireless communications*, 2021, 20(1): 421 - 439. DOI: 10.1109/TWC.2020.3024887
- [9] CHEN W C, WEN C-K, LI X, et al. Channel customization for limited feedback in RIS-assisted FDD systems [J]. *IEEE transactions on wireless communications*, 2023, 22(7): 4505 - 4519. DOI: 10.1109/TWC.2022.3226442
- [10] MU X D, LIU Y W, GUO L, et al. Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications [J]. *IEEE transactions on wireless communications*, 2022, 21(5): 3083 - 3098. DOI: 10.1109/TWC.2021.3118225
- [11] WEI L, HUANG C W, ALEXANDROPOULOS G C, et al. Channel estimation for RIS-empowered multi-user MISO wireless communications [J]. *IEEE transactions on communications*, 2021, 69(6): 4144 - 4157. DOI: 10.1109/TCOMM.2021.3063236
- [12] LIU C, LIU X M, NG D W K, et al. Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications [J]. *IEEE transactions on wireless communications*, 2022, 21(2): 898 - 912. DOI: 10.1109/TWC.2021.3100148
- [13] WEI L, HUANG C W, GUO Q H, et al. Joint channel estimation and signal recovery for RIS-empowered multiuser communications [J]. *IEEE transactions on communications*, 2022, 70(7): 4640 - 4655. DOI: 10.1109/TCOMM.2022.3179771
- [14] JIN Y, ZHANG J Y, ZHANG X D, et al. Channel estimation for semi-passive reconfigurable intelligent surfaces with enhanced deep residual networks [J]. *IEEE transactions on vehicular technology*, 2021, 70(10): 11083 - 11088. DOI: 10.1109/TVT.2021.3109937
- [15] HU J F, YIN H F, BJÖRNSON E. MmWave MIMO communication with semi-passive RIS: a low-complexity channel estimation scheme [C]// *Proc. IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021: 1 - 6. DOI: 10.1109/GLOBECOM46510.2021.9685434
- [16] LI M Y, ZHANG S, GE Y, et al. Joint channel estimation and data detection for hybrid RIS aided millimeter wave OTFS systems [J]. *IEEE transactions on communications*, 2022, 70(10): 6832 - 6848. DOI: 10.1109/TCOMM.2022.3199019
- [17] ALEXANDROPOULOS G C, VLACHOS E. A hardware architecture for reconfigurable intelligent surfaces with minimal active elements for explicit channel estimation [C]// *Proc. 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020: 9175 - 9179. DOI: 10.1109/ICASSP40776.2020.9053976
- [18] ZHANG H Y, SHLEZINGER N, ALAMZADEH I, et al. Channel estimation with simultaneous reflecting and sensing reconfigurable intelligent metasurfaces [C]// *Proc. IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2021: 536 - 540. DOI: 10.1109/SPAWC51858.2021.9593172
- [19] SCHROEDER R, HE J G, BRANTE G, et al. Two-stage channel estimation for hybrid RIS assisted MIMO systems [J]. *IEEE transactions on communications*, 2022, 70(7): 4793 - 4806. DOI: 10.1109/TCOMM.2022.3176654
- [20] TAHA A, ALRABEIAH M, ALKHATEEB A. Enabling large intelligent surfaces with compressive sensing and deep learning [J]. *IEEE access*, 2021, 9: 44304 - 44321. DOI: 10.1109/ACCESS.2021.3064073
- [21] YANG S J, LYU W T, WANG D L, et al. Separate channel estimation with hybrid RIS-aided multi-user communications [J]. *IEEE transactions on vehicular technology*, 2023, 72(1): 1318 - 1324. DOI: 10.1109/TVT.2022.3205370
- [22] HEATH R W, GONZÁLEZ-PRELCIC N, RANGAN S, et al. An overview of signal processing techniques for millimeter wave MIMO systems [J]. *IEEE journal of selected topics in signal processing*, 2016, 10(3): 436 - 453. DOI: 10.1109/JSTSP.2016.2523924
- [23] FANG J, LI X J, LI H B, et al. Low-rank covariance-assisted downlink training and channel estimation for FDD massive MIMO systems [J]. *IEEE transactions on wireless communications*, 2017, 16(3): 1935 - 1947. DOI: 10.1109/TWC.2017.2657513
- [24] HAN Y, HSU T H, WEN C K, et al. Efficient downlink channel reconstruction for FDD transmission systems [C]// *Proc. 27th Wireless and Optical Communication Conference (WOCC)*. IEEE, 2018: 1 - 5. DOI: 10.1109/WOCC.2018.8372727
- [25] HAN Y, HSU T H, WEN C K, et al. Efficient downlink channel reconstruction for FDD multi-antenna systems [J]. *IEEE transactions on wireless communications*, 2019, 18(6): 3161 - 3176. DOI: 10.1109/TWC.2019.2911497
- [26] HAN Y, LIU Q, WEN C-K, et al. Tracking FDD massive MIMO downlink channels by exploiting delay and angular reciprocity [J]. *IEEE journal of selected topics in signal processing*, 2019, 13(5): 1062 - 1076. DOI: 10.1109/JSTSP.2019.2935320

- [27] SELVAN K T, JANASWAMY R. Fraunhofer and Fresnel distances: unified derivation for aperture antennas [J]. *IEEE antennas and propagation magazine*, 2017, 59(4): 12 – 15. DOI: 10.1109/MAP.2017.2706648
- [28] SHI X, WANG J, SUN Z, et al. Spatial-chirp codebook-based hierarchical beam training for extremely large-scale massive MIMO [EB/OL]. (2022-10-07) [2023-08-15]. <https://arxiv.org/abs/2210.03345>
- [29] SOUTHWELL W H. Validity of the Fresnel approximation in the near field [J]. *Journal of the optical society of America*, 1981, 71(1): 7. DOI: 10.1364/josa.71.000007
- [30] CUI M Y, DAI L L. Channel estimation for extremely large-scale MIMO: far-field or near-field? [J]. *IEEE transactions on communications*, 2022, 70(4): 2663 – 2677. DOI: 10.1109/TCOMM.2022.3146400
- [31] YANG J, ZENG Y, JIN S, et al. Communication and localization with extremely large lens antenna array [J]. *IEEE transactions on wireless communications*, 2021, 20(5): 3031 – 3048. DOI: 10.1109/TWC.2020.3046766
- [32] LU Z Z, HAN Y, JIN S, et al. Near-field channel reconstruction and user localization for ELAA systems [C]//Proc. International Symposium on Wireless Communication Systems (ISWCS). IEEE, 2022: 1 – 6. DOI: 10.1109/ISWCS56560.2022.9940362
- [33] MAMANDIPOOR B, RAMASAMY D, MADHOW U. Newtonized orthogonal matching pursuit: frequency estimation over the continuum [J]. *IEEE transactions on signal processing*, 2016, 64(19): 5066 – 5081. DOI: 10.1109/TSP.2016.2580523
- [34] OZTURK C, KESKIN M F, WYMEERSCH H, et al. RIS-aided near-field localization under phase-dependent amplitude variations [J]. *IEEE transactions on wireless communications*, 2023, 22(8): 5550 – 5566. DOI: 10.1109/TWC.2023.3235306
- [35] FENG K M, LI X, HAN Y, et al. Joint beamforming optimization for reconfigurable intelligent surface-enabled MISO-OFDM systems [J]. *China communications*, 2021, 18(3): 63 – 79. DOI: 10.23919/JCC.2021.03.006
- [36] LU Z, HAN Y, JIN S, et al. Near-filed localization and channel estimation for ELAA systems [J]. *IEEE transactions on wireless communications*, 2023, Early Access. DOI: 10.1109/TWC.2023.3336328

Biographies

LU Zhizheng received his BS degree in communications engineering from Nanjing University of Science and Technology, China in 2021. Currently, he is pursuing his PhD degree in information and communications engineering from Southeast University, China. His research interests include extremely large-scale multiple-input multiple-output, near-field channel estimation, and hybrid reconfigurable intelligence surface.

HAN Yu (hanyu@seu.edu.cn) received her BS degree in communications engineering from Hangzhou Dianzi University, China in 2012, and MS and PhD degrees in information and communications engineering from Southeast University, China in 2015 and 2020, respectively. She was a postdoctoral fellow with Singapore University of Technology and Design, Singapore till 2022. Currently, she is an associate professor with Southeast University. Her research interests include extra large-scale MIMO and reconfigurable intelligent surface.

JIN Shi (jinshi@seu.edu.cn) received his BS degree in communications engineering from Guilin University of Electronic Technology, China in 1996, MS degree from Nanjing University of Posts and Telecommunications, China in 2003, and PhD degree in information and communications engineering from Southeast University, China in 2007. From June 2007 to October 2009, he was a research fellow with the Adastral Park Research Campus, University College London, UK. He is currently with the National Mobile Communications Research Laboratory, Southeast University. His research interests include wireless communications, random matrix theory, and information theory. He is serving as an area editor for the *Transactions on Communications* and *IET Electronics Letters*. He was an associate editor for the *IEEE Transactions on Wireless Communications*, *IEEE Communications Letters*, and *IET Communications*. Dr. JIN and his co-authors have been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory, the IEEE Vehicular Technology Society 2023 Jack Neubauer Memorial Award, a 2022 Best Paper Award and a 2010 Young Author Best Paper Award by the IEEE Signal Processing Society. He is an IEEE fellow.

Degree of Freedom Analysis for Holographic MIMO Based on a Mutual-Coupling-Compliant Channel Model



SUN Yunqi^{1,2}, JIAN Mengnan^{1,2}, YANG Jun^{1,2},
ZHAO Yajun^{1,2}, CHEN Yijian^{1,2}

(1. State Key Laboratory of Mobile Network and Mobile Multimedia
Technology, Shenzhen 518055, China;
2. ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTECOM.202401005

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240307.1433.002>, published
online March 9, 2024

Manuscript received: 2023-12-14

Abstract: Degree of freedom (DOF) is a key indicator for spatial multiplexing layers of a wireless channel. Traditionally, the channel of a multiple-input multiple-output (MIMO) half-wavelength dipole array has a DOF that equals the antenna number. However, recent studies suggest that the DOF could be less than the antenna number when strong mutual coupling is considered. We utilize a mutual-coupling-compliant channel model to investigate the DOF of the holographic MIMO (HMIMO) channel and give an upper bound of the DOF with strong mutual coupling. Our numerical simulations demonstrate that a dense array can support more DOF per unit aperture as compared with a half-wavelength MIMO system.

Keywords: channel model; degree of freedom; holographic MIMO; mutual coupling

Citation (Format 1): SUN Y Q, JIAN M N, YANG J, et al. Degree of freedom analysis for holographic MIMO based on a mutual-coupling-compliant channel model [J]. *ZTE Communications*, 2024, 22(1): 34 – 40. DOI: 10.12142/ZTECOM.202401005

Citation (Format 2): Y. Q. Sun, M. N. Jian, J. Yang, et al., “Degree of freedom analysis for holographic MIMO based on a mutual-coupling-compliant channel model,” *ZTE Communications*, vol. 22, no. 1, pp. 34 – 40, Mar. 2024. doi: 10.12142/ZTECOM.202401005.

1 Introduction

Holographic MIMO (HMIMO) is a new concept of communication technology, which has attracted increasing interest from researchers^[1-5]. There are two major types of HMIMO systems. One of them is communicating through a transmitting array with possibly uncountable antennas^[6]. The other refers to collecting and forming the phase information through the holographic principle^[7-9]. Both types of HMIMO communication systems offer increased capability for electromagnetic field manipulation, therefore indicating an increase in the degree of freedom (DOF) of the spatial channel.

The DOF can be defined in three ways: the rank definition, the spatial multiplexing gain definition, and the bandwidth definition. The rank definition defines the rank of the MIMO channel matrix as the DOF of the system^[10]. The spatial multiplexing gain definition is the limit of the channel capacity over the logarithmic signal-to-noise ratio (SNR) as the SNR grows to infinity^[11]. This definition is based on the fact that the

MIMO channel capacity grows linearly with logarithmic SNR and the smaller number of antennas between the transmitting and receiving side when the SNR is large enough. The bandwidth definition is a direct expansion of the time-frequency bandwidth by applying the spatial Fourier transform to the spatial domain^[12-14]. The conjugate domain is the so-called wavenumber domain, which is band-limited because the radiating electromagnetic waves are bounded by their vacuum wavenumber, such that,

$$\mathbf{k} = \mathbf{k}_x + \mathbf{k}_y + \mathbf{k}_z, \quad (1)$$

$$k^2 = k_x^2 + k_y^2 + k_z^2, \quad (2)$$

where \mathbf{k} is the wavevector of the electromagnetic wave, and $k = |\mathbf{k}|$; \mathbf{k}_x , \mathbf{k}_y , and \mathbf{k}_z are the wavevectors in xyz directions, $k_x = |\mathbf{k}_x|$, $k_y = |\mathbf{k}_y|$, and $k_z = |\mathbf{k}_z|$.

The spatial bandwidth is defined as the area in the wavenumber domain supported by the possible wavevectors satisfying the restrictions. Related research shows that this restriction forms a circular area denoted by $k_x^2 + k_y^2 \leq k^2$, which is essentially the propagating modes of the electromagnetic waves

This work was supported in part by National Key Research and Development Program of China under Grant No. 2020YFB1807600.

and is clearly band-limited.

Recent DOF analysis can be conducted with a wavenumber domain HMIMO model for the spatial bandwidth definition^[6, 14]. The DOF of the wavenumber model is defined by the number of discrete plane wave modes, which gives a desirable approximation in the wavenumber domain. Since the HMIMO channel in the wavenumber domain is band-limited, the DOF of the HMIMO channel is defined for a uniform linear array (ULA) and uniform planar array (UPA) respectively:

$$\eta_1 = \frac{2}{\lambda} l, \text{ for ULA with length } l, \quad (3)$$

$$\eta_2 = \frac{\pi}{\lambda^2} l_0 l_1, \text{ for UPA with size } l_0 \times l_1, \quad (4)$$

where η_1 and η_2 stand for the DOF of the one-dimensional and two-dimensional HMIMO transmitting arrays, and λ is the wavelength.

Recent DOF analysis of HMIMO can also follow the rank definition. For example, the dyadic Green function HMIMO model studies the physical nature of the radiation and makes use of the dyadic Green function to determine the vector solution of the spatial electrical field. Due to the vector nature of the electrical field, researchers claim that HMIMO could support three independent orthogonal polarizations^[15], thereby increasing the size of the channel matrix and the DOF.

Moreover, researchers have also studied the influence of mutual coupling on the HMIMO channel DOF. The close antenna spacing and the continuous current surface are very popular among the current configuration of HMIMO antennas. Such configuration would inevitably cause the single port transmitting signal to form a current distribution on the transmitting surface because of electromagnetic laws, which is the physical illustration of mutual coupling. To study the influence of the mutual coupling on the rank of the channel matrix, an effective DOF was proposed to evaluate the channel performance in Ref. [16], which is

$$\eta_e = \frac{\left(\sum_i \sigma_i\right)^2}{\sum_i \sigma_i^2}, \quad (5)$$

where σ_i is the i -th singular value of the channel matrix. The mutual coupling can also be included into the DOF analysis by modifying the spatial correlation matrix^[17]. This model successfully reveals the effective DOF of HMIMO under mutual coupling from the transmitting side.

While there are many studies on HMIMO, it is still important to establish a mutual coupling compliant channel model and examine the impact of mutual coupling on the HMIMO channel. In this paper, we focus on the DOF analysis for HMIMO with strong mutual coupling for both transmitting and

receiving sides, and reveal the potential DOF of HMIMO for communicating with sufficient receiving antennas. In Section 2, we introduce the mutual-coupling-compliant channel model of a communication system with strong mutual coupling in its transmitting array. In Section 3, we introduce the DOF analysis method based on popular HMIMO models and highlight several novel DOF features observed by the mutual-coupling-compliant model. Finally, we summarize and discuss the DOF analysis under mutual coupling and its potential application in the last section.

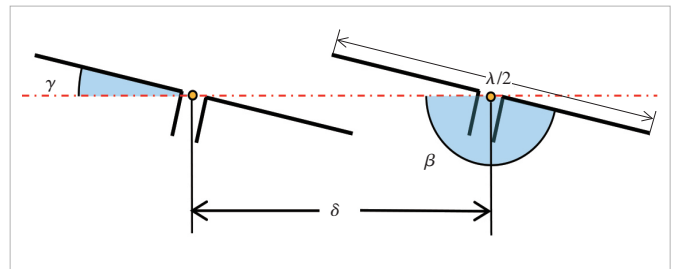
2 DoF Analysis Based on a Mutual-Coupling-Compliant Channel Model

2.1 Circuit Model for HMIMO

The channel model is one of the most fundamental problems in studying HMIMO communication. In an HMIMO system, a tightly-coupled transmitting array (Tx) and a weakly-coupled receiving array (Rx) communicate at a central frequency f_0 . We assume that the Tx and Rx are parallel to each other, and the distance between them is much greater than the wavelength.

We utilize a mutual-coupling-compliant channel model proposed by AKROUT et al.^[18] to analyze the DOF features in the HMIMO system. The basic concept of the circuit model is to simplify the complicated electromagnetic problem to a simple current-voltage equation, which is Ohm's law^[19]. The circuit model introduces an impedance to describe the radiation efficiency and the path loss simultaneously. And the impedance is calculated by the vector potential method with a simplified boundary conditions of the antenna's current. For common MIMO systems, the antennas are spaced far enough so that the mutual coupling is negligible, resulting in a diagonal impedance matrix. However, for HMIMO systems, the close spacing of the antennas introduces the strong mutual coupling inside the array and introduces off-diagonal entries in the impedance matrix. For instance, we consider two half-wavelength dipoles radiating at wavelength $\lambda = cf_0$, where c is the speed of light, and the antenna spacing is denoted by δ . The configuration of the two dipoles is shown in Fig. 1.

The impedance of two antennas defined as Z_{TR} is given



▲ Figure 1. Configuration of two dipoles. The two dipoles are separated by δ at their center, and the tilt angles with respect to the connecting line of their centers are γ and β (denoted as the blue area in the figure)

by an electrical-law-driven model denoted by $M(\delta, \beta, \gamma)$, such that,

$$\begin{aligned} Z_{TR} = & -3\sqrt{R[Z_R]R[Z_T]} \left[\frac{1}{2} \sin\beta \sin\gamma \left(\frac{1}{jk_0\delta} + \frac{1}{(jk_0\delta)^2} + \right. \right. \\ & \left. \left. \frac{1}{(jk_0\delta)^3} \right) + \cos\beta \cos\gamma \left(\frac{1}{jk_0\delta} + \frac{1}{(jk_0\delta)^2} + \frac{1}{(jk_0\delta)^3} \right) \right] \\ & \exp(-jk_0\delta) \triangleq M(\delta, \beta, \gamma), \end{aligned} \quad (6)$$

where $k_0 = 2\pi/\lambda$ is the wavenumber in vacuum and j is the imaginary unit; β and γ are the tilting angles mentioned in Fig. 1. Note that Z_T and Z_R are the self-impedance of the transmitting and receiving antennas and can be easily evaluated by a classical antenna impedance theory from CHU^[20]. Theoretically, the impedance matrix of the Tx and Rx arrays can be determined by calculating all the mutual-impedance between antennas, forming a Tx self-impedance matrix Z_T and Rx self-impedance matrix Z_R . By altering the spacing into the mutual coupling region (i.e. $\delta < \lambda/2$), this model can be extended to an HMIMO system where the Tx's strong mutual coupling should be considered.

2.2 Spatial Correlation of Impedance Matrix

We employ the rank definition to extract the DOF property of the circuit model for HMIMO. The channel model follows the work by AKROUT et al.^[18] The explicit expression of the channel matrix is given by

$$H(f) = \beta_0 R_{in} P Z_{RT} Q, \quad (7)$$

where β_0 is a readout amplification coefficient, R_{in} is the internal resistance of the receiving antenna, and P and Q are two auxiliary impedance matrices defined as

$$P \triangleq (Z_R + R_{in} I_{M \times M})^{-1}, \quad (8)$$

$$Q \triangleq (Z_T + R I_{N \times N} - Z_{TR} P Z_{RT})^{-1}, \quad (9)$$

where $I_{M \times M}$ and $I_{N \times N}$ are identity matrices with M and N rows respectively.

Note that the circuit model decomposes the channel matrix into a product of the self-impedance matrices (Z_T and Z_R) and the trans-impedance matrix (Z_{TR}). The rank of the channel matrix is then determined by the minimum rank of P , Z_{RT} , and Q . Recall the basic assumptions of HMIMO system that the Rx has weak coupling. Therefore, Z_R is a diagonal matrix and of full rank. To study the properties of Z_T and Z_{TR} , we make another assumption that the differences in the current distribution caused by the feed position are negligible. The configuration of the array indicates that the mutual impedance model $M(\delta, \beta, \gamma)$ has two fixed components β and γ . If we set $\beta = \pi$

and $\gamma = 0$, the entries of Z_T is

$$Z_{T,mn} \propto - \left(\frac{1}{(jk_0\delta_{mn})^2} + \frac{1}{(jk_0\delta_{mn})^3} \right) \exp(-jk_0\delta_{mn}), \quad (10)$$

where $k_0 = 2\pi/\lambda$ is the wavenumber in vacuum. The impedance approaches 0 asymptotically if the spacing is bigger than $\lambda/2$, which is the condition of a traditional weak coupling array.

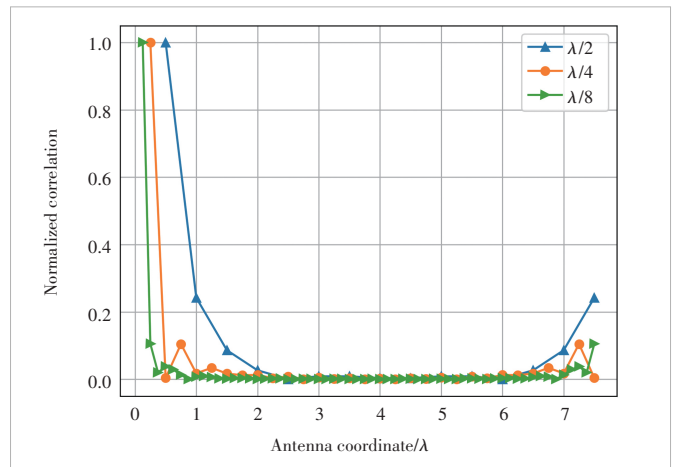
We then derive a formula for estimating the rank of the impedance matrix Z_T for a ULA. Since the impedance matrix is only a function of the distance between the i -th and j -th antennas, the row vectors of Z_T have cyclic symmetry. Therefore, the rank of Z_T is bounded by the auto-correlation of the row vectors. The auto-correlation of the first row vector Z_{T0} is given by

$$\begin{aligned} \Re[Z_{T0} Z_{T0}^*] & \propto \\ & \sum_{m \neq 0} \sum_{n \neq 0} \left(\frac{1}{(jk_0\delta_{0m})^2} + \frac{1}{(jk_0\delta_{0m})^3} \right) \left(\frac{1}{(jk_0\delta_{0n})^2} + \frac{1}{(jk_0\delta_{0n})^3} \right) \\ & \exp(-jk_0(\delta_{0m} + \delta_{0n})), \end{aligned} \quad (11)$$

where δ_{0m} denotes the distance between the first and m -th antenna.

The numerical evaluation (Fig. 2) shows that the mutual coupling increases the spatial correlation between the exciting antenna and its neighboring antenna. This is an exact answer to the DOF degradation for fixed antenna numbers when the mutual coupling is considered. This is in agreement with the spatial correlation matrix reported by YUAN et al.^[17]

However, increasing the number of antennas within a fixed aperture would provide additional DOF compared to the half-wavelength weak coupling array due to the rapid decrease of the correlation for antennas with spacing bigger than $\lambda/2$. The asymptotic DOF provided by the mutual coupling for ULA can be estimated by subtracting the $\lambda/2\delta$ antennas from the total



▲ Figure 2. Spatial correlation for uniform linear array (ULA) with spacing $\lambda/2$, $\lambda/4$ and $\lambda/8$. The correlation is normalized to their corresponding maximum

antennas, which is $\eta \leq \lfloor N - \lambda/\delta \rfloor$, where N is the number of the antennas in the ULA with spacing δ . The same approach can be applied to the UPA, yielding a DOF of $\eta \leq \lfloor N - (\pi\lambda^2)/(4\delta^2) \rfloor$.

3 Numerical Simulations of DOF

3.1 ULA and UPA Configuration

We use the circuit model to analyze the DOF of ULA and UPA with mutual coupling under a Rayleigh scattering environment. The channel DOF is defined as the significant singular values of the channel matrix. The configurations of ULA and UPA can be found in Fig. 3.

As shown in Fig. 3, the two tilt angles should meet $\beta + \gamma = \pi$, where γ should avoid $n\pi/2$ to make the tight-coupled dipoles physically feasible. Note that the spacing δ , the tilt angle γ , and the distance D fully determine the array configuration.

When considering the NLoS scenario with the Rayleigh fading model, the trans-impedance matrix can be given by

$$\mathbf{Z}_{RT}^{\text{Rayleigh}} = \frac{c}{2\pi f D^{\alpha/2}} \Re[\mathbf{Z}_R(f)]^{\frac{1}{2}} \mathbf{F} \Re[\mathbf{Z}_T(f)]^{\frac{1}{2}}, \quad (12)$$

where α is the path loss parameter, and \mathbf{F} is a small-scale fading matrix with $\mathbf{F}_{ij} \sim CN(0, 1/2)$.

3.2 Numerical DOF Analysis

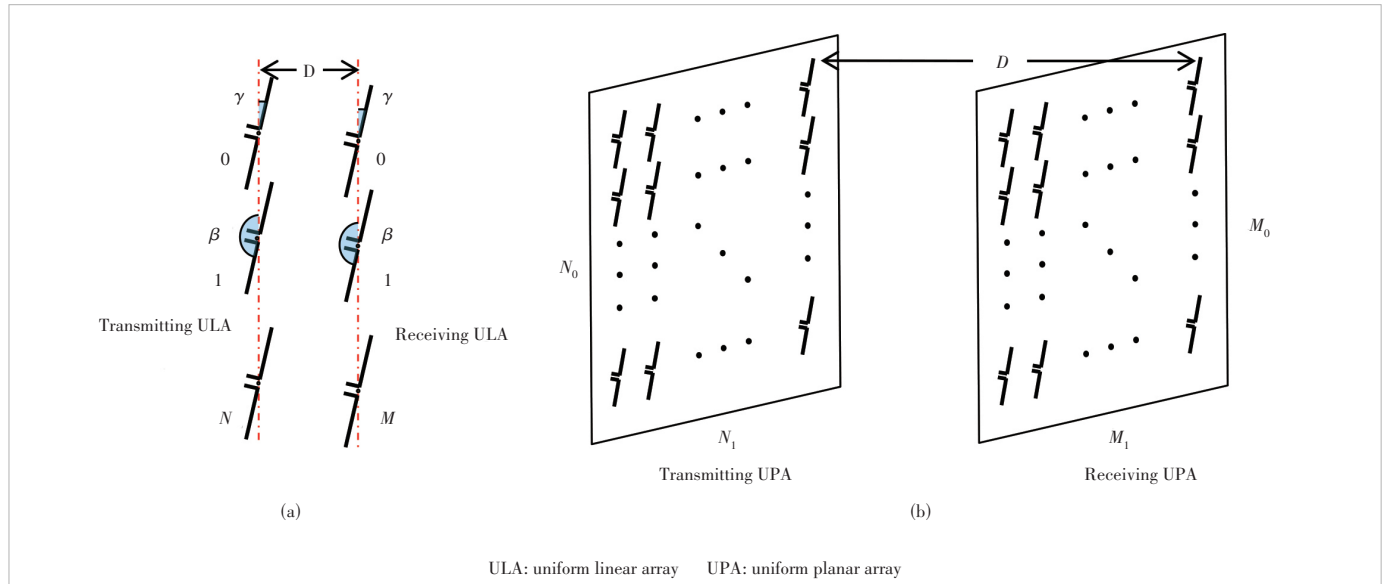
We define the rank of the channel matrix as the DOF of the system. We aim to investigate the DOF of the ULA and UPA

communication system under various scenarios. These scenarios consider the near-field and the far-field, the antenna spacing, and the distance between the Tx and Rx respectively. These scenarios are listed in Table 1 with indexes 1 – 16. Note that we only change the antenna numbers in the transmitting array and fix the receiving ULA and UPA with 16 half-wavelength antennas and 4×4 half-wavelength antennas respectively. To study the statistical properties of

▼Table 1. ULA and UPA configurations for mutual coupling simulation

Configuration Index	Configuration Type	Distance/ λ	Spacing/ λ	Antenna Number	Aperture Area
1	ULA_FF	200	0.5	16	7.5λ
2	ULA_NF	10	0.5	16	7.5λ
3	ULA_FF_MC	200	0.25	16	3.75λ
4	ULA_NF_MC	10	0.25	16	3.75λ
5	ULA_FF_DEN	200	0.25	31	7.5λ
6	ULA_NF_DEN	10	0.25	31	7.5λ
7	ULA_FF_UDEN	200	0.1	76	7.5λ
8	ULA_NF_UDEN	200	0.1	76	7.5λ
9	UPA_FF	200	0.5	16	$1.5\lambda \times 1.5\lambda$
10	UPA_NF	10	0.5	16	$1.5\lambda \times 1.5\lambda$
11	UPA_FF_MC	200	0.25	16	$0.75\lambda \times 0.75\lambda$
12	UPA_NF_MC	10	0.25	16	$1.5\lambda \times 1.5\lambda$
13	UPA_FF_DEN	200	0.25	49	$1.5\lambda \times 1.5\lambda$
14	UPA_NF_DEN	10	0.25	49	$1.5\lambda \times 1.5\lambda$
15	UPA_FF_UDEN	200	0.1	256	$1.5\lambda \times 1.5\lambda$
16	UPA_NF_UDEN	200	0.1	256	$1.5\lambda \times 1.5\lambda$

DEN: dense
FF: far-field
MC: mutual-coupling
NF: near-field
UDEN: ultra dense
ULA: uniform linear array
UPA: uniform planar array



▲Figure 3. ULA and UPA configuration, with the transmitting and receiving arrays spaced by D . (a) ULA configuration: The transmitting and receiving arrays are parallel to each other and have N dipoles and M dipoles respectively; the dipoles are parallel to each other with spacing δ . (b) UPA configuration: The transmitting and receiving arrays are parallel to each other and have $N_0 \times N_1$ dipoles and $M_0 \times M_1$ dipoles respectively; the dipoles are parallel to each other with spacing δ

the Rayleigh fading channel, we generate 100 sets of \mathbf{F} and calculate the corresponding channel matrices for each scene. The singular values are an average of 100 random channel matrices.

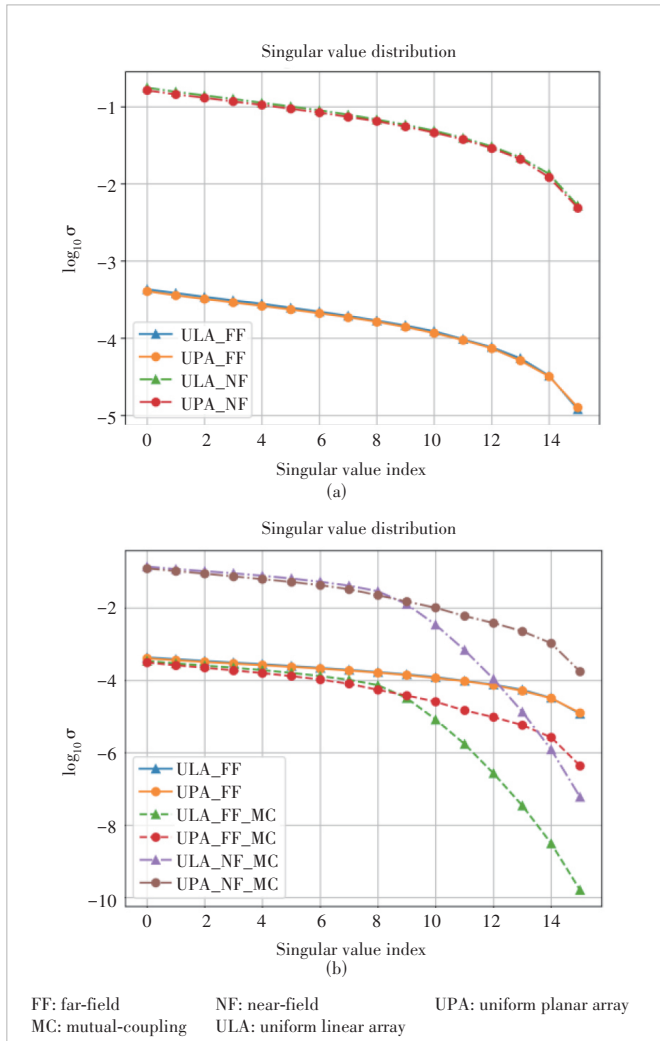
First, we study the DOF of a ULA system and a UPA system with $\lambda/2$ spacing and 200λ distance as a benchmark. In this scenario, the ULA system and the UPA system are weak coupling arrays in the far field. The DOF distribution is shown in Fig. 4(a) for a far-field channel as well as a near-field channel at the distance of 10λ . Additionally, simulations considering the mutual coupling are shown in Fig. 4(b) for $\lambda/4$ spacing with the same antenna configuration.

In Fig. 4(a), it can be seen that the ULA and UPA have similar singular value distributions in both the far-field and the near-field. Furthermore, the singular values for the simulated near-field channel are much bigger than those of the

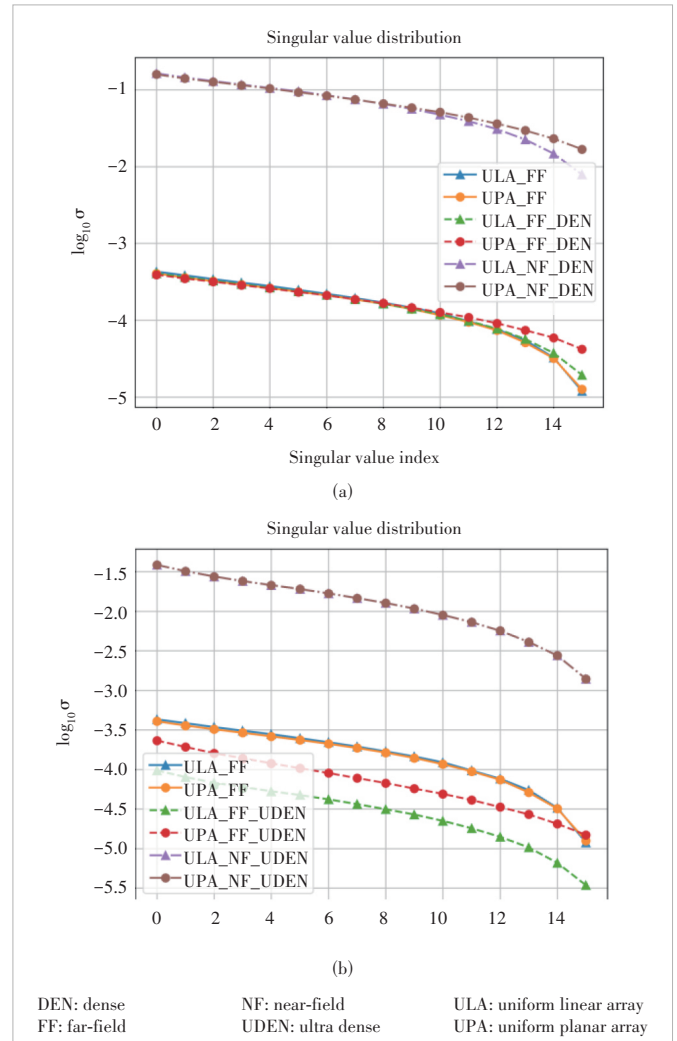
far-field channel. This can be clearly interpreted as a gain of the path loss. Since $D_{\text{farfield}} = 20D_{\text{nearfield}}$, the singular values should decrease by a factor of 400 in the far-field, which is exactly the inverse-square law. The simulations in Fig. 4(b) supports this similarity in modest mutual coupling. Also, the rich scattering channel greatly increases the available DOF of the system, thereby improving spatial multiplexing and channel performance.

Second, we study the influence of planting more antennas into an array with a fixed aperture to the channel matrix. Specifically, we analyze eight antenna configurations with a fixed 7.5λ ULA and $1.5\lambda \times 1.5\lambda$ UPA aperture listed in Table 1 with indexes 5 - 8 and 13 - 16.

Fig. 5 shows that the ULA and UPA share similar singular values in the near- and far-fields of a dense array configuration and the near field of a super-dense array configuration.



▲ Figure 4. Singular values in non-mutual coupling and modest mutual coupling: (a) singular values for $\lambda/2$ spacing ULA and UPA at the far-field (200λ) and near-field (10λ); (b) singular values for $\lambda/4$ spacing ULA and UPA at the far-field (200λ) and near-field (10λ). See Table 1 for the detailed configurations denoted by the legends



▲ Figure 5. Singular values for dense arrays with fixed aperture: (a) The dense array has 31 antennas ULA and 7×7 elements UPA, and antenna spacing is $\lambda/4$; (b) the super-dense array has 76 antennas ULA and 16×16 elements UPA, and antenna spacing is $\lambda/10$. See Table 1 for the detailed configurations denoted by the legends

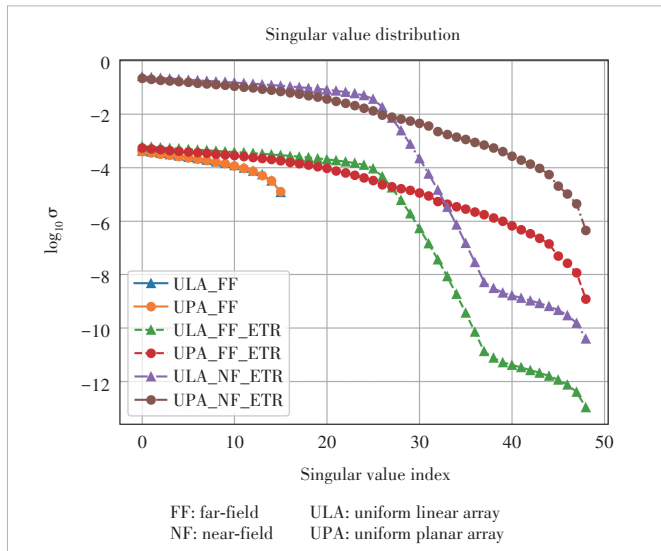
However, the singular values of the super-dense arrays in the far-field are smaller than the benchmark, indicating that the channel's DOF for both ULA and UPA may be degraded due to strong mutual coupling.

3.3 Potential DOF

To further investigate the vanishing singular values truncated by the limited antenna numbers in the receiving array, we select the scenarios where Tx has $\lambda/4$ spacing. We then increase the number of antennas in the corresponding Rx to match that of the Tx. The results for the singular values with increased receiving arrays are shown in Fig. 6.

The singular values in the dense ULA and UPA with the same antenna number have significant differences in their decaying modes compared to the fixed receiving array with only 16 antennas. The singular values of the ULA decrease rapidly at about half of the index, while those of UPA decrease more slowly. Both ULA and UPA have more significant singular values than the benchmark condition and are decreasing much slower than communicating with 16 Rx antennas. This feature indicates that the mutual coupling could support more DOF as the number of receiving antennas increases.

We define the number of the singular values of a mutual coupling channel that is bigger than the biggest singular value of a $\lambda/2$ dipole array with the same aperture as the potential DOF. This can be represented by the equation $\eta_p = \text{num}(\sigma_\delta > \max(\sigma_{\lambda/2}))$, where num indicates a numerate operation, σ_δ and $\sigma_{\lambda/2}$ are singular values of an array with a fixed aperture, and δ stands for the antenna spacing. It is worth noting that if $\delta = \lambda/2$, the potential DOF is 0. From Figs. 5(a) and 6, we can see that the increasing antennas contribute more DOF for the



▲ Figure 6. Singular values for ULA (49 dipoles) and UPA (7×7 dipoles) with the same antenna number at both Tx and Rx, denoted as ETR in the legends. See Table 1 for the detailed configurations denoted by the legends

channel.

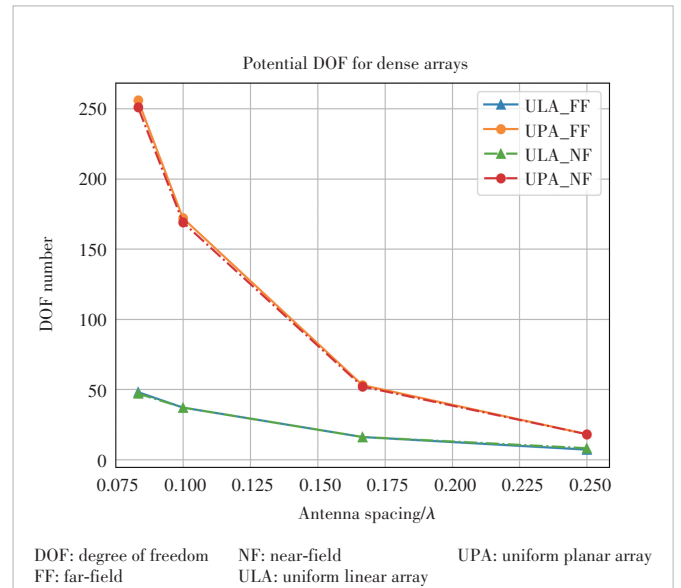
Fig. 7 demonstrates how the decreasing spacing affects the potential DOF. As the antenna spacing decreases, the potential DOF increases significantly. Moreover, the UPAs harvest more potential DOF than ULAs both in the near-field and the far-field. However, the difference in a specific antenna configuration between the near-field and the far-field is negligible. It is worth noting that the spatial correlation between the antennas restricts the DOF to reach the antenna number, as mentioned in Section 3.

4 Conclusions

In this paper, we introduce the DOF analysis method for HMIMO systems based on a mutual-coupling-compliant channel model. We demonstrate that the localized spatial correlation of the impedance matrix contributes to the additional channel DOF. We further verify the spatial correlation with the mutual-coupling-compliant model and highlight the potential DOF for a dense array with a fixed aperture in the presence of abundant receiving antennas. The numerical simulations indicate that the potential DOF of a dense array with fixed aperture increases with the number of the antennas despite of the mutual coupling. Therefore, dense arrays are practicable for harvesting more channel DOF and improving the communication channel quality (e.g. capacity) within a fixed antenna aperture.

Acknowledgement

The authors would like to thank Dr. DUAN Xiangyang, Dr. ZONG Baiqing and Dr. YU Hongkang from ZTE Corporation for their kind support.



▲ Figure 7. Potential DOF for antenna spacing $\delta = \lambda/12$, $\delta = \lambda/10$, $\delta = \lambda/6$, and $\delta = \lambda/4$

References

- [1] PIZZO A, MARZETTA T L, SANGUINETTI L. Spatially-stationary model for holographic MIMO small-scale fading [J]. *IEEE journal on selected areas in communications*, 2020, 38(9): 1964 - 1979. DOI: 10.1109/JSAC.2020.3000877
- [2] HUANG C W, HU S, ALEXANDROPOULOS G C, et al. Holographic MIMO surfaces for 6G wireless networks: opportunities, challenges, and trends [J]. *IEEE wireless communications*, 2020, 27(5): 118 - 125. DOI: 10.1109/MWC.001.1900534
- [3] AN J C, YUEN C, HUANG C W, et al. A tutorial on holographic MIMO communications: part I: channel modeling and channel estimation [J]. *IEEE communications letters*, 2023, 27(7): 1664 - 1668. DOI: 10.1109/LCOMM.2023.3278683
- [4] AN J C, YUEN C, HUANG C W, et al. A tutorial on holographic MIMO communications: part II: performance analysis and holographic beamforming [J]. *IEEE communications letters*, 2023, 27(7): 1669 - 1673. DOI: 10.1109/LCOMM.2023.3278682
- [5] AN J C, YUEN C, HUANG C W, et al. A tutorial on holographic MIMO communications: part III: open opportunities and challenges [J]. *IEEE communications letters*, 2023, 27(7): 1674 - 1678. DOI: 10.1109/LCOMM.2023.3278676
- [6] PIZZO A, MARZETTA T, SANGUINETTI L. Holographic MIMO communications under spatially-stationary scattering [C]//Proc. 54th Asilomar Conference on Signals, Systems, and Computers. IEEE, 2020: 702 - 706. DOI: 10.1109/IEEECONF51394.2020.9443506
- [7] ZHANG H B, ZHANG H L, DI B Y, et al. Holographic integrated sensing and communication [J]. *IEEE journal on selected areas in communications*, 2022, 40(7): 2114 - 2130. DOI: 10.1109/JSAC.2022.3155548
- [8] DENG R Q, ZHANG Y T, ZHANG H B, et al. Reconfigurable holographic surfaces for ultra-massive MIMO in 6G: practical design, optimization and implementation [J]. *IEEE journal on selected areas in communications*, 2023, 41(8): 2367 - 2379. DOI: 10.1109/JSAC.2023.3288248
- [9] ZHU J A, LIU K Z, WAN Z, et al. Sensing RISs: enabling dimension-independent CSI acquisition for beamforming [J]. *IEEE transactions on information theory*, 2023, 69(6): 3795 - 3813. DOI: 10.1109/TIT.2023.3243836
- [10] GOLDSMITH A. *Wireless communications* [M]. New York: Cambridge University Press, 2005
- [11] ALI JAFAR S, FAKHEREDDIN M J. Degrees of freedom for the MIMO interference channel [J]. *IEEE transactions on information theory*, 2007, 53(7): 2637 - 2642. DOI: 10.1109/TIT.2007.899557
- [12] TSE D, VISWANATH P. *Fundamentals of wireless communication* [M]. Cambridge: Cambridge University Press, 2005
- [13] PIZZO A, DE JESUS TORRES A, SANGUINETTI L, et al. Nyquist sampling and degrees of freedom of electromagnetic fields [J]. *IEEE transactions on signal processing*, 2022, 70: 3935 - 3947. DOI: 10.1109/TSP.2022.3183445
- [14] PIZZO A, MARZETTA T L, SANGUINETTI L. Degrees of freedom of holographic MIMO channels [C]//Proc. IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2020: 1 - 5. DOI: 10.1109/SPAWC48557.2020.9154219
- [15] WEI L, HUANG C W, ALEXANDROPOULOS G C, et al. Tri-polarized holographic MIMO surfaces for near-field communications: channel modeling and precoding design [J]. *IEEE transactions on wireless communications*, 2023, 22(12): 8828 - 8842. DOI: 10.1109/TWC.2023.3266298
- [16] SHEN Y S, HE Z, SHA W, et al. Electromagnetic effective-degree-of-freedom prediction with parabolic equation method [J]. *IEEE transactions on antennas and propagation*, 2023, 71(4): 3752 - 3757. DOI: 10.1109/TAP.2022.3233487
- [17] YUAN S, CHEN X M, HUANG C W, et al. Effects of mutual coupling on degree of freedom and antenna efficiency in holographic MIMO communications [J]. *IEEE open journal of antennas and propagation*, 2023, 4: 237 - 244. DOI: 10.1109/OJAP.2023.3245667
- [18] AKROUT M, SHYANOV V, BELLILI F, et al. Super-wideband massive MIMO [J]. *IEEE journal on selected areas in communications*, 2023, 41(8): 2414 - 2430. DOI: 10.1109/JSAC.2023.3288269
- [19] ORFANIDIS S J. *Electromagnetic waves and antennas* [M]. Rutgers: Rutgers University, 2002
- [20] CHU L J. Physical limitations of omni-directional antennas [J]. *Journal of applied physics*, 1948, 19(12): 1163 - 1175. DOI: 10.1063/1.1715038

Biographies

SUN Yunqi (sun.yunqi@zte.com.cn) received his BE and PhD degrees from Tsinghua University, China in 2017 and 2022, respectively. He is currently an engineer with ZTE Corporation. His research interests include massive MIMO, holographic MIMO, antenna theory, and radio sensing.

JIAN Mengnan received her BS degree in information engineering from Beijing Institute of Technology, China in 2016 and MS degree from Tsinghua University, China in 2019. She is currently an engineer with ZTE Corporation. Her research interests include massive MIMO, reconfigurable intelligent surfaces, and holographic MIMO.

YANG Jun received his BS degree in geophysics from the China University of Geosciences (Wuhan) in 2011 and joint-training PhD degree in geophysics from the University of Science and Technology of China and the University of North Carolina at Charlotte, USA in 2016. He is currently an algorithm engineer with ZTE Corporation. His research interests include reconfigurable intelligent surfaces, MIMO channel modeling, and computational electromagnetics.

ZHAO Yajun received his BE, MS, and PhD degrees. Since 2010, he has assumed the role of Chief Engineer at the Wireless and Computing Product R&D Institute, ZTE Corporation. Prior to this, he contributed to wireless technology research at the Wireless Research Department, Huawei. Currently, his primary focus centers on 5G standardization technology and the advancement of future mobile communication technology, particularly 6G. His research pursuits encompass a broad spectrum, including reconfigurable intelligent surfaces (RIS), spectrum sharing, flexible duplex, CoMP, and interference mitigation. He has played an instrumental role in founding the RIS Tech Alliance (RISTA) and currently holds the position of Deputy Secretary General within the organization. Additionally, he is a founding member of the RIS task group under the purview of the China IMT-2030 (6G) Promotion Group, where he serves as the deputy leader.

CHEN Yijian graduated from Central South University, China. He is currently working at ZTE Corporation. His research interests include reconfigurable intelligent metasurfaces, extremely large-scale MIMO technology, and electromagnetic information theory.



Near-Field Beam Training for Holographic MIMO Communications: Typical Methods, Challenges and Future Directions

SHEN Jiayu¹, YANG Jun², ZHU Chen³, DENG Zhiji⁴,
HUANG Chongwen¹

(1. Zhejiang University, Hangzhou 310058, China;
2. ZTE Corporation, Shenzhen 518055, China;
3. Polytechnic Institute, Zhejiang University, Hangzhou 310015, China;
4. Zhejiang Dahua Technology Co., Ltd., Hangzhou 310053, China)

DOI: 10.12142/ZTECOM.202401006

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240314.1348.002.html>,
published online March 14, 2024

Manuscript received: 2023-12-19

Abstract: Holographic multiple-input multiple-output (HMIMO) has become an emerging technology for achieving ultra-high frequency spectral efficiency and spatial resolution in future wireless systems. The increasing antenna aperture leads to a more significant characterization of the spherical wavefront in near-field communications in HMIMO scenarios. Beam training as a key technique for wireless communication is worth exploring in this near-field scenario. Compared with the widely researched far-field beam training, the increased dimensionality of the search space for near-field beam training poses a challenge to the complexity and accuracy of the proposed algorithm. In this paper, we introduce several typical near-field beam training methods: exhaustive beam training, hierarchical beam training, and multi-beam training that includes equal interval multi-beam training and hash multi-beam training. The performances of these methods are compared through simulation analysis, and their effectiveness is verified on the hardware testbed as well. Additionally, we provide application scenarios, research challenges, and potential future research directions for near-field beam training.

Keywords: holographic multiple-input multiple-output (HMIMO); beam training; near-field; equal interval multi-beam (EIMB) training; hash multi-beam (HMB) training

Citation (Format 1): SHEN J Y, YANG J, ZHU C, et al. Near-field beam training for holographic MIMO communications: typical methods, challenges and future directions [J]. *ZTE Communications*, 2024, 22(1): 41 - 52. DOI: 10.12142/ZTECOM.202401006

Citation (Format 2): J. Y. Shen, J. Yang, C. Zhu, et al., "Near-field beam training for holographic MIMO communications: typical methods, challenges and future directions," *ZTE Communications*, vol. 22, no. 1, pp. 41 - 52, Mar. 2024. doi: 10.12142/ZTECOM.202401006.

1 Introduction

Holographic multiple-input multiple-output (HMIMO), known as the advanced MIMO technology^[1-2], and large intelligent surfaces^[3] are utilized by the next generation of wireless networks to meet stringent requirements for data rates, connectivity, and reliability. The spatially continuous aperture in HMIMO schemes enhances both spatial resolution and beamforming gains by utilizing compact radiating elements^[4]. In the HMIMO setting, high-frequency communication is an important feature and application. High-frequency signals, like the millimeter-wave (mmWave), are susceptible to obstruction and attenuation by obstacles during transmission, and the beams are narrow and directional. So, the beamforming technique is extremely important, by which the signal can be transmitted in the form of a directional beam^[5-6]. By leveraging this technology, it is possible to provide positioning services in cellular networks over ultra-wide bandwidths with unprecedented accuracy, making it a promising service for

6G networks^[7-8]. Meanwhile, for communication issues, beamforming brings higher spatial resolution and beam directivity, making the energy concentrated in a certain direction to realize high signal-to-noise ratio (SNR) transmission, but it requires precise knowledge of the angle of arrival (AoA) and angle of departure (AoD) of the propagation channel. Typically, AoA and AoD are obtained during the beam training process before the data transmission. Beam training overhead is a major burden that reduces spectral efficiency. The beam management procedure of the 802.11ad millimeter-wave standard consists of alternating base stations sweeping the entire angular space, and its beam training overhead is proportional to the spatial resolution, the number of base stations, and the degree of dynamics of the propagation environment^[9]. As a result, the beam training overhead increases with narrower beamwidths brought about by large-scale antenna arrays and more users. Therefore, beam training methods that combine low complexity and high accuracy are needed.

It is worth pointing out that in HMIMO scenarios, the antenna aperture is generally very large, and thus the Rayleigh distance $Z = 2D^2/\lambda$, which separates the near and far fields of electromagnetic radiation, is no longer negligible^[10]. For example, if we consider that the array aperture of the base station (BS) is 0.1 m and the carrier frequency is 28 GHz, we can get that the Rayleigh distance is only 1.8 m according to the formula; if we consider that the array aperture of the BS is 0.7 m, the Rayleigh distance will be 91.5 m. Therefore, users are more likely to occur in the near-field region of the BS antenna array, where far-field beam training suffers a significant loss of accuracy because of the more significant properties of spherical waves in the near-field region. To solve these difficulties, we need to focus on near-field beam training and find efficient beam training methods.

Currently, there are many beam training algorithms for the far field^[11-14]. However, due to the spherical wave characteristics of near-field communication, the existing far-field beam training algorithms cannot be directly applied to the near field. Specifically, traditional far-field beamforming directs the beam of a planar wavefront to a specific angle^[15-16], whereas in the near-field region, the beam of a spherical wavefront can achieve beam focusing, concentrating the beam energy in terms of the angle and range, thus providing a new degree of freedom (DoF), i.e., the distance dimension, for controlling the spatial energy distribution. This brings challenges to the research of beam training algorithms for the near field.

We will introduce the typical algorithms for near-field beam training in HMIMO scenarios and compare their performance through simulation analysis. Moreover, we will discuss the application scenarios, main difficulties and possible future research directions of near-field beam training. The rest of the paper is organized as follows. In Section 2, this paper introduces the system model of the interest scenario and the codebook forms of beam training. Several typical methods and application scenarios of near-field beam training are introduced in Section 3. In Section 4, simulation results and configuration of the testbed are given. Finally, we analyze the research challenges and possible future research directions for near-field beam training in Section 5.

2 System Model

As shown in Fig. 1, we consider a downlink HMIMO communication scenario where K BSs and U users are distributed in the 3D space. The BS is deployed in the xz -plane, which employs a hybrid precoding architecture that equips V radio-

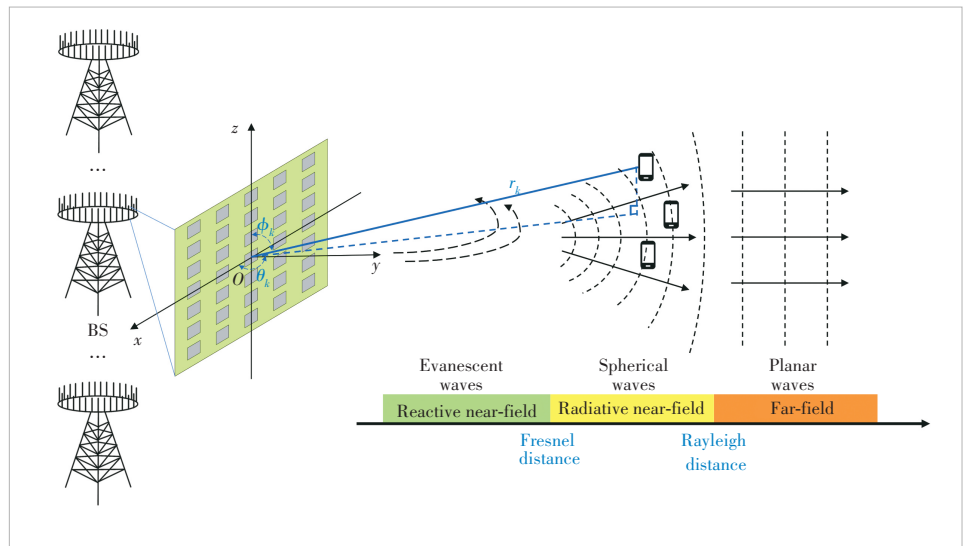
frequency (RF) chains and an $N_x \times N_z$ antenna uniform planar array (UPA), where $V \ll N_x N_z$. Note that if $N_z = 1$, it is the case of uniform linear array (ULA). Each user device is equipped with a single antenna. The central wavelength, horizontal/vertical antenna spacing, and operating frequency are λ_c , d_x/d_z and f_c , respectively. The coordinate of the (m, n) -th antenna element of the k -th BS is $(x_{k,n}, y_k, z_{k,m})$, where $x_{k,n} = r_k \cos\theta_k \sin\phi_k + nd_x$, $y_k = r_k \sin\theta_k \sin\phi_k$, and $z_{k,m} = r_k \cos\phi_k + md_z$, $n = 1 - \frac{N_x + 1}{2}, \dots, N_x - \frac{N_x + 1}{2}$, $m = 1 - \frac{N_z + 1}{2}, \dots, N_z - \frac{N_z + 1}{2}$. r_k, θ_k and ϕ_k denote the distance, azimuth angle, and the elevation angle of the link from the original O of the k -th BS. For beam training, K BSs transmit training symbols in several directions while the users listen to the channel in all directions using a quasi-omnidirectional beam. For each user, the alignment of the BS side requires only its received multi-BS superimposed signal power. Thus, we can perform downlink beam training for different users simultaneously without interference.

Let $h_k \in \mathbb{C}^{N_x N_z \times 1}$ denote the channel from the k -th BS to the user, the received signal y at the user can be expressed by

$$y = \sum_{k=1}^K h_k^H F_{RF_k} f_{BB_k} x + n_k, \quad (1)$$

where $f_{BB_k} \in \mathbb{C}^{V \times 1}$ and $F_{RF_k} \in \mathbb{C}^{N_x N_z \times V}$ denote the digital precoder and analog beamformer at the k -th BS side, x denotes the transmit symbol with power P_0 , and $n_k \sim \mathcal{CN}(0, \sigma^2)$ denotes Gaussian additive white noise.

In the HMIMO communication scenario, it is assumed that all users are in the near-field region, which is different from the normal MIMO scenario. Based on the spherical wave assumption, the near-field channel can be represented as



▲ Figure 1. Downlink Holographic multiple-input multiple-output (HMIMO) communication scenario

$$h_k = \sqrt{N_x N_z} \sum_{l=1}^L \beta_l e^{-j\psi_l} g_l(\theta, \phi, r), \quad (2)$$

where L is the number of paths, β_l and ψ_l denote the complex path gain and phase shift of the l -th path. In the line of sight (LoS) link dominated channel, $\beta_1 \gg \beta_l, l \neq 1$, thus

$$h_k = \sqrt{N_x N_z} \beta_1 e^{-j\psi_1} g_1(\theta, \phi, r), \quad (3)$$

where $g_1(\theta, \phi, r)$ denotes the near-field steering vector, which is given by

$$g_1(\theta, \phi, r) = \frac{1}{\sqrt{N_x N_z}} \left[e^{-j\frac{2\pi}{\lambda_c} \left(D \left(\frac{N_x-1}{2}, \frac{N_z-1}{2} \right) - r_k \right)}, \dots, \right. \\ \left. e^{-j\frac{2\pi}{\lambda_c} (D(m,n) - r_k)}, \dots, e^{-j\frac{2\pi}{\lambda_c} \left(D \left(\frac{N_x-1}{2}, \frac{N_z-1}{2} \right) - r_k \right)} \right]^H, \quad (4)$$

where $D(m, n)$ is the distance between the (m, n) -th antenna of the k -th BS and the user.

Fully implementing digital beamforming is challenging because of the massive size of the antenna array and the high expenses and power consumption of the RF chain^[17]. The solution here is hybrid analog/digital precoding, where the required precoding and beamforming are in the analog and digital domains. The digital precoding layer adds more DoF to the precoding design problem than the pure analog beamforming solution. This allows hybrid precoding to achieve better performance, especially in multi-stream and multi-user transmissions. It is important to note that analog beamforming with finite RF chains and constant amplitude constraints relies heavily on predefined codebooks. Let $a(\theta, \phi, r) = F_{RF} f_{BB}$ denote the beam-aligned observation vector, also known as the codeword. Codebook-based beam training is to find the optimal beam-aligned observation vector \mathbf{a} from the codebook $\mathcal{C} = [a(\theta_1, \phi_1, r_1); \dots; a(\theta_N, \phi_N, r_N)]$ to maximize the beam gain at the receiver. It can be easily obtained that the optimal beam-aligned observation vector is $\mathbf{a}(\theta, \phi, r) = \mathbf{g}_1(\theta, \phi, r)$. Without loss of generality, it is similar to the ULA scenario where the users are all in the yz -plane. The near-field channel is modeled as:

$$h_k = \sqrt{N_x N_z} \beta_1 e^{-j\psi_1} b_1(\theta, r), \quad (5)$$

where the near-field steering vector considers two dimensions.

A qualitative change in HMIMO compared with MIMO systems is that HMIMO cell elements are placed denser and denser to form almost spatially continuous apertures. At the same time, the number of antenna units is increasing (N_x and N_z are larger in HMIMO systems). In this case, the neighboring cell elements in between will act. Mutual coupling, which is considered harmful in conventional communication systems, can be mitigated in MIMO antenna arrays by placing

neighboring antenna elements within a half-wavelength distance of each other. However, proper utilization of mutual coupling in HMIMO surfaces has the potential to achieve super directionality, which can be more apparent when the number of cell elements becomes larger. In addition, this feature can potentially improve the received SNR, thus extending the coverage of the structure. In conclusion, it is important to study the mutual coupling effects and devise convenient mathematical models for coupling-aware wireless design. In addition to mutual coupling, spatially continuous apertures offer another important potential, that is signal processing can be performed entirely on the EM domain instead of the traditional digital domain. To this end, EM-inspired analyses and techniques can be designed to revolutionize the existing frameworks for wireless operation into one based on the hybrid digital EM domain, or even hybrid digital and EM-based frameworks, paving the way for a high degree of integration. Flexible processing aims for high spatial resolution and low latency wireless communication.

3 Near-Field Beam Training

This section focuses on beam training in near-field communication. First, three major existing methods are introduced: exhaustive beam training^[18], hierarchical beam training^[18-19] and multi-beam training^[20]. Then, we will analyze the potential application scenarios of near-field beam training.

3.1 Exhaustive Beam Training

A straightforward approach to this problem is conducting an exhaustive search for the various dimensions of the constructed codebook. The performance of the exhaustive search method depends mainly on the constructed codebook.

Considering the scenario under HMIMO, where the users are located in a three-dimensional space, the simplest way to construct the codebook is a codebook with equally spaced samples in each of the three dimensions, i.e., polar coordinate (θ, ϕ, r) or cartesian coordinate (x, y, z) . Subsequently, all codewords in the designed near-field codebook are traversed sequentially in exhaustive search training^[18]. Similarly, if we consider the scenario where the users are all in the yz -plane, it is a 2D exhaustive search for θ and r dimensions or the x and y dimensions. In addition to this, we can consider the polar-domain near-field codebook^[21], which is given by $\mathbf{C} = \{\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_N\}$ with $\mathbf{C}_n = \{c_{n,0}, c_{n,1}, \dots, c_{n,S_n-1}\}, \forall n \in \mathcal{N}$. Each codeword $c_{n,s_n} \in \mathcal{S}_n \triangleq \{0, 1, \dots, S_n - 1\}$ corresponds to a beam pointing to a specific direction-and-distance pair, (θ_n, r_{n,s_n}) , which is given by $c_{n,s_n} = b(\theta_n, r_{n,s_n})$. The codebook is constructed in the following manner. Initially, the two-dimensional spatial domain is divided into N sectors with a central angle θ_n . In the distance domain, the distance for each sector is divided into S_n distinctive non-uniform regions that correspond to each central angle θ_n . The distance sampling in-

interval increases with distance. This is expected since the effect of distance on the beamforming design decreases with distance and eventually disappears in the far-field region. The advantage of constructing the codebook in this way is that the coherence between different code words in the codebook is small, which improves the efficiency of the exhaustive search.

Obviously, the narrower the sampling interval during codebook construction, the greater the codebook's spatial resolution, and the outcomes are more precise following exhaustive beam training. However, this training process results in increased delay and overhead^[22]. It should be pointed out that the exhaustive search method can always find the optimal solution by constructing a reasonable codebook, therefore making it a suitable benchmark for comparing with other schemes.

3.2 Hierarchical Beam Training

The hierarchical codebook scheme has been proposed to solve the beam training problem in the far field^[13-14]. The hierarchical codebook is inspired by the fact that the overhead of beam training mainly depends on the size of the codebook, but the size of the codebook and the accuracy of the training are a tradeoff. Therefore, the hierarchical codebook considers the design of different levels of sub-codebooks, each of which has different sampling intervals and numbers of sampling points. During the training process, the codebook is searched sequentially layer by layer to simultaneously reduce the codebook space and ensure the training accuracy. Of course, the design of hierarchical codebooks should have corresponding principles^[23]. Firstly, the codebook needs to cover the entire sampling space, which is to ensure the accuracy of the results. At the same time, in order to reduce the steps in training, the lower codebook is a subset of the upper codebook, so that when searching layer by layer, only the corresponding subset needs to be searched in turn. Following are two hierarchical beam training schemes for near-field beam training.

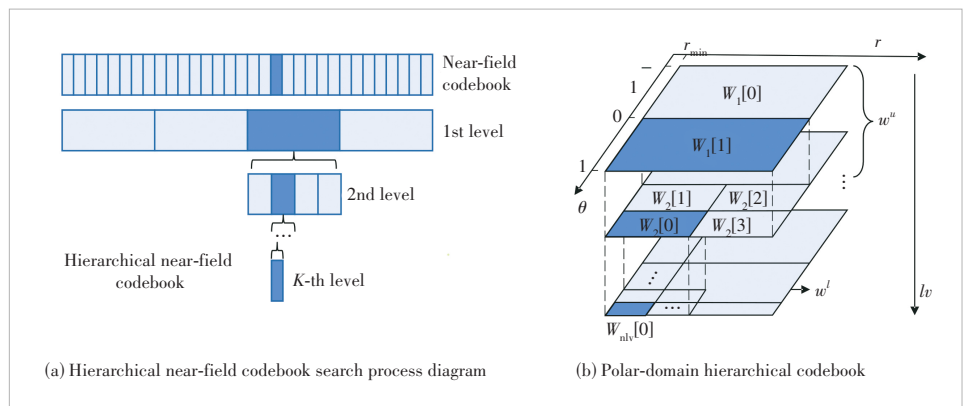
As shown in Fig. 2(a), similar to the layered codebook in the far-field beam search, a similar effect can be achieved in the near-field by setting smaller and smaller sampling intervals in the layer-by-layer codebook. The difference is that the layered codebook for the far-field is one-dimensional, while the layered codebook for the near-field is two- or three-dimensional. Based on the hierarchical near-field codebook described above, the basic idea of the hierarchical near-field beam training scheme is to search from the 1st level sub-codebook to the K -th level sub-codebook^[18]. In addition to this, there is another hierarchical beam training scheme based on the polar domain representation^[19]. In order to adapt to the characteristics

of near-field communication, the hierarchical codebook is designed as in Fig. 2(b). The codebook $\mathcal{W} = \{\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_{n_l}\}$ consists of n_l levels of codebooks. The upper-layer codebook $\mathcal{W}^u = \{\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_{n_l-1}\}$ is designed to reduce the beam searching steps. While the lower-layer codebook $\mathcal{W}^l = \{\mathbf{W}_{n_l}\}$ aims to provide high beamforming gain for receivers in the Fresnel region. To reduce the beam searching steps, the upper-layer codebook \mathcal{W}^u is designed. In the upper-layer codebook, the i -th codeword of the l -th layer codebook $\mathbf{W}_l[i]$ is assigned to a subset of codewords \mathcal{W}_{l+1}^i in the next layer: $\mathcal{W}_{l+1}^i = \{\mathbf{W}_{l+1}[i_1], \mathbf{W}_{l+1}[i_2], \dots, \mathbf{W}_{l+1}[i_k]\}$. Beam relocation and beam rotation methods are used for the design of the upper codebook. In the polar domain, beam rotation and beam relocation can be used to introduce a rotation angle and relocation of a specific distance from an existing codeword, respectively. As a result, a hierarchical codebook can be obtained for each layer. The lower-layer codebook \mathcal{W}^l is designed to match the strongest multipath components (MPCs) in the channel and provide high beamforming gains as many as possible. After the construction of a codebook, a top-down level by level codebook search remains.

The advantage of hierarchical beam training is the reduced training overhead, but there are some disadvantages. The use of wide beams early in the hierarchical beam training leads to a decrease in beamforming gain, thus the correct wide beam may not be recognized at a low SNR and ultimately the optimal beam direction may not be detected. Secondly, layered beam training suffers from serious inherent error propagation problems^[24]. Moreover, the hierarchical training results in additional delay due to the need to compare and discriminate the recognition results at each level to determine the training sub-codebook for the next stage.

3.3 Multi-Beam Training

In order to reduce the time of conventional single-beam training, multi-beam training methods for HMIMO communication are introduced^[20].



▲ Figure 2. Hierarchical codebook in the near-field communication

3.3.1 Equal Interval Multi-Beam (EIMB) Training

First, the (horizontal) antennas or reflective elements are divided into M subarrays, each consisting of $L \triangleq N_x/M$ neighboring elements. For each subarray $m \in \mathcal{M} \triangleq \{1, 2, \dots, M\}$, we equip it with a separate codebook, \mathbf{W}_m , which consists of $J = N_x$ codewords covering the same sampling direction as the single-beam codebook, i.e., $\{\alpha(j), j \in \mathcal{J}\}$. Thus, we have

$$\mathbf{W}_m = \{\mathbf{w}_m(1), \mathbf{w}_m(2), \dots, \mathbf{w}_m(J)\}, \forall m \in \mathcal{M}, \quad (6)$$

where $\mathbf{w}_m(j) \in \mathbb{C}^{L \times 1}$, $j \in \mathcal{J}$ represents the codeword of subarray m and uses only L antennas or reflecting elements to direct the reflected beam in the direction $\alpha(j)$. Based on the single-beam codeword $\mathbf{C}(j)$, we construct $\mathbf{w}_m(j)$ as:

$$\mathbf{w}_m(j) \triangleq [\mathbf{C}(j)]_{(m-1)L+1:mL}. \quad (7)$$

Thus, when all M subarrays turn the reflected beam in the same direction $\alpha(j)$, the composite multi-beam encoding $[\mathbf{w}_1^T(j), \mathbf{w}_2^T(j), \dots, \mathbf{w}_M^T(j)]^T$ is equivalent to its single-beam counterpart $\mathbf{C}(j)$. Compared with the full-array codeword $\mathbf{C}(j)$, each subarray codeword $\mathbf{w}_m(j)$ has a wider beam width (i.e., $2M/N_x$ versus $2/N_x$) and a smaller beam gain (i.e., N_x/M versus N_x). Next, for the proposed multi-beam training, the BS is allowed to steer the subarray beams in several different directions simultaneously, usually according to the multi-beam codebook \mathbf{W}_m , $m \in \mathcal{M}$ varying over the duration of the training symbols. By appropriately designing the beam directions of the BS subarrays on different training symbols, the optimal beam direction for each user can be found with high probability by a simple comparison of the received signal power/SNR. For ease of exposition, we consider a typical case where $M = 2^R$, $R \in \mathbb{Z}$, and L are even numbers, $N_x = ML$. The proposed fast beam training method consists of two phases, namely beam scanning and beam identification, as described below.

1) The beam scanning phase consists of $1 + \log_2 M$ rounds in which the set of BS subarray beam directions sends multiple training symbols. For each round $r \in \{1, 2, \dots, \log_2 M + 1\}$, we denote $B(r, b)$ as the collection of subarray beam directions (in ascending order) during the b -th training symbol. For any set of beam directions $\mathcal{A} \subseteq \mathcal{J}$, we define its intra-set distance as

$$d_s(\mathcal{A}) = \min_{p, q \in \mathcal{A}; p \neq q} |p - q|. \quad (8)$$

Thus, a larger in-set distance indicates that the beams indexed by \mathcal{A} are farther apart in the spatial domain. For $r = 1$, the N_x directions

are mapped into L bins, each bin including M directions. In order to separate the beam directions in each bin as much as possible to minimize inter-beam interference, the bins are set to

$$B(1, b) = \{b, b + L, \dots, b + (M - 1)L\}, \forall b \in \{1, 2, \dots, L\}, \quad (9)$$

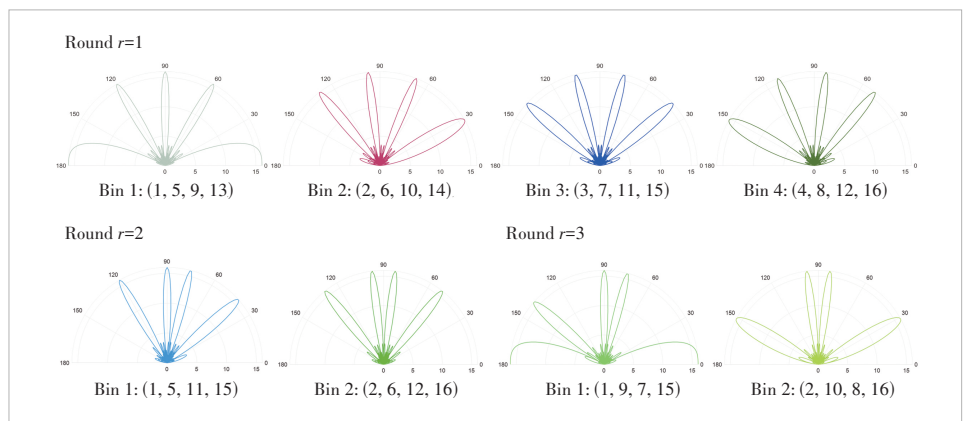
which maximizes the minimum bin distance among all bins, $d_s(B(1, b)) = L$, $\forall b \in \{1, 2, \dots, L\}$, as shown in Fig. 3. Then, for each group ℓ , an odd subset of $B(1, \ell)$, and an even subset of ℓ $B(1, \ell + L/2)$ are extracted to construct a new bin. In addition, for each round of beam sweeping, the same (maximum) intra-bin distance for different bins can be numerically derived as

$$d_s(B(r, b)) = \begin{cases} L, & \text{if } r \in \{1, 2\}, \\ L/2, & \text{if } r \in \{3, 4, \dots, \log_2 M + 1\}. \end{cases} \quad (10)$$

Finally, the total number of training symbols is

$$T_t^{(fa)} = L + \frac{L(\log_2 M)}{2} = \frac{N_x}{M} \left(1 + \frac{\log_2 M}{2}\right). \quad (11)$$

2) Beam identification: after the scanning phase, each user can independently identify its optimal BS beam direction based on its own received power/SNR in the first phase. Consider an arbitrary user k . Let $P_k(r, b)$ denote its received power at bin b in the r -th sweep round and $\hat{I}_k(r)$ denote the set of candidate directions for its optimal beam direction after the r -th sweep round. For $r = 1$, $\hat{I}_k(1)$ is set to be the best bin with the maximum received power, i. e., $\hat{I}_k(1) = B(1, b_k^*)$, $b_k^* = \arg \max_{b \in \{1, 2, \dots, L\}} P_k(1, b)$. For each subsequent round $r \in \{2, 3, \dots, \log_2 M + 1\}$, the user only needs to check a bin that has a common direction with $\hat{I}_k(1)$, since it is the only one that may contain the optimal beam direction, denoted by $B(r, b_k(r))$, where $b_k(r) = \{b \in \{1, 2, \dots, L/2\} \mid B(r, b) \cap \hat{I}_k(1) \neq \emptyset\}$. For this bin, since the expected received power of the corresponding



▲ Figure 3. Multi-beam schematic of the hierarchical search process

multibeam covering/not covering the best direction is approximated by (ignoring receiver noise and any inter-beam interference) $P_k(1, b_k^*)$ and 0, the binary decision threshold for the received power is set as

$$P_k^{(th)} \triangleq (P_k(1, b_k^*) + 0)/2 = P_k(1, b_k^*)/2. \quad (12)$$

It is used to determine whether the optimal beam direction is included by $B(r, b_k(r))$. Thus, for each $r \in \{2, 3, \dots, \log_2 M + 1\}$, combined with the binary decision of $\hat{I}_k(r-1)$, the new candidate direction for the r -round is determined as follows

$$\hat{I}_k(r) = \begin{cases} \hat{I}_k(r-1) \cap B(r, b_k(r)), & \text{if } P_k(r, b) \geq P_k^{(th)}, \\ \hat{I}_k(r-1) \setminus B(r, b_k(r)), & \text{if } P_k(r, b) < P_k^{(th)}, \end{cases} \quad \forall k \in \mathcal{K}. \quad (13)$$

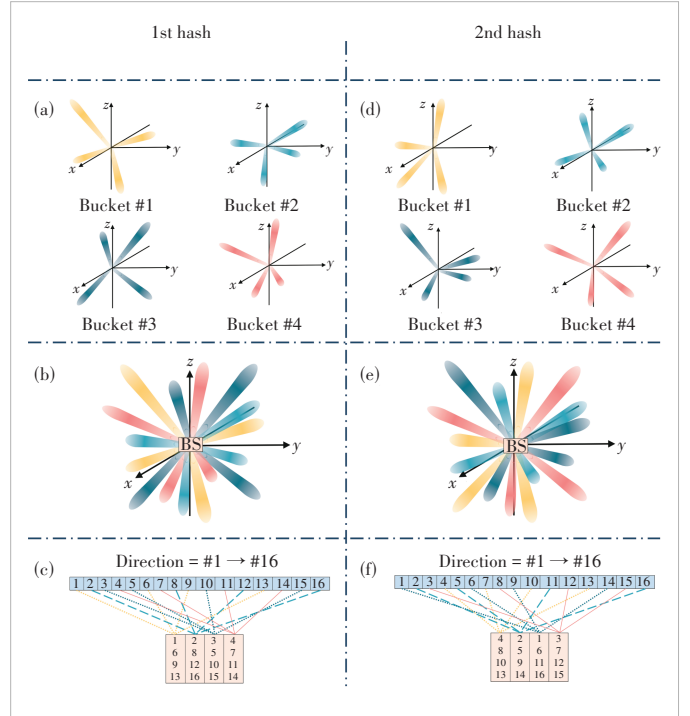
3.3.2 Hashing Multi-Beam Training

To address the challenges of high complexity and low identification accuracy in the aforementioned beam training techniques, particularly in many-to-many near-field communication scenarios, an efficient hashing multi-beam (HMB) training scheme is proposed.

Firstly, we discuss how to generate HMB. HMB training leverages the structure of the sparse Fourier transform and select codewords from the single-beam codebook \mathcal{C} to form multi-beams. Define the universe of keys with a total order as $\mathcal{U} = 0, 1, \dots, N_c - 1$, where N_c is the total number of codewords in the codebook. The hash function is denoted as $h: \mathcal{U} \rightarrow \mathcal{T}$, where $\mathcal{T} = 0, 1, \dots, B - 1$ represents the set of interpreted hash values. The parameter B signifies the number of available hash values. In this framework, the universe of keys remains fixed, and the HMB method randomly selects hash functions from a family $\mathcal{H} = \{h_1, h_2, \dots, h_{|\mathcal{H}|}\}$ of hash functions. Here, $|\mathcal{H}|$ represents the number of distinct functions present in the family \mathcal{H} . Fig. 4 illustrates the process of two times hashing, where the codewords representing 16 different directions are uniformly hashed into 4 multi-beams. Each multi-beam covers a total of $N_c/B = 4$ different directions.

Moreover, the proposed method involves joint design of the response of all antennas to generate the multi-beam codebook $\tilde{\mathcal{C}}$ for training, which offers enhanced performance compared with the simple antenna partitioning approach.

The training process in this method consists of two phases: the scanning phase and the voting phase. During the scanning phase, a total of L rounds of hashing mapping are performed to accurately train the beams. In each round, the BS randomly selects a distinct hash function from the family \mathcal{H} . Simultaneously, the BS transmits training symbols using its designed multi-beam codeword. This approach significantly reduces the complexity compared with traditional methods that require alternate scanning of all BSs. By using the hash functions and



▲ Figure 4. Process of two times hashing

transmitting training symbols with the designed codewords, the scanning phase achieves accurate beam training while minimizing the number of time slots required. Specifically, B hash functions are randomly selected from the family \mathcal{H} over L rounds, resulting in a total of $Q = BL$ time slots needed for the scanning phase. During these time slots, the BS transmits training symbols with the corresponding multi-beam codewords, resulting in Q received signal power measurements.

Following the scanning phase, the voting phase is carried out to align the signals from different BSs. Since the signals from multiple BSs are superimposed, a demultiplexing algorithm and the voting mechanism are then employed to separate the signals from each BS.

Specifically, by sorting the power measurements in descending order, the soft decision assigns the L time slots with the $((k-1)L+1)$ -th to kL -th largest values in the received signal power to the k -th strongest BS m_k . Thus, the time slot $q_{m_k} = \arg \max_{(k-1)L+1:kL} \text{descend}(\mathbf{P})$ is considered to contain the aligned beam direction between the BS m_k and the user. This step distinguishes the received signal power of different BSs. Afterward, based on the demultiplexed time slot q_{m_k} , the next step is to vote on the directions in the multi-beams used in the q_{m_k} -th time slot. The direction that receives the highest number of votes is considered the correctly aligned direction for BS m_k .

Fig. 5 illustrates an example of the voting process in a scenario involving two BSs. The received signal power values are known and arranged in descending order. The time slots with the first L highest received signal power is highlighted in orange,

		BS 1				BS 2				BS 1				BS 2			
$L=1$	$b=1$	13	25	24	5	19	6	1	21	13	25	24	5	19	6	1	21
	$b=2$	26	8	9	21	22	13	23	16	26	8	9	21	22	13	23	16
	$b=3$	10	7	11	6	27	29	20	28	10	7	11	6	27	29	20	28
	$b=4$	32	28	30	1	4	10	30	14	32	28	30	1	4	10	30	14
	$b=5$	12	4	31	20	17	8	18	24	12	4	31	20	17	8	18	24
	$b=6$	18	14	17	29	25	3	12	31	18	14	17	29	25	3	12	31
	$b=7$	16	19	15	3	5	26	15	2	16	19	15	3	5	26	15	2
	$b=8$	27	22	23	2	7	32	9	11	27	22	23	2	7	32	9	11
$L=2$	$b=1$	25	3	26	18	17	26	29	15	25	3	26	18	17	26	29	15
	$b=2$	5	31	21	13	9	27	22	2	5	31	21	13	9	27	22	2
	$b=3$	9	16	23	11	11	23	12	6	9	16	23	11	11	23	12	6
	$b=4$	7	24	32	17	31	8	32	4	7	24	32	17	31	8	32	4
	$b=5$	28	19	4	8	3	14	19	28	28	19	4	8	3	14	19	28
	$b=6$	6	15	20	22	10	21	5	1	6	15	20	22	10	21	5	1
	$b=7$	30	1	10	12	20	13	18	16	30	1	10	12	20	13	18	16
	$b=8$	2	27	14	29	30	7	24	25	2	27	14	29	30	7	24	25
...																	

▲ Figure 5. Voting process in a scenario involving two BSs

while the time slots with the $(L + 1)$ -th to $2L$ -th highest power are marked in blue. In Fig. 5(a), when voting is performed on the multi-beams corresponding to the orange markers, it is observed that direction 26 in BS 1 receives a significant number of votes, while the votes in BS 2 are more scattered. Consequently, the direction with the highest number of votes across both BSs is direction 26 in BS 1, denoted as $m_1 = 1$. Subsequently, in Fig. 5 (b), voting is conducted on the remaining BS 2 using the multi-beams marked in blue. This voting process leads to the determination of $m_2 = 2$ and direction 5.

Here we discuss the advantages of HMB training. Compared with the single-beam, multi-beam can scan multiple directions simultaneously, greatly reducing the number of beams to cover the space. Both the hashing and the equal interval approaches use multi-beam. However, the EIMB method uses a predetermined multi-beam combination approach as well as a hard decision of threshold comparison, while the proposed hash method utilizes hashing multi-beam, multi-round voting, and soft decision, which are the main aspects that make the difference.

1) HMB training employs randomly selected hash functions, which can be perceived as introducing a perturbation to the deterministic combination approach, thereby rendering the combination random. This randomness ensures that the sub-beams are sufficiently dispersed, thereby reducing the impact of leakage interference between sub-beams, particularly after multiple rounds of voting. As in Fig. 4, if there is interference from beam 1 to 9 in the first round, but 1 and 9 are not in the same multi-beam in subsequent rounds, there will be no further interference from beam 1.

2) By employing an L -round voting mechanism, each round carries an equal weight, thus mitigating the influence of errors from individual rounds when L is relatively large. Moreover, since L is complexity dependent, we indicate that L is on the order of $O(\log M_s)$ when the identification error is equal to

$1/M_s$. That is, we sacrifice some accuracy compared with the exhaustive approach but greatly reduce the training overhead to $\text{Blog}MN$.

3) Another distinction lies in the use of the soft decision. We consider the relative value, unlike the equal interval method that considers the absolute value. This approach makes it less susceptible to noise and variations in practical power strength. Hence, our method achieves higher accuracy and greater robustness.

3.4 Potential Application Scenarios

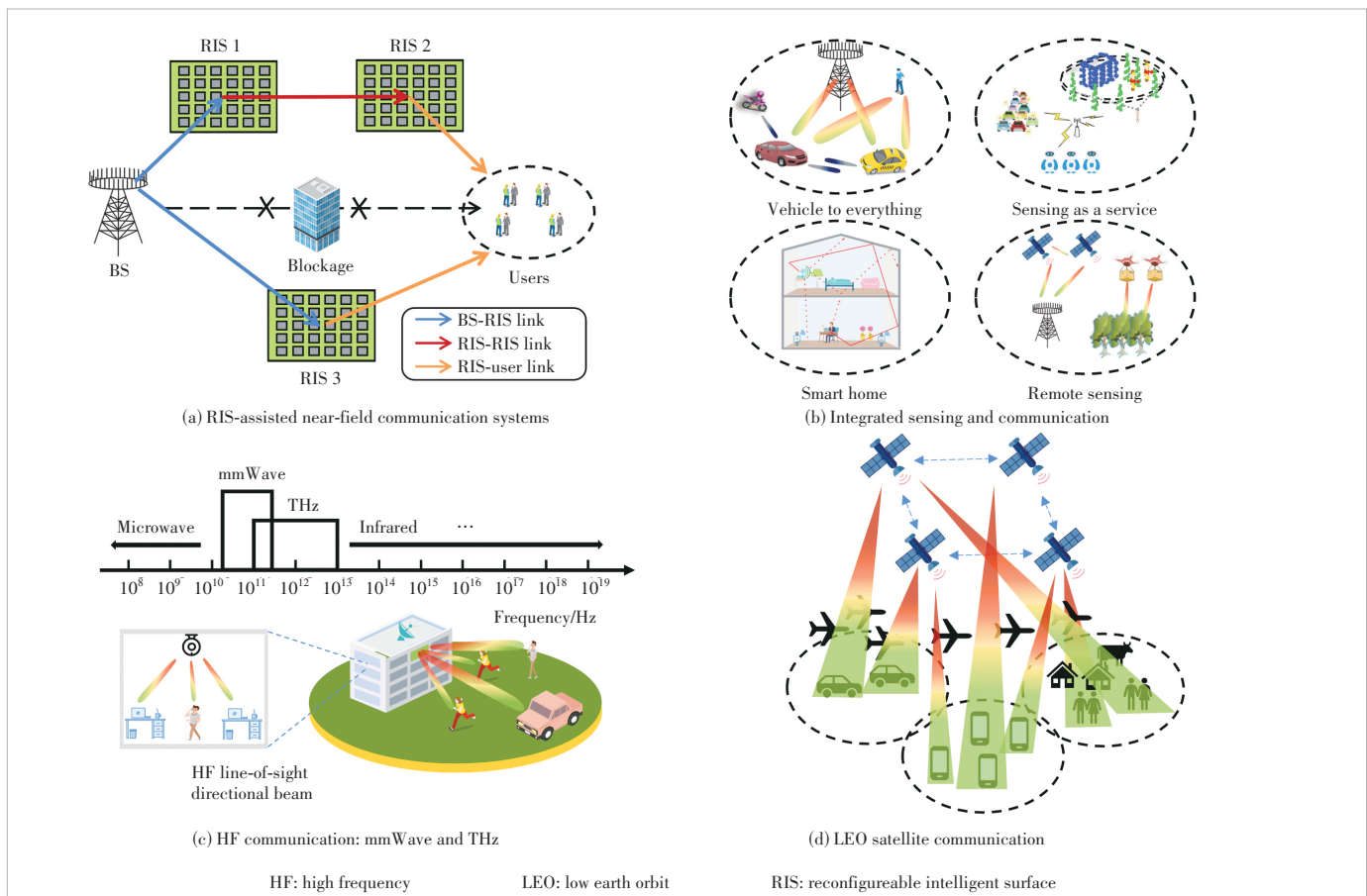
There are many potential application scenarios for near-field beam training, and four categories are highlighted here: reconfigure able intelligent surface (RIS) assisted near-field communication systems^[25], integrated sensing and communication (ISAC) systems^[26], high frequency (HF) communication^[27] and low earth orbit (LEO) satellite communication^[28], as shown in Fig. 6.

3.4.1 RIS-Assisted Near-Field Communication

As a key technology for future wireless communications, RIS can manipulate electromagnetic waves, thereby improving the efficiency of spectrum and energy^[29]. In particular, with a large number of passive reflective units, RIS can be programmed to control the propagation environment of wireless communications to achieve signal enhancement or interference suppression. For high-frequency communications in HMIMO scenarios such as millimeter wave communications, the direct path between the base station and the user can be easily blocked by obstacles. By deploying one or more RISs^[30], a virtual LoS BS-RIS-user link can be provided to improve its communication performance. As shown in Fig. 6(a), three kinds of links exist in the RIS-assisted communication scenario: the BS-RIS link, the RIS-RIS link, and the RIS-user link. In order to obtain a large beamforming gain, the corresponding beam training is required between all three paths to establish the link with high SNR. Because of the high frequencies and the fact that both BS and RIS can be considered as large aperture antenna arrays, all three links tend to be in near-field communication. Therefore, beam training methods in near-field communication can be applied to extend the coverage of the communication and improve the signal-to-noise ratio (SNR) of the paths.

3.4.2 ISAC

ISAC, where both sensing and communication functionalities share the same frequency bands and hardware, has emerged as a crucial technology for future wireless systems^[31].



▲ Figure 6. Potential application scenarios

Many application scenarios for future communications, such as vehicle to everything, sensing as a service, smart home, remote sensing, and other future wireless application scenarios shown in Fig. 6(b), require high performance sensing and wireless communication technologies. Meanwhile, HMIMO, as a key technology in future communication, enables future communication signals to have higher resolution in both time and angular domains, which provides a technical basis for ISAC. In future ISAC communication scenarios, extremely large-scale antenna arrays at the BS or techniques incorporating RIS will introduce the characteristics of near-field regions and high-frequency communication^[32]. Therefore, near-field based beam scanning training can be applied here and used to improve the signal quality for sensing and communication, and to enhance the overall ISAC sensing and wireless communication capabilities.

3.4.3 HF Communication: MmWave and THz

Terahertz and mmWave bands can be used for next-generation mobile communications to enhance mobile broadband because of their abundant spectrum resources. Through the technology of multiple antennas or multiple reflection units in HMIMO, the control of high-frequency signals is more

stable and reliable. To summarize, mmWave and terahertz can achieve point-to-point transmission at a high speed and short distances according to their characteristics^[33]. Many applications, such as mmWave and terahertz BS, can be arranged indoors and outdoors as shown in Fig. 6(c). High-bandwidth wireless communication is achieved through beam management and allocation based on the distance and angle information between the terminal and the BS. However, the signal can be easily blocked by objects such as buildings, trees, and rainwater due to its high frequency, which can result in signal attenuation and degradation of transmission quality. Therefore, when performing point-to-point transmission, the transmission path of the signal and environmental factors need to be taken into account to avoid excessive obstruction of the signal. At this time, the application of beam training methods in the near field can better manage and align the beams of high frequency and improve the stability of wireless communication.

3.4.4 LEO Satellite Communication

LEO communications have a wide range of applications in many fields^[28], as shown in Fig. 6(d). LEO satellites can be used in the field of aviation to provide wide-area coverage of communication services for aircraft. It can support applica-

tions such as in-flight Internet, flight information transmission, communication and navigation, improving the safety of aircraft and the comfort of passengers. LEO satellites can also cover marine areas and provide communication services for ships at sea. This is of great significance for communication between ships, remote monitoring and rescue. LEO satellites can make up for the lack of communication infrastructure in remote areas and provide people with communication services such as telephone, Internet and data transmission. This is also important for residents, researchers and rescue teams in remote areas. Communication links between LEO satellites and terminals usually face problems such as fading and multipath interference^[34]. In order to obtain lower delay and larger bandwidth, fast beam training is very important to accurately point the beam direction of the antenna towards the terminal, thus enhancing the signal strength, improving the quality of the received and transmitted signals, and further reducing the interference and improving the spectral efficiency. Meanwhile, LEO satellite systems usually have multiple beams and can provide services to multiple users at the same time, so the fast multi-beam training algorithm has a wide range of application prospects here as well.

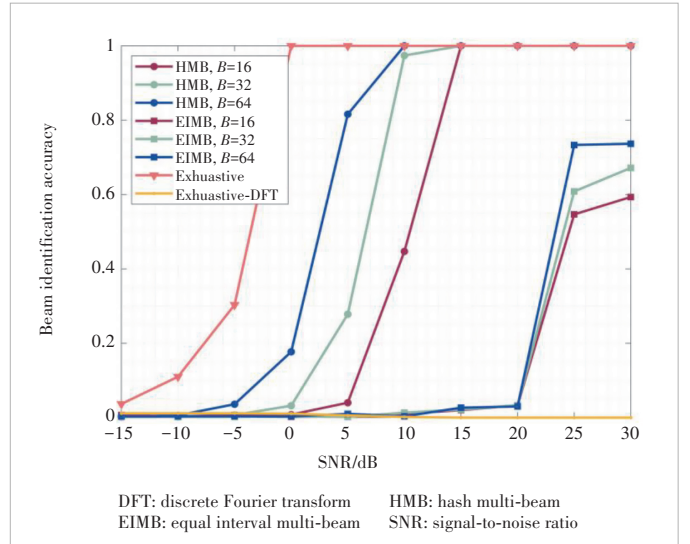
4 Simulation Results and Testbed

We now use the simulation results to evaluate the performance of the typical beam training methods previously presented in Section 3. It is worth stating that we focus on the EIMB and HMB methods because the multi-beam scheme is more efficient in terms of beam training overhead. The number of BSs and the operating frequency are set to $K = 5$ and $f_c = 28$ GHz respectively, and the signal wavelength is $\lambda_c = 0.01$ m. The planar antenna array of BS contains $N_x = 4, N_z = 128$ antennas, and the spacing between the antennas is $d_x = d_z = \lambda_c/2$. The reference SNR is $\gamma = \frac{P_0 N_x N_z \rho_0}{r_0^2 \sigma^2}$ with $\rho_0 = -72$ dB, $P_0 = 15$ dBm, and $\sigma^2 = -70$ dBm. And the achievable rate in $\text{bit} \cdot \text{s}^{-1} \cdot \text{Hz}^{-1}$ is given by $R = \log_2 \left(1 + \gamma \left| f_{BB}^T F_{RF}^T \mathbf{g}_1 \right|^2 \right)$.

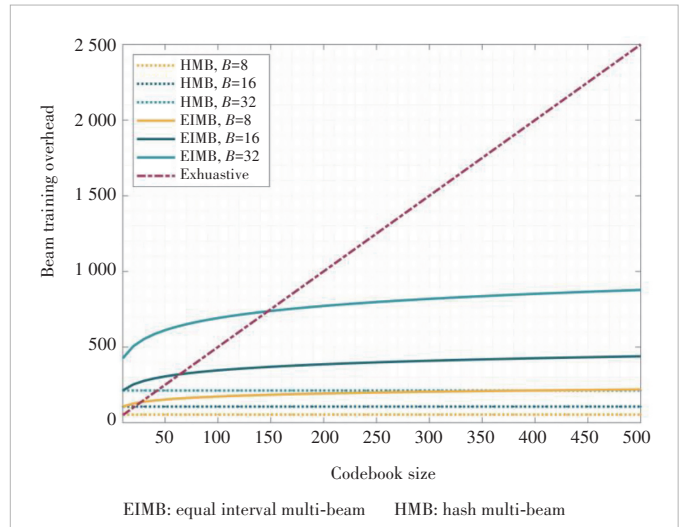
Fig. 7 plots the effect of SNR on the identification accuracy. With the same simulation setup, we use EIMB training with the near-field codebook and exhaustive training with the discrete Fourier transform (DFT) codebook (“Exhaustive-DFT”) as the baseline. Firstly, it can be seen that with increasing SNR, the influence of noise becomes smaller and the identification accuracy of all beam training methods gradually increases; under the exhaustive beam training, the accuracy with the near-field codebook converges to 1, while that with the far-field codebook is significantly lower, which confirms the effectiveness of the designed codebook in near-field conditions. In addition, to compare the performance of different beam training methods, the performance of EIMB is quite

lower compared with the exhaustive and HMB approaches when utilizing near-field codebooks. Moreover, when the SNR is relatively low, the identification accuracy is considerably improved compared with EIMB training. However, it can be noticed that as the number of multi-beams B decreases, the accuracy gradually decreases, since the number of sub-beams R increases as B decreases, which makes the leakage interference between sub-beams have a greater effect on the identification result.

Fig. 8 plots the effect of different codebook sizes on the beam training overhead. Beam training overhead is determined by the size of the space scanned by each beam training algorithm. To meet the actual situation without affecting the results of the situation, increased single search overhead weights and other parameters, such as the exhaustive algorithm’s overhead, are proportional to the size of the codebook. The exhaustive approach traverses the entire beam space and



▲ Figure 7. Success beam identification accuracy versus SNR



▲ Figure 8. Training overhead versus codebook size

the training time is proportional to the codebook size, resulting in a very high training overhead. HMB training is at the logarithmic level with a training time of $Q = BL = O(B \log N_x N_z)$, which can significantly reduce the training overhead compared with exhaustive beam training. Based on the simulation results, we can initially judge the superiority of the hash beam training algorithm in terms of training overhead and accuracy.

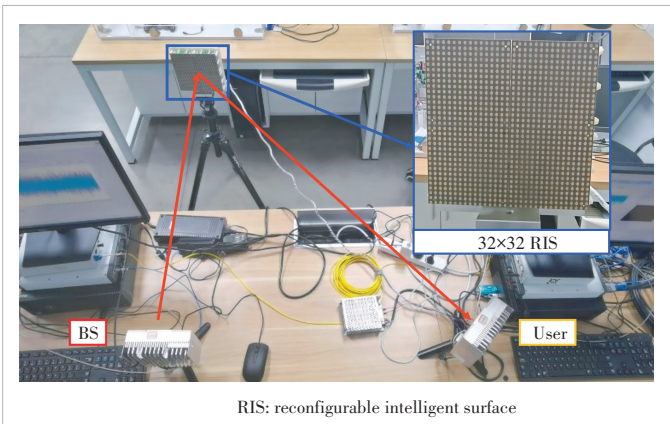
We conduct practical testing of the near-field beam training algorithm in a scenario involving RIS-assisted mmWave communication. Fig. 9 shows the test scenario, where the hardware equipment for the transmitter and receiver consists of a software radio USRP X410, a mmWave RF front-end, and a RIS, respectively. By applying the aforementioned algorithm, we are able to estimate the angle and distance from the RIS to the User, assuming that we have prior knowledge of the angle and distance from the BS to the RIS. Note that the codewords designed here are phase-encoded codewords for RIS, which is reasonable because large-scale RIS is also an application scenario for near-field communication under HMIMO. We verified the effectiveness of the above near-field beam training method through real hardware tests. In the LabVIEW front panel of Fig. 10, the simple implementation of hash beam scanning is shown, which includes the RIS codebook for each time slot tested, the corresponding beam direction maps of the codebook, the polling results corresponding to different hash functions, and the merged polling results. Due to the limitations of the hardware equipment, it is currently not possible to verify the time complexity of comparing the existing methods.

5 Research Challenges and Opportunities

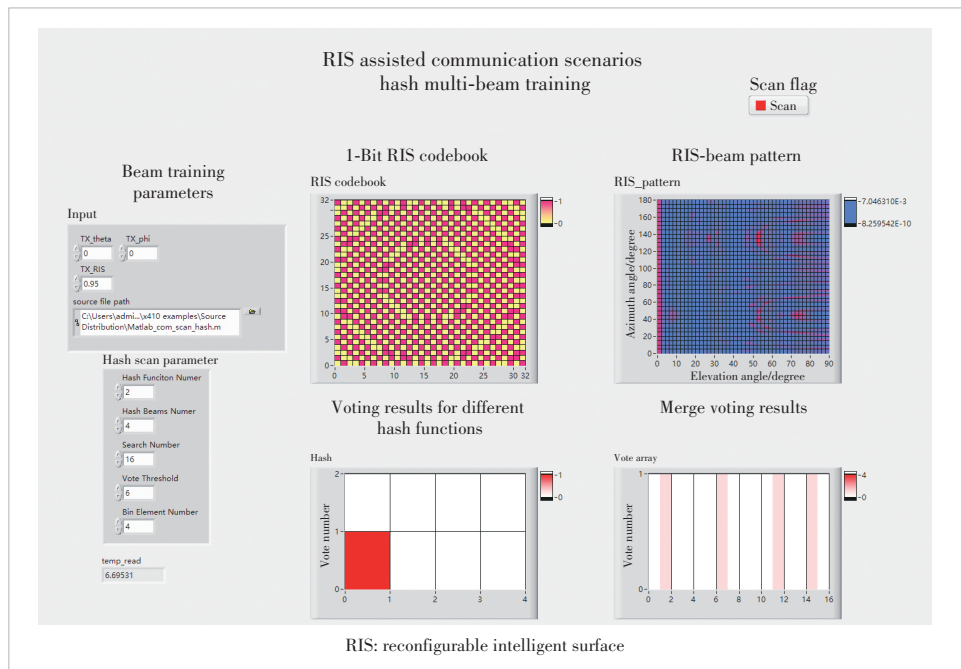
The most crucial difference between near-field beam training and far-field beam training lies in the change of search dimension and search space, which is also the challenge of near-field beam training. The increase in the search dimension and in the search space makes higher demands on the time complexity and convergence of the beam training algorithm. Due to the inherent unity of the problem, we can generalize possible future research directions in far-field beam training to near-field beam training. Also, possible research directions can be planned based on the characteristics of near-field communication itself. In addition to the independent near-field and far-field

beam training, there are also complex scenarios that require hybrid beam training when the user distance is unknown. The following are some potential future research directions and opportunities.

1) Machine learning based beam training: By using machine learning methods^[35-36], learning the data and interacting with the environment continuously, a corresponding beam training method can be obtained. For example, it is possible^[37] to use a database containing past beam measurements at a specific location to determine the direction of non-line-of-sight (NLoS) link beam formation. A deep learning network^[38] is used to predict beam formation vectors using features that carry information about the surrounding environment. A multi-armed bandit framework^[39] is proposed which can train multiple beam pairs in each attempt. In the multi-armed bandit



▲ Figure 9. RIS-assisted mmWave communication testbed



▲ Figure 10. Schematic diagram of the software interface

setup, for weapons with high probability of getting high rewards, more attempts are required to get a certain cumulative reward. Combining the adaptive learning capabilities of machine learning with the domain knowledge of wireless communications allows interaction with changing environments^[40], thus enabling algorithms to adapt effectively to new environments.

2) Near-field characteristics-based beam training: The characteristics of near-field communication are the challenges and possible future research directions for studying near-field beam training methods. For example, for HMIMO near-field communication systems, the near-field beam splitting effect poses a new challenge to beam training, where beams generated at different frequencies are distorted to different locations. To overcome this challenge, an effective broadband rainbow beam training scheme is proposed^[41], which utilizes a true time delay (TTD) device to flexibly control the near-field beam splitting effect in order to achieve fast beam training. Also, based on the distance-domain and angle-domain dimensional challenges due to the near-field spherical wave properties, a 2D polar-domain codebook is proposed^[21], where each beam codeword points to a specific location with a target angle and distance. In terms of angle and distance sampling, the sampling can be designed to minimize the column coherence of neighboring code words.

6 Conclusions

In this paper, we provide a comprehensive overview of near-field beam training for HMIMO communication scenarios. Specifically, we first introduce the characteristics of near-field communication in HMIMO scenarios, and the differences between beam training in near-field and far-field conditions. Then, based on the system model of the interest scenario and the codebook forms for beam training, we introduce several typical methods for near-field beam training: exhaustive beam training, hierarchical beam training and multi-beam training. Also, we discuss some potential application scenarios of near-field beam training. Our performance analysis compares the effectiveness of these methods and confirms their validity on the hardware testbed. Moreover, we discuss the research challenges of near-field beam training and propose several possible future research directions. In the future, we will propose more efficient near-field beam training methods for HMIMO scenarios.

References

- [1] PIZZO A, SANGUINETTI L, MARZETTA T L. Fourier plane-wave series expansion for holographic MIMO communications [J]. *IEEE transactions on wireless communications*, 2022, 21(9): 6890 – 6905. DOI: 10.1109/TWC.2022.3152965
- [2] HUANG C W, HU S, ALEXANDROPOULOS G C, et al. Holographic MIMO surfaces for 6G wireless networks: opportunities, challenges, and trends [J]. *IEEE wireless communications*, 2020, 27(5): 118 – 125. DOI: 10.1109/MWC.001.1900534
- [3] HU S, RUSEK F, EDFORS O. Beyond massive MIMO: the potential of data transmission with large intelligent surfaces [J]. *IEEE transactions on signal processing*, 2018, 66(10): 2746 – 2758. DOI: 10.1109/TSP.2018.2816577
- [4] DARDARI D, DECARLI N. Holographic communication using intelligent surfaces [J]. *IEEE communications magazine*, 2021, 59(6): 35 – 41. DOI: 10.1109/MCOM.001.2001156
- [5] CHEN Z, MA X Y, ZHANG B, et al. A survey on terahertz communications [J]. *China communications*, 2019, 16(2): 1 – 35. DOI: 10.12676/j.cc.2019.02.001
- [6] WANG J Y, LAN Z, PYO C W, et al. Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems [J]. *IEEE journal on selected areas in communications*, 2009, 27(8): 1390 – 1399. DOI: 10.1109/JSAC.2009.091009
- [7] HUANG Y, WU Q Q, LU R, et al. Massive MIMO for cellular-connected UAV: challenges and promising solutions [J]. *IEEE communications magazine*, 2021, 59(2): 84 – 90. DOI: 10.1109/MCOM.001.2000552
- [8] SAAD W, BENNIS M, CHEN M Z. A vision of 6G wireless systems: Applications, trends, technologies, and open research problems [J]. *IEEE network*, 2020, 34(3): 134 – 142. DOI: 10.1109/MNET.001.1900287
- [9] VA V, CHOI J, HEATH R W. The impact of beamwidth on temporal channel variation in vehicular channels and its implications [J]. *IEEE transactions on vehicular technology*, 2017, 66(6): 5014 – 5029. DOI: 10.1109/TVT.2016.2622164
- [10] SELVAN K T, JANASWAMY R. Fraunhofer and Fresnel Distances: Unified derivation for aperture antennas [J]. *IEEE antennas and propagation magazine*, 2017, 59(4): 12 – 15. DOI: 10.1109/MAP.2017.2706648
- [11] HUR S, KIM T, LOVE D J, et al. Millimeter wave beamforming for wireless backhaul and access in small cell networks [J]. *IEEE transactions on communications*, 2013, 61(10): 4391 – 4403. DOI: 10.1109/TCOMM.2013.090513.120848
- [12] HE T, XIAO Z Y. Suboptimal beam search algorithm and codebook design for millimeter-wave communications [J]. *Mobile networks and applications*, 2015, 20(1): 86 – 97. DOI: 10.1007/s11036-015-0568-5
- [13] XIAO Z Y, HE T, XIA P F, et al. Hierarchical codebook design for beamforming training in millimeter-wave communication [J]. *IEEE transactions on wireless communications*, 2016, 15(5): 3380 – 3392. DOI: 10.1109/TWC.2016.2520930
- [14] QI C H, CHEN K J, DOBRE O A, et al. Hierarchical codebook-based multiuser beam training for millimeter wave massive MIMO [J]. *IEEE transactions on wireless communications*, 2020, 19(12): 8142 – 8152. DOI: 10.1109/TWC.2020.3019523
- [15] HEADLAND D, MONNAI Y, ABBOTT D, et al. Tutorial: terahertz beamforming, from concepts to realizations [J]. *APL photonics*, 2018, 3(5): 051101. DOI: 10.1063/1.5011063
- [16] WU Z D, DAI L L. Multiple access for near-field communications: SDMA or LDMA? [J]. *IEEE journal on selected areas in communications*, 2023, 41(6): 1918 – 1935. DOI: 10.1109/JSAC.2023.3275616
- [17] ALKHATEEB A, MO J H, GONZALEZ-PRELCIC N, et al. MIMO precoding and combining solutions for millimeter-wave systems [J]. *IEEE communications magazine*, 2014, 52(12): 122 – 131. DOI: 10.1109/MCOM.2014.6979963
- [18] WEI X H, DAI L L, ZHAO Y J, et al. Codebook design and beam training for extremely large-scale RIS: far-field or near-field? [J]. *China communications*, 2022, 19(6): 193 – 204. DOI: 10.23919/JCC.2022.06.015
- [19] CHEN J W, GAO F F, JIAN M N, et al. Hierarchical codebook design for near-field mmWave MIMO communications systems [J]. *IEEE wireless communications letters*, 2023, 12(11): 1926 – 1930. DOI: 10.1109/LWC.2023.3299354
- [20] YOU C S, ZHENG B X, ZHANG R. Fast beam training for IRS-assisted multiuser communications [J]. *IEEE wireless communications letters*, 2020, 9(11): 1845 – 1849. DOI: 10.1109/LWC.2020.3005980

- [21] CUI M Y, DAI L L. Channel estimation for extremely large-scale MIMO: far-field or near-field? [J]. *IEEE transactions on communications*, 2022, 70(4): 2663 – 2677. DOI: 10.1109/TCOMM.2022.3146400
- [22] WANG P L, FANG J, ZHANG W Z, et al. Beam training and alignment for RIS-assisted millimeter-wave systems: state of the art and beyond [J]. *IEEE wireless communications*, 2022, 29(6): 64 – 71. DOI: 10.1109/MWC.006.2100517
- [23] YOU C, ZHANG Y, WU C, et al. Near-field beam management for extremely large-scale array communications [EB/OL]. (2023-06-28) [2023-12-12]. <https://arxiv.org/abs/2306.16206>
- [24] NOH S, ZOLTOWSKI M D, LOVE D J. Multi-resolution codebook based beamforming sequence design in millimeter-wave systems [C]//*Global Communications Conference (GLOBECOM)*. IEEE, 2015: 1 – 6. DOI: 10.1109/GLOCOM.2015.7417848
- [25] HUANG C W, ZAPPONE A, ALEXANDROPOULOS G C, et al. Reconfigurable intelligent surfaces for energy efficiency in wireless communications [J]. *IEEE transactions on wireless communications*, 2019, 18(8): 4157 – 4170. DOI: 10.1109/TWC.2019.2922609
- [26] LIU A, HUANG Z, LI M, et al. A survey on fundamental limits of integrated sensing and communication [J]. *IEEE communications surveys & tutorials*, 2022, 24(2): 994 – 1034. DOI: 10.1109/COMST.2022.3149272
- [27] SONG H J, NAGATSUMA T. Present and future of terahertz communications [J]. *IEEE transactions on terahertz science and technology*, 2011, 1(1): 256 – 263. DOI: 10.1109/TTHZ.2011.2159552
- [28] SU Y T, LIU Y Q, ZHOU Y Q, et al. Broadband LEO satellite communications: architectures and key technologies [J]. *IEEE wireless communications*, 2019, 26(2): 55 – 61. DOI: 10.1109/MWC.2019.1800299
- [29] BASAR E, DI RENZO M, DE ROSNY J, et al. Wireless communications through reconfigurable intelligent surfaces [J]. *IEEE access*, 2019, 7: 116753 – 116773. DOI: 10.1109/ACCESS.2019.2935192
- [30] LIU M B, LI X, NING B Y, et al. Deep learning-based channel estimation for double-RIS aided massive MIMO system [J]. *IEEE wireless communications letters*, 2023, 12(1): 70 – 74. DOI: 10.1109/LWC.2022.3217294
- [31] LIU F, CUI Y H, MASOUIROS C, et al. Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond [J]. *IEEE journal on selected areas in communications*, 2022, 40(6): 1728 – 1767. DOI: 10.1109/JSAC.2022.3156632
- [32] WANG Z L, MU X D, LIU Y W. STARS enabled integrated sensing and communications [J]. *IEEE transactions on wireless communications*, 2023, 22(10): 6750 – 6765. DOI: 10.1109/TWC.2023.3245297
- [33] EL AYACH O, RAJAGOPAL S, ABU-SURRA S, et al. Spatially sparse precoding in millimeter wave MIMO systems [J]. *IEEE transactions on wireless communications*, 2014, 13(3): 1499 – 1513. DOI: 10.1109/TWC.2014.011714.130846
- [34] YOU L, LI K X, WANG J H, et al. Massive MIMO transmission for LEO satellite communications [J]. *IEEE journal on selected areas in communications*, 2020, 38(8): 1851 – 1865. DOI: 10.1109/JSAC.2020.3000803
- [35] ZHU F H, WANG B H, YANG Z H, et al. Robust millimeter beamforming via self-supervised hybrid deep learning [C]//*The 31st European Signal Processing Conference (EUSIPCO)*. IEEE, 2023: 915 – 919. DOI: 10.23919/EUSIPCO58844.2023.10289989
- [36] WANG X Q, ZHU F H, ZHOU Q Y, et al. Energy-efficient beamforming for RISs-aided communications: gradient based meta learning [EB/OL]. (2023-11-12) [2023-12-10]. <https://arxiv.org/abs/2311.06861>
- [37] AVILES J C, KOUKI A. Position-aided mm-wave beam training under NLOS conditions [J]. *IEEE access*, 2016, 4: 8703 – 8714. DOI: 10.1109/ACCESS.2016.2631222
- [38] ALKHATEEB A, ALEX S, VARKEY P, et al. Deep learning coordinated beamforming for highly-mobile millimeter wave systems [J]. *IEEE access*, 2018, 6: 37328 – 37348. DOI: 10.1109/ACCESS.2018.2850226
- [39] CHEN W, WANG Y J, YUAN Y. Combinatorial multi-armed bandit: general framework, results and applications [C]//*International Conference on Machine Learning*. ACM, 2013: 151 – 159. DOI: 10.5555/3042817.3042836
- [40] ZHANG J J, HUANG Y M, WANG J H, et al. Intelligent interactive beam training for millimeter wave communications [J]. *IEEE transactions on wireless communications*, 2021, 20(3): 2034 – 2048. DOI: 10.1109/TWC.2020.3038787
- [41] CUI M Y, DAI L L, WANG Z C, et al. Near-field rainbow: wideband beam training for XL-MIMO [J]. *IEEE transactions on wireless communications*, 2023, 22(6): 3899 – 3912. DOI: 10.1109/TWC.2022.3222198

Biographies

SHEN Jiayu is currently pursuing his BS degree with the College of Information Science and Electronic Engineering at Zhejiang University, China. His research interests focus on near-field communications.

YANG Jun received his BS degree in geophysics from the China University of Geosciences (Wuhan), China in 2011 and the joint-training PhD degree in geophysics from the University of Science and Technology of China and the University of North Carolina, USA in 2016. He is currently a senior algorithm engineer with ZTE Corporation. His research interests include reconfigurable intelligent surfaces, MIMO communications and computational electromagnetics.

ZHU Chen received his BS degree from North University of China in 2010, and MS degree from Zhejiang University of Technology, China in 2013. He is currently engaged in teaching and research at the College of Engineering, Zhejiang University, China. His main research interests include general sense computing integration, machine learning, image processing, and cloud-edge collaborative computing.

DENG Zhiji received his BS degree in communication engineering from Xiamen University in 2008. He is now working in Zhejiang Dahua Technology Co., Ltd., as the president of Central Research Institute/Future Communication Research Institute. His main research interests include global multidimensional sensing and large-scale video networking. He has hosted and participated in eight major scientific research projects at the national and ministerial levels. He has won multiple first and second prizes in science and technology awards at the provincial, ministerial, and national levels. He has won the titles of “High Level Talents” of Zhejiang Province, “Model Worker” of Binjiang, “Outstanding Invention Talent” of the Provincial Invention Association, “Outstanding Engineer” of the China Software Industry Association, etc. He has dominated four international standards of ITU-T and 10 industry standards, and has applied for 270 patents.

HUANG Chongwen (chongwenhuang@zju.edu.cn) received his BS degree from the Binhai College, Nankai University, China in 2010, and MS degree from the University of Electronic Science and Technology of China in 2013. He had worked with the Institute of Electronics, Chinese Academy of Sciences (IECAS) as a research engineer since July 2013. From September 2015, he had started his PhD journey with Singapore University of Technology and Design (SUTD), and CentraleSupélec University, France under the supervision of Prof. Chau YUEN and Prof. Mérouane DEBBAH. From October 2019 to September 2020, he was a post-doctoral researcher at SUTD. Since September 2020, he has been working with Zhejiang University, China as a tenure-track young professor. His main research interests include holographic MIMO surface/reconfigurable intelligent surface, B5G/6G wireless communications, mmWave/THz communications, and deep learning technologies for wireless communications. He was a recipient of the IEEE Marconi Prize Paper Award in wireless communications in 2021. He received the Singapore Government PhD Scholarship and PHC Merlion PhD Grant (2016 – 2019) for studying in CentraleSupélec, France. In addition, he has served as the chair of several wireless communications flagship conferences, including the session chair of 2021 IEEE WCNC, 2021 IEEE VTC-Fall, and the symposium chair of IEEE WCSP 2021.



Near-Field Wireless Power Transfer, Sensing and Communication with Bessel Beams

CAO Xinghan¹, YIN Huarui¹, YOU Changsheng²

(1. University of Science and Technology of China, Hefei 230026, China;

2. Southern University of Science and Technology, Shenzhen 518055, China)

DOI: 10.12142/ZTECOM.202401007

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240308.1551.004>, published online March 11, 2024

Manuscript received: 2024-01-19

Abstract: The Bessel beam, characterized by its unique non-diffracting properties, holds promising applications. In this paper, we provide a detailed introduction and investigation into the theory and research of the Bessel beam, with a special focus on its generation and applications in the near-field region. We provide an introduction to the concepts, properties, and foundational theories of the Bessel beam. Additionally, the current study on generating Bessel beams and their applications is categorized and discussed, and potential research challenges are proposed in this paper. This review serves as a solid foundation for researchers to understand the concept of the Bessel beam and explore its potential applications.

Keywords: Bessel beams; near-field; non-diffractive beams; beam synthesis technology

Citation (Format 1): CAO X H, YIN H R, YOU C S. Near-field wireless power transfer, sensing and communication with Bessel beams [J]. *ZTE Communications*, 2024, 22(1): 53 - 61. DOI: 10.12142/ZTECOM.202401007

Citation (Format 2): X. H. Cao, H. R. Yin, and C. S. You, "Near-field wireless power transfer, sensing and communication with Bessel beams," *ZTE Communications*, vol. 22, no. 1, pp. 53 - 61, Mar. 2024. doi: 10.12142/ZTECOM.202401007.

1 Introduction

The next generation of mobile communication technology is advancing towards higher frequency bands^[1]. Meanwhile, new concepts such as the Internet of Everything (IoE) and dense networking are considered as the future direction of mobile communications^[2]. These trends collectively imply that devices will communicate over shorter wavelengths in closer proximity, shifting the challenges of communication and sensing from the far-field region to the near-field region^[3]. One of the significant differences between the far-field region and near-field region is that both angle and distance simultaneously affect signal transmission between transmitting and receiving devices. As a result, more factors need to be precisely considered when modeling signals^[4]. This presents new opportunities and challenges for the design of beamforming.

The Bessel beam, as a type of beam confined to the near-field region, was first introduced by DURNIN et al. in 1987^[5]. Starting from the Helmholtz equation in free space, a non-singular solution, namely, the Bessel beam, was proposed for describing a non-diffracting scalar field, in which "non-diffracting" means that the intensity pattern of the beam in a transverse plane remains unchanged along the propagation direction and the term "Bessel" is derived from the matching of the expression of the beam with Bessel functions. The Bessel beam exhibits unique

and intriguing non-diffracting and self-healing properties, enabling it to maintain a stable intensity and lateral shape along the propagation direction. Additionally, the Bessel beam exhibits a series of excellent properties, including a small main lobe size, great directionality, high intensity, and long depth of focus, making it a competitive candidate for various applications. However, an ideal Bessel beam carries infinite energy and requires an infinite generation aperture, which is physically unattainable. In practical applications, Bessel beams with finite energy can only be generated on truncated apertures, which we refer to as "pseudo-Bessel beams". The propagation distance of such beams is limited by the aperture size and axicon angle. Consequently, the generation of pseudo-Bessel beams is restricted to the near-field region. In this paper, the term "the Bessel beam" mainly refers to "the pseudo-Bessel beam", and we no longer distinguish between these two concepts.

Due to the constraints imposed by their near-field characteristic, practically generated Bessel beams are often confined to very short distances. This has confined the study of Bessel beams primarily to experimental generation for a long time. There have been numerous attempts to generate Bessel beams in the fields of optics, acoustics, and microwaves. These studies explore methods for generating Bessel beams with different devices and materials, such as phase-shift surface antennas^[6-7],

fish-eye lenses^[8], metasurfaces^[9-11], metamaterial lenses^[12], phased array antennas^[13], radial line slot array (RLSA)^[14-15], and leaky radial waveguides^[16]. However, in recent years, with the increase in communication and sensing scenarios of higher frequency bands in the near-field region and the in-depth study of Bessel beams, the bright application potential of the Bessel beam is gradually being discovered and explored in areas such as wireless power transfer (WPT), sensing, and communication, especially in the near-field region. These studies have helped demonstrate the significance of researching the generation of Bessel beams and further explore the unique properties of Bessel beams.

In this paper, we conduct a survey on the theory and research related to the Bessel beam. A brief review of the concept and basic theory of Bessel beams is provided in Section 2. Approaching from the perspectives of generation and applications, the current research on Bessel beams are categorized and discussed in Sections 3 and 4, respectively. Additionally, insightful discussions on the challenges existing in research on Bessel beams are provided. This review aims to aid researchers in understanding the concepts of Bessel beams and exploring their potential applications.

2 Theory of Bessel Beam

The concept of the Bessel beam is inspired by the scalar wave equation in free space. When considering only the intensity of the electric field and neglecting its directionality, the wave equation can be expressed by the following scalar expression^[5]:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right)E(x, y, z, t) = 0. \quad (1)$$

The Bessel beam, as an exact non-singular solution to the above equation for scalar fields propagating into the source-free region $z \geq 0$, can be expressed in the axisymmetric form as follows:

$$E(x, y, z, t) = \exp[i(\beta z - \omega t)] \int_0^{2\pi} \exp[i\alpha(x \cos\phi + y \sin\phi)] \frac{d\phi}{2\pi} = \exp[i(\beta z - \omega t)] J_0(\alpha\rho), \quad (2)$$

where x , y and z represent the three-dimensional coordinates, and $\rho^2 = x^2 + y^2$; β and α represent the longitudinal and transverse wave vectors, respectively, and $\beta^2 + \alpha^2 = (\omega/c)^2$; J_0 represents the zero-order Bessel function of the first kind. The most remarkable feature of the Bessel beam is that, when the wave number β is real, it generates a non-diffracting field, that is, its time-averaged intensity profile is independent of z :

$$I(x, y, z \geq 0) = \frac{1}{2} |E(x, y, z, t)|^2 = I(x, y, z = 0). \quad (3)$$

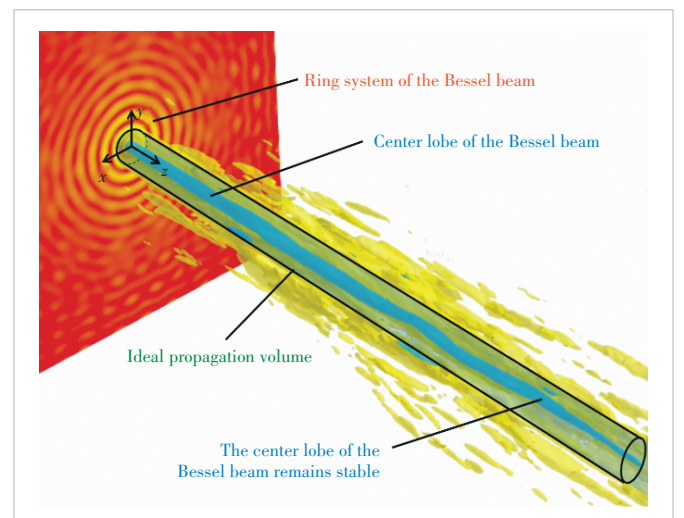
For reference, an ideal Bessel beam profile is illustrated in Fig. 1^[17]. Compared with other forms of beam focusing, such as

linear beamforming and focused beams, the key distinction of Bessel beams lies in their objective to achieve focusing along a line. The Bessel beam can be conceptually understood as “a beam propagating along a line”, spreading in space as a cylinder. On the contrary, the traditional phased array linearly shifted beam can be understood as “a beam pointing in a direction”, spreading in space as a cone, and the focused beam can be considered as a “beam pointing to a point”, achieving spatial concentration at a single point. Therefore, compared with other beamforming techniques, Bessel beams have a larger focusing range and the advantage of not diffusing in the propagation direction, allowing for their self-healing property.

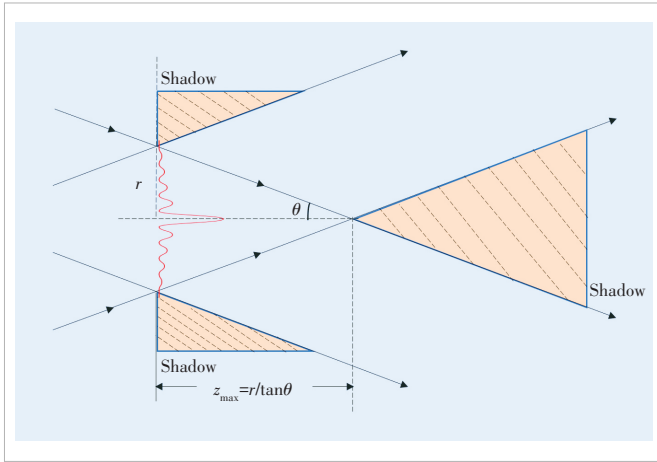
However, one can easily find from Eq. (2) that an ideal Bessel beam would require an infinite generating aperture and infinite energy because it is not square integrable. The actually generated Bessel beam is represented by a truncated version of Eq. (2) with a finite aperture at the xOy plane, given by the following expression:

$$E(x, y, z = 0, t) = \begin{cases} \exp[i(\beta z - \omega t)] \int_0^{2\pi} \exp[i\alpha(x \cos\phi + y \sin\phi)] \frac{d\phi}{2\pi} = \\ \exp[i(\beta z - \omega t)] J_0(\alpha\rho) (\rho \leq R) \\ 0 \quad (\rho > R) \end{cases}. \quad (4)$$

The truncated Bessel beam is illustrated in Fig. 2. Based on Fig. 2 and Eq. (4), it can be observed that the truncated Bessel beam can be regarded as a superposition of a series of plane waves with the same amplitude and the same angle relative to the z -axis, but different azimuthal angles distributed in the range of 0 to 2π . Such a series of plane waves will produce a geometrical shadow zone, i.e., a region where the truncated Bessel beam cannot maintain its non-diffracting characteristics. Therefore, the propagation range of the truncated Bessel beam is given as:



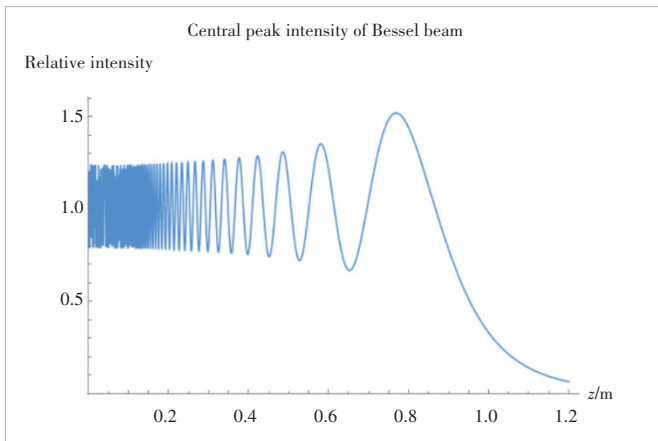
▲ Figure 1. Illustration of an ideal Bessel beam^[17]



▲ Figure 2. Illustration of a truncated Bessel beam

$$z_{\max} = r/\tan\theta = r \left[(2\pi/\alpha\lambda)^2 - 1 \right]^{1/2}, \quad (5)$$

where $\theta = \sin^{-1}(\alpha\lambda/2\pi)$ represents the traveling angle relative to the z -axis and R represents the radius of the aperture. In comparison with the definitions of the reactive near-field and the radiative near-field, Eq. (5) appears to be independent of frequency. However, for practically generated Bessel beams, their frequency and axion angle are constrained. Generally speaking, the propagation distance of Bessel beams is smaller than the Rayleigh distance, and their range is closer to the radiation near-field. Despite the limited propagation distance, the truncated Bessel beam still maintains excellent non-diffracting characteristics within the range of z_{\max} . To investigate the field distribution and characteristics of the truncated Bessel beam in the near-field region, the scalar diffraction theory is adopted, i.e., the Rayleigh-Sommerfeld formula or the Fresnel diffraction integral. The field at each point in free space is described as a weighted sum of the fields on the generating aperture. According to the Rayleigh-Sommerfeld formula, the relationship between the intensity of the truncated Bessel beam and the propagation distance is given in Fig. 3^[18]. It can be observed that the



▲ Figure 3. Central peak intensity of a truncated Bessel beam^[18]

main lobe intensity of the truncated Bessel beam undergoes some fluctuations and rapidly decreases beyond the limit distance given by z_{\max} . In addition, the angular spectrum theory can also be used to analyze the field distribution of truncated Bessel beams. As the research on the Bessel beam has advanced, more complex methods such as the vector wave method and leaky-wave theory have been employed to analyze the effects of Bessel beams^[19]. These methods, beyond the scope of this paper, provide additional insights into the characteristics of Bessel beams.

3 Generation of Bessel Beam

The generation of Bessel beams that are novel and distinctive structured beams has been a key area of research in this field. In this section, we categorize the research on generating Bessel beams into two types: based on the scalar wave theory and based on the vector wave theory. We provide an insightful review and discussion of these studies.

3.1 Based on Scalar Wave Theory

The earliest concept of the Bessel beam was derived based on the scalar wave theory. Consequently, most of the early research on generating Bessel beams originated from the scalar wave theory. The objective of these studies can be summarized to achieve a specific phase distribution on the generation aperture that satisfies the requirements for the Bessel beam: the phase at any point on the aperture linearly increases with its distance from the center, as shown in the following equation:

$$\varphi = k_0 \sqrt{x^2 + y^2} \sin\theta, \quad (6)$$

where $k_0 = 2\pi/\lambda$, and θ represents the axion angle. Metasurfaces possess the excellent capability to continuously alter the phase distribution. In Ref. [9], a transmission metasurface was designed for the generation of the Bessel beam, enabling beam steering by adjusting the size of the cells with great transmission efficiency. By altering the length of the annular grooves etched on the dielectric substrate to meet the phase distribution of Bessel beams in the generation aperture, a similar transmissive metasurface design can be found in Ref. [10].

In the field of optics, lenses are often used to manipulate the phase for the formation of Bessel beams. In Ref. [8], a planar half Maxwell fisheye lens and a planar flat lens were combined to generate Bessel beams. The former is employed to transfer the cylindrical like surface waves generated by a point source to plane surface waves, while the latter achieves the modulation of the required phase distribution for the Bessel beam on the plane wave by altering the refractive index distribution of the medium. Such a dual-lens design reduces the demands on the refractive index and thickness of the plane lens surface. By employing gradient index metamaterials, a broadband lens capable of transforming spherical beams emitted from feed into Bessel beams was proposed in Ref. [12], operating in the frequency band of 12 – 18 GHz. Similar to the principles of lenses, phase-

shifting surface (PSS) antennas can also serve as beam converters. Such research can be found in Refs. [6] and [7], where a PSS antenna with broadband and variable polarization characteristics was designed and optimized through maximizing the depth of field and maximizing the transmission distance and an antenna with the ability to freely change the beam direction is proposed, respectively. By combining two PSSs, a near-field rotatable Bessel beam generator with a simple feeding network was proposed in Ref. [20].

In contrast to the methods mentioned above for continuous control of the phase distribution, two-dimensional antenna arrays are used for the Bessel beam generation through sub-sampled distributions, as studied in Ref. [13]. By modeling antenna elements with spherical waves and optimizing the array response on an observation synthesis plane using the least mean square error approach to approximate the expected response of a Bessel beam, the researchers obtained a phase distribution very similar to Eq. (6). However, sub-sampled distributions resulted in deviations in the amplitude distribution on the aperture and the Bessel function. The impact of antenna element spacing and distribution on the generated Bessel beam was also discussed in the research, providing guidance for generating Bessel beams through conventional large antenna arrays.

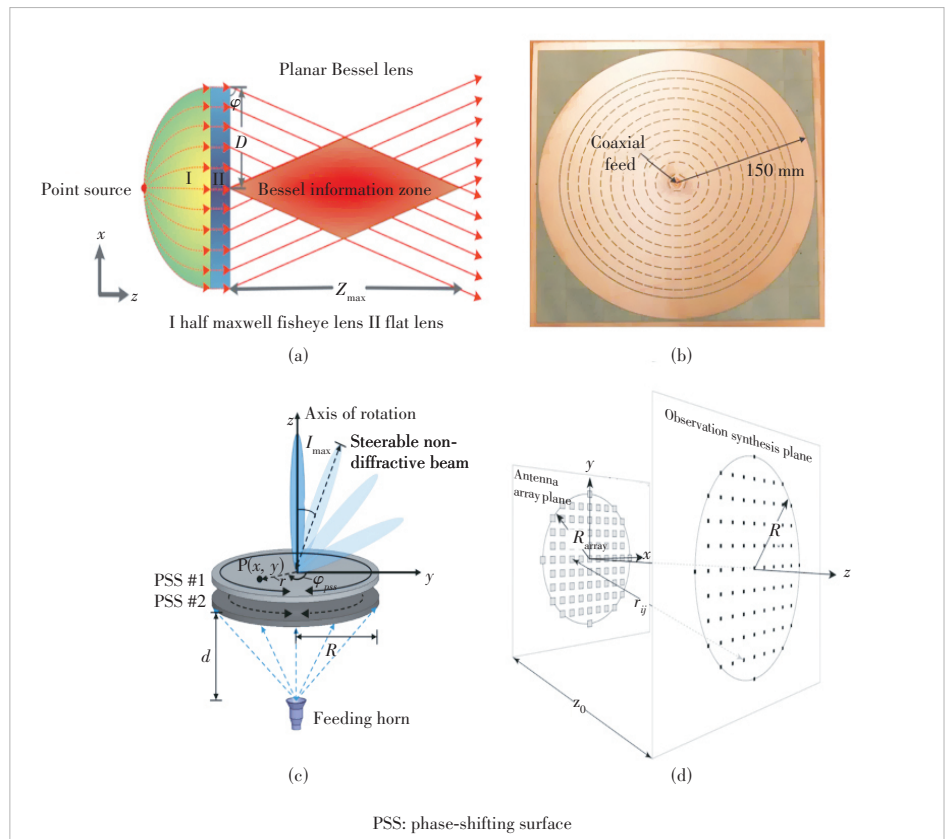
The designs mentioned above require the feed point to be precisely located at the focus of the converter, posing challenges in terms of device integration^[19]. RLSA antennas have been employed to address this issue. In Ref. [14], a holographic technique was applied to synthesize the field distribution on the antenna aperture to generate the desired pseudo-Bessel beam in the near-field region. The RLSA consists of a radial parallel plate waveguide in which a radial propagating mode is excited to feed a slot distribution etched on the upper metallic plate, achieving propagation distances of several hundred wavelengths with very high radiation efficiency (95%) and focusing efficiency (80%). In Ref. [15], the conclusion showed that it was possible to generate Bessel beams by defining an azimuthally invariant inward traveling wave distribution on the RLSA.

Additionally, other studies explore the generation of variations in Bessel beams. In Ref. [21], a shared-aperture strategy was proposed to generate dual asymmetric Bessel beams, propagating in different directions using metasurfaces. In Ref. [22], a metasurface was designed to

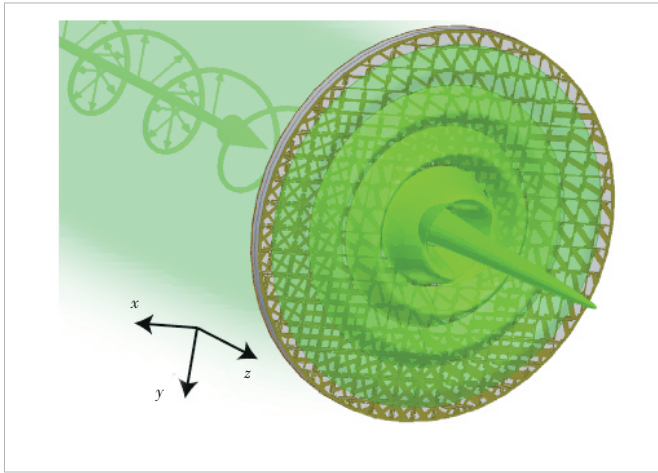
achieve the superposition of two Bessel beams at different axicon angles, providing a more uniform energy distribution for wireless power transfer applications. Some studies also suggest that modulating the excitation function can reduce the intensity oscillations of truncated Bessel beams, thereby achieving a more uniform energy distribution^[23-24]. Fig. 4 shows different structures to generate Bessel beams.

3.2 Based on Vector Wave Theory

However, the scalar wave theory is only applicable for cases where the aperture size is much larger than the operating wavelength. For generating narrow beamwidths, the paraxial approximation is no longer suitable^[25]. Some studies have begun to explore the use of vector wave theory to guide the generation of Bessel beams. This approach necessitates the use of vector potential to derive the distribution of the internal field within the material and the conditions required for generating the desired electric field distribution. The polarization modes of the field, i.e., transverse magnetic (TM), transverse electric (TE), and transverse electromagnetic (TEM), also need to be taken into consideration. In Ref. [11], two metasurfaces with the simultaneous control of wavefront phase and polarization capability were combined to convert linearly and circularly polarized Gaussian beams into vector Bessel beams, demonstrating low loss and high efficiency (Fig. 5).



▲ Figure 4. Different structures to generate Bessel beams: (a) the surface-wave Bessel lens^[8], (b) the top view of the radial line slot array (RLSA) antenna^[15], (c) the steerable antenna^[7], and (d) the antenna array^[13]

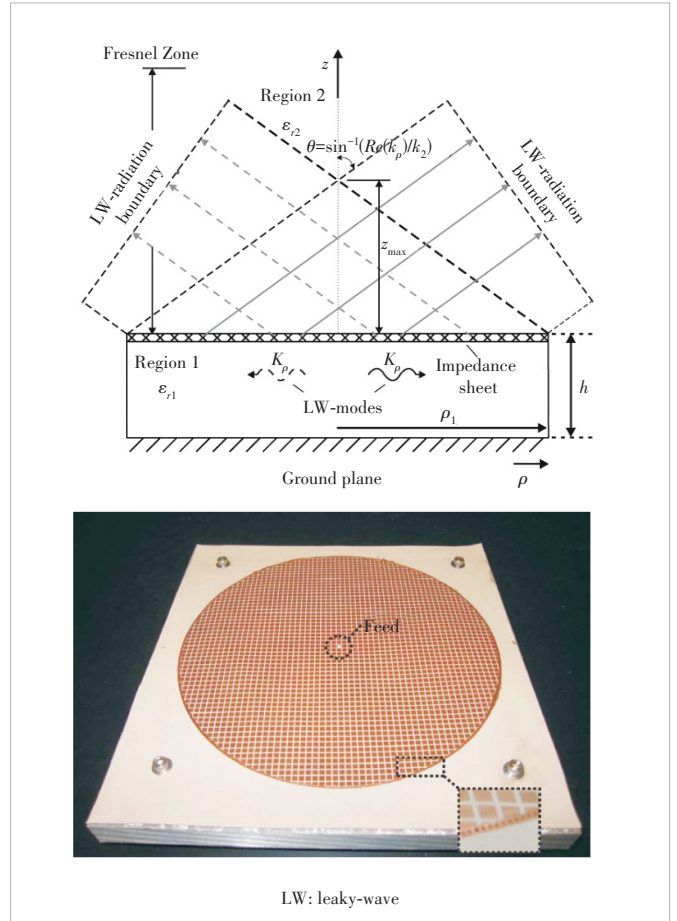


▲ Figure 5. A metasurface transforms a circularly polarized Gaussian beam into a vector Bessel beam with high efficiency^[11]

Other works primarily achieve this through leaky waveguides. In 2011, leveraging the completeness and orthogonality of circular waveguide modes, thin circular loop antennas placed coaxially inside the waveguide were used to generate Bessel beams in the microwave domain^[26]. The study specifically discussed the TE polarization. ETTORRE et al. expanded the possibilities of generating Bessel beams through leaky-wave modes. They proposed the method of generating Bessel beams through a leaky radial waveguide using vector potential analysis in Ref. [25] and experimentally validated it in Ref. [27] (Fig. 6). The proposed structure can generate arbitrary zeroth-order propagating Bessel beams at microwave and millimeter-wave frequencies. In Ref. [16], a TE-polarized Bessel antenna operating at 5.8 GHz, designed by the principle of duality with the TM mode, was proposed using leaky-wave modes. Other studies can be found in Refs. [28] and [29], where a radial parallel plate waveguide (PPW) loaded with annular slots operating at 30 GHz with longitudinal polarization and transverse polarization were proposed respectively. It is worth noting that, although the studies mentioned above focus on the 3 – 30 GHz frequency range, the researchers in Refs. [25] and [27] have indicated that their proposed structures can also be applied to millimeter waves. In other frequency bands, such as acoustic waves, millimeter waves, terahertz, and optical waves, the research on generating Bessel beams can be found in Refs. [30 – 34], and their principles and methods are generally similar.

3.3 Research Challenges

Currently, the efficient and convenient generation of Bessel beams remains a topic of research in the field of Bessel beam studies. There are several research issues worth exploring. Firstly, reducing device complexity and enhancing versatility pose significant challenges in generating Bessel beams. As mentioned earlier, the generation of Bessel beams requires precise phase control, which is currently difficult to achieve with existing devices. Approaches such as using metasurfaces or lenses



▲ Figure 6. Illustration of the Bessel beam leaky radial waveguide^[25] and its experimental prototype^[27]

often involve the integration of multiple devices, leading to additional costs. Furthermore, some phase-controlled designs necessitate highly intricate feed networks. Investigating how to synthesize Bessel beams on existing antenna structures or easily scalable devices is a research proposition worth exploring. Secondly, it is worth investigating how to manipulate the characteristics of generated Bessel beams. For instance, altering the phase and intensity distribution on the generating aperture can change properties such as direction, depth of field, main lobe width, intensity distribution, and others. Lastly, research on generating Bessel beams through vector wave theory is still in its early stages and deserves further exploration. In conclusion, the diffraction analysis and near-field characteristics of Bessel beams necessitate more refined antenna designs. In the future, antenna arrays with higher integration and continuous phase modulation capabilities are expected to address this issue.

4 Application of Bessel Beam

As mentioned earlier, the unique properties of Bessel beams, including non-diffractive characteristics, self-healing properties, a small main lobe size, great directionality, high intensity, and a long depth of focus, have sparked researchers' interests in ex-

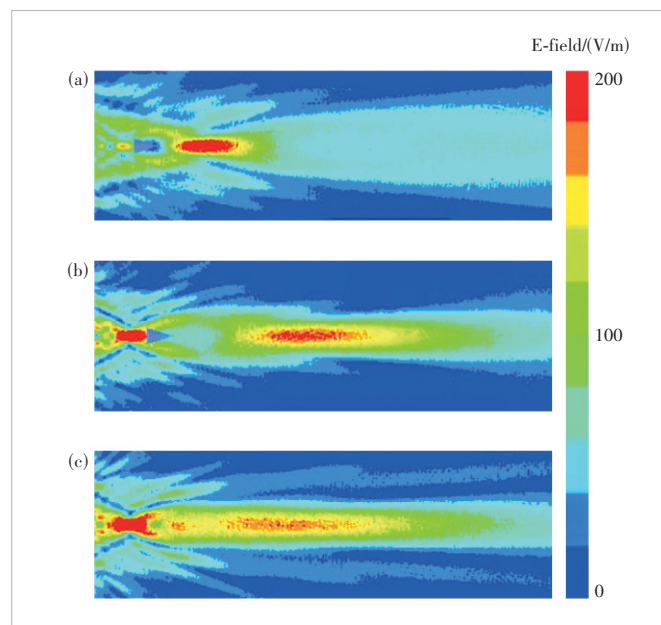
ploring their potential applications. In addition, the development of mobile communication technologies toward the near-field region has unlocked more possibilities for the practical application of this kind of short-range beams. In this section, we will categorize current applications of Bessel beams into three classes for discussion: wireless power transfer, sensing, and communication, and provide an introduction to the challenges that exist in current research in each category.

4.1 Wireless Power Transfer

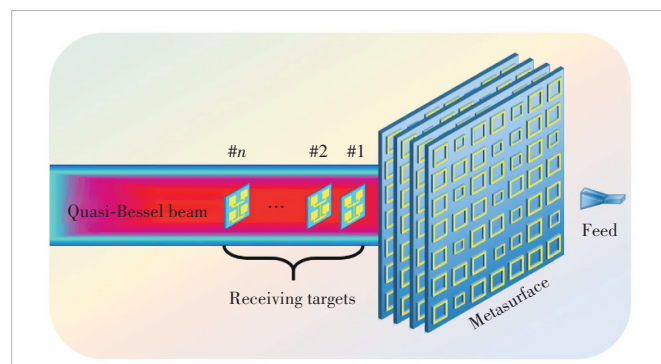
Wireless power transfer holds promising prospects in the near-field range, providing a convenient and efficient means of electricity delivery for applications such as sensor networks and mobile terminals. The efficiency is one of the most crucial properties in WPT systems. To enhance transfer efficiency, researchers are exploring ways to focus energy at the transmitting end. There are many types of beams designed for energy focus, such as focusing beams and supergain beams. Among these, Bessel beams offer a compromise, providing a solution that does not depend on high-precision positioning but still provides excellent focusing performance. Due to their non-diffracting characteristics and excellent directionality, Bessel beams can ensure that the emitted energy remains focused without dispersion within an effective range along a line. Additionally, Bessel beams have a long focal depth, which makes it possible to charge multiple devices simultaneously.

In Ref. [10], a transmissive metasurface unit was designed to compare the power transmission performance of Bessel beams with focused beams and high-gain directional beams, as shown in Fig. 7. The experimental results validated that Bessel beams could consistently maintain excellent focusing performance within their propagation range, making them an excellent choice for short-range wireless power transfer. Additionally, the possibility of simultaneously charging multiple devices using the long focal depth, large energy focusing range and self-healing characteristics of Bessel beams was preliminarily explored in this study. This work was further expanded in Ref. [22] in the following year. The researchers designed a metasurface with high transmission coefficient units by controlling and superimposing two Bessel beams. They achieved a Bessel beam with a half-power beam length (HPBL) exceeding 10%. This design enables a uniform energy distribution over a wide range along a single line, allowing simultaneous power transfer to multiple receiving devices. A WPT system with five receiving targets was designed (Fig. 8), and the scenarios under different apertures of receiving antennas were discussed. Experimental results demonstrated that the receiving targets achieved similar and high WPT efficiency, showcasing the excellent capability of Bessel beams in multiple-target energy transfer. In Ref. [35], the aperture field of the Bessel beam was modulated by introducing Gaussian functions. The generated Gaussian-Bessel beam was compared with the TEM mode of a coaxial cable and the Bessel beam for power transmission efficiency between two radiating apertures in the

near-field zone. The designed transmitter efficiently transferred radiofrequency power over a range of several wavelengths, achieving high efficiency (exceeding 50% at 30 wavelengths). The experimental results demonstrated that the Gaussian-Bessel beam was the optimal choice in this scenario due to its narrow spectrum and reduced diffraction of the radiated beam. In Ref. [36], the coupling coefficient between focused apertures supporting two types of cylindrical vector beams (Laguerre-Gaussian and Bessel-Gaussian) was derived in closed form. The results demonstrated that within the Fresnel zone, a finite aperture Bessel beam closely approaches the optimal field profile, exhibiting optimal wireless energy transfer performance. In Ref. [37], two coupled transverse magnetic (TM)-polarized, transverse electric (TE)-polarized, or hybrid-polarized Bessel beams launchers were evaluated and compared for WPT efficiency in the near-field radiation region. The study demonstrated the excellent performance of Bessel beams for WPT application in the micro-



▲ Figure 7. E-field intensity distributions of (a) high-gain directional beam, (b) focused beam and (c) Bessel beam^[10]



▲ Figure 8. Schematic of the multitarget wireless power transfer (WPT) system^[22]

wave and millimeter-wave frequency bands.

4.2 Sensing

Thanks to their extended axial length and invariant lateral resolution, along with their excellent focusing performance, Bessel beams have a wide range of applications in various fields, including medical imaging, terahertz imaging, and ground-penetrating radars (GPRs). In 1990, a study compared the main lobe characteristics of Bessel beams with those of Gaussian beams^[38]. The lateral resolution of Bessel beams was provided in closed form. The study designed a broadband non-diffracting Bessel beam transducer, which can enhance imaging frame rates, resulting in clearer and sharper medical ultrasound images due to its long focal depth (>200 mm). The potential of these non-diffracting beams in the field of medical imaging is tremendous. In Refs. [39] and [40], higher resolution and more information were achieved by utilizing Bessel beams instead of Gaussian beams in endoscopic imaging and cellular imaging scenarios, respectively. In Ref. [41], focused Bessel beams were employed in biomedical analysis, effectively achieving pathology-based measurements. To mitigate the impact of Bessel beam side lobes on imaging, a method called complementary beam subtraction (CBS) was proposed in Ref. [42], where a complementary beam is constructed for the Bessel beam. By scanning both the Bessel beam and its complementary beam and subtracting the two images, high axial resolution images can be restored. The imaging performance was experimentally validated using a spatial light modulator.

Moreover, Bessel beams can be used to overcome the trade-off between lateral resolution and depth of focus in traditional terahertz imaging, showing excellent potential for terahertz penetration imaging. A study conducted in Ref. [43] explored the possibility of manufacturing Bessel beam generation components for terahertz imaging applications through 3D printing technology. Additionally, regarding the penetration capability of Bessel beams, their propagation properties after passing through absorbing layers were studied in Ref. [44], which provides preliminary exploration for their potential applications in GPRs^[29]. In the near-field range, the acoustic version of Bessel beams was used to achieve microphone localization through beam scanning^[45]. However, to the best of our knowledge, research on localization using Bessel beams is currently quite limited.

4.3 Communication

Current research on Bessel beams in the field of communication mainly focuses on optics and terahertz domains. Free-space optical (FSO) communication, as a powerful complement to existing wireless communication technologies, contributes to achieving ultra-high-speed wireless communication in many important applications. In FSO systems, Bessel beams are used for interference resistance because of their self-healing characteristics. In 2011, leveraging the non-diffractive and self-healing properties of Bessel beams, GATTO et al. achieved and experi-

mentally tested an apparatus for an orbital angular momentum (OAM)-multiplexed FSO communication system^[46]. In 2016, researchers established a model for the transmission of Bessel-Gauss (BG) beams with OAM modes, deriving analytical expressions for the channel information capacity of free-space optical communication systems^[47]. This demonstrated that BG beams could effectively mitigate the impact of turbulence on OAM mode transmission. In Ref. [48], a practical anti-turbulence free-space optical communication scheme was proposed for visible light communication, based on an easily tunable zero-order Bessel beam. Compared with traditional Gaussian beams, this scheme enhances link stability and provides a power compensation of 2 dB to 3 dB, addressing the turbulence challenges encountered in visible light communication. Bessel beams have also been found to enhance physical layer security in FSO systems. The potential of using Bessel-Gauss beams to increase secure capacity and improve channel stability is explored in Refs. [49] and [50].

Terahertz communication is considered one of the most promising breakthrough technologies for the next generation of wireless standards. However, it suffers from extremely high propagation losses, which limits communication distances to the near-field range. Bessel beams offer a solution for near-field communication by resisting obstacles and providing a stable communication link. In Ref. [51], researchers investigated near-field phenomena in THz communication, explored link budget requirements, and proposed a promising solution using Bessel beams for near-field THz communication. Under experimental conditions for single-carrier near-field sub-THz communications, they utilized Bessel beams and a 16-QAM modulation scheme, achieving 40 Gbit/s with a bandwidth of 20 GHz. This study was further extended in Ref. [52], confirming that Bessel beams, in comparison to focused beams, can maintain a higher and more stable signal-to-noise ratio over a larger communication range.

4.4 Research Challenges

The research on the application of Bessel beams is still in the exploration stage, primarily due to challenges in their generation and limitations in propagation distance. Relatively, the application prospects of Bessel beams in wireless power transfer have been explored more extensively. However, for WPT applications, Bessel beams face an inevitable trade-off: increasing the transmission distance requires enlarging the generation aperture, which can lead to more sidelobe energy leakage. To enhance their transmission distance for practical applications, methods to reduce sidelobes must be considered. For instance, the use of broad-spectrum Bessel beams (known as X-waves) can enhance transmission efficiency. For the application of Bessel beams in sensing, current research is primarily focused on the field of medical imaging, with relatively fewer studies in other domains. Meanwhile, the application of Bessel beams in communication, to counter interference, improve communication stability and enhance security, is still in the early stage of exploration. Cur-

rently, the research is primarily focused on their interference resistance in FSO communication. In other high-frequency ranges, such as terahertz and millimeter waves, as well as under near-field conditions, the excellent directionality and self-healing characteristics of Bessel beams have the potential to achieve high-speed and stable transmission links with the ability to penetrate obstacles, making them a promising alternative technology for beamforming or beam focusing. However, research in these aspects is still relatively lacking. Finally, designing Bessel beam generation devices that meet practical application requirements is an inevitable challenge in their applied research.

5 Conclusions

This paper provides a comprehensive review of the Bessel beam concept and the research related to it. Starting from the scalar wave theory, we introduce the properties and fundamental theory of Bessel beams. The current focus of research on Bessel beams is their generation, with some exploration of their applications. The generation methods are categorized into two types: based on the scalar wave theory and based on the vector wave theory. We also review the current implementations of Bessel beam generation and explore their applications from the perspectives of wireless power transfer, sensing, and communication. In conclusion, Bessel beams require a more precise description and study of their properties through near-field theories, akin to the challenges faced in traditional communication and sensing in the near-field region. Simultaneously, the application of Bessel beams imposes higher requirements on the precise phase modulation capability and integration of emitting devices, introducing new challenges and opportunities for beamforming and antenna design. In the future, the enormous application potential of Bessel beams is expected to be further explored in higher communication frequency bands and shorter-distance scenarios.

References

- [1] AKYILDIZ I F, HAN C, NIE S. Combating the distance problem in the millimeter wave and terahertz frequency bands [J]. *IEEE communications magazine*, 2018, 56(6): 102 - 108. DOI: 10.1109/MCOM.2018.1700928
- [2] MIRAZ M H, ALI M, EXCELL P S, et al. A review on Internet of Things (IoT), Internet of Everything (IoE) and Internet of Nano Things (IoNT) [C]// *Proc. Internet Technologies and Applications (ITA)*. IEEE, 2015: 219 - 224. DOI: 10.1109/ITeA.2015.7317398
- [3] ZHANG H Y, SHLEZINGER N, GUIDI F, et al. Beam focusing for near-field multiuser MIMO communications [J]. *IEEE transactions on wireless communications*, 2022, 21(9): 7476 - 7490. DOI: 10.1109/TWC.2022.3158894
- [4] FRIEDLANDER B. Localization of signals in the near-field of an antenna array [J]. *IEEE transactions on signal processing*, 2019, 67(15): 3885 - 3893. DOI: 10.1109/TSP.2019.2923164
- [5] DURNIN J. Exact solutions for nondiffracting beams I the scalar theory [J]. *Journal of the optical society of America A*, 1987, 4(4): 651 - 654. DOI: 10.1364/josaa.4.000651
- [6] ZHONG Y C, CHENG Y J. Ka-band wideband large depth-of-field beam generation through a phase shifting surface antenna [J]. *IEEE transactions on antennas and propagation*, 2016, 64(12): 5038 - 5045. DOI: 10.1109/TAP.2016.2618849
- [7] ZHONG Y C, CHENG Y J. Generating and steering quasi-nondiffractive beam by near-field planar risley prisms [J]. *IEEE transactions on antennas and propagation*, 2020, 68(12): 7767 - 7776. DOI: 10.1109/TAP.2020.2998872
- [8] LI Y B, CAI B G, WAN X, et al. Diffraction-free surface waves by metasurfaces [J]. *Optics letters*, 2014, 39(20): 5888 - 5891. DOI: 10.1364/OL.39.005888
- [9] ZENG T, DENG L, LI S F. Generation of non-diffractive beam by using metasurface for wireless power transfer (WPT) applications [C]// *Proc. IEEE Asia-Pacific Microwave Conference (APMC)*. IEEE, 2019: 4 - 6. DOI: 10.1109/APMC46564.2019.9038329
- [10] XUE H, WU X N, CUI X W, et al. Wireless power transfer based on Bessel beam [C]// *Proc. IEEE International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM)*. IEEE, 2021: 1 - 3. DOI: 10.1109/iWEM53379.2021.9790507
- [11] PFEIFFER C, GRBIC A. Controlling vector Bessel beams with metasurfaces [J]. *Physical review applied*, 2014, 2(4): 044012. DOI: 10.1103/PhysRevApplied.2.044012
- [12] QI M Q, TANG W X, CUI T J. A broadband Bessel beam launcher using metamaterial lens [J]. *Scientific reports*, 2015, 5: 11732. DOI: 10.1038/srep11732
- [13] LEMAITRE-AUGER P, ABIELMONA S, CALOZ C. Generation of Bessel beams by two-dimensional antenna arrays using sub-sampled distributions [J]. *IEEE transactions on antennas and propagation*, 2013, 61(4): 1838 - 1849. DOI: 10.1109/TAP.2012.2232263
- [14] MAZZINGHI A, BALMA M, DEVONA D, et al. Large depth of field pseudo-bessel beam generation with a RLSA antenna [J]. *IEEE transactions on antennas and propagation*, 2014, 62(8): 3911 - 3919. DOI: 10.1109/TAP.2014.2324557
- [15] ETTORRE M, PAVONE S C, CASALETTI M, et al. Experimental validation of Bessel beam generation using an inward Hankel aperture distribution [J]. *IEEE transactions on antennas and propagation*, 2015, 63(6): 2539 - 2544. DOI: 10.1109/TAP.2015.2419261
- [16] LU P, VOYER D, BRÉARD A, et al. Design of TE-polarized Bessel antenna in microwave range using leaky-wave modes [J]. *IEEE transactions on antennas and propagation*, 2018, 66(1): 32 - 41. DOI: 10.1109/TAP.2017.2768584
- [17] FAHRBACH F O, ROHRBACH A. Propagation stability of self-reconstructing Bessel beams enables contrast-enhanced imaging in thick media [J]. *Nature communications*, 2012, 3: 632. DOI: 10.1038/ncomms1646
- [18] ANDRIOTTY M S. *Nondiffracting Waves in 2D and 3D* [D]. Williamsburg: College of William and Mary, 2017
- [19] LU P, WAGIH M, GOUSSETIS G, et al. A comprehensive survey on transmitting antenna systems with synthesized beams for microwave wireless power transmission [J]. *IEEE journal of microwaves*, 2023, 3(4): 1081 - 1101. DOI: 10.1109/JMW.2023.3285825
- [20] ZHONG Y C, CHENG Y J. Wideband quasi-nondiffraction beam with accurately controllable propagating angle and depth-of-field [J]. *IEEE transactions on antennas and propagation*, 2017, 65(10): 5035 - 5042. DOI: 10.1109/TAP.2017.2736536
- [21] TANG H, DING S, ZHU Z J. Generation of dual asymmetric non-diffracting beams based on transmission metasurface in microwave-frequency region [C]// *Proc. International Conference on Microwave and Millimeter Wave Technology (ICMMT)*. IEEE, 2022: 1 - 3. DOI: 10.1109/ICMMT55580.2022.10022831
- [22] XUE H, WU X N, CUI X W, et al. Multitarget wireless power transfer system using metasurface for quasi-bessel beams with large half power beam length [J]. *IEEE transactions on microwave theory and techniques*, 2022, 70(10): 4449 - 4462. DOI: 10.1109/TMTT.2022.3197619
- [23] COX A J, D'ANNA J. Constant-axial-intensity nondiffracting beam [J]. *Optics letters*, 1992, 17(4): 232 - 234. DOI: 10.1364/ol.17.000232

- [24] HERMAN R M, WIGGINS T A. Apodization of diffractionless beams [J]. *Applied optics*, 1992, 31(28): 5913 – 5915. DOI: 10.1364/AO.31.005913
- [25] ETTORRE M, GRBIC A. Generation of propagating Bessel beams using leaky-wave modes[J]. *IEEE transactions on antennas and propagation*, 2012, 60(8): 3605 – 3613. DOI: 10.1109/TAP.2012.2201088
- [26] SALEM M A, KAMEL A H, NIVER E. Microwave Bessel beams generation using guided modes [J]. *IEEE transactions on antennas and propagation*, 2011, 59(6): 2241 – 2247. DOI: 10.1109/TAP.2011.2143683
- [27] ETTORRE M, RUDOLPH S M, GRBIC A. Generation of propagating Bessel beams using leaky-wave modes: experimental validation [J]. *IEEE transactions on antennas and propagation*, 2012, 60(6): 2645 – 2653. DOI: 10.1109/TAP.2012.2194674
- [28] PAVONE S C, ETTORRE M, ALBANI M. Analysis and design of Bessel beam launchers: longitudinal polarization [J]. *IEEE transactions on antennas and propagation*, 2016, 64(6): 2311 – 2318. DOI: 10.1109/TAP.2016.2550049
- [29] PAVONE S C, ETTORRE M, CASALETTI M, et al. Analysis and design of Bessel beam launchers: transverse polarization [J]. *IEEE transactions on antennas and propagation*, 2021, 69(8): 5175 – 5180. DOI: 10.1109/TAP.2021.3060049
- [30] CHILLARA V K, DAVIS E S, PANTEA C, et al. Ultrasonic Bessel beam generation from radial modes of piezoelectric discs [J]. *Ultrasonics*, 2019, 96: 140 – 148. DOI: 10.1016/j.ultras.2019.02.002
- [31] VAQUERO Á F, PINO M R, ARREBOLA M, et al. Bessel beam generation using dielectric planar lenses at millimeter frequencies [J]. *IEEE access*, 2020, 8: 216185 – 216196. DOI: 10.1109/ACCESS.2020.3040898
- [32] FENG P Y, QU S W. Generation of Bessel beams at millimeter-wave band using 3-D printed axicon lenses [C]//*Proc. International Symposium on Antennas and Propagation (ISAP)*. IEEE, 2018: 1 – 2
- [33] YANG C F, WU G B, CHEN B J, et al. Terahertz Bessel beam scanning enabled by dispersion-engineered metasurface [J]. *IEEE transactions on microwave theory and techniques*, 2023, 71(8): 3303 – 3311. DOI: 10.1109/TMTT.2023.3234610
- [34] CHU X X, SUN Q, WANG J, et al. Generating a Bessel-Gaussian beam for the application in optical engineering [J]. *Scientific reports*, 2015, 5: 18665. DOI: 10.1038/srep18665
- [35] PAKOVIĆ S, ZHOU S Y, GONZÁLEZ-OVEJERO D, et al. Bessel-gauss beam launchers for wireless power transfer [J]. *IEEE open journal of antennas and propagation*, 2021, 2: 654 – 663. DOI: 10.1109/OJAP.2021.3078234
- [36] ALSOLAMY F, ALOMAR W A, GRBIC A. Cylindrical vector beams for wireless power transfer [J]. *IEEE transactions on antennas and propagation*, 2021, 69(3): 1716 – 1727. DOI: 10.1109/TAP.2020.3026428
- [37] NEGRI E, FUSCALDO W, BURGHIGNOLI P, et al. Leaky-wave analysis of TM-, TE-, and hybrid-polarized aperture-fed bessel-beam launchers for wireless-power-transfer links [J]. *IEEE transactions on antennas and propagation*, 2023, 71(2): 1424 – 1436. DOI: 10.1109/TAP.2022.3231086
- [38] LU J Y, GREENLEAF J F. Ultrasonic nondiffracting transducer for medical imaging [J]. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 1990, 37(5): 438 – 447. DOI: 10.1109/58.105250
- [39] LEE K S, ROLLAND J P. Bessel beam spectral-domain high-resolution optical coherence tomography with micro-optic axicon providing extended focusing range [J]. *Optics letters*, 2008, 33(15): 1696. DOI: 10.1364/OL.33.001696
- [40] PLANCHON T A, GAO L, MILKIE D E, et al. Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination [J]. *Nature methods*, 2011, 8(5): 417 – 423. DOI: 10.1038/nmeth.1586
- [41] VUILLEMIN N, MAHOU P, DÉBARRE D, et al. Efficient second-harmonic imaging of collagen in histological slides using Bessel beam excitation [J]. *Scientific reports*, 2016, 6(1): 29863. DOI: 10.1038/srep29863
- [42] JIA H, YU X H, YANG Y L, et al. Axial resolution enhancement of light-sheet microscopy by double scanning of Bessel beam and its complementary beam [J]. *Journal of biophotonics*, 2019, 12(1): e201800094. DOI: 10.1002/jbio.201800094
- [43] WEI X L, LIU C M, NIU L T, et al. Generation of arbitrary order Bessel beams via 3D printed axicons at the terahertz frequency range [J]. *Applied optics*, 2015, 54(36): 10641 – 10649. DOI: 10.1364/AO.54.010641
- [44] MUGNAI D, SPALLA P. Electromagnetic propagation of Bessel-like localized waves in the presence of absorbing media [J]. *Optics communications*, 2009, 282(24): 4668 – 4671. DOI: 10.1016/j.optcom.2009.09.020
- [45] HASEGAWA K. Indoor self localization of a single microphone based on asynchronous scanning of modulated Bessel beams [C]//*Proc. 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. IEEE, 2019: 142 – 147
- [46] GATTO A, TACCA M, MARTELLI P, et al. Free-space orbital angular momentum division multiplexing with Bessel beams [J]. *Journal of optics*, 2011, 13(6): 064018. DOI: 10.1088/2040-8978/13/6/064018
- [47] CHENG M J, GUO L X, LI J T, et al. Channel capacity of the OAM-based free-space optical communication links with Bessel-Gauss beams in turbulent ocean [J]. *IEEE photonics journal*, 2016, 8(1): 7901411. DOI: 10.1109/JPHOT.2016.2518865
- [48] LU Z H, GUO Z F, FAN M S, et al. Tunable Bessel beam shaping for robust atmospheric optical communication [J]. *Journal of lightwave technology*, 2022, 40(15): 5097 – 5106. DOI: 10.1109/JLT.2022.3172134
- [49] WANG T L, DJORDJEVIC I B. Physical-layer security of a binary data sequence transmitted with bessel-gaussian beams over an optical wiretap channel [J]. *IEEE photonics journal*, 2018, 10(6): 7908611. DOI: 10.1109/JPHOT.2018.2878912
- [50] WANG T L, GARIANO J A, DJORDJEVIC I B. Employing Bessel-Gaussian beams to improve physical-layer security in free-space optical communications [J]. *IEEE photonics journal*, 2018, 10(5): 7907113. DOI: 10.1109/JPHOT.2018.2867173
- [51] SINGH A, REDDY I V A K, BODET D, et al. Bessel beams for 6G—a performance analysis [C]//*Proc. 56th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2022: 658 – 664. DOI: 10.1109/IEEECONF56349.2022.10052090
- [52] SINGH A, PETROV V, JORNET J M. Utilization of Bessel beams in wide-band sub terahertz communication systems to mitigate beamsplit effects in the near-field [C]//*Proc. 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023: 1 – 5. DOI: 10.1109/ICASSP49357.2023.10096275

Biographies

CAO Xinghan received his BS degree from the School of Electronics and Information Technology, Sun Yat-sen University, China in 2022. He is currently working toward an MS degree at University of Science and Technology of China. His research interests include near-field sensing and near-field communications.

YIN Huarui received his PhD degree from the School of Information Science and Technology, University of Science and Technology of China in 2006. He is currently an associate professor with University of Science and Technology of China. His current research interests include short-range wireless communications, throughput of quasi symmetric multicast wireless networks, and multiple access of mass connections for large scale antennas wireless network.

YOU Changsheng (yous@sustech.edu.cn) received his PhD degree in electrical and electronic engineering from the University of Hong Kong, China in 2018. He is currently an assistant professor with Southern University of Science and Technology, China. His current research interests include edge computing and edge intelligence, intelligent reflecting surface communication, massive MIMO communication, and unmanned aerial vehicle communication technology.

Recent Advances in Video Coding for Machines Standard and Technologies



ZHANG Qiang^{1,2}, MEI Junjun^{1,2}, GUAN Tao^{1,2},
SUN Zhewen³, ZHANG Zixiang³, YU Li³

(1. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China;
2. ZTE Corporation, Shenzhen 518057, China;
3. Huazhong University of Science and Technology, Wuhan 430074, China)

DOI: 10.12142/ZTECOM.202401008

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240203.1937.002.html>,
published online February 5, 2024

Manuscript received: 2023-07-02

Abstract: To improve the performance of video compression for machine vision analysis tasks, a video coding for machines (VCM) standard working group was established to promote standardization procedures. In this paper, recent advances in video coding for machine standards are presented and comprehensive introductions to the use cases, requirements, evaluation frameworks and corresponding metrics of the VCM standard are given. Then the existing methods are presented, introducing the existing proposals by category and the research progress of the latest VCM conference. Finally, we give conclusions.

Keywords: video coding for machines; VCM; video compression

Citation (Format 1): ZHANG Q, MEI J J, GUAN T, et al. Recent advances in video coding for machines standard and technologies [J]. *ZTE Communications*, 2024, 22(1): 62 - 76. DOI: 10.12142/ZTECOM.202401008

Citation (Format 2): Q. Zhang, J. J. Mei, T. Guan, et al., "Recent advances in video coding for machines standard and technologies," *ZTE Communications*, vol. 22, no. 1, pp. 62 - 76, Mar. 2024. doi: 10.12142/ZTECOM.202401008.

1 Introduction

Vision is an important way for human beings to perceive the external environment. As a digital form of visual information, video has become an indispensable part of everyone's daily life. With the explosive growth of video data, more and more video has been used for machine vision analysis in industrial or intelligent scenes, while only a small portion is watched by humans. With the continuous development of multimedia technology, machines process video more automatically and intelligently, and even interact with each other directly without human participation. Therefore, this type of communication is often referred to as machine-to-machine (M2M) communication.

Traditional video coding frameworks, such as H.264/Advanced Video Coding (AVC), H.265/high efficiency video coding (HEVC), and H.266/Versatile Video Coding (VVC), were developed for the Human Visual System (HVS), achieving higher compression efficiency while maintaining good subjective quality. The essence of achieving efficient compression is to remove as much redundancy as possible. For example, human eyes have a high perception of brightness and a low per-

ception of chroma, so the redundant chrominance information will be discarded appropriately during encoding. Human eyes are less sensitive to high-frequency information than to low-frequency information, hence the high-frequency information can be compressed by transformation and quantization. In addition, pixel values change little in adjacent space and time, therefore spatial and temporal redundancy can be removed through intra-frame and inter-frame prediction, respectively. However, while removing redundancy is effective for compressing visual information in human vision, it does not necessarily apply to machine vision. Traditional codecs are designed to compress images and videos in a way that exploits the physiological properties of human eyes, making the compression less perceptible to human viewers. However, in machine vision task scenes, the compressed information will be provided to machines for consumption instead of humans. Machines do not possess the same visual characteristics as humans, and therefore, there is a need to develop a new coding standard specifically tailored for machine vision tasks. The development of machine learning/deep learning and the widespread use of image and video sensors have led to massive data demands in various intelligent platforms, including the Internet of Things (IoT), Internet of Vehicles (IoV), video surveillance, and smart cities. In this era of AI-driven video intelligence, many researchers are exploring ways to achieve

This work was partially supported by ZTE Industry-University-Institute Cooperation Funds.

higher compression efficiency while maintaining good performance on machine vision tasks, considering the characteristics of the machine system. As a result, there is an urgent need to develop new standards for video compression in machine vision tasks.

Machine vision tasks are different from human vision tasks, which have different purposes and evaluation metrics. For traditional human vision tasks, the goal is to rebuild input images or videos to their original subjective quality, with evaluation metrics used to measure the degree of pixel domain distortion. In the field of machine vision tasks, the VCM standard identifies and summarizes common machine tasks. It specifically outlines five representative tasks, namely object detection, instance segmentation, object tracking, human pose estimation and action recognition. These tasks are crucial in machine vision applications and play a significant role in advancing the capabilities of video coding for machines. In order to compare the effects of different methods, new evaluation metrics are introduced to measure the performance of machine vision tasks instead of pixel domain distortion for human viewing. The pursuit of high-efficiency compression in machine vision, while simultaneously mitigating redundancy, has garnered significant attention from scholars. However, the challenge lies in reducing the bitrate while preserving the accuracy of machine vision tasks. Therefore, the international standard organization named Moving Picture Experts Group (MPEG) established the VCM group, focusing on achieving efficient compression for machine vision. The work of the VCM group can be summarized as follows: 1) introducing use cases and related requirements for machine vision and hybrid human-machine vision compression tasks; 2) collecting data sets with evaluation metrics and ground truth; 3) collecting technical evidence of feature compression and human-machine hybrid compression; 4) developing an evaluation framework to compare different technology options; 5) developing the standards for VCM.

The VCM standard has gained significant attention from scholars and experts in both academia and industry, resulting in notable progress. In this paper, we will give a comprehensive summary of the usage scenarios, representative technical proposals of VCM, and the latest developments of VCM standards.

This paper is organized as follows. We first provide an introduction of the VCM standard, including but not limited to use cases, requirements, and processing pipelines. And then we introduce two different technical routes in the standard respectively, and divide related technical proposals into several categories for introduction. In addition, the experimental performance and standardization process corresponding to these two routes are also introduced. Finally, we conclude with a discussion.

2 Related Works

This section provides a brief overview of the development of VCM standards. To improve the performance and efficiency of VCM, some early research has been devoted to extracting key information that can characterize machine vision tasks from images, that is, visual features. These visual features are stored in a highly compact representation, which can effectively address various visual analysis tasks with significantly smaller data volume compared with compressed video. For this method of using compact representation to assist machine vision tasks, MPEG standardized it as compact descriptors for visual search (CDVS)^[1] and for video analysis (CDVA)^[2] in 2015 and 2019, respectively. In CDVS, the methods based on interest point detection and local feature description are used to calculate the compact expression of image features. The normalization process has made significant progress in reducing the number of image feature data and computational complexity. On the basis of CDVS, CDVA utilizes deep learning technology to shorten the descriptor length while simultaneously enhancing its ability to describe videos.

Recently, with the development of deep neural networks, learning-based methods have been widely used in image and feature compression. LE et al.^[3] proposed an image codec through end-to-end training and balanced trade-offs among the image distortion, analysis accuracy and bitrate. LE et al.^[4] further proposed an inference-time content-adaptive strategy to optimize the latent representation of the end-to-end image codec. This method improves the compression efficiency of machine-consumption videos. TU et al.^[5] used an end-to-end method to compress image signals and semantic features of different granularity, and a cross-layer context model was proposed to reduce the information redundancy among them. CHOI et al.^[6] studied the deep feature compression performance between the mobile device and the cloud, and proposed a method that improved the accuracy of tasks under lossy feature compression. They also investigated latent-space scalability for multi-task collaborative intelligence in Ref. [7].

In terms of VCM standards, for example, DUAN et al.^[8] introduced and analyzed VCM in combination with the new fields generated in the emerging MPEG standardization work, giving the definition, formulation, and paradigm of VCM. This literature focuses more on the theoretical framework of video coding and provides a paradigm of collaborative video and feature compression, which has been submitted to the VCM group as a piece of evidence. GAO et al.^[9] summarized the work of the MPEG VCM group, and elaborated on the specific content and requirements of VCM standards. Additionally, with the rapid development of standards, many new methods have been proposed during the standardization process.

In addition, China's video coding standard organizations including the Audio Video coding Standard (AVS) and China Electronics Standardization Institute (CESI) have launched research on vision coding for machines. The data coding for ma-

chine (DCM) intelligence frameworks was jointly proposed by CESI and related enterprises in September 2020. The DCM group primarily elucidates the fundamental content of machine vision coding, organizes the technical requirements and application scenario of machine vision coding, outlines the future development direction of machine vision coding, and proposes potential technical routes for machine vision coding.

In this paper, we will mainly introduce and analyze the proposed proposals during the latest VCM standard development. We will also classify the existing proposals and provide detailed introductions to representative proposals.

3 Background of VCM Standard

The VCM standard has been attracting extensive interest since it was proposed, and more than 100 experts from about 30 research institutes around the world have devoted themselves to the MPEG VCM standard. In July 2019, the VCM working group was established. In July 2020, the formulation of requirements was completed and the first version of the VCM standard evaluation framework was determined. The first version of evidence collection and general test conditions was released in January 2021. In October 2021, according to the collected evidences, the VCM standard working group decided to divide the current coding scheme for machine vision into two technical routes, namely Track-1 and Track-2, where Track-1 stands for feature coding and Track-2 represents compression and then analysis (CTA). The Track-2 Call for Proposals (CfP) and the Track-1 Call for Evidences (CfE) were released in April 2022 and July 2022, respectively. The MPEG 140th meeting that just ended has already collected a series of response proposals for Track-1 CfE and Track-2 CfP.

To facilitate the development of a new standard, it is essential to clearly define the application scenario and establish a specific evaluation process. Such an approach allows for the comparison and evaluation of various solutions within the same evaluation system, which guides the direction of standard development. These details are described as follows.

3.1 Use Cases

As communication transmission technology advances, particularly with the widespread adoption of 5G, various industries are undergoing digital and intelligent transformations. VCM has corresponding usage scenarios in many places in daily life, including smart cities, intelligent transportation, intelligent surveillance, smart IoT, smart IoV, automatic driving, information security, smart factories, intelligent content, and so on, which have huge market and development potential. The current proposals on VCM have been set for several specific use cases, which can be roughly divided into the following categories.

3.1.1 Intelligent Transportation

A typical application of machine vision tasks is intelligent

transportation. As vehicle traffic increases, there is a growing demand for intelligent transportation and vehicle intelligence to address traffic congestion and reduce accidents. The cameras and sensors deployed on cars can generate a large number of data, which can be used for intelligent transportation tasks. In intelligent transportation systems, cars may need to communicate with each other to perform different tasks. Sensors in urban infrastructure can transfer visual information to different vehicles, and then the information is used to complete machine vision tasks. The interconnection of vehicles is a key aspect of intelligent transportation systems. The IoV is expected to play a crucial role in future urban transportation systems, as it has the potential to improve safety, productivity, accessibility, and road efficiency. Vehicles equipped with multiple cameras can capture their surroundings and use deep neural networks (DNN) to detect and recognize objects such as cars, pedestrians, and street furniture, as well as events such as traffic jams or accidents. In addition, the use of infrared and non-visible light imaging technologies like radar can provide enhanced vision capabilities in low-light and adverse weather conditions, such as fog or rain. Intelligent transportation has vast potential, while the significant requirement for M2M communication and video compression among IoV is an urgent challenge that needs to be addressed.

3.1.2 Intelligent Surveillance

Another common application scenario is the surveillance system. The industry has undergone decades of development and is now relatively mature with a large market size. Surveillance technology has been widely used in smart cities, traffic control, medical services, personal usage, etc. Currently, surveillance systems can use neural networks to deal with different tasks, such as object detection and tracking. However, considering the number of sensors and the resolution, frame rate, and length of video to be transmitted, current surveillance systems usually occupy a large amount of transmission bandwidth. Moreover, the increase in the number of front-end cameras adds a huge computational load to the back-end servers that perform machine vision tasks. Surveillance scenarios have many applications in reality. For example, behavior detection can detect potential hazards and object detection can analyze the crowd. In addition to the intelligent tasks mentioned above, tasks such as fire detection and hazardous chemical leak detection can be performed. Some of these tasks can share the same upstream task module or pre-processing process, enabling the reuse of resources. In this case, the descriptor is extracted on the front-end device and sent to the back-end server along with the video/feature encoded by the VCM codec to accomplish the machine vision tasks of the back-end.

3.1.3 Intelligent Content

With the popularity of short video Apps, people are used to

uploading videos to various social platforms, which makes protecting certain groups of people, teenagers for example, from inappropriate content a big problem. At the same time, traditional manual review is time-consuming and labor-intensive, making it difficult to process massive data. Machine vision technology enables the intelligent review, rating, processing and distribution of images, videos and social media in real-time. In this case, it is necessary to propose intelligent content. Intelligent content can detect objects in videos through pre-trained neural networks, which automatically identify, classify, and grade them. For example, when a video containing harmful information or dangerous scenes is played by the account of a teenager or a child, this technology can block the dangerous information, and only the identified adult audience is allowed to view these videos. Additionally, closed-circuit television (CCTV) is widely used nowadays to prevent crime or provide useful information, such as area or traffic information. However, invasion of personal privacy has always been a big problem in the security field. There is a need to prevent misuse of personal information in public and drone imagery. One solution is to automatically identify the private areas in the video, such as faces and license plates, and store their location information. Only when the viewer has the corresponding permission can the original image be displayed, otherwise the image in the private area will not be displayed, thus protecting the privacy of the parties involved.

3.2 Requirements of VCM Standard

For the above application scenarios, the VCM standard should meet the following key requirements:

1) Efficient compression performance. Compared with VVC, VCM improves the Bjontegaard Delta (BD)^[10] gain by at least 30% in machine vision tasks, and supports various bit rates and delays. In addition, the complexity of the VCM should also be kept within a reasonable range.

2) Multi-task capability. The VCM can support video encoding for various types of visual analysis tasks, and the bit stream can be decompressed to perform various machine vision tasks. In the VCM standard, different vision tasks can use the same module, for example, backbone and feature compression modules can be shared, while different back-end networks are used to implement different machine vision tasks.

3) Various forms of input. Different types of sensor modules and processing performances vary in reality, and VCM needs to adapt to image/video signals generated by various sensors. The input signal includes monochrome, RGB, and YUV in the color format, and the input signal also includes non-visible light images such as infrared images and laser radars. VCM codec should adapt to various types of input.

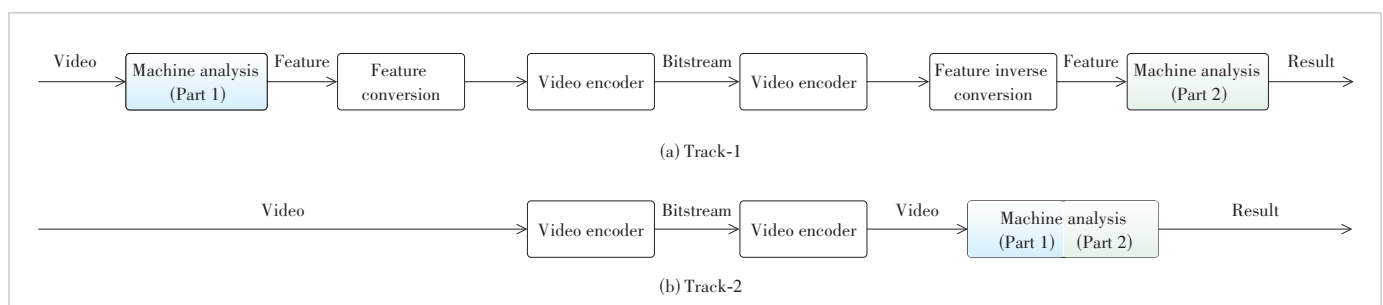
3.3 Processing Pipeline

The current mainstream VCM solutions can be roughly divided into Track-1 and Track-2 according to the compressed content. The content compressed in Track-1 is the feature map. The input image or video will be sent into a backbone network to get the P-layer feature map, and then these feature maps will be compressed, transmitted, and reconstructed. Finally, the reconstructed feature maps are fed into different back-end machine task networks (such as object detection and instance segmentation network models). The method of VCM Track-1 only compresses and reconstructs the feature map instead of the original image or video, which means it does not take the visual requirements of human eyes into consideration. The flowchart of Track-1 is shown in Fig. 1(a), where the blue module extracts feature maps from the input video, and the green module implements machine vision tasks through feature maps.

The content compressed in Track-2 is the original images or videos, and the processing flow of this method is consistent with that of the common method of compressing first and then analyzing the video. The reconstructed output will be sent to the machine vision module to perform corresponding tasks. The method of VCM Track-2 also aims to accomplish machine vision tasks, while additionally achieving reconstruction of the original input in the pixel domain. This approach provides the advantage of addressing the requirements of both human viewing and machine vision analysis tasks simultaneously. It should be noted that the support for human eye consumption is not required by the VCM standard but an additional extension. The flowchart of Track-2 is shown in Fig. 1(b).

4 Evaluation Framework

To compare the pros and cons of different schemes, an evaluation framework is needed to measure the performance of differ-



▲ Figure 1. Flowcharts of Track-1 and Track-2

ent technologies, including machine vision tasks, datasets, evaluation metrics, evaluation pipeline, and anchor generation.

4.1 Machine Vision Tasks

There are many possible machine vision tasks in practice, and the following five typical machine vision tasks are proposed during the VCM standardization process for real-world needs: 1) object detection, 2) instance segmentation, 3) object tracking, 4) action recognition and 5) pose estimation. The object of action recognition and pose estimation is humans, and human joint detection is required before performing the machine vision task.

The relevant information about machine vision tasks is shown in Table 1. To facilitate the comparison of the performance of different methods, a unified evaluation framework is proposed for different tasks, including network architecture and data sets. Since the difficulty of supporting multi-task optimization and the related data sets are not easy to obtain, most of the methods are aimed at object detection, instance segmentation, and object tracking. Therefore, the MPEG group narrows the task scope from five categories to three.

4.2 Datasets

There are a variety of public datasets for object detection and instance segmentation. However, due to issues like copyright licensing, most datasets are only allowed for non-commercial use. Among the remaining datasets, considering the use of license terms and other restrictions, the following five datasets for VCM are used.

1) OpenImageV6: OpenImageV6 has more than 20 000 images with annotation boxes and labels for object detection and instance segmentation. VCM selects 5 000 pictures to form a new dataset OpenImageV6-5K.

2) FLIR: FLIR provides annotated thermal imaging datasets and corresponding unannotated RGB images. The data are acquired by RGB cameras and thermal imaging cameras installed on the car. The objects in the dataset are mostly pedes-

trians and various vehicles, which is suitable for machine vision tasks in driving scenarios. Related experiments show that infrared images have better detection performance than RGB images in low light.

3) HiEve: The HiEve dataset is mainly aimed at human behavior, including a large number of gestures (>1 million), complex event action labels (>56 000), and long-term continuous trajectories (average trajectory length>480). This dataset is mainly used for object tracking, action recognition and pose estimation.

4) TVD: The TVD dataset contains 86 video sequences covering different contents. Each sequence has a resolution of 3 840×2 160 and has 65 frames. For the object detection task, TVD provides 166 images of 1 920×1 080 in RGB24 format, and the bounding box is calibrated. For the instance segmentation task, segmentation mask annotations are also provided for these 166 images. For the target tracking task, 3 videos and the corresponding labels are provided.

5) SFU-HW-Object-v1: The SFU-HW-Object-v1 dataset provides object detection annotations for 18 types of objects in video sequences used in the development of the High Efficiency Video Coding (HEVC) standard.

4.3 Evaluation Metrics

For different machine vision tasks, different evaluation metrics are used to measure the performance of various technical solutions.

The cost of storing and transmitting the bit stream is an important factor. For image datasets, bits per pixel (BPP) shall be used, which refers to the effective number of bits occupied by each pixel (ignoring the channel), and the calculation is

$$\text{BPP} = \frac{\text{total bits}}{\text{total pixels}}, \quad (1)$$

where total pixels refer to the total number of pixels of overall images at their original resolution and total bits refer to the number of bits occupied by the compressed code stream. For video sequences, the bitrate shall be measured in kilo bits per second (kbit/s), which is defined as:

$$\text{bitrate} = \frac{\text{total bits} \times \text{fps}}{1\,000 \times \text{frames}}, \quad (2)$$

where fps denotes the number of frames per second and frames denote the number of encoded frames.

Mean average precision (mAP) is used to measure the performance of the object detection and instance segmentation tasks. When an object is categorized, there are four cases of the prediction results, including true positive $TP(T_{IoU})$, false positive $FP(T_{IoU})$, false negative $FN(T_{IoU})$, and true negative $TN(T_{IoU})$, which are defined with an Intersection over Union (IoU) threshold T_{IoU} for that category. Among them, true/false

▼Table 1. Information about machine vision tasks

Tasks	Network Architecture	Training Dataset
Object detection	Faster R-CNN with ResNeXt-101 backbone ^[11]	COCO train2017 ^[12]
		OpenImageV6 ^[13]
		FLIR ^[14]
		TVD ^[15]
		SFU-HW-Object-v1 ^[16]
Instance segmentation	Faster R-CNN with ResNeXt-101 backbone	OpenImageV6 TVD
Object tracking	JDE-1088x608 ^[17]	HiEve ^[18] TVD
Action recognition*	SlowFast ^[19]	HiEve
Pose estimation*	HRNet ^[20]	COCO train2017 MPII Human Pose ^[21] HiEve

*Considering the difficulty of implementation, such as the limited dataset and usage scenarios, these tasks are not considered mandatory.

R-CNN: region with convolutional neural networks

represents the output of the neural network, and positive/negative represents the label in the ground truth.

The recall and precision of the given IoU threshold are defined as follows:

$$\text{recall}(T_{\text{IoU}}) = \frac{\text{TP}(T_{\text{IoU}})}{\text{TP}(T_{\text{IoU}}) + \text{FN}(T_{\text{IoU}})}, \quad (3)$$

$$\text{precision}(T_{\text{IoU}}) = \frac{\text{TP}(T_{\text{IoU}})}{\text{TP}(T_{\text{IoU}}) + \text{FP}(T_{\text{IoU}})}. \quad (4)$$

The neural network of detection or segmentation can pair the recall and precision values with specific IoU thresholds and different confidence levels. For each recall value r in the pairs, we define $p(r)$ as the maximum precision value in all precision values where the corresponding recall values are above the given recall value r :

$$p(r) = \max_{\tilde{r} \geq r} \text{precision}(\tilde{r}). \quad (5)$$

The average precision (AP) of a given category of an object is defined as the average value of $p(r)$ for all recall values provided by the neural network, which can characterize the area of the entire precision-recall (PR) curve. The mAP means the average AP of all categories. By default, mAP is equivalent to mAP@0.5, where 0.5 represents the threshold of IoU. In addition, mAP@[0.5: 0.95] represents the average mAP on different IoU thresholds (from 0.5 to 0.95, with a step size of 0.05).

For object tracking, multiple object tracking accuracy (MOTA) is used to measure the capability of the algorithm to continuously track the target, which can intuitively measure the performance of the tracker and is not affected by the accuracy of the detector. The calculation formula of MOTA is:

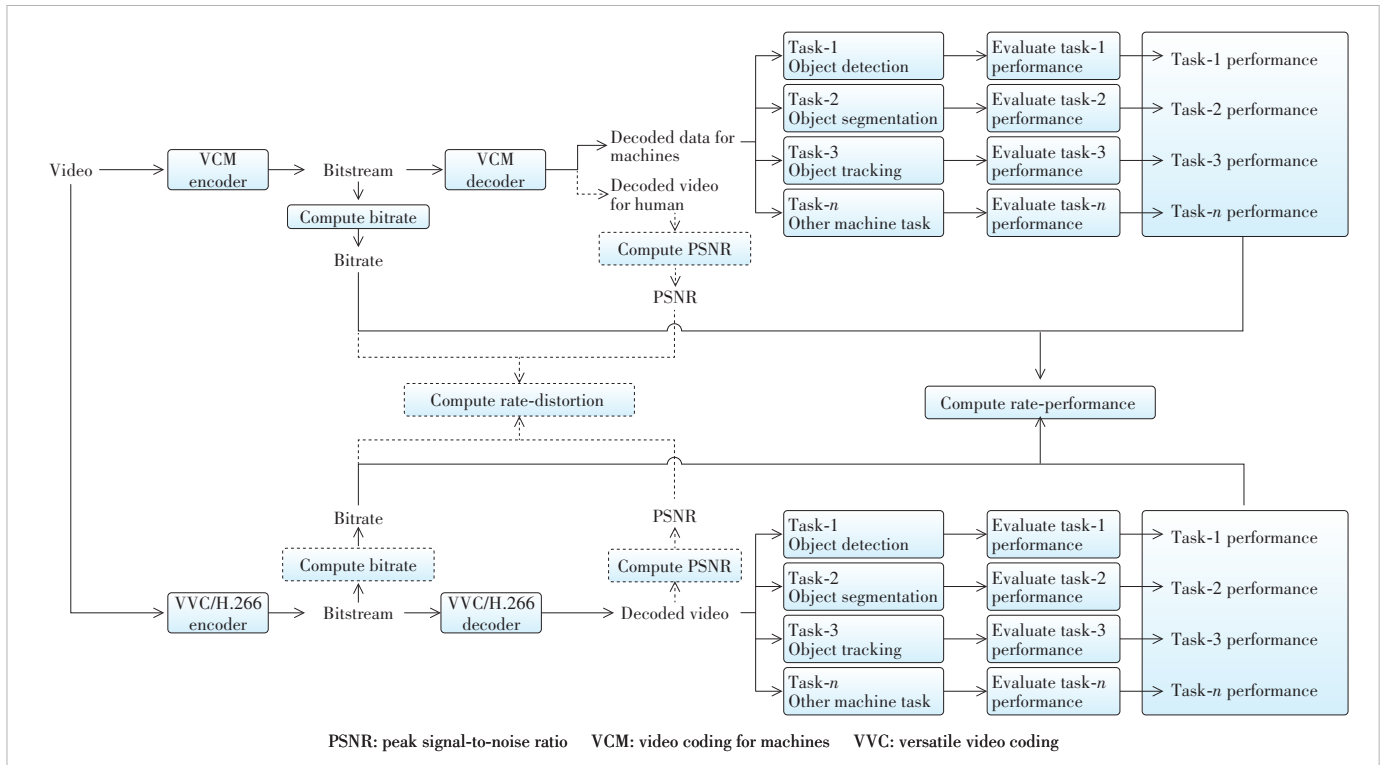
$$\text{MOTA} = 1 - \frac{\sum(\text{FN} + \text{FP} + \text{IDsw})}{\sum \text{GT}}, \quad (6)$$

where IDsw is the number of ID changes assigned by the real label during the tracking process. GT is ground truth, which represents the true value in the labels of all tracking targets in the current frame.

In traditional video encoding methods, to facilitate a fair comparison of compression performance, the bitrate is often converted to obtain the modified Bjontegaard delta bit-rate (BD-rate) under the same encoding quality. Similarly, this approach can be adopted for machine vision coding, with a slight modification. Instead of using the original peak signal-to-noise ratio (PSNR) metric to measure image quality, the accuracy rate, which serves as a performance measurement for machine vision tasks, can be used to calculate the BD rate. Similarly, the accuracy of machine vision tasks can be represented by BD-mAP or BD-MOTA, which shows the task's accuracy while maintaining a consistent bitrate.

4.4 Evaluation Pipeline

Fig. 2 shows the pipeline of the evaluation framework for



▲ Figure 2. Pipeline of VCM evaluation framework, where the dotted lines represent reconstruction quality assessment for human vision

the VCM standard. The encoding and decoding processes of using VVC and VCM codecs are roughly similar. They both send the video to the encoder to compress it into a code stream, decode the code stream, and send it to different machine vision tasks to obtain the performance of different tasks. However, in machine consumption, the code stream size and machine vision task performance are evaluated, and the decoded data of the code stream are not necessarily restored to the original input, as long as the code rate is reduced while ensuring task performance. In the method for human consumption, the code stream size and the quality of the reconstructed video are evaluated, and the code stream needs to be restored and reconstructed for humans to watch.

4.5 Anchor Generation

In an image/video encoding process, the original input is usually scaled to a smaller resolution (100%, 75%, 50% and 25% of the original size) before compression and scaled back to the original resolution. In the process of anchor generation, the input image/video will be cropped by down-sampling, then compressed and decompressed by VTM^[22], using up-sampling methods to restore the original input. After that, it will be sent to the task network to perform machine vision tasks and calculate the accuracy with ground truth. Among them, the open-source software FFmpeg 4.2.2^[23] is used for the up-down sampling and format conversion of the image. In addition, for the case where the zoom ratio is 100%, instead of directly inputting the original image, padding is performed to fill the edge.

5 Proposals for VCM Standard

Existing methods for machine vision coding can be roughly divided into two technical approaches: Track-1 and Track-2. The methods of VCM Track-1 are feature map coding methods while the methods of VCM Track-2 compress videos first and then reconstruct them as the input of the machine vision tasks. The representative methods of these two approaches are introduced in this section. In addition, this chapter will elaborate on the performance results comparison between different Track-1 methods and the standardization of the Track-2 framework.

5.1 Methods for Track-1

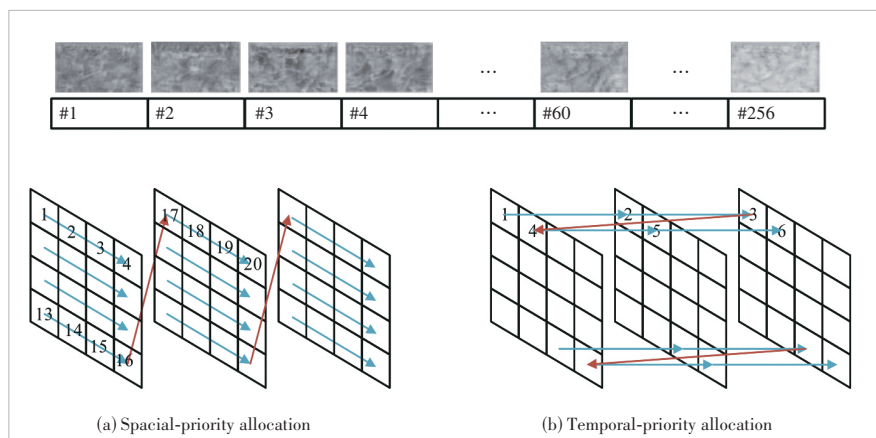
The methods of VCM Track-1 utilize the backbone network to extract feature maps from images or videos, which are then compressed, transmitted, and reconstructed before being sent to the back-end analysis network for machine vision tasks. This type of compression method only compresses the feature maps required for machine vision tasks. The primary advantage of this method is its capability to significantly re-

duce the number of data that need to be transmitted. However, it cannot realize the reconstruction of the original input signal. These methods can be mainly divided into the following categories in terms of specific implementation.

5.1.1 Feature Map Rearrangement and Compression by VTM

VVC uses traditional coding methods such as transform and quantization to compress images/videos, and the feature map can be regarded as a sequence of images to some extent. Therefore, an intuitive idea is to rearrange the obtained feature maps, and then treat them as ordinary images for compression and transmission through the traditional codec VVC. The feature map with 256 channels was regarded as 256 images in Ref. [24], where they were stitched, rearranged, and then quantized into 10-bit YUV 4:0:0 grayscale images, using VTM for compression. There might be not only spatial redundancy but also temporal redundancy between feature maps, and a method was proposed for reordering based on the gradient between feature map channels^[25]. Using the spatial-priority assignment and temporal-priority allocation strategy for feature maps reduces BPP while maintaining mAP performance. The schematic diagram is shown in Fig. 3, where the feature maps of 256 channels are rearranged by the method of space priority and time priority, which explores the spatial redundancy and temporal redundancy of feature maps, taking both of them into account to achieve efficient feature map compression.

This kind of method is relatively simple, intuitive and easy to implement. The feature map can be easily converted into a grayscale image for storage. The process of compression, transmission and reconstruction is also accomplished by existing codec software, without complicated convolution operation or model training process. However, since the data volume of the feature map is much larger than that of the original image, and there is no obvious advantage in using VTM to compress the feature map compared to directly compressing the input image/



▲ Figure 3. Spatial-priority and temporal-priority allocation, where the numbers in the diamond blocks represent the feature map numbers, and the arrows represent the rearrangement order

video, this type of method appears less in the proposals.

5.1.2 Multi-Scale Feature Compression Framework

In the feature extraction module, a multi-scale feature compression (MSFC) method is proposed, obtaining several feature maps of different sizes from the input signal. The number of channels of these feature maps is the same, and the size is halved from the low-level features to the high-level features one by one (for example, the length and width of the P2 layer are 1/4 of the input, the length and width of the P5 layer are 1/32 of the input, and the number of channels of the feature map of the P-layer is 256). Due to the huge number of channels, the data volume of low-level feature maps is much higher than that of the original image, so it is necessary to perform size reduction on large-scale feature maps. In addition, adjusting all feature maps to the same size can also make the feature map compression process more convenient. Therefore, some approaches to multi-scale feature compression have been proposed by scholars.

An MSFC method was proposed to compress the P-layer feature map in Ref. [26], as shown in Fig. 4, which specifically included three modules: a multi-scale feature fusion (MSFF) module, single-scale feature compression (SSFC) and multi-scale feature reconstruction (MSFR). In this method, the MSFF module scaled and fused the P2 – P5 feature maps to obtain a new single-scale feature map. The SSFC module included an encoder and a decoder to compress, transfer and reconstruct the single-scale feature map. In the MSFR module, the fused single-scale feature map would be restored by super-resolution, convolution and other operations to obtain the reconstruction of the input P2 – P5 feature maps.

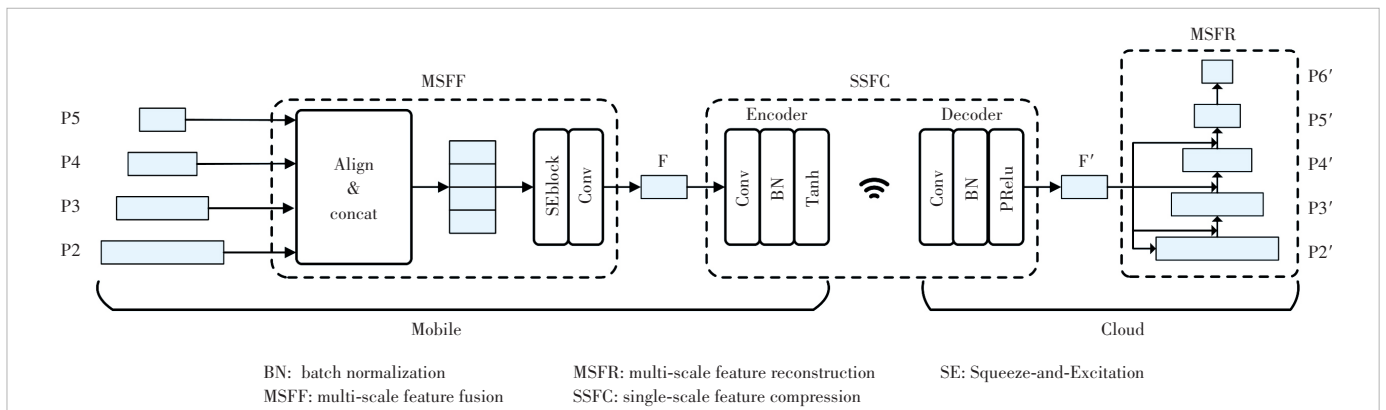
Based on the framework of MSFC, an MSFF module was optimized in Ref. [27], reducing the number of channels of the original feature map from 256 to 144 to further compress the P-layer features. Compared with the original MSFC method, this method achieved a huge bit rate gain while maintaining the accuracy of the back-end machine vision task. Further improvements^[28] were made based on the proposal in Ref. [27].

On the one hand, they supplemented the experimental results of further reducing the number of channels after P-layer compression to 64. On the other hand, they proposed a bottom-up structural optimization MSFF module. The experimental results showed that the compression efficiency of P-layer features could be further improved based on Ref. [27], which used the optimized MSFF module by further integrating the low-level semantic features and high-level semantic features, achieving a more compact expression of multi-scale features. This method reduced the fused feature dimension to 64, which finally achieved a performance gain of 84.98% compared with the image Anchor. The model was also trained on the COCO dataset and tested on the OpenImagesV6 5K dataset, and the back-end task model is Faster R-CNN. A dual-scale MSFC was proposed in Ref. [29]. Specifically, P2 and P3 were divided into one group, and P4 and P5 were grouped into one for compressed transmission respectively. Dual-scale MSFC provided higher fidelity for tasks that required higher spatial precision. Specifically, for high-quality operating points, dual-scale MSFC was used; while for other rate points, regular MSFC was used.

The above are several mainstream multi-scale feature compression methods. It is not difficult to find out that the core idea is to compress large-size feature maps into smaller feature maps and concatenate them together, compressing and transmitting smaller-size feature maps to achieve bit-rate saving. These methods are more intuitive and easier to implement.

5.1.3 End-to-End Feature Compression Methods

With the continuous development of deep learning, end-to-end image compression algorithms are becoming more and more mature. In addition to VTM software for encoding and decoding, some methods use learning-based end-to-end compression models for the compressed transmission of feature maps. Feature maps were compressed with a learning-based end-to-end compression model^[30]. In the process of end-to-end feature compression, the output features of the backbone layer of the target detection network were quantized and sent to the



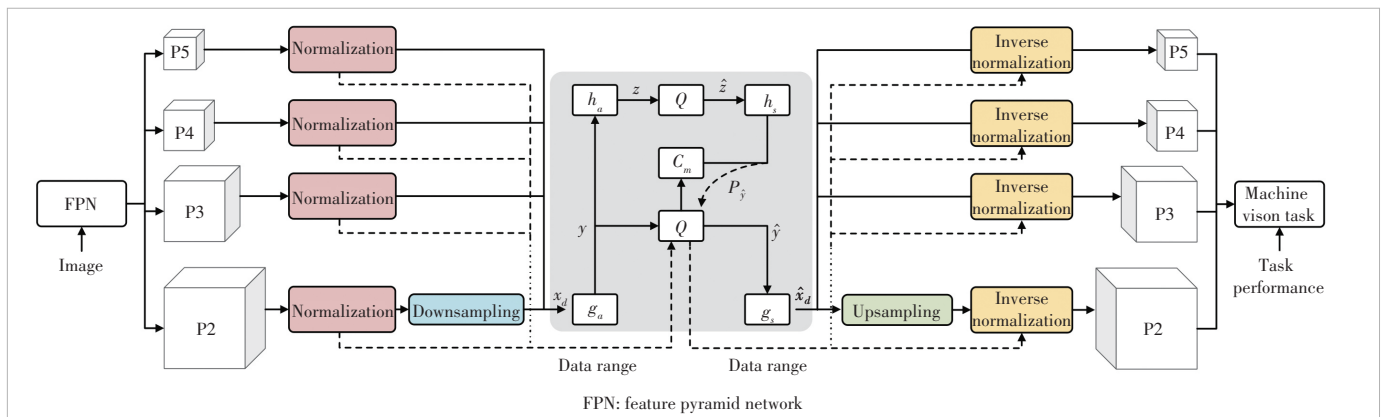
▲ Figure 4. Flowchart of multi-scale feature compression (MSFC), where MSFF scales and fuses different feature maps into one, SSFC realizes the compressed transmission of feature maps, and MSFR is used to reconstruct feature maps of different sizes

compression framework, CompressAI^[31], using bmshij2018-hyperprior^[32] to retrain the model, and the loss function was the bit-rate and mean squared error (MSE) loss. The test set of the experiment was the COCO dataset, and the compression performance was slightly worse than that of the image Anchor. ZHANG et al.^[33] used cheng2020^[34], an end-to-end SOTA model for image compression, to compress P-layer feature maps, as shown in Fig. 5. The model was trained on the OpenImagesV6 dataset and tested on the OpenImagesV6 5K dataset, using Mask R-CNN-X101-feature pyramid networks (FPN) as the back-end task model. In this method, the P2 – P5 layer features were first normalized to 0 – 1, and the real data range was additionally transmitted as auxiliary information. Considering that the P2 layer features had the least impact on task accuracy while containing the largest number of data, it was down-sampled to 1/2 of the original size. The end-to-end compression model used SOTA cheng2020 integrated with the CompressAI framework. In order to adapt to the input feature map, the input dimension received by the model was changed from 3 to 256, and the down-sampling in the compression network was removed. The experimental results showed that compared with the image anchor, the proposed scheme had a performance gain of 69.08% under the BD-rate-mAP@0.5 index. A feature compression architecture was proposed based on the cheng2020 model^[35], adding a feature fusion module to accommodate multi-layer inputs from FPN. The encoder was trained in an end-to-end method, and the losses were MSE and BPP. When processing a video, the key frame feature encoder operated similarly to the image feature encoder, while for non-key frames, a motion encoder, an inter-prediction module, and an encoder were used.

5.1.4 Redesigned Feature Compression Framework

The methods above are based on existing mature frameworks such as VVC, the end-to-end image compression model, and the MSFC model. The schemes introduced below do not follow the existing ones while redesigning a set of feature compression-oriented coding schemes. Their frameworks all

contain resampling, transformation, quantization and entropy coding processes. The stage-two output feature^[36] of MobileNet^[37] was encoded, which had the same size as the original input image. The scheme flow is shown as follows. First, add an offset to the feature value of the output feature of stage two, so that the center of the value is 0, and then convert all the features into a two-dimensional matrix with a size of N rows and 16 columns. For each 4×4 small block, perform discrete cosine transformation (DCT), then quantize the transformation coefficient, transmit the maximum value of the transformation coefficient as auxiliary information, and encode and compress the quantized value and auxiliary information using an entropy encoder, thus completing the entire feature compression process. The DCT transformation^[38] was changed to the complex wavelet transformation (DWT), and other settings were similar. The schemes given in the two proposals were relatively preliminary, and the test results were not clear. The method in Ref. [39] was refined based on Ref. [36], and at the same time adopted the back-end task model Faster R-CNN-X101-FPN specified in the VCM standard to compress the features of the backbone layer. This method uses a larger transform kernel size (8×8) and divides the transform coefficients into direct current (DC) components and alternate current (AC) components. An 8×8 block has 64 transform coefficients, including 1 DC coefficient and 63 AC coefficients. Different quantization methods are used for DC and AC coefficients. For the binary arithmetic coding part, the design scheme is consistent with the traditional coding algorithm, which is also based on the block sequence. Considering that the context model of the entropy coder, context adaptive binary arithmetic coding (CABAC), used in traditional coding standards such as H.264 is very complicated, this proposal simplifies the design of the context model and determines the context model based only on the symbol position. Compared with other solutions, it can realize a single-model variable bit rate, but as of the time of proposal, it has not yet realized a single-model wide-range adjustable bit rate, and the achieved effect has 23.78% perfor-



▲ Figure 5. Flowchart of end-to-end feature compression for P-layer feature maps, using an end-to-end image compression model instead of single-scale feature compression (SSFC) in multi-scale feature compression (MSFC)

mance compared with image anchor gain. A feature compression framework based on the principal component analysis (PCA) and DeepCABAC was proposed in Ref. [40]. First, perform PCA transformation on the output of the P-layer features by Faster R-CNN X101-FPN. Considering the number of P2 layer feature data is large, down-sampling is used on the P2 layer features and will be down-sampled and merged into the P3 layer feature map according to the channel dimension. Then divide P4, P5 and P3 layer feature maps combined with the P2 layer into three groups to perform PCA transformation separately, use VVC to encode the main basis vector and mean feature, and use DeepCABAC to perform entropy encoding on the transformation coefficient. This method is tested on the OpenImagesV6 5K dataset, and its performance is better than that of the P-layer feature compression anchor and worse than that of the image compression anchor. The PCA technology to process the P-layer features of the FPN output was used in Ref. [41]. Larger sizes P2 and P3 are scaled and concatenated by down-sampling, while P4 and P5 are processed separately. The PCA transformation is performed by introducing a predefined mean and bias. This PCA reveals redundancies between channels rather than the spatial features of a given image expressed in the feature domain. The data are then quantized, rearranged and compressed with VTM.

5.1.5 Other Feature Compression Methods

The methods above have been intensively studied, while there are still some exploration schemes beyond the mainstream that are difficult to categorize. These methods will be summarized in this section to give researchers some new inspiration. In the MSFC method, the reduced-size feature map will be up-sampled by convolution, and the low-level features will be gradually restored from the high-level features. In terms of size, this step can be regarded as super-resolution. There is a certain redundancy between feature maps of different layers. Therefore, one method is to directly perform super-resolution on high-level features to obtain low-level features. It was pointed out that the P-level features had different sensitivities to back-end visual tasks, which was reflected in the fact that the bottom-level features have less influence and the top-level features have a greater impact^[42]. Hence, this method only

transmits some of the top-level features and zooms them to obtain bottom-level features with a super-resolution technique. The experimental results show that the effect of approximating the image anchor can be achieved by only transmitting a part of the features. VTM is used for feature compression.

The method of feature compression is oriented to machine vision tasks, and the reconstructed features can only be used for machine vision tasks and cannot be used to reconstruct images. However, some proposals consider the feasibility of reconstructing images based on feature compression. In Ref. [43], the original image is encoded through feature extraction, and the reconstructed features can be used for back-end machine vision tasks. At the same time, the reconstructed features can also be used to reconstruct the original input through the designed deep neural network. The experimental results show that the optimal peak signal-to-noise ratio (PNSR) of the reconstructed image can reach 31.139 dB, which can restore the input image well while realizing the machine vision task.

5.2 Track-1 CfE Responses

Since the method in VCM Track-1 started later than that in Track-2 and progressed relatively slowly, the previous meetings mainly collected relevant evidence for Track-1 to verify its effectiveness. At present, in the 140th meeting of MPEG, the reports of the objective results received in response to CfE have been sorted out and summarized, and the performance of proposals with better effects is shown in Table 2.

At present, a total of 6 responses to feature compression evidence proposals with complete and effective tests have been received, 4 of which have better performance than video anchors, and all 6 proposals have better performance than image anchors. According to the test results, the VCM standard group believes that the current evidence proposals are sufficient enough, requesting MPEG WG2 to carry out the drafting of the Call for Proposal. In recent meetings, several new proposals were put forward, showcasing remarkable performance. Among them, the most outstanding result was achieved^[48], which has a -92.21% BD-rate gain compared with the image anchor for object detection tasks on the OpenImageV6 dataset. Compared with the feature anchor, there is a -94.70% per-

▼Table 2. Proposal summary results on TVD-overall

Overall	Instance Segmentation		Overall	Object Detection		Overall	Object Detection	
	BD-rate over video	BD-rate over feature		BD-rate over video	BD-rate over feature		BD-rate over video	BD-rate over feature
Ref. [44]	-87.44%	-97.58%	Ref. [44]	-79.21%	-95.56%	Ref. [44]	-81.11%	-94.15%
Ref. [41]	63.69%	-74.43%	Ref. [41]	-47.46%	-89.48%	Ref. [41]	-54.51%	-85.06%
Ref. [45]	-80.18%	-97.09%	Ref. [45]	-93.04%	-98.60%	Ref. [45]	-94.46%	-98.34%
Ref. [35]	218.93%	-33.01%	Ref. [46]	-19.35%	-83.38%	Ref. [29]	-70.39%	-91.14%
Ref. [29]	-77.40%	-95.84%	Ref. [29]	-78.11%	-95.84%			
Ref. [47]	-64.94%	-92.17%	Ref. [47]	-69.08%	-92.30%			

BD: Bjontegaard Delta

formance gain. Similarly, for instance segmentation tasks, the BD-rate gains are -83.35% and -96.01% respectively compared with image anchor and feature anchor. This proposal holds substantial academic value and exhibits promising research prospects.

5.3 Methods for Track-2

The methods of VCM Track-2 compress and transmit image/video first, and then perform back-end machine vision tasks. Compared with the methods of VCM Track-1, they complete the reconstruction of the input signal, which can be viewed by human eyes instead of only machine tasks. These methods generally use the end-to-end compression framework. Since the back-end machine vision task depends on reconstructing image/video frames, the consistent approach of the Track-2 proposal with better performance is to change the loss function of the end-to-end compression network, adding task loss weighting to the loss function. The end-to-end compression model is retrained when fixing the network parameters to obtain better compression performance on specific tasks. This type of method can be mainly divided into the following categories in terms of a specific implementation.

5.3.1 Learning-Based Joint Optimization for Image Compression and Vision Tasks

At present, the effective proposals in VCM Track-2 are all based on the combination of learning-based image coding and back-end machine vision tasks. Taking the object detection task as an example, the entire model is composed of neural networks, where the encoder and decoder use relatively mature end-to-end image compression models, such as cheng2020^[34], bmsbj2018-hyperprior^[32], and mbt2018-mean^[49]. And then the parameters of the back-end machine vision task model are fixed, introducing them as a loss function to the training process. Most of the current methods^[50-53] in Track-2 adopt this idea. It is obvious that the core of such proposals is actually how to train a more efficient image compression network under the condition of changing the loss function so that it can perform better on specific back-end machine vision tasks.

5.3.2 Video Coding For Machine Vision Based on Region of Interest

When a video is encoded, not all regions of the picture are of an equal value, since some regions contain important information that attracts the attention from humans or machines, while others are ignored. Based on this situation, some researchers have proposed a video coding method based on the region of interest (RoI). Simply speaking, these methods allocate more bit rate resources to the region with a high degree of interest for refined coding, while appropriately reducing coding quality for the non-interesting area, which achieves overall code rate savings. Currently, there are several representative methods as follows.

Ericsson proposed a preliminary idea implemented based on VTM^[54], and the compressed video can be decoded using any compatible VVC decoder. This method performs object detection on the input video sequence, using the detection box as the RoI area, and adjusts the offset of the quantization parameter (QP) according to whether the coding tree units (CTU) contain detection boxes. If the CTU contains a detection box, it will be further distinguished depending on the size of the object. In terms of compression efficiency, this method achieves the same machine vision performance as the anchor at a lower bit rate.

KALVA et al.^[55] used YOLOv7^[56] to detect the RoI area. Since the detection boxes of different targets overlapped, they were superimposed and spliced to obtain an irregular polygonal RoI area, and then the irregular area was divided into rectangles of different sizes which would be scaled, rearranged, and packed into a new frame. This method could separately encode and decode the RoI area and the non-interested area, thereby reducing the overall bit rate. In Ref. [57], a method based on machine attention calculation was proposed, which calculated the machine attention value of the input signal at the encoding end, and divided the coding unit (CU) block in one frame into high attention area (HAR) and low attention area (LAR), taking different rate control strategies respectively. At the decoding end, these two areas were decoded separately and reassembled into a complete frame to reconstruct the input signal. This method had a certain gain on the image dataset, and the performance fluctuated obviously according to the sequence and different bit rate points on the video dataset. The above methods reduce the bit rate while maintaining the accuracy of the machine vision task.

5.3.3 Machine Vision-Oriented Video Coding Based on Spatial-Temporal Up-and-Down Sampling

A video signal can be viewed as several consecutive images. Different from a single picture, adjacent video frames are similar in the time domain, and there is a large temporal redundancy. However, many methods simply treat video as multiple pictures without considering their temporal correlation, and there is still room for further optimization.

At present, some researchers compress video in the time domain, usually using the method of frame extraction and frame complementation. A joint adaptive spatial-temporal (JAST) system was proposed in Ref. [58], which converted the input signal into a form more suitable for encoding through machine vision-oriented pre-analysis and pre-processing. The adaptive temporal up-and-down sampling based on texture complexity could be accomplished by frame extraction and interpolation technology to reduce temporal redundancy, and spatial up-sampling could be realized by image resizing and super-resolution technology to reduce spatial redundancy. In addition, since VVC was used as the core codec, JAST naturally inherited many capabilities of VVC, which were universal and

could be widely deployed. This technology was also used similarly in Ref. [59], where a content-adaptive resampling (CAR) network was proposed to achieve spatial down-sampling and was restored through bicubic interpolation, using frame extraction to perform down-sampling in the time domain. At the decoding end, the video frame interpolation network based on the optical flow was used for restoration.

5.3.4 Video Coding for Machines Optimized by End-to-End Video Compression Framework

Similar to learning-based end-to-end image compression algorithms, such methods can also be used for video compression. Traditional encoders adopt different coding strategies for I-frames and P-frames of video. This feature was considered and a learning-based video codec was proposed in Ref. [60], including feature extraction, feature encoding and decoding, image reconstruction, and neural network modules. Besides, different training network models were adopted for the I-frame and P-frame, and the reconstruction result of the compressed code stream of the I-frame was used to assist in the reconstruction of the P-frame.

5.3.5 Hybrid Learning-Based Model and Traditional Codec to Boost Video Coding for Machines

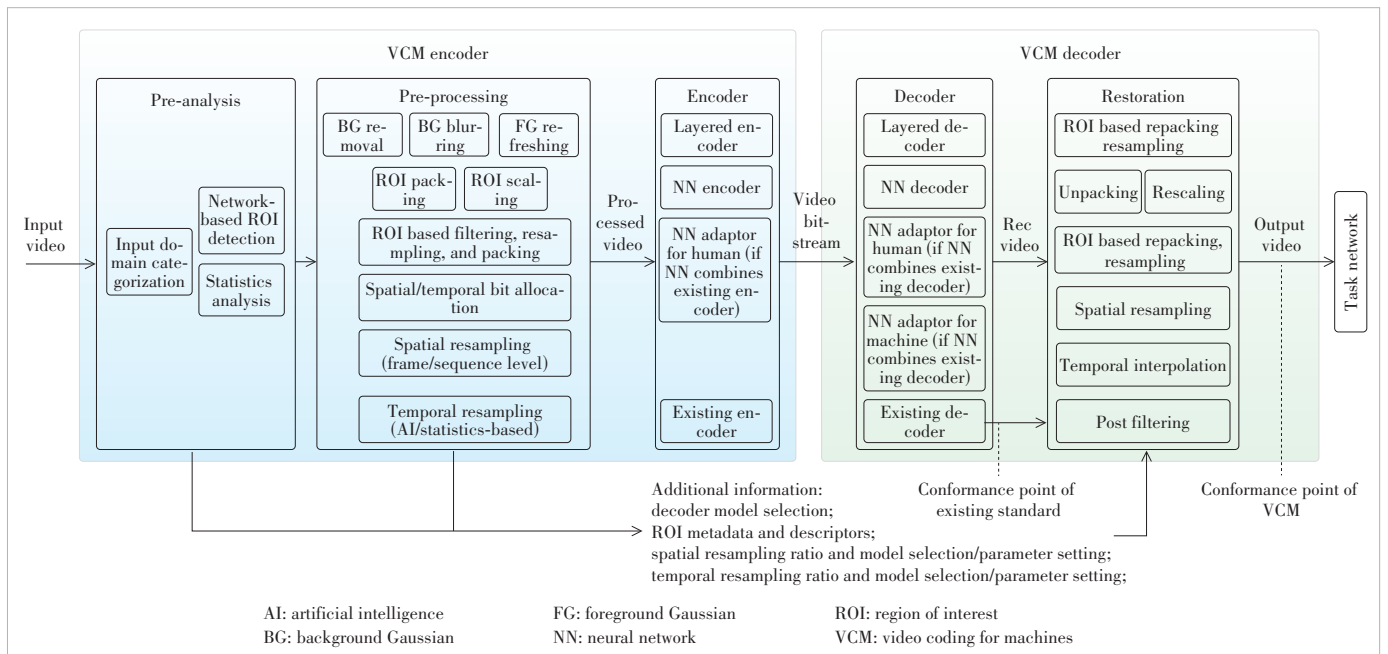
The learning-based end-to-end compression method has a high compression performance, but its time complexity is high, and it will be time-consuming to compress all frames end-to-end. In order to solve this problem, some proposals have adopted hybrid video coding methods that only adopt learning-based compression algorithms for the important frames. Learning-based image coding was only used for I-frames and optical flow-based inter-frame coding for P-

frames^[61]. In application, only the ROI area and auxiliary target frame position information were encoded for image coding. The ROI area was determined by the object detection network. For video coding, only in the smallest ROI region that all objects in a group of pictures (GOP) were coded. The network optimized the loss function for image reconstruction MSE. This method performed well on object detection and instance segmentation but had a negative gain on object tracking. In Ref. [62], learning-based image coding was used for the I-frame, and the traditional encoder VVC was used for the P-frame, which achieved good results in object detection, instance segmentation and object tracking.

5.4 Framework of VCM Track-2 Standardization

The method in VCM Track-2 started earlier and received many valuable proposals during several meetings after the CfP was proposed. In the 140th meeting of MPEG, 17 Track-2 proposal responses were officially received, involving technical directions such as end-to-end video compression, traditional block-based video compression, hybrid end-to-end and traditional video compression. In the current solution, there is a maximum performance improvement of 55% for object tracking, a maximum performance improvement of 45% for instance segmentation, and a maximum performance improvement of 40% for object detection. Based on the current effect, MPEG continues to promote the VCM standardization work and officially transfers Track-2 from the demand group to the video group. In the most recent meetings, the main work content of VCM Track-2 was to promote standardization and develop the test software platform of VCM.

Fig. 6 shows the overall framework of the Track-2 proposal,



▲ Figure 6. An overall framework of the Track-2 proposal

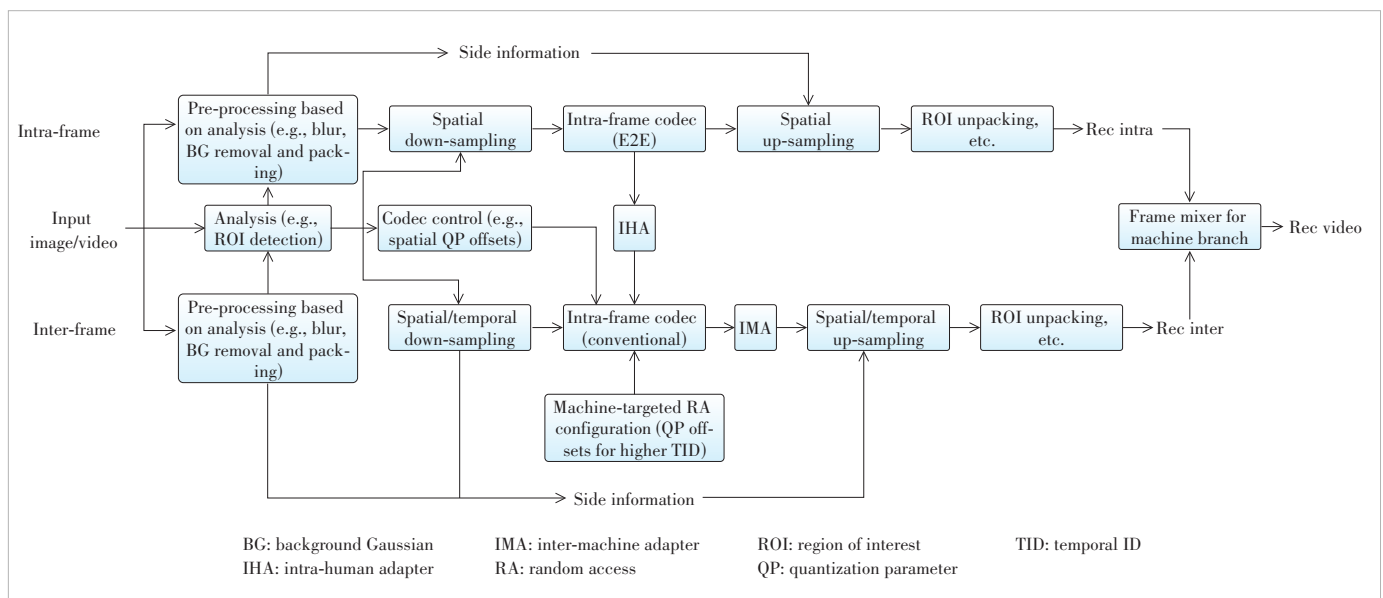
which includes five parts: pre-analysis, pre-processing, encoder, decoder, and restoration. The pre-analysis part includes network-based ROI detection to identify regions of interest, in addition to input region classification and statistical analysis modules. The pre-processing part includes ROI-based packing, resampling, rearranging, and resampling of temporal and spatial dimensions. The codecs can use existing legacy codecs, layered codecs, neural network-based codecs, and network-based adapters for humans. In the restoration part, ROI-based packing and scaling are reversed and spatio-temporal resampling is used to reconstruct the input video and fed into the back-end machine vision task network. In addition to the above main route, some additional information, including ROI metadata, descriptors and sampling ratios, is also transmitted for the restoration of the input signal.

In addition, the test software platform of VCM called Test Model Under Verification (TMUV) 1.0 is also developed and verified, and its flow chart is shown in Fig. 7. Among them, different strategies are adopted for the input video frame according to intra-frame and inter-frame. The intra-frame is pre-processed and spatially down-sampled and then compressed by an end-to-end codec for the intra-frame, utilizing spatial up-sampling and ROI unpacking to restore. For inter-frames, the overall process is similar, while the up-and-down sampling part is different from that in intra-frames, using up-and-down sampling based on time dimension and space dimension. In addition, the traditional codec is used for compression, and the compression result of the intra-frame is processed by the intra-frame human eye adapter to assist in the encoding and decoding of the inter-frame, and then converted into a code stream suitable for machine vision by the inter-frame machine adapter for subsequent frame-to-frame reconstruction. Finally, intra-frames and inter-frames are mixed and assembled, and

sent to subsequent machine vision tasks.

6 Conclusions

We introduce the current development status, background and evaluation framework of the VCM standard, in addition to some mainstream implementation methods and the latest advances. VCM is a research hotspot with numerous scholars and companies deeply involved. The VCM standardization organization has established guidelines for use cases, requirements, test frameworks, datasets, and evaluation metrics. Additionally, a series of proposals and evidence are collected to facilitate the formation of an international standard. At present, the two main routes (Track-1 and Track-2) have achieved significant performance gains, and related methods are further explored and improved. However, there are still some problems in the existing schemes, such as insufficient support for multitask, high complexity of learning-based schemes, and difficulties in simultaneously achieving efficient compression for human vision and machine vision. As standardization advances and more in-depth research is conducted, it is expected that the VCM standard can achieve efficient machine vision-oriented compression and facilitate M2M communication in the wild. Our aim in writing this paper is to promote a deeper understanding of VCM among scholars and researchers, and to encourage greater involvement in this field to drive its rapid development. Currently, there is a growing expectation for the convergence of research areas between two types of video coding standard organizations: NNVC, which focuses on neural network-based approaches, and JVET, which is centered on traditional codecs. The merging of expertise aims to encompass the existing VCM standard and foster the development of video coding techniques specifically tailored for machine vision applications. As research progresses in the field



▲ Figure 7. Flowchart of VCM's test software platform TMUV 1.0

of VCM standards, it is anticipated that the standard will continually improve its performance and enable the emergence of new and innovative applications. The integration of knowledge and efforts between neural network-based approaches and traditional codec methods holds promise for advancing the capabilities of machine vision-oriented video coding.

References

- [1] DUAN L Y, LOU Y H, BAI Y, et al. Compact descriptors for video analysis: the emerging MPEG standard [J]. *IEEE multimedia*, 2019, 26(2): 44 - 54. DOI: 10.1109/MMUL.2018.2873844
- [2] DUAN L Y, CHANDRASEKHAR V, CHEN J, et al. Overview of the MPEG-CDVS standard [J]. *IEEE transactions on image processing*, 2016, 25(1): 179 - 194. DOI: 10.1109/TIP.2015.2500034
- [3] LE N, ZHANG H L, CRICRI F, et al. Image coding for machines: an end-to-end learned approach [C]//*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021: 1590 - 1594. DOI: 10.1109/ICASSP39728.2021.9414465
- [4] LE N, ZHANG H L, CRICRI F, et al. Learned image coding for machines: a content-adaptive approach [C]//*IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2021: 1 - 6. DOI: 10.1109/ICME51207.2021.9428224
- [5] TU H Y, LI L, ZHOU W G, et al. Semantic scalable image compression with cross-layer priors [C]//*The 29th ACM International Conference on Multimedia*. ACM, 2021: 4044 - 4052. DOI: 10.1145/3474085.3475533
- [6] CHOI H, BAJIĆ I V. Deep feature compression for collaborative object detection [C]//*The 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018: 3743 - 3747. DOI: 10.1109/ICIP.2018.8451100
- [7] CHOI H, BAJIĆ I V. Latent-space scalability for multi-task collaborative intelligence [C]//*IEEE International Conference on Image Processing (ICIP)*. IEEE, 2021: 3562 - 3566. DOI: 10.1109/ICIP42928.2021.9506712
- [8] DUAN L Y, LIU J Y, YANG W H, et al. Video coding for machines: a paradigm of collaborative compression and intelligent analytics [J]. *IEEE transactions on image processing*, 2020, 29: 8680 - 8695. DOI: 10.1109/TIP.2020.3016485
- [9] GAO W, LIU S, XU X Z, et al. Recent standard development activities on video coding for machines [EB/OL]. (2021-05-26) [2023-06-06]. <https://arxiv.org/abs/2105.12653>
- [10] BJONTEGAARD G. Calculation of average PSNR differences between RD-curves [J]. *Computer science*, 2001
- [11] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(6): 1137 - 1149. DOI: 10.1109/TPAMI.2016.2577031
- [12] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context [M]. *Computer Vision*. Cham: Springer International Publishing, 2014: 740 - 755. DOI: 10.1007/978-3-319-10602-1_48
- [13] Open images dataset. Open images dataset v6 [EB/OL]. [2023-06-06]. <https://storage.googleapis.com/openimages/web/download.html>
- [14] TELEDYNE FLIR. Free FLIR thermal dataset for algorithm training [EB/OL]. [2023-06-06]. <https://www.flir.com/oem/adas/adas-dataset-form>
- [15] ROY A, GUINAUDEAU C, BREDIN H, et al. TVD: a reproducible and multiply aligned TV series dataset [C]//*The 9th International Conference on Language Resources and Evaluation*. ELRA, 2014: 418 - 425
- [16] BAJIĆ I, CHOI H, HOSSEINI E, et al. Sfu-hw-objects-v1 [EB/OL]. (2020-06-24) [2023-06-06]. <https://www.frdr-dfrd.ca/repo/dataset/59931535-9ffd-4cc3-a3c2-4b06d06603d1>
- [17] WANG Z D, ZHENG L, LIU Y X, et al. Towards real-time multi-object tracking [M]. *Computer Vision*. Cham: Springer International Publishing, 2020: 107 - 122. DOI: 10.1007/978-3-030-58621-8_7
- [18] LIN W Y, LIU H B, LIU S Z, et al. Human in events: a large-scale benchmark for human-centric video analysis in complex events [EB/OL]. (2020-05-09) [2023-06-06]. <https://arxiv.org/abs/2005.04490>
- [19] FAN H, LI Y, XIONG B, et al. Pyslowfast [EB/OL]. [2023-07-06]. <https://github.com/facebookresearch/slowfast>
- [20] WANG J D, SUN K, CHENG T H, et al. Deep high-resolution representation learning for visual recognition [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2021, 43(10): 3349 - 3364. DOI: 10.1109/TPAMI.2020.2983686
- [21] ANDRILUKA M, PISHCHULIN L, GEHLER P, et al. 2D human pose estimation: New benchmark and state of the art analysis [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2014: 3686 - 3693. DOI: 10.1109/CVPR.2014.471
- [22] VTM. VVC Test Model [EB/OL]. [2023-06-06]. <https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftwareVTM/-/tags>
- [23] TOMAR S. Converting video formats with FFmpeg [J]. *Linux journal*, 2006, 146: 93 - 94
- [24] ISO/IEC MPEG. The test results of compressing P-layer feature maps on the Mask R-CNN network: m59942 [S]. 2022
- [25] ISO/IEC MPEG. A feature map compression method by using generated feature frames for object segmentation: m57958 [S]. 2021
- [26] ISO/IEC MPEG. Investigation on deep feature compression framework for multi-task: m58772 [S]. 2022
- [27] ISO/IEC MPEG. Compression of FPN multi-scale features for object detection using VVC: m59562 [S]. 2022
- [28] ISO/IEC MPEG. Performance of the enhanced MSFC with bottom-up MSFF: m60197 [S]. 2022
- [29] ISO/IEC MPEG. Response to CfE on video coding for machine from canon: m60821 [S]. 2022
- [30] ISO/IEC MPEG. Experimental results of feature compression using CompressAI: m56716 [S]. 2021
- [31] BÉGAINT J, RACAPÉ F, FELTMAN S, et al. CompressAI: a PyTorch library and evaluation platform for end-to-end compression research [EB/OL]. (2020-11-05) [2023-06-06]. <http://arxiv.org/abs/2011.03029>
- [32] BALLÉ J, MINNEN D, SINGH S, et al. Variational image compression with a scale hyperprior [EB/OL]. (2018-02-01) [2023-06-06]. <http://arxiv.org/abs/1802.01436>
- [33] ZHANG G Y, LIU Z, XU X, et al. Learning-based feature compression for instance segmentation: m60240 [S]. Geneva: ISO/IEC MPEG, 2022
- [34] CHENG Z X, SUN H M, TAKEUCHI M, et al. Learned image compression with discretized Gaussian mixture likelihoods and attention modules [C]//*Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2020: 7936 - 7945. DOI: 10.1109/CVPR42600.2020.00796
- [35] ISO/IEC MPEG. An end-to-end video feature compressing method with feature fusion modules: m60803 [S]. 2022
- [36] ISO/IEC MPEG. A DCT based feature compression algorithm: m58000 [S]. 2020
- [37] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications [EB/OL]. (2017-04-17) [2023-06-06]. <https://arxiv.org/abs/1704.04861>
- [38] ISO/IEC MPEG. A DWT based feature compression algorithm: m58001 [S]. 2020
- [39] ISO/IEC MPEG. DCT based feature compression: m60257 [S]. 2022
- [40] ISO/IEC MPEG. Advanced feature map compression based on optimal transformation with VVC and DeepCABAC: m58787 [S]. 2022
- [41] ISO/IEC MPEG. Response to CfE: a transformation-based feature map compression method: m60788 [S]. 2022
- [42] ISO/IEC MPEG. Improvement of a feature map compression using predicted p-layer feature map for object detection: m59512 [S]. 2022
- [43] ISO/IEC MPEG. End-to-end learning-based image reconstruction framework for hybrid machine and human vision: m58994 [S]. 2022
- [44] ISO/IEC MPEG. Response from Hanbat National University and ETRI to CfE on video coding for machines: m60761 [S]. 2022
- [45] ISO/IEC MPEG. Response to VCM CfE: multi-scale feature compression with QP-adaptive feature channel truncation: m60799 [S]. 2022
- [46] ISO/IEC MPEG. An end-to-end image feature compressing method with feature fusion module: m60802 [S]. 2022
- [47] ISO/IEC MPEG. Response to VCM call for evidence from Tencent and Wuhan University - a learning-based feature compression framework: m60925 [S]. 2022

- [48] ISO/IEC MPEG. Hybrid loss training based on reversed SIMO in feature compression for object detection and instance segmentation: m63174 [S]. 2023
- [49] MINNEN D, BALLÉ J, TODERICI G. Joint autoregressive and hierarchical priors for learned image compression [C]//The 32nd International Conference on Neural Information Processing Systems. ACM, 2018: 10794 – 10803. DOI: 10.5555/3327546.3327736
- [50] ISO/IEC MPEG. ZJU's response to VCM CfE: deep learning-based compression for machine vision: m56445 [S]. 2021
- [51] ISO/IEC MPEG. End-to-end image compression towards machine vision for object detection: m57500 [S]. 2021
- [52] ISO/IEC MPEG. End-to-end learning-based compression for object detection: m58165 [S]. 2021
- [53] ISO/IEC MPEG. Object detection and instance segmentation results with recent NN codecs: m58050 [S]. 2021
- [54] ISO/IEC MPEG. Response to call for proposals from Ericsson: m60757 [S]. 2022
- [55] KALVA H, ADZIC V, KRAUSE B F A, et al. Response to VCM CfP from the Florida Atlantic University and OP solutions, LLC: m60743 [S]. Geneva: ISO/IEC MPEG, 2022
- [56] WANG C Y, BOCHKOVSKIY A, LIAO H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//The IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2023: 7464 – 7475. DOI: 10.1109/cvpr52729.2023.00721
- [57] ISO/IEC MPEG. Response to VCM CfP: video coding with machine-attention: m60378 [S]. 2022
- [58] ISO/IEC MPEG. Video coding for machines CfP response from Alibaba and City University of Hong Kong: m60737 [S]. 2022
- [59] ISO/IEC MPEG. Response to VCM call for proposals – an EVC based solution: m60779 [S]. 2022
- [60] ISO/IEC MPEG. Response to the CfP on video coding for machine from Zhejiang University: m60741 [S]. 2022
- [61] ISO/IEC MPEG. Response to VCM call for proposals from Tencent – an end-to-end learning based solution: m60777 [S]. 2020
- [62] ISO/IEC MPEG. Response to the CfP of the VCM by Nokia: m60753 [S]. 2022

Biographies

ZHANG Qiang is the Director of the Big Video Committee of ZTE Corporation. His research interests include computer vision, audio and video codec, transmission, and network architecture.

MEI Junjun is the Chief R&D engineer in the audio and video field of ZTE Corporation, engaged in the evolution of the integrated video cloud network architecture, computer vision, audio and video codec, and transmission technologies.

GUAN Tao is the senior system architect of ZTE Corporation, mainly engaged in the architecture design and algorithm research of video systems and industrial digital systems. He has participated in standard organizations, initiated and compiled the formulation of a number of communication standards, and applied for more than 20 national invention patents.

SUN Zhewen is a student at the School of Electronic Information and Communications, Huazhong University of Science and Technology, China, currently engaged in research on video coding.

ZHANG Zixiang is pursuing his master's degree at the School of Electronic Information and Communications, Huazhong University of Science and Technology, China. He is currently engaged in research on video coding.

YU Li (hustlyu@hust.edu.cn) received her PhD degree from Huazhong University of Science and Technology (HUST), China in 1999, where she is currently a professor with the School of Electronic Information and Communications, HUST. Her research interests include image and video coding, multimedia communications, and artificial intelligence.

RIS-Assisted Cell-Free MIMO: A Survey



ZHAO Yaqiong¹, KE Hongqin¹, XU Wei^{1,2}, YE Xinquan³,
CHEN Yijian³

(1. National Mobile Communications Research Laboratory, Southeast University, Nanjing 211189, China;
2. Purple Mountain Laboratories, Nanjing 211111, China;
3. ZTE Corporation, Shenzhen 518507, China)

DOI: 10.12142/ZTECOM.202401009

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240311.1430.002>, published online March 12, 2024

Manuscript received: 2023-07-15

Abstract: Cell-free (CF) multiple-input multiple-output (MIMO) is a promising technique to enable the vision of ubiquitous wireless connectivity for next-generation network communications. Compared to traditional co-located massive MIMO, CF MIMO allows geographically distributed access points (APs) to serve all users on the same time-frequency resource with spatial multiplexing techniques, resulting in better performance in terms of both spectral efficiency and coverage enhancement. However, the performance gain is achieved at the expense of deploying more APs with high cost and power consumption. To address this issue, the recently proposed reconfigurable intelligent surface (RIS) technique stands out with its unique advantages of low cost, low energy consumption and programmability. In this paper, we provide an overview of RIS-assisted CF MIMO and its interaction with advanced optimization designs and novel applications. Particularly, recent studies on typical performance metrics such as energy efficiency (EE) and spectral efficiency (SE) are surveyed. Besides, the application of RIS-assisted CF MIMO techniques in various future communication systems is also envisioned. Additionally, we briefly discuss the technical challenges and open problems for this area to inspire research direction and fully exploit its potential in meeting the demands of future wireless communication systems.

Keywords: mmWave; beyond 5G (B5G); Internet of Everything (IoE); cell-free MIMO; RIS; unmanned aerial vehicle (UAV); physical layer security (PLS); wireless energy transfer (WET)

Citation (Format 1): ZHAO Y Q, KE H Q, XU W, et al. RIS-assisted cell-free MIMO: a survey [J]. *ZTE Communications*, 2024, 22(1): 77 – 86. DOI: 10.12142/ZTECOM.202401009

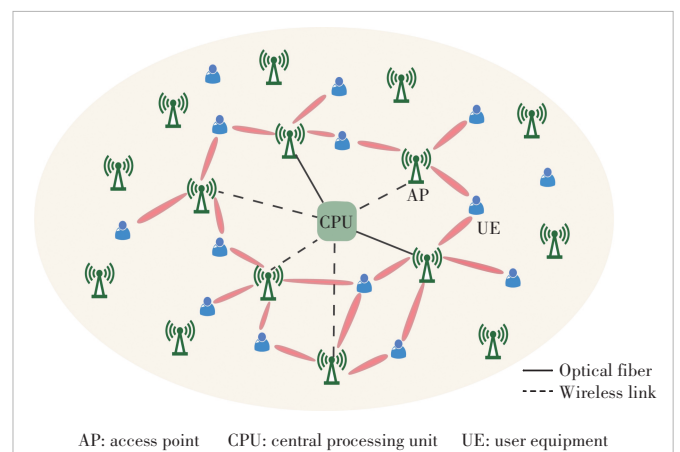
Citation (Format 2): Y. Q. Zhao, H. Q. Ke, W. Xu, et al., “RIS-assisted cell-free MIMO: a survey,” *ZTE Communications*, vol. 22, no. 1, pp. 77 – 86, Mar. 2024. doi: 10.12142/ZTECOM.202401009.

1 Introduction

With the promising new technologies such as the millimeter wave (mmWave), massive multiple-input multiple-output (MIMO), and ultra-dense network (UDN), 5G wireless communication and beyond can provide increased data rates, as well as reduced latency, higher reliability, and greater connectivity, which further supports novel applications such as telemedicine and remote driving^[1-3]. However, there are new challenges in practical deployment of these technologies. For example, mmWave suffers from low penetration and is vulnerable to interference, leading to limited signal coverage. Besides, massive MIMO and UDN require the intensive deployment of more co-located antennas and base stations (BSs). This incurs high costs and power consumption. Furthermore, as cell density increases, inter-cell interference becomes dominant, which negatively affects the system performance^[4].

This work was supported in part by ZTE Industry-University-Institute Cooperation Funds.

Recently, a novel user-centric network mechanism called cell-free (CF) MIMO was proposed to address the issue^[5-7]. In CF MIMO systems, geographically distributed access points (APs) cooperate to serve all user equipment (UE) without cell boundaries through coherent transmission, as shown in Fig. 1.



▲ Figure 1. Cell-free multiple-input multiple-output (MIMO) network

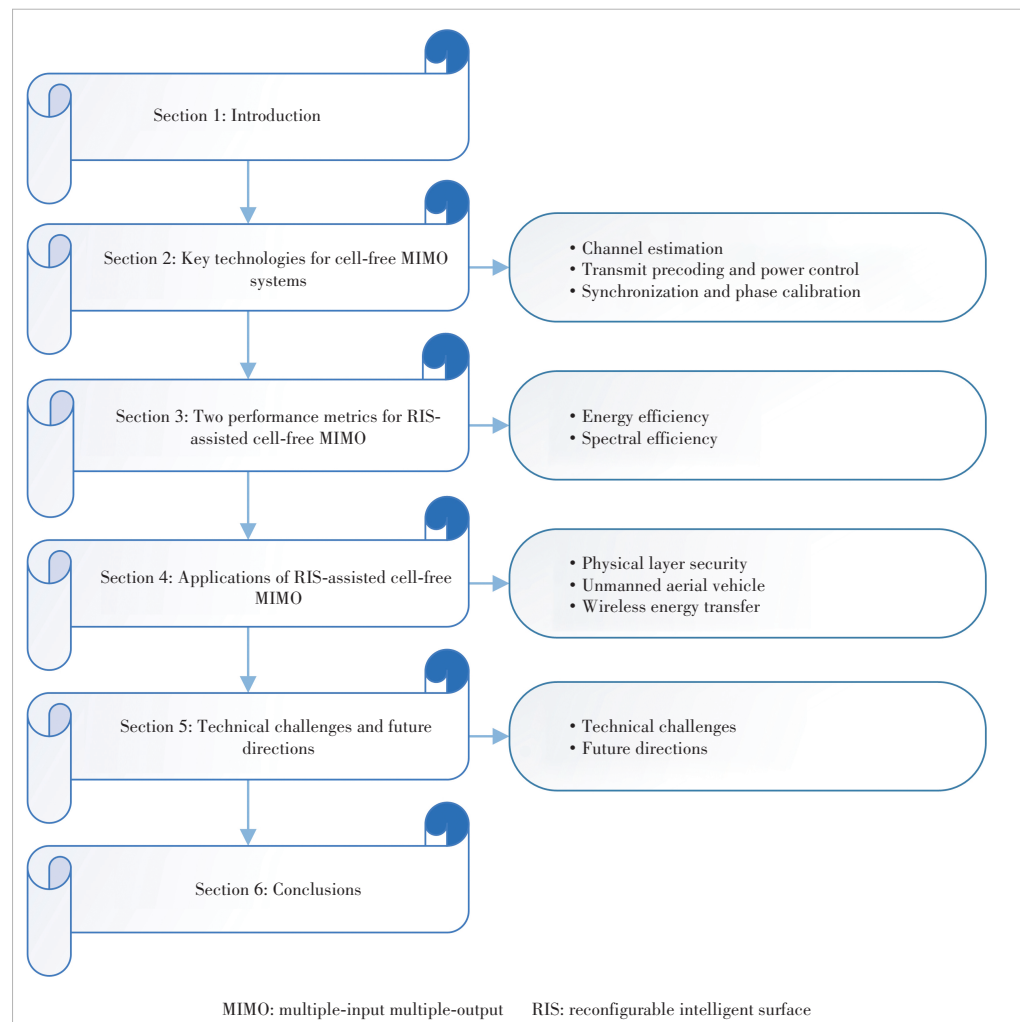
This is achieved under the control of a central processing unit (CPU) via high-speed fiber or wireless backhaul/fronthaul links. The biggest difference between CF MIMO and conventional massive MIMO systems is the absence of clear boundary demarcation between cells, which allows inter-cell interference to be effectively cancelled. According to Refs. [8 – 10], the CF MIMO system outperforms the traditional co-located antenna systems, such as massive MIMO, in terms of 95%-likely per-user spectral efficiency (SE). This is because the received signal at each UE suffers from disparate large-scale fading components from distinct APs, leading to high diversity gain. Besides, good quality of service (QoS) can be guaranteed to the cell-edge UE by deploying more APs in their vicinity, and thus the CF MIMO system often provides wider coverage than its counterpart. Due to these appealing advantages, the CF MIMO technique has been regarded as a promising candidate for future wireless networks^[11], and has attracted increasing attention for its wide application in channel estimation^[12], transmit beamforming^[13] and resource allocation^[14].

Nevertheless, to harvest the promising performance gains of the CF MIMO technique, more distributed APs are required, which brings about high costs and power consumption in practical implementations. This issue becomes even more important when considering green and sustainable communications. Among various candidates, the emerging meta-surface technique called reconfigurable intelligent surface (RIS) stands out with its unique capabilities of low cost, low energy consumption and programmability^[15–19]. Made up of a large number of passive components, RIS can boost communication by reprogramming incident signals and reflecting them in a certain direction under the control of an intelligent controller. In addition, RIS can be handily integrated into existent communication scenarios with low cabling costs and can be densely deployed regardless of interference management among multiple RISs. The above benefits have motivated extensive study interests on the various

applications of RISs in channel capacity improvement, coverage extension, and power saving.

Motivated by the above discussions, we try to provide a thorough survey on RIS-assisted CF MIMO for next-generation wireless communications and future network applications in this paper. Specifically, we focus on recent contributions to typical performance metrics such as energy efficiency and spectral efficiency to understand the role of RISs in CF-MIMO systems from the viewpoint of communication optimization. Furthermore, some novel emerging applications of the RIS-assisted CF MIMO technique within various future networks are investigated. Besides, some open problems and technical challenges are pointed out.

The organization of this survey is provided in Fig. 2, where key technologies of CF MIMO are first introduced in Section 2. The interplay between RISs and CF MIMO systems is then introduced in terms of main communication performance metrics and optimization frameworks in Section 3. Emerging applications of RIS-assisted CF MIMO in wireless networks are further discussed in Section 4. Finally, the technical challenges



▲ Figure 2. Structure of this paper

and future directions are pointed out in Section 5 before the concluding remarks in Section 6.

2 Key Technologies for Cell-Free MIMO Systems

In this section, we introduce the main technologies of CF MIMO for 5G and beyond in terms of two communication protocols, i.e., time division duplexing (TDD) and frequency division duplexing (FDD). More specifically, we emphasize the importance of channel estimation, transmit precoding and power control, synchronization, and phase calibration in CF MIMO systems. These techniques differ from those used in traditional massive MIMO systems when considering the distributed antenna configurations and practical fronthaul constraints. These differences play an important role in the system implementation and parameter design.

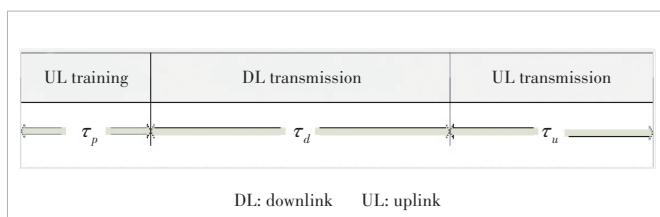
2.1 Channel Estimation

Channel estimation plays a vital role in transmit beamforming and signal detection, as it helps suppress inter-user interference and improve the network capacity. Based on different communication protocols, such as TDD and FDD, various channel estimation techniques have been applied to CF MIMO systems.

Under TDD protocols, existing works mainly focus on the pilot-based channel estimation technique. Specifically, users send orthogonal pilot sequences of length τ_p to all the APs as shown in Fig. 3, and the APs then estimate the corresponding user channel according to different criteria after receiving the training signals^[20]. Typical criteria include the least square (LS) method, the minimum mean square error (MMSE) method, the maximum likelihood (ML) method, and so on.

For FDD systems with channel non-reciprocity, the channel estimation problem needs more attention because the feedback overhead of channel state information (CSI) scales linearly with the number of RISs and APs. To solve this issue, the angle-reciprocity principle of the multipath components is utilized, which assumes that if the gap between uplink and downlink frequencies is less than several GHz, the uplink angle-of-arrival and downlink angle-of-departure are similar and the large-scale fading components remain unchanged^[21-23].

CF MIMO is essentially a single-cell massive MIMO sys-



▲ Figure 3. Pilot and data transmission under time division duplexing (TDD) communication protocol

tem, with antennas distributed over a wide geographic area^[24]. As a result, the joint channel from the AP to the user is strongly correlated in space, with some APs being closer to the user than others. Therefore, spatial correlation between different APs can be used for channel estimation in CF MIMO, reducing pilot and improving estimation efficiency when compared with traditional MIMO channel estimation.

2.2 Transmit Precoding and Power Control

In the transmission phase, the APs utilize the estimated channel coefficients in the channel estimation phase to perform data precoding and power control. In CF MIMO systems, distributed APs are connected to a CPU via the dedicated fronthaul link for payload data and power control coefficients exchange. For this reason, a low-complexity distributed precoding technique is suggested. The most commonly used precoding schemes are maximum-ratio transmit precoding^[5] and conjugate precoding^[13]. In both the schemes, the APs perform transmit precoding locally in a distributed manner, reducing the burden of CPU. However, since these distributed signal processing techniques only exploit their local CSI, they perform badly for anti-interference. To reduce the inter-user interference and further improve the system performance, some centralized precoding techniques, e.g., zero-forcing (ZF) precoding and MMSE precoding, are advocated at the cost of an increased fronthaul burden^[6,21]. This is because the prerequisite to implementing centralized precoding is the availability of all related channels at the CPU. In general, the optimal transmit precoding scheme should strike a balance between the fronthaul traffic and the network.

Different from the above precoding techniques in TDD systems, angle-based precoding schemes are more commonly used in FDD protocols. Specifically, these angle-based precoding techniques use the estimated multi-path components such as the large-scale fading factors and Angle-of-Arrival (AoA), along with the angle reciprocity principle, to determine the transmit vectors according to different precoding options, e.g., the angle-based MMSE precoding, the angle-based conjugate precoding, and the angle-based ZF precoding^[21,25].

Advanced power control is pivotal for enhancing the network capacity or QoS and simple equal power allocation may result in poor system performance. A simple but efficient way to perform power control is to allocate power to different users at each AP proportional to the corresponding channel gain, which has been used by many papers that aim at improving achievable data rates^[13,26]. On the other hand, to provide uniformly good service for different users, the max-min power control technique is suggested^[27-28]. By maximizing the minimum spectral efficiency of all users under the constraints of per-AP transmit power, all users have almost equal spectral efficiency with max-min power control. In FDD systems, the above power control techniques, namely equal power, water-filling power, and max-min power allocation techniques, can

also be utilized under angle-based precoding according to different requirements^[21].

While the above precoding techniques can achieve optimal antenna gain, they demand the same number of radio frequency (RF) chains as that of the transmit antennas. This causes high hardware cost and power consumption, especially in CF MIMO systems equipped with a vast number of APs. To this end, hybrid analog and digital precoding, for which multiple antennas are connected to the same RF chain, is a more attractive alternative^[29]. However, due to the unit-modulus constraint of the analog phase shifters, it is hard to obtain the optimal solution of the hybrid precoding. Up to now, the existing studies on the design of hybrid precoding schemes in CF MIMO systems are mainly divided into two kinds. One exploits the block coordinate descent (BCD) algorithm to alternatively optimize the analog and digital domain precodings^[30], and the other resorts to heuristic solutions, where analog precoding is designed by maximizing the received signal power and traditional digital precoding schemes such as ZF can be utilized based on the equivalent channel for multiuser interference cancellation. Besides, advanced machine learning-based solutions have also shown great potential in the design of hybrid precoding^[31].

2.3 Synchronization and Phase Calibration

In CF MIMO systems with massive geographically distributed APs and users, the signals from different APs generally experience different propagation delays, leading to timing synchronization among the received signals. To realize coherent transmission, the APs need to maintain relative signal time synchronization and phase calibration^[32]. Furthermore, the synchronization accuracy requirement of CF MIMO networks becomes more demanding as more users become time-sensitive. To address this issue, many time synchronization schemes are proposed, such as one-way message dissemination, two-way message exchange, and receiver-only synchronization. Generally, after the clock offset and clock skew are estimated, the above timing misalignment management schemes are adopted^[33]. Besides, to ensure accurate CSI at the APs, phase calibration should be performed on the uplink transmitted signals. To this end, a reference signal is needed to obtain the over-the-air reciprocity calibration^[34].

3 Two Performance Metrics For RIS-Assisted Cell-Free MIMO

With the help of RISs, CF MIMO can enable better communication to user terminals than typical massive MIMO and pure CF MIMO. To harvest the promised performance gains provided by the RIS-assisted CF MIMO technique, it is necessary to optimize the resource allocation of the distributed network. The optimization objective can be different QoS metrics based on the actual need. In this section, we investigate the interplay between the RIS-assisted MIMO technique and wire-

less resource allocation optimization in terms of various communication metrics.

3.1 Energy Efficiency

RISs are highly recommended in CF MIMO systems due to their low cost and energy consumption. Various studies have been devoted to deeply investigating how RIS acts as an energy-efficient technique in CF MIMO systems. By iteratively optimizing the digital beamforming at the APs and the RIS-based analog beamforming, the energy efficiency of the RIS-assisted CF MIMO system can beat that of the conventional distributed antenna system and the no-RIS case by orders of magnitude in Ref. [35]. Furthermore, it is observed that by replacing some of the APs with RISs, a two-fold performance gain in terms of energy efficiency can be harvested compared with the conventional CF MIMO scheme without RIS, which indicates the validity of RISs. Similarly, the authors in Ref. [36] proposed an iterative alternating algorithm based on a fractional programming method and successive lower-bound maximization to maximize the energy efficiency. The simulation also illustrates that with the default settings, the proposed design can offer an improvement of at least 168% higher energy efficiency than traditional user-centric CF MIMO systems without RIS assistance.

However, the above research has not considered practical constraints such as limited backhaul capacity. As a remedy, Ref. [37] devised an energy efficiency maximization design for RIS-aided CF MIMO networks under the constraints of per-AP transmit power, limited backhaul capacity, nonconvexity of RIS reflection, and minimum achievable rate requirement. As expected, the performance of the proposed RIS-aided CF MIMO system outperforms that of the co-located network with or without RIS.

3.2 Spectral Efficiency

As a counterpart of energy efficiency, spectral efficiency is also a main communication metric. Since RIS can construct a favorable propagation environment by artificially creating a virtual line-of-sight (LoS) link, the signals of users at a dead zone or cell edge can be improved, thus bringing about spectral efficiency enhancement to the CF MIMO system.

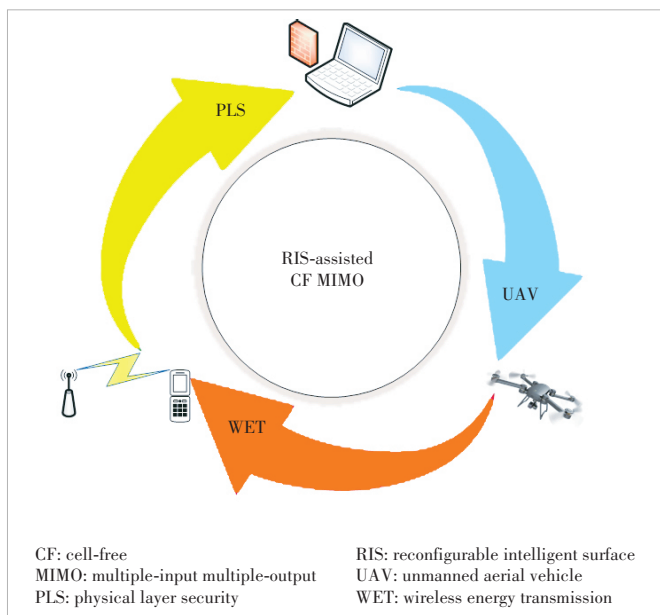
As the first attempt, the authors in Ref. [38] proposed the concept of RIS-assisted CF MIMO to improve the network capacity or the spectral efficiency of the distributed MIMO system. Specifically, the authors considered a general framework with multiple antennas, multiple APs, multiple users, multiple RISs and multiple carriers. At the beginning of a large coherent timescale, each user is matched with specified RISs by the proposed linear conic relaxation-based matching method. In the rest of the coherent timescale, only the channels of the matched user-RIS pairs are estimated and exploited for joint active and passive precoding design, and those of the unmatched RISs are modeled as random noise. This process is

repeated in the next coherent timescale until the end of data transmission. Once the RIS-UE association is fixed, the joint optimization problem is decoupled into two sub-problems through Lagrangian dual reformulation and multidimensional complex quadratic transform. The two decoupled sub-problems, i.e., the active transmit precoding at the APs and the passive reflection precoding at the RISs, are converted to two quadratically constrained quadratic program problems again, which can be solved by existing optimization tools. By iteratively optimizing the two sub-problems alternatively until converge, a sub-optimal solution to the weighted spectral efficiency of the system is obtained.

The authors in Ref. [39] also considered the burden on the fronthaul network, and proposed a decentralized cooperative joint active and passive precoding framework for RIS-assisted CF MIMO systems based on the alternating direction method of multipliers to maximize the weighted downlink spectral efficiency. Since only a few variables are incrementally updated and transmitted to the next AP, the backhaul overhead can be significantly reduced compared to the case with the whole CSI exchange among APs.

4 Applications of RIS-Assisted Cell-Free MIMO

This section presents various novel applications of the RIS-assisted CF MIMO technique, e.g., physical layer security (PLS), unmanned aerial vehicle (UAV) networks, and wireless energy transmission (WET), as depicted in Fig. 4. In fact, integrating RIS and CF MIMO with these emerging technologies can offer opportunities for beyond 5G networks. In particular, we provide an overview of state-of-the-art contributions on integrating each technology with RIS-assisted CF MIMO to pin-



▲ Figure 4. Applications of RIS-assisted CF MIMO

point the main techniques that can be adopted for developing future wireless network services and novel applications in the following section.

4.1 Physical Layer Security

With the development of wireless communication and the internet of everything (IoE), handheld devices such as mobile phones have carried more and more important/private information such as identity information, credit card passwords, ehealth data, and chat history. However, these devices are more vulnerable to security threats or attacks due to their limited memory capacity, battery life, computational power, and network bandwidth. Hence, physical layer security has become one of the top concerns of future wireless communication systems, especially for CF MIMO systems.

RIS-assisted secure issues have gained great attention recently and a large number of works have shown that by deploying RISs in the vicinity of the user and the eavesdropper, the physical layer security of the CF MIMO systems can be improved. Specifically, in Ref. [38], the maximum weight sum secrecy rate problem is considered by jointly optimizing the active cooperating precodings at multiple APs and the passive cooperating precodings at multiple RISs. The non-convex problem is decomposed into two separate sub-problems, and each sub-problem is solved by semi-definite relaxation and successive convex approximation techniques. With the aim of reducing channel estimation and feedback overhead, the authors further proposed a two-timescale transmission, where only partial CSI of the corresponding RISs that are selected by the assigning scheme is needed in each small coherence time, and the whole CSI is needed only at the beginning of the large coherence time.

Similarly, the work in Ref. [40] also studied the role of RISs in CF MIMO secure issues in the presence of active eavesdroppers. Channel estimation errors caused by pilot contamination, spoofing attacks and reuse are considered. To manage the inter-user interference, ZF precoding is applied to the CPU. Then by iteratively optimizing the RIS phase shifts and the downlink power allocation, the information leakage to active eavesdroppers is minimized while the legitimate user's performance is guaranteed.

4.2 Unmanned Aerial Vehicle

Due to their immense potential in logistics transportation, agricultural cultivation, delivery of medical supplies and telecommunications, UAVs have been recommended as a key enabling technology in both commercial and civilian fields. Specifically, a UAV acts as a flying AP that can dynamically change its location to enhance the network coverage and thus improve the network capacity. Since there are fewer obstacles in the airspace, it is easier to establish an air LoS link by deploying UAVs rather than a ground LoS link by deploying APs.

In traditional CF MIMO systems, a UAV usually serves as

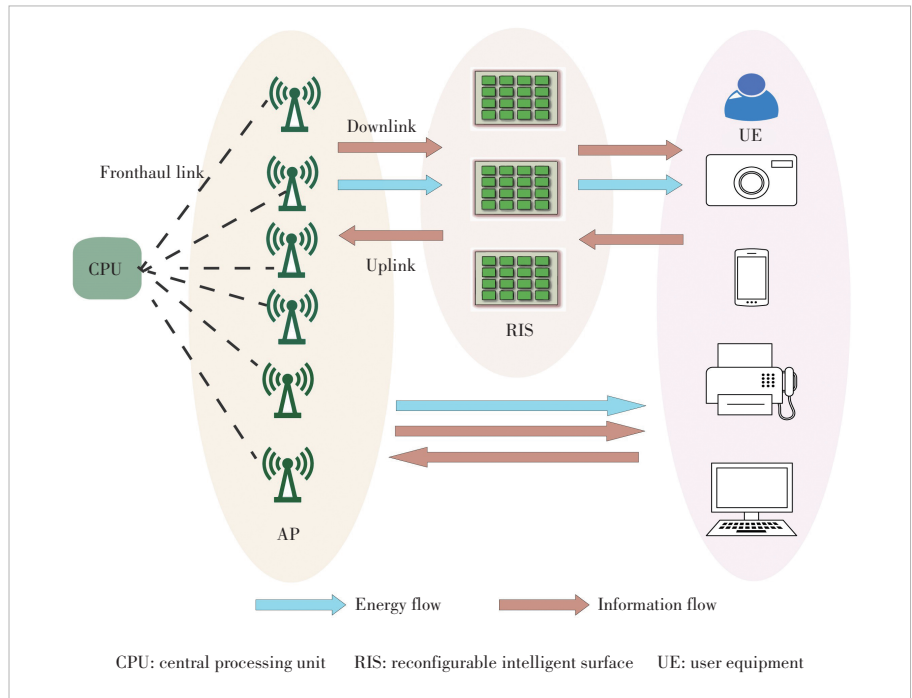
an active relay to assist the transmission from APs to users. However, since the AP antennas are generally tilted downwards to serve the ground users, communication with UAV usually suffers from poor signal strength. One way to deal with this problem is by tilting the AP antenna upwards or allocating more transmit power to the UAV. However, this would weaken the communication with ground users. Motivated by the emerging RIS technology, some researchers proposed to utilize RIS to strengthen the UAV communication while maintaining or even improving the downlink rate at the ground users^[41]. By judiciously designing the reflection coefficients of the RIS and the power allocation, the achievable rate of both the UAV and the ground users can be improved without changing the AP antenna tilt.

4.3 Wireless Energy Transfer

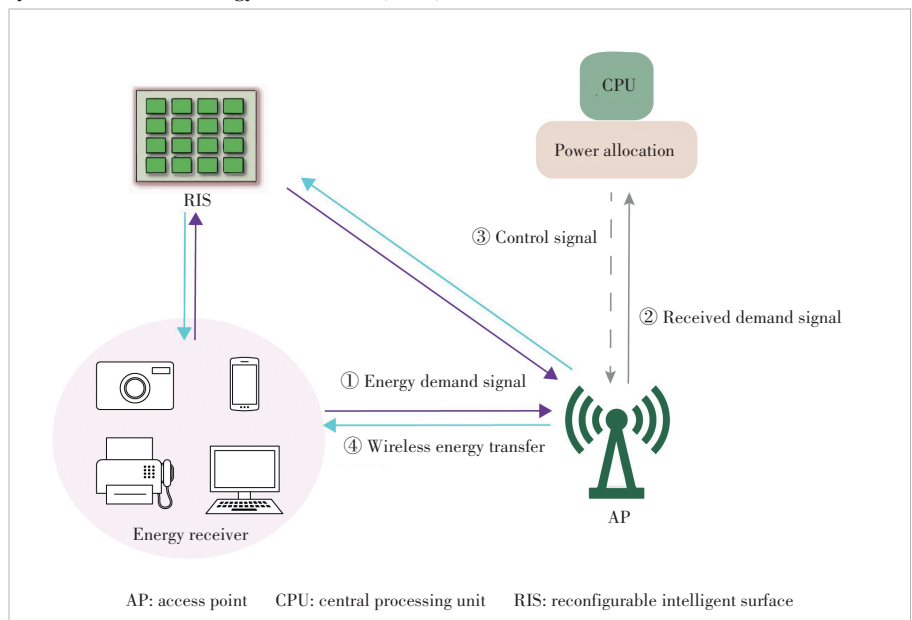
In practice, most IoE devices will be battery- or energy-constrained due to their limited sizes. Although frequent battery replacement and replenishing may offer a temporary solution to this problem, in future communication systems with multitudinous ubiquitous mobile access devices such as CF MIMO, this comes with dramatically high labor and material costs. As such, advanced energy supplement techniques are urgently needed to improve the energy shortage challenges.

Among various candidates, wireless energy transfer has been regarded as a promising technology to address these critical energy replenishing problems. Specifically, the radio frequency (RF) energy signal radiated by the AP can be acquired at the energy receiver by utilizing the far-field radiation properties of electromagnetic waves. After that, the AP transforms the received energy signal into energy for potential use. Unfortunately, despite the appealing merits of WET, it still faces some practical problems such as severe loss attenuation and signal obstruction. As a remedy, RIS has exactly the ability to smartly enhance the channel conditions via software-controlled signal reflection. Figs. 5 and 6 illustrate the specific transmission flow and the system architecture of the RIS-assisted CF MIMO system with

WET respectively. As shown in Fig. 5, for the proposed RIS-assisted CF MIMO systems, multiple cooperative APs transmit both information signals and energy-carrying signals to multiple energy receivers and information receivers, where RISs are deployed to enhance the corresponding communication links. In Fig. 6, the system architecture can be divided into four parts. In the first part, the energy receivers transmit the energy demand signal to the APs through direct and RIS-



▲ Figure 5. Transmission flow of RIS-assisted cell-free (CF) multiple-input multiple-output (MIMO) system with wireless energy transmission (WET)



▲ Figure 6. Architecture of RIS-assisted cell-free multiple-input multiple-output (CF MIMO) system with wireless energy transmission (WET)

assisted cascaded channels. Second, the APs transmit the demand signal received to the CPU. Then, the CPU allocates power and transmits a control signal to APs based on the signal transmitted by APs. Finally, APs transmit wireless energy signals to the energy receiver via downlink.

Recently, some works have considered the WET in RIS-assisted CF MIMO systems and analyzed its opportunities and challenges. For example, the authors of Ref. [42] answered the question of how the WET technology could be integrated into RIS-assisted CF MIMO systems from the main application scenarios, the four-stage transmission procedure, the specific deployment and hardware design, and the operation modes. The weighted sum rate maximization problem for RIS-assisted CF MIMO systems was studied in Ref. [43], under the total power constraints, the energy harvesting constraints and the unit-modulus constraints of the RISs. In Ref. [44], the authors proposed to substitute APs with UAVs to construct a CF MIMO-mount UAV assisted by RISs for RF energy transfer, which combines the benefits of CF MIMO, RISs and UAV. This framework where the distributed UAVs act as APs and operate in a cell-free fashion can provide a strong LoS signal and thus harvest more energy than other benchmarks.

5 Technical Challenges and Future Directions

Although several aspects of the RIS-assisted CF MIMO technique have been discussed in this paper, there are still some remaining challenges and unidentified areas in this field. In the following, we will point out some of the major challenges and future directions for implementing this technology in future wireless networks to realize its full potential.

5.1 Technical Challenges

To achieve the promised performance gains brought by the RIS-assisted CF MIMO technique, exact channel estimation is indispensable, which is rather difficult due to the following reasons. First, it is challenging to achieve perfect channel estimation for RISs, if not impossible. Due to their passive nature, RISs do not have any capability of signal processing, which makes it an arduous task to obtain individual channels of each hop of the links as in traditional relay systems. Second, RISs are usually composed of hundreds or even thousands of reflecting elements. Consequently, the channel coefficients that need to be estimated are tremendous, leading to an unaffordable overhead.

Due to hardware limitations of RIS elements, achieving optimal phase shift on all frequencies simultaneously in RIS-assisted MIMO systems is impractical, resulting in the so-called beam squint effect^[45]. Specifically, the application of a large system bandwidth and intelligent reflecting surface causes a non-negligible propagation delay of electromagnetic waves and thus leads to a frequency-dependent antenna ar-

ray response, i.e., beam pointing direction varies with frequency. However, ignoring the beam squint effects can lead to severe beam gain loss and seriously degrade system performance, especially for RIS-assisted wideband mmWave systems. Therefore, the impact of beam squint needs to be taken into consideration, and effective measures should be implemented to mitigate its impact in RIS-assisted systems.

Another critical technical challenge of the RIS-assisted CF MIMO system is low-complexity signal processing. On the one hand, to provide better QoS for future networks with massive access devices, it is necessary to deploy more APs and RISs with numerous antennas and reflecting elements. On the other hand, the problem of jointly optimizing the active precodings at APs and passive precodings at RISs is always non-convex due to the unit-modulus constraints of RISs, which always comes with a vast number of time-consuming iterations, leading to high calculation complexity and huge time latency. Meanwhile, due to the coupling between the optimization variables, the problem is hard to solve^[38]. While many existing researchers have focused on and proposed various sub-optimal solutions to this problem, considering practical system implementations such as channel estimation and hardware impairments may make it even more sophisticated to solve^[37]. Thus, there is an urgent need for more low-complexity signal processing techniques.

5.2 Future Directions

Learning-based techniques have attracted increasing attention from both academia and industry due to their ability to process a vast number of data in emerging wireless applications^[46]. However, only a few studies have investigated the learning-enabled RIS-assisted CF MIMO networks. In fact, learning can play an essential role in channel estimation and power allocation for RIS-aided systems and CF MIMO systems. For example, some results have shown that the neural network-based estimation method can outperform the traditional estimation methods even with a few pilots, and the deep learning-based resource allocation can approximate the globally optimal performance for max-min power control. As such, more in-depth studies should be conducted in learning-enabled RIS systems and CF MIMO systems to fully exploit the potential of these techniques.

Near-field communication is an indispensable part of future 6G networks, which will bring several challenges to wireless communications. The low cost and low power consumption enable RIS to be equipped with a large number of reflective elements^[47]. Meanwhile, increasing the number of antennas in the BS is also an important trend for future communication systems. However, large-scale RIS panels and large antenna arrays in the BS can lead to the near-field effect, especially in RIS-assisted CF MIMO wireless communication systems. Most traditional techniques designed for the far field suffer severe performance losses in the near-field re-

gion, such as channel estimation and beamforming^[48]. Thus, it is practically important to efficiently adapt these traditional techniques to the near-field domain in IRS-aided CF MIMO systems and explore the possibilities of improving capacity by exploiting the near-field spherical wavefront.

Another emerging feature of 6G networks is the integration of sensing and communications, which enables the exploitation of dense sensing objects, computing resources and complicated communication resources to construct a perceptive network^[49]. In CF MIMO systems, distributed array antennas can be used to significantly increase the effective array aperture, thus improving the sensing accuracy. Moreover, the position, speed and other information obtained from integrated sensing and communication can be utilized to optimize transmission performance and simplify computation complexity in CF MIMO systems. Therefore, the integrated design of communication and sensing based on CF MIMO systems is a potential technology for the exponentially growing demands and application requirements of 6G.

Most existing research has focused on CF MIMO systems with low-mobility users moving at a velocity of less than 10 km/h. However, in practical implementations with a large number of distributed users, there exist some users with high movement rates, although few, but vital. This is because in high mobility scenarios, the combination effect of Doppler frequency offset, timing offset and phase noise may trigger the inter-carrier interference and therefore deteriorate the system performance. Therefore, it is crucial to consider transmit precoding and power allocation algorithms under high vehicular speeds.

Besides, hardware constraints and imperfections are also an interesting topic that needs to be investigated. In the initial research on RIS-assisted CF MIMO wireless communication systems, most works have mainly assumed the ideal RIS reflection model for the ease of communication optimization design and performance analysis. However, in practice, RISs always suffer from hardware imperfections/impairments such as limited precision of reflection and phase-dependent amplitude that could seriously decrease the system performance. Hence, it is necessary to build a hardware-constrained system model that can accurately capture the hardware imperfections of RISs to cater to practical use.

The association between APs, users and RISs is also a difficult task that remains to be solved. Prior research has been conducted in a scenario where all APs jointly transmit their signals to all users through all RISs. This is however both power- and performance-inefficient since only the users near APs and RISs can benefit from the pre-designed transmit precoding. A more practical and efficient way is to dynamically allocate a group of RISs and APs that have the best channel conditions to serve each individual user, and this kind of matching schemes has been verified to have better performance compared with the all-AP scheme. This AP-RIS-UE

matching approach is still in its fancy and needs to be investigated especially when considering passive RISs.

Finally, the space-air-ground-sea integrated networks (SAGSIN) is a network architecture for 6G, which has been widely envisioned as a promising solution to complete, multi-angle and high-speed communication coverage worldwide^[50]. SAGSIN is divided into four segments: the space network, air network, ground network, and sea network. It is a comprehensive integration of systems, technologies and applications, rather than just connecting different communication networks. Therefore, the RIS-assisted CF MIMO network is well suited for SAGSIN because of its significant network coverage capabilities and extraordinarily high data transfer speeds. However, given the heterogeneity of multi-layer networks among SAGSIN, it is significant to thoroughly investigate how to integrate the RIS-assisted CF MIMO network with other networks to satisfy user QoS and achieve complementary benefits among various networks.

6 Conclusions

In this paper, we provide a comprehensive survey of RIS-assisted CF MIMO system by reviewing the recent achievements in this area. In particular, we point out the key performance metrics, such as spectral efficiency and energy efficiency, and introduce the corresponding communication optimization techniques to achieve these metrics. Moreover, we discuss the potential applications of RIS-assisted CF MIMO in novel scenarios in B5G systems such as PLS, UAV and WET. Finally, the open problems and potential directions are studied to promote in-depth investigations and developments of RIS-assisted CF MIMO in future wireless networks.

References

- [1] Samsung. The vision of 6G: bring the next hyper-connected experience to every corner of life [EB/OL]. [2023-03-22]. https://cdn.codeground.org/nsr/downloads/researchareas/20201201_6G_Vision_web.pdf
- [2] WONG V W S, SCHÖBER R, NG D W K, et al. Key technologies for 5G wireless systems [M]. Cambridge, UK: Cambridge University Press, 2017
- [3] ANDREWS J G, BUZZI S, CHOI W, et al. What will 5G be? [J]. IEEE journal of selected areas in communications, 2014, 32(6): 1065 - 1082. DOI: 10.1109/JSAC.2014.2328098
- [4] LOZANO A, HEATH R W, ANDREWS J G. Fundamental limits of cooperation [J]. IEEE transactions on information theory, 2013, 59(9): 5213 - 5226. DOI: 10.1109/TIT.2013.2253153
- [5] NGO H Q, ASHIKHMIN A, YANG H, et al. Cell-free massive MIMO versus small cells [J]. IEEE transactions on wireless communications, 2017, 16(3): 1834 - 1850. DOI: 10.1109/TWC.2017.2655515
- [6] NAYEBI E, ASHIKHMIN A, MARZETTA T L, et al. Precoding and power optimization in cell-free massive MIMO systems [J]. IEEE transactions on wireless communications, 2017, 16(7): 4445 - 4459. DOI: 10.1109/TWC.2017.2698449
- [7] NAYEBI E, ASHIKHMIN A, MARZETTA T L, et al. Cell-free massive MIMO systems [C]//Proc. 49th Asilomar Conference on Signals, Systems

- and Computers. IEEE, 2015: 695 – 699. DOI: 10.1109/ACSSC.2015.7421222
- [8] BJÖRNSON E, SANGUINETTI L. Making cell-free massive MIMO competitive with MMSE processing and centralized implementation [J]. IEEE transactions on wireless communications, 2020, 19(1): 77 – 90. DOI: 10.1109/TWC.2019.2941478
- [9] LIU P, LUO K, CHEN D, et al. Spectral efficiency analysis of cell-free massive MIMO systems with zero-forcing detector [J]. IEEE transactions on wireless communications, 2020, 19(2): 795 – 807. DOI: 10.1109/TWC.2019.2948841
- [10] PAPAZAFEIROPOULOS A, KOURTESSIS P, DI RENZO M, et al. Performance analysis of cell-free massive MIMO systems: a stochastic geometry approach [J]. IEEE transactions on vehicular technology, 2020, 69(4): 3523 – 3537. DOI: 10.1109/TVT.2020.2970018
- [11] ELHOUSHY S, IBRAHIM M, HAMOUDA W. Cell-free massive MIMO: a survey [J]. IEEE communications surveys & tutorials, 2022, 24(1): 492 – 523. DOI: 10.1109/COMST.2021.3123267
- [12] JIN Y, ZHANG J Y, JIN S, et al. Channel estimation for cell-free mmwave massive MIMO through deep learning [J]. IEEE transactions on vehicular technology, 2019, 68(10): 10325 – 10329. DOI: 10.1109/TVT.2019.2937543
- [13] ATTARIFAR M, ABBASFAR A, LOZANO A. Modified conjugate beamforming for cell-free massive MIMO [J]. IEEE wireless communications letters, 2019, 8(2): 616 – 619. DOI: 10.1109/LWC.2018.2890470
- [14] MOSLEH S, ALMOSA H, PERRINS E, et al. Downlink resource allocation in cell-free massive MIMO systems [C]/Proc. International Conference on Computing, Networking and Communications (ICNC). IEEE, 2019: 883 – 887. DOI: 10.1109/ICCNC.2019.8685542
- [15] ZHAO Y Q, XU W, SUN H, et al. Cooperative reflection design with timing offsets in distributed multi-RIS communications [J]. IEEE wireless communications letters, 2021, 10(11): 2379 – 2383. DOI: 10.1109/LWC.2021.3100989
- [16] ZHAO Y Q, XU W, YOU X H, et al. Cooperative reflection and synchronization design for distributed multiple-RIS communications [J]. IEEE journal of selected topics in signal processing, 2022, 16(5): 980 – 994. DOI: 10.1109/JSTSP.2022.3173749
- [17] ZHAO Y Q, XU W, HUANG Y M, et al. Joint transceiver and passive beamforming optimization for RIS-assisted MIMO systems [C]/Proc. 13th International Conference on Wireless Communications and Signal Processing (WCSP). IEEE, 2021: 1 – 5. DOI: 10.1109/WCSP52459.2021.9613174
- [18] ZHU F S, ZHAO Y Q, XU W, et al. CSIT-free model aggregation for multi-RIS-assisted over-the-air computation [C]/Proc. International Symposium on Wireless Communication Systems (ISWCS). IEEE, 2022: 1 – 5. DOI: 10.1109/ISWCS56560.2022.9940351
- [19] SHI W, XU W, YOU X H, et al. Intelligent reflection enabling technologies for integrated and green internet-of-everything beyond 5G: communication, sensing, and security [J]. IEEE wireless communications, 2023, 30(2): 147 – 154. DOI: 10.1109/MWC.018.2100717
- [20] BUZZI S, D'ANDREA C, ZAPPONE A, et al. User-centric 5G cellular networks: resource allocation and comparison with the cell-free massive MIMO approach [J]. IEEE transactions on wireless communications, 2020, 19(2): 1250 – 1264. DOI: 10.1109/TWC.2019.2952117
- [21] ABDALLAH A, MANSOUR M M. Efficient angle-domain processing for FDD-based cell-free massive MIMO systems [J]. IEEE transactions on communications, 2020, 68(4): 2188 – 2203. DOI: 10.1109/TCOMM.2020.2969351
- [22] ABDALLAH A, MANSOUR M M. Angle-based multipath estimation and beamforming for FDD cell-free massive MIMO [C]/Proc. IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2019: 1 – 5. DOI: 10.1109/SPAWC.2019.8815420
- [23] KIM S, SHIM B. FDD-based cell-free massive MIMO systems [C]/Proc. IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2018: 1 – 5. DOI: 10.1109/SPAWC.2018.8446023
- [24] BJÖRNSON E, HOYDIS J, SANGUINETTI L. Massive MIMO networks: spectral, energy, and hardware efficiency [J]. Foundations and trends in signal processing, 2017, 11(3/4): 154 – 655. DOI: 10.1561/2000000093
- [25] KIM S, SHIM B. AoD-based statistical beamforming for cell-free massive MIMO systems [C]/Proc. IEEE 88th Vehicular Technology Conference (VTC-Fall). IEEE, 2018: 1 – 5. DOI: 10.1109/VTCFall.2018.8690646
- [26] BJÖRNSON E, SANGUINETTI L. Scalable cell-free massive MIMO systems [J]. IEEE transactions on communications, 2020, 68(7): 4247 – 4261. DOI: 10.1109/TCOMM.2020.2987311
- [27] INTERDONATO G, KARLSSON M, BJÖRNSON E, et al. Local partial zero-forcing precoding for cell-free massive MIMO [J]. IEEE transactions on wireless communications, 2020, 19(7): 4758 – 4774. DOI: 10.1109/TWC.2020.2987027
- [28] CHAKRABORTY S, BJÖRNSON E, SANGUINETTI L. Centralized and distributed power allocation for max-min fairness in cell-free massive MIMO [C]/Proc. 53rd Asilomar Conference on Signals, Systems, and Computers. IEEE, 2019: 576 – 580. DOI: 10.1109/IEEECONF44664.2019.9048903
- [29] KASAM J, CASTANHEIRA D, SILVA A, et al. Distributed hybrid equalization for cooperative millimeter-wave cell-free massive MIMO [J]. IEEE transactions on communications, 2022, 70(8): 5300 – 5316. DOI: 10.1109/TCOMM.2022.3186980
- [30] FENG C H, SHEN W Q, AN J P, et al. Weighted sum rate maximization of the mmwave cell-free MIMO downlink relying on hybrid precoding [J]. IEEE transactions on wireless communications, 2022, 21(4): 2547 – 2560. DOI: 10.1109/twc.2021.3113458
- [31] YETIS C M, BJÖRNSON E, GISELSSON P. Joint analog beam selection and digital beamforming in millimeter wave cell-free massive MIMO systems [J]. IEEE open journal of the communications society, 2021, 2: 1647 – 1662. DOI: 10.1109/OJCOMS.2021.3094823
- [32] ROGALIN R, BURSALIOGLU O Y, PAPAPOPOULOS H, et al. Scalable synchronization and reciprocity calibration for distributed multiuser MIMO [J]. IEEE transactions on wireless communications, 2014, 13(4): 1815 – 1831. DOI: 10.1109/twc.2014.030314.130474
- [33] LI B Q, ZHU X, JIANG Y F, et al. Cooperative time synchronization and parameter estimation via broadcasting for cell-free massive MIMO networks [C]/Proc. IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2022: 2100 – 2105. DOI: 10.1109/WCNC51071.2022.9771548
- [34] CAO Y, WANG P, ZHENG K, et al. Experimental performance evaluation of cell-free massive MIMO systems using COTS RRU with OTA reciprocity calibration and phase synchronization [J]. IEEE journal on selected areas in communications, 2023, 41(6): 1620 – 1634. DOI: 10.1109/JSAC.2023.3276057
- [35] ZHANG Y T, DI B Y, ZHANG H L, et al. Beyond cell-free MIMO: energy efficient reconfigurable intelligent surface aided cell-free MIMO communications [J]. IEEE transactions on cognitive communications and networking, 2021, 7(2): 412 – 426. DOI: 10.1109/TCCN.2021.3058683
- [36] LU H C, ZHAO D, WANG Y Z, et al. Joint power control and passive beamforming in reconfigurable intelligent surface assisted user-centric networks [J]. IEEE transactions on communications, 2022, 70(7): 4852 – 4866. DOI: 10.1109/TCOMM.2022.3174071
- [37] LE Q N, NGUYEN V D, DOBRE O A, et al. Energy efficiency maximization in RIS-aided cell-free network with limited backhaul [J]. IEEE communications letters, 2021, 25(6): 1974 – 1978. DOI: 10.1109/LCOMM.2021.3062275
- [38] ZHANG Z J, DAI L L. A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network [J]. IEEE transactions on signal processing, 2021, 69: 4085 – 4101. DOI: 10.1109/TSP.2021.3088755
- [39] HUANG S C, YE Y, XIAO M, et al. Decentralized beamforming design for intelligent reflecting surface-enhanced cell-free networks [J]. IEEE

- wireless communications letters, 2021, 10(3): 673 – 677. DOI: 10.1109/LWC.2020.3045884
- [40] ELHOUSHY S, IBRAHIM M, HAMOUDA W. Exploiting RIS for limiting information leakage to active eavesdropper in cell-free massive MIMO [J]. *IEEE wireless communications letters*, 2022, 11(3): 443 – 447. DOI: 10.1109/LWC.2021.3130169
- [41] AL-NAHHAS B, CHAABAN A, HOSSAIN M J. Improving UAV communication in cell free MIMO using a reconfigurable intelligent surface [C]// *Proc. IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2022: 1152 – 1157. DOI: 10.1109/GCWkshps56602.2022.10008748
- [42] SHI E Y, ZHANG J Y, CHEN S F, et al. Wireless energy transfer in RIS-aided cell-free massive MIMO systems: opportunities and challenges [J]. *IEEE communications magazine*, 2022, 60(3): 26 – 32. DOI: 10.1109/MCOM.001.2100671
- [43] YANG Z Y, ZHANG Y. Beamforming optimization for RIS-aided SWIPT in cell-free MIMO networks [J]. *China communications*, 2021, 18(9): 175 – 191. DOI: 10.23919/JCC.2021.09.014
- [44] KHALIL A A, SELIM M Y, RAHMAN M A. CURE: enabling RF energy harvesting using cell-free massive MIMO UAVs assisted by RIS [C]// *Proc. IEEE 46th Conference on Local Computer Networks (LCN)*. IEEE, 2021: 533 – 540. DOI: 10.1109/LCN52139.2021.9524984
- [45] SUN H R, ZHANG S, MA J P, et al. Time-delay unit based beam squint mitigation for RIS-aided communications [J]. *IEEE communications letters*, 2022, 26(9): 2220 – 2224. DOI: 10.1109/LCOMM.2022.3184074
- [46] XU W, YANG Z, NG D W K, et al. Edge learning for B5G networks with distributed signal processing: semantic communication, edge computing, and wireless sensing [J]. *IEEE journal of selected topics in signal processing*, 2023, 17(1): 9 – 39. DOI: 10.1109/JSTSP.2023.3239189
- [47] PAN Y J, PAN C H, JIN S, et al. Joint channel estimation and localization in the near field of RIS enabled mmwave/subTHz communications [EB/OL]. (2022-08-24)[2023-03-30]. <https://arxiv.org/abs/2208.11343>
- [48] CUI M Y, WU Z D, LU Y, et al. Near-field MIMO communications for 6G: fundamentals, challenges, potentials, and future directions [J]. *IEEE communications magazine*, 2023, 61(1): 40 – 46. DOI: 10.1109/MCOM.004.2200136
- [49] LIU F, CUI Y H, MASOUIROS C, et al. Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond [J]. *IEEE journal on selected areas in communications*, 2022, 40(6): 1728 – 1767. DOI: 10.1109/JSAC.2022.3156632
- [50] GUO H Z, LI J Y, LIU J J, et al. A survey on space-air-ground-sea integrated network security in 6G [J]. *IEEE communications surveys & tutorials*, 2022, 24(1): 53 – 87. DOI: 10.1109/COMST.2021.3131332

Biographies

ZHAO Yaqiong received her BE degree in electrical engineering from Southeast University, China in 2018. She is currently working toward a PhD degree with the School of Information Science and Engineering, National Mobile Communications Research Laboratory, Southeast University. Her current research interests include massive MIMO, mmWave communications, and reconfigurable intelligent surfaces for wireless communications.

KE Hongqin received his BE degree in electrical engineering from Xiamen University, China in 2022. He is currently working toward an MS degree with the School of Information Science and Engineering, National Mobile Communications Research Laboratory, Southeast University, China. His current research interests include massive MIMO, integrated sensing and communication, and reconfigurable intelligent surfaces for wireless communications.

XU Wei (wxu@seu.edu.cn) received his BSc degree in electrical engineering and MS and PhD degrees in communication and information engineering from Southeast University, China in 2003, 2006 and 2009, respectively. Between 2009 and 2010, he was a post-doctoral research fellow with the Department of Electrical and Computer Engineering, University of Victoria, Canada. He is currently a professor at the National Mobile Communications Research Laboratory, Southeast University. He was an adjunct professor of the University of Victoria, Canada from 2017 to 2020, and a Distinguished Visiting Fellow of the Royal Academy of Engineering, UK in 2019. He is also with the Purple Mountain Laboratory and has served as a project leader since 2020. He has co-authored over 100 refereed journal papers in addition to 36 domestic patents and four US patents granted. His research interests include information theory, signal processing and machine learning for wireless communications. He was an editor of *IEEE Communications Letters* from 2012 to 2017. He is currently an editor of *IEEE Transactions on Communications* and a senior editor of *IEEE Communications Letters*. He received the Best Paper Awards from a number of prestigious IEEE conferences including IEEE Globecom/ICCC, etc. He received the Science and Technology Award for Young Scholars of the Chinese Institute of Electronics in 2018.

YE Xinquan received his MS degree from Xidian University, China in 2018. He is currently an engineer with ZTE Corporation. His research interests include cell-free massive MIMO and reconfigurable intelligent surface.

CHEN Yijian received his BS degree from Central South University, China in 2006. He is currently a senior engineer with ZTE Corporation. His current research interests include massive MIMO, coordinated multi-point transmission, high-frequency communications, and channel modeling.



Research on Multi-Core Processor Analysis for WCET Estimation

LUO Haoran¹, HU Shuisong¹, WANG Wenyong¹,
TANG Yuke², ZHOU Junwei²

(1. University of Electronic Science and Technology of China, Chengdu
611731, China;
2. ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTECOM.202401010

<https://kns.cnki.net/kcms/detail/34.1294.TN.20240226.1212.002>, published
online February 26, 2024

Manuscript received: 2023-06-21

Abstract: Real-time system timing analysis is crucial for estimating the worst-case execution time (WCET) of a program. To achieve this, static or dynamic analysis methods are used, along with targeted modeling of the actual hardware system. This literature review focuses on calculating WCET for multi-core processors, providing a survey of traditional methods used for static and dynamic analysis and highlighting the major challenges that arise from different program execution scenarios on multi-core platforms. This paper outlines the strengths and weaknesses of current methodologies and offers insights into prospective areas of research on multi-core analysis. By presenting a comprehensive analysis of the current state of research on multi-core processor analysis for WCET estimation, this review aims to serve as a valuable resource for researchers and practitioners in the field.

Keywords: real-time system; worst-case execution time (WCET); multi-core analysis

Citation (Format 1): LUO H R, HU S S, WANG W Y, et al. Research on multi-core processor analysis for WCET estimation [J]. *ZTE Communications*, 2024, 22(1): 87 - 94. DOI: 10.12142/ZTECOM.202401010

Citation (Format 2): H. R. Luo, S. S. Hu, W. Y. Wang, et al., "Research on multi-core processor analysis for WCET estimation," *ZTE Communications*, vol. 22, no. 1, pp. 87 - 94, Mar. 2024. doi: 10.12142/ZTECOM.202401010.

1 Introduction

For intelligent devices with real-time embedded systems, the safety-critical nature of tasks is of paramount importance. Programs are required to accomplish tasks within stringent time constraints, which not only imposes significant demands on the accuracy of program logic, but also has a consequential impact on the safety-critical performance of the system. For instance, automotive control systems must carry out various tasks within specified time limits, such as engine control, braking systems, and in-car entertainment. These systems must respond quickly to driver inputs and adapt accordingly to road and traffic conditions. Any untimely response could potentially lead to serious traffic accidents.

In the design of embedded systems, predictability and reliability are crucial factors that must be taken into account. Specifically, the predictability of timing behavior is of utmost importance. To address this issue, the concept of the worst-case execution time (WCET) is a key point.

WCET refers to the maximum time a program takes to execute under the most unfavorable conditions in real-time systems. Accurately estimating WCET of a task is critical to guarantee that the task is completed within its deadline without in-

curred any delays. This estimation can be utilized to optimize resource allocation and scheduling strategies in the system, thereby making sure that tasks are executed within acceptable time intervals.

WCET multi-core analysis can help designers evaluate the performance of applications in multi-core processor systems and provide guidance for scheduling algorithms and resource allocation. Through WCET multi-core analysis, application performance in multi-core processor systems can be optimized, response time can be reduced, and system reliability and efficiency can be improved.

However, due to the increasing application demands and the increasingly complex design of systems, higher performance requirements are being placed on system performance, leading to more and more desktop processor features being deployed on the real-time systems. The diversity of complex processor features has made the execution time of instructions in the processor uncertain, which has led to uncertainty in WCET analysis.

Therefore, the results of WCET analysis require both safety and accuracy. The actual execution time of the program should not exceed the estimated time, and the analysis results should be as close as possible to the actual maximum execution time of the program, which is referred to as the true value. The true value is theoretically existent, and under the current research, the requirements for WCET also differ based on the

This work was supported by ZTE Industry-University-Institute Cooperation Funds under Grant No. 2022ZTE09.

classification of soft real-time systems and hard real-time systems. Soft real-time systems do not always have safety requirements, but hard real-time systems require that the program’s WCET should never be underestimated. Additionally, this serves as the fundamental metric for assessing the dependability criteria of WCET analysis outcomes^[1].

According to LYU’s survey^[2], the execution time of a task usually follows some distribution, which is influenced by various factors such as processor status, data input, and program logic. Under the coverage of all possible execution paths, there exist theoretical maximum and minimum values, representing the best-case execution time (BCET) and WCET of the program, respectively.

As embedded real-time operating systems become more powerful, their underlying hardware structures are also becoming increasingly complex. Nowadays, the industry frequently runs real-time tasks on chips with multi-core processor architecture. Therefore, studying the WCET of a task in the context of multi-core systems has significant practical importance. Correspondingly, this paper aims to explore WCET analysis techniques in the field of multi-core systems and provide an overview of their key frameworks and modeling methods.

The rest of this paper is organized as follows. Section 2 offers an overview of timing analysis techniques for WCET. Section 3 summarizes the existing analysis framework, analysis content, and key areas in the field of multi-core platforms. Section 4 lists the primary analysis tools for multi-core timing analysis of WCET. Section 5 summarizes the paper, highlights the challenges of static and dynamic timing analysis in the context of multi-core systems, and discusses possible development trends and the future of WCET analysis methods.

2 Overview of Timing Analysis Techniques

With the development of computer systems and the increasing complexity of processor architectures, more efficient techniques have been developed based on traditional static and dy-

namic methods for WCET analysis, such as analyzing WCET problems from a probabilistic perspective.

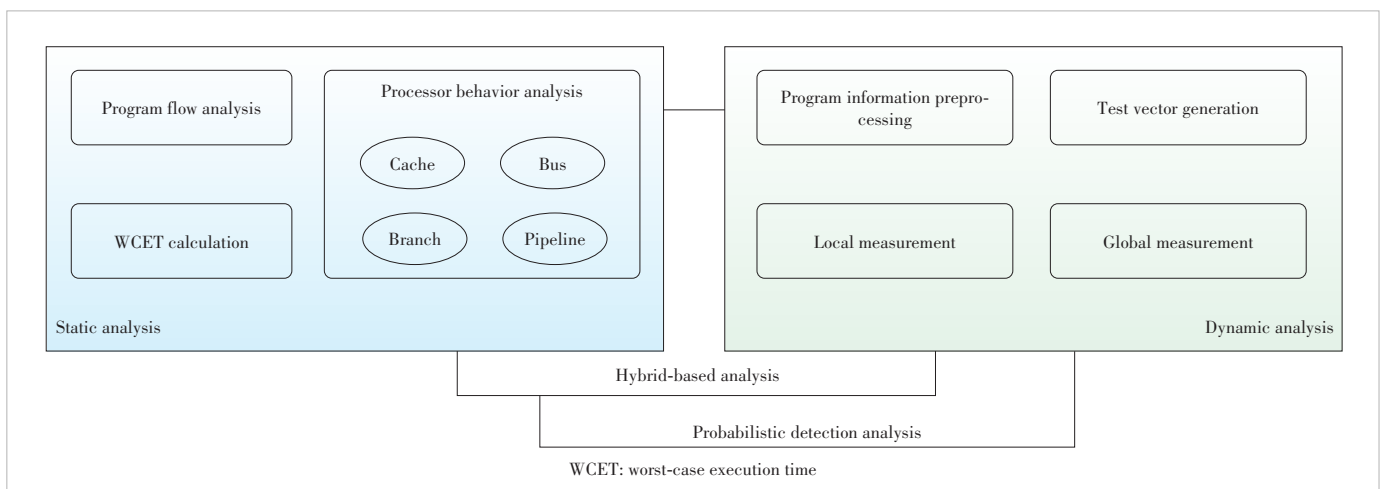
2.1 Static Analysis

Static analysis involves analyzing the pre-information indicated by the program and hardware prior to its execution on the hardware platform. As shown in Fig. 1, static analysis usually consists of the following three steps.

1) Processor behavior analysis^[3-14]: The main purpose of this step is to analyze the execution time of instructions and local programs on a specific processor. The behavior of a processor can be influenced by a few factors, including cache behavior, instruction pipelining, branch prediction, and instruction level parallelism. Each of these factors can affect the execution time of a program and must be taken into account when analyzing the WCET.

2) Program flow analysis^[15-17]: Program flow directly affects the WCET of a program. It includes control flow graphs (CFGs), data flow analysis, and abstract interpretation. CFGs are graphical representations of a program’s control flow, showing the order in which program statements are executed. Data flow analysis tracks the flow of data through a program and provides information about the changes in data during program execution by performing a static analysis of the program to infer the relationships between variable definitions and their uses. Abstract interpretation involves analyzing a program’s abstract semantics, rather than its concrete behavior. By abstracting the concrete execution details, abstract interpretation provides a simplified yet sound approximation of the program’s behavior.

3) WCET calculation: This step involves calculating the program’s execution time and obtaining the program’s execution time on the longest execution path. The selection of the calculation methodology employed has a direct impact on both program flow analysis and processor behavior analysis, underscoring its fundamental significance among the three primary techniques^[1].



▲ Figure 1. Overview of timing analysis architecture

2.2 Dynamic Analysis

Dynamic analysis, in contrast to static analysis, involves analyzing a program during its runtime to determine its WCET. Dynamic analysis is particularly useful in real scenarios. It can be tailored to specific scenarios such as the use of a particular processor pipeline model, cache size, and branch prediction strategy to analyze the program of interest. Essential data and multiple repetitions of execution are required to be recorded during dynamic timing analysis to increase the sample space and reduce the likelihood of incidental events, resulting in more accurate and reliable results. For example, recording the program’s execution time by observing wall-clock time and performing multiple repetitions of the action will increase the results’ credibility, as demonstrated by mathematical analysis.

Both static and dynamic analysis techniques have their own merits and drawbacks, and conducting a systematic comparison can aid in comprehending their distinctions. Table 1 outlines the primary facets of comparison and the corresponding features of the two analysis methodologies^[2].

3 Related Research of WCET Analysis on Multi-Core Platforms

3.1 Static Timing Analysis on Multi-Core Platforms

In the field of multi-core computing, researchers have proposed various techniques to estimate WCET accurately. One such approach is the static analysis of WCET by focusing on shared resources. Shared cache and shared bus domains have emerged as the primary research areas for static timing analysis in multi-core systems. Among these, the shared cache domain has attracted significant attention from researchers who have developed diverse approaches^[5,8,18-21] and models to address cache-related issues.

Additionally, CHATTOPADHYAY et al. proposed a framework for multi-core platform analysis^[22], which underscores the importance of considering shared resources in WCET estimation. These advancements in multi-core research have enhanced our ability to estimate WCET accurately and create more reliable real-time systems.

3.1.1 Shared Cache Analysis

As real-time embedded systems become increasingly complex, multi-core shared cache architectures are becoming

more prevalent in modern embedded devices. Cache conflicts in these architectures stem primarily from memory access request interference of parallel tasks in the shared cache. If a task running on one core fails, it can replace certain cache lines in the shared cache, which may cause another parallel task on another core to require the data in those same cache lines. This leads to interference between the two tasks. Therefore, in multi-core shared cache architectures, the execution time of a task may be significantly affected by other concurrently running tasks, and hence WCET analysis must consider the interference between parallel tasks that share the cache.

YAN and ZHANG proposed an address mapping-based conflict analysis method to address this issue^[5]. This approach abstractly analyzes the hits and misses of tasks in both the L1 cache and the L2 shared cache, and examines whether the task being analyzed in the L2 shared cache is affected by tasks on other cores. A conflict is identified if the task being analyzed maps to the same cache line as a task on another core. Based on the conflict analysis results, the method adjusts the hit rate of the task analyzed in the L2 cache and finally obtains the WCET of the task under multi-core conflict situations by using implicit path enumeration techniques. Moreover, this method proposes a shared cache analysis method based on abstract interpretation, with a comprehensive consideration of safety, which provides a way to increase the accuracy of WCET analysis for cache in multi-core shared resources.

LYU et al. proposed a method for time-series modeling and analyzing cache conflicts based on model checking^[6]. The authors used the UPPAAL model checker, which is based on the theory of timed automata, to generate automata for each task, L1 cache behavior, and shared L2 cache behavior. During the execution of the task automata, a global clock is maintained to update the L2 cache automata. Conflicts are detected by analyzing the time spent accessing the L2 cache task and the established simulation time. This model checking-based method provides an effective means of analyzing cache conflicts for time-sensitive applications.

CHEN et al. proposed a method for modeling and analyzing cache conflicts based on execution graphs^[21]. Using execution graphs, they constructed the timing of tasks in both the non-interference state and maximum interference state (assuming all access to the shared L2 cache misses) and identified the earliest conflict timing, Earliest, in the non-interference state, which represents the earliest fetch time of the access request

▼Table 1. Comparison of static and dynamic WCET analysis methods

Comparative Content	Static WCET Analysis	Dynamic WCET Analysis
Safety of analysis results	Safe	Unsafe
Analysis complexity	High	Low
Adaptability to new architectures	Bad	Good
Analytical accuracy	Mainly depending on the analysis technique	Mainly depending on the test vector
User assistance that improves the quality of analysis	Providing additional control flow information	Providing better test vectors

WCET: worst-case execution time

that hits the shared cache. Similarly, the authors identified the latest conflict timing, Latest, in the interference state, which represents the latest fetch time of the access request that hits the shared cache. By analyzing the relationship between Earliest and Latest for different access requests of tasks, they determined whether conflicts existed. If the earliest conflict timing of one access request is greater than the latest conflict timing of another access request, or the latest conflict timing of one access request is less than the earliest conflict timing of another access request, no conflicts occur. This method provides an effective means of modeling and analyzing cache conflicts based on execution graphs. It can be used to identify and resolve conflicts in complex multi-core embedded systems.

The comparison of different shared cache conflict analysis methods is shown in Table 2.

CHATTOPADHYAY et al. proposed subdividing shared cache analysis into instruction cache analysis, address analysis, data cache analysis, cache access classification, and unified cache analysis^[22]. They also optimized the shared cache analysis.

ZHU et al. addressed this issue by proposing a multilevel coherence protocol that defines a coherence domain consisting of intra-domain access and cross-domain access in a multi-core processor^[7]. The cache state update functions within and across the coherence domains are obtained based on cache read/write strategies and the modified exclusive shared and invalid (MESI) coherence protocol. Consequently, this enables WCET analysis under the nested multilevel coherence protocol. This approach provides a way to analyze cache interference in multi-core processors and improves the accuracy of WCET analysis for real-time systems.

Although many approaches have been proposed for analyzing shared caches, there remain some challenges in the existing methods. For instance, LYU et al.'s approach based on automata may lead to state explosion, while CHEN et al.'s method may produce excessively conservative WCET results. Due to the crucial significance of accurate analysis of shared caches, improving the accuracy of conflict analysis is a significant concern. Researchers are exploring the hardware implementation of code execution to develop more precise techniques for analyzing shared cache conflicts. As this area is continuously evolving and developing, further research is necessary to improve the existing methods and to develop new and more accurate techniques for analyzing shared cache conflicts.

3.1.2 Shared Bus Analysis

In embedded multi-core processors, interconnects and com-

munication between on-chip cores are typically established using buses. However, this bus allocation scheme can introduce interference that in turn affects the WCET of an analysis task. Researchers have proposed several solutions to this problem.

Regarding the on-chip bus allocation strategies, namely TDMA or Round Robin, CHATTOPADHYAY et al. proposed a method for calculating the bus waiting delay using relative offsets, assuming that the first memory access request of each basic block in the task always occurs at the start of the corresponding processor core's bus time slice^[8]. By combining the memory access request sequence obtained based on abstract interpretation analysis within the basic block and the hit/miss status in L1 and L2 caches, the relative offset Δ of the first memory access request can be calculated, and thus the bus waiting time, Wait (Δ), for other memory access requests within the basic block can be computed. The theoretical foundation of this method is that if a program commences during its running core's bus time slice, its WCET estimation is unaffected by the choice of bus slot. Consequently, the approach can consistently provide a reliable WCET estimation and enhances the precision of bus waiting delay analysis. However, when handling loop structures within a task, the method must be aligned for each loop iteration, and the cumulative error grows as the number of loops increases.

KELTER et al.^[9] proposed a bus wait delay analysis method based on data flow analysis techniques to address the shortcomings of the method in Ref. [8]. They extended the basic blocks in the control flow graph (CFG) and classified the instruction sequences into one of three types: never accessing the bus, always accessing the bus, or possibly accessing the bus. They calculated the relative offset on the bus based on the instruction sequences and used data flow analysis to divide the calculation results into two types of offset sets: collection and interval. The algorithm utilizes the data flow analysis on the offset sets to identify whether a stable state can be achieved and then select the maximum value as the bus offset. The algorithm iteratively analyzes the collection and interval offset sets until a fixed point is reached in the offset set, or the loop limit is reached, avoiding the loop alignment requirement present in other approaches. This feature provides an advantage to the algorithm and further improves the accuracy of the WCET estimation.

On-chip bus delay analysis faces a significant challenge due to the absence of precise timing details on memory access requests. To address this issue, CHEN et al.^[20] proposed a method to establish timing based on the execution order of memory access requests and the temporal order of basic

▼Table 2. Comparison of different shared cache conflict analysis methods

Comparative Content	Analysis Based on Address Mapping	Analysis Based on Logical Order	Analysis Based on Model Checking	Analysis Based on Time Series Category
Analysis accuracy	Low	Higher	High	High
Analysis complexity	Low	High	High	Higher
Analytical method	Address mapping analysis	Cache conflict graph and logic analysis	Time automaton	Block-based temporal modeling

blocks. They determined the upper and lower bounds for the timing of every memory access request, used the information to compute the bus waiting time, and scrutinized the limits. This method produces relatively accurate results, but its complexity increases significantly with the increase of the task size.

In conclusion, analyzing on-chip shared bus usage remains a challenging problem for researchers. Scholars attempt to analyze conflicts by delving deep into the hardware usage logic of the bus to develop an analytical mathematical model that incorporates bus constraints on WCET to reduce the pessimism of the WCET analysis result. This approach aims to improve the precision of the WCET analysis and provide a more accurate understanding of on-chip shared bus performance.

3.2 Dynamic Timing Analysis of Multi-Core Platforms

In the field of multi-core research, commercial scenarios generally utilize dynamic methods for WCET estimation. Due to the complexity of static analysis methods, using WCET upper bounds with some degree of unsafety may be accepted to reduce the computation complexity. In such cases, measurement-based methods can be employed to estimate the WCET by executing the program's code on a real measured processor and using a combination of hardware and software methods. Dynamic WCET analysis methods need to address issues such as testing tools, testing scope (local/global), and testing vectors.

DEVERGE et al.^[10] proposed a secure measurement-based method that utilizes structural testing to generate input test data and analyzes WCET based on the hardware structure attributes of the platform where the code is executed. The authors presented partial test results on PowerPC 7450 and proposed an implementation method for dynamic testing. To obtain a more secure WCET estimation, hardware conditions, such as cache and branch prediction buffers, must be manually configured before testing to satisfy the safety criteria of WCET. This manual configuration is necessary to ensure the accuracy of the analysis and the safety of the system.

RIEDER et al.^[11] proposed a dynamic measurement-based WCET analysis framework that uses the concept of local measurement, which measures each code component separately. The authors also introduced a mechanism that accesses at least two pins on the hardware platform to measure the code execution time.

WENZEL et al.^[12] proposed a method for generating test vectors based on the model checking techniques that build upon the dynamic measurement approach mentioned above. The approach involves partitioning the code into control flow graphs (CFGs), generating test vectors through software-based execution that is guided by the test vectors, and using hardware to perform actual testing.

While the results obtained through dynamic WCET methods may not be completely safe, for soft real-time systems, obtaining safe WCET values for the code under analysis is not al-

ways necessary. Additionally, dynamic methods have relatively low complexity, making them valuable in practical situations. Dynamic WCET analysis methods largely focus on innovation and exploration in the areas of testing tools and test vector generation.

3.3 Other Analysis Methods of Multi-Core Platforms

In addition to conventional static and dynamic analysis techniques, hybrid methods and probabilistic detection approaches are increasingly being used to calculate WCET in multicore systems and research efforts are expanding beyond traditional static and dynamic analysis techniques, leading to the emergence of more innovative and effective techniques. Hybrid-based timing analysis and probabilistic detection timing analysis are among typical methods used in this regard.

3.3.1 Hybrid-Based Timing Analysis

Hybrid-based timing analysis used in multi-core systems involves combining static and dynamic analysis techniques by relaxing certain indicators to achieve a more accurate WCET estimation.

The hybrid method integrates static and dynamic approaches to overcome the shortcomings of each method. For instance, static analysis techniques can determine the control flow graph, avoiding the issue of identifying the worst-case path input values in dynamic analysis methods. In contrast, dynamic methods can ensure that all sub-paths are measured, achieving higher structural coverage with less time overhead than complete path coverage. A typical implementation of hybrid analysis involves estimating the program's WCET time by observing the time of the shortest sub-path WCET in the code and combining program structure information and other data obtained through static analysis.

However, hybrid methods still face challenges. On high-performance hardware, context sensitivity and execution history dependence can affect the accuracy of WCET estimation when dynamic measurement-based methods are used to measure sub-paths. Therefore, the hybrid method that uses this measurement method will inherit errors caused by these problems. While hybrid methods cannot guarantee the upper bound of WCET, their estimated results are much more accurate than fully dynamic measurement methods.

3.3.2 Probabilistic Detection Timing Analysis

In probabilistic detection timing analysis of multi-core systems, the probabilistic worst-case execution time (pWCET) distribution represents the minimum upper bound of the execution time distribution of a program under valid operational scenarios. The scenario of operation is defined as an infinite repetition of input states and initial hardware state sequences, which describes the feasible ways in which the program can be repeatedly executed to a large extent.

There are two methods of probabilistic detection timing analysis.

1) The static probabilistic timing analysis (SPTA) method, as its name suggests, is similar to static analysis in its approach. However, unlike traditional static analysis methods, SPTA considers uncertain behaviors in the hardware, software, and program as stochastic events and uses probability distribution methods to analyze the upper bound of the pWCET distribution, rather than the WCET upper bound^[23]. Therefore, the SPTA method must determine certain dynamic attributes of the program before executing the code. For those attributes that cannot be accurately determined, such as cache replacement and shared cache conflicts in multi-core environments, the results may be overly pessimistic due to conservative estimates. Thus, the main challenge of SPTA is to accurately construct the model and determine the necessary information for the model to ensure the accuracy of the SPTA analysis. As far as we are aware, SPTA techniques have not been extended to address multi-level caches or multi-core systems with shared caches.

2) The measurement-based probabilistic timing analysis (MBPTA) method is based on the extreme value theory (EVT) and uses the extreme value method to estimate the pWCET distribution of a program^[23]. Current MBPTA methods first use the idea of dynamic detection to execute the program under dynamic measurement methods and obtain a sample of execution time observations, which is called a measurement protocol. Under this measurement protocol, the program is executed multiple times in one or more scenarios, and in a fixed initial state, the pWCET distribution is finally derived by EVT. Constructing an appropriate measurement protocol is critical in this method, and the initial state must be a sample that is likely to occur in the scenario being studied. Similarly, this method encounters significant challenges due to the diversity of scenarios and initial states. In practical operation, there may not exist a single input pattern or hardware state.

4 Multi-Core WCET Analysis Tools

4.1 Commercial Tools

Table 3 provides a comparison of the functional support and model methods used by some comprehensive tools in the field of WCET analysis. In the table, aiT, which is a commercial

tool developed by AbsInt, uses static analysis techniques to create control flow graphs and performs value analysis, cache analysis, pipeline analysis, path analysis, loop and recursion analysis, and many more. It supports the analysis of both instruction and data cache and offers support for various cache replacement policies. It also supports both in-order and out-of-order pipelines and can analyze certain PCI buses^[24]. Additionally, this software supports multiple chips with different architectures, including PowerPC and ARM. In conclusion, aiT is a powerful and feature-rich commercial tool that is widely applicable to WCET analysis.

RapiTime, another commercial tool developed by Rapita Systems, uses a hybrid testing approach that combines static analysis with testing to generate a not overly pessimistic WCET upper bound. It is not restricted to specific platforms or processors, but requires certain means to access embedded devices to collect partial testing information and analyze the WCET^[25]. This tool is a direct manifestation of dynamic methods.

XMOS Timing Analyzer, a commercial tool developed by XMOS Ltd., is primarily designed for XMOS processors. The tool uses a dynamic measurement approach where the code is annotated with specific markers and binary tools are utilized to measure the time between two endpoints. This tool still employs a dynamic testing approach^[26].

In commercial application scenarios, WCET analysis of hard real-time embedded systems, such as aiT, still uses static analysis. However, in some soft real-time embedded system scenarios, dynamic analysis or hybrid-based analysis is often used when the feasibility is taken into account.

4.2 Open-Source Tools

Table 4 provides a comparison of open-source multi-core WCET analysis tools. In the table, Chronos is a WCET analysis tool developed by the National University of Singapore. It uses a static analysis approach implemented on a SimpleScalar simulator that emulates an instruction set like MIPS (Microprocessor without interlocked piped stages architecture). This tool supports the analysis of instruction cache, data cache, hierarchical cache, and shared cache. It also provides support for microarchitecture modeling of superscalar, in-order, out-of-order, and dynamic branch prediction for pipe-

▼Table 3. Comparison of commercial multi-core WCET analysis tools

Tool	Company	Method	Target Processors	Core Area		
				Cache	Pipeline	Periphery
aiT	AbsInt	Static analysis	Am486, IntelDX4, ARM, PowerPC, etc.	I-Cache, D-Cache, direct/SA, LRU, PLRU, pseudo round robin	In-order/ out-of-order	PCI bus
RapiTime	Rapita Systems	Hybrid-based	-	-	-	-
XMOS Timing Analyzer	XMOS Ltd.	Measurement-based	XMOS processors	-	-	-

ARM: Advanced RISC Machine
D-cache: data cache

I-Cache: instruction cache
LRU: least recently used

PCI: Peripheral Component Interconnect
PLRU: pseudo-LRU

SA: set associative cache
WCET: worst-case execution time

▼Table 4. Comparison of open-source multi-core WCET analysis tools

Tool	Institution	Method	Target Processors	Core Area		
				Cache	Pipeline	Periphery
Chronos	National University of Singapore	Static analysis	SimpleScalar	I-cache, D-cache, Unified cache, Direct/SA, LRU	Multi-issue superscalar, in-order, out-order, dynamic branch prediction	
Bound-T	Tidorum	Static analysis	SPARC/ERC32, etc.	-	-	IPET
Heptane	IRISA	Static analysis	-	D-cache, I-cache,	Branch prediction	
SWEET	Mälardalen Real-Time Research Centre	Static analysis	NECV850E ARM9	-	-	Flow analysis

ARM: Advanced RISC Machine
D-cache: data cache
I-Cache: instruction cache
IPET: Implicit Path Enumeration Technique

IRISA: The Institute for Research in IT and Random Systems
LRU: least recently used
PCI: Peripheral Component Interconnect
PLRU: pseudo-LRU

SA: set associative cache
SWEET: Swedish Execution Time Tool
WCET: worst-case execution time

line analysis^[27]. In addition, Chronos provides SimpleScalar instruction-defined formats, which allows it to be ported to different instruction sets and similar simple processor architectures. This tool is feature-rich and comprehensive and has significant scalability.

Bound-T is a static analysis tool that was originally commercial software but later became an open-source tool. It is designed to analyze processors with relatively simple architectures. While this tool also handles microarchitecture analysis directly, it does not consider pipeline superscalar and other complex features. Instead, Bound-T calculates the time for basic blocks through the program control flow graph^[28].

Heptane^[29] is an open-source WCET analysis tool developed by the IRISA team. It uses a static analysis method to estimate the WCET and implements Implicit Path Enumeration Technique (IPET). The tool includes cache analysis techniques for many cache architectures, such as multiple replacement policies, cache hierarchy, and shared caches among multiple cores. In addition, Heptane supports various instruction sets, including MIPS, ARM, MSP430, and RISC-V and is periodically maintained and updated. The tool has a stable prototype development and can support various instruction sets.

SWEET^[30] is an open-source tool for static WCET analysis that is known for its program flow analysis implementation. It is designed to analyze the WCET time bounds for two different processors using three distinct techniques: a fast path-based method, a global IPET method, and a hybrid clustering method.

These tools can be parallelized to run on multi-core platforms, but doing so may introduce some errors. After analyzing and comparing the WCET analysis tools mentioned above in both commercial and academic contexts, we can conclude that static analysis and hybrid dynamic analysis methods are commonly employed in commercial scenarios. On the other hand, academic researchers always attempt to propose various models and methods based on static analysis to improve accuracy

and safety.

In the long run, the WCET research is ongoing, and as processor architectures become increasingly complex, new challenges are posed to the development of effective analysis tools. However, the ultimate objectives remain the same: to guarantee the safety of WCET upper-bound analysis and explore methods for enhancing its accuracy.

5 Conclusions

This paper reviews the calculation of WCET for programs running on multi-core processors. It highlights that the calculation of WCET is a critical aspect of timing analysis in hard real-time systems and can be accomplished by various methods, such as static and dynamic timing analysis along with specialized modeling of the hardware system. A comprehensive analysis of existing techniques for WCET estimation is also provided, including both static and dynamic analysis methods for multi-core processors. Furthermore, it introduces traditional techniques for WCET estimation using both static and dynamic analysis, which extends the analysis to encompass multi-core processors beyond the scope of single-core processors. Finally, it discusses the major challenges arising from different program execution scenarios on multi-core platforms.

In the future, the accuracy of WCET estimation will play an increasingly critical role in multi-core systems. Therefore, continued research is necessary for guaranteeing the dependable and timely operation of hard real-time systems. For safety reasons, static analysis is still the mainstream method in the field of the hard real-time. From the perspective of hardware characteristics and software protocols, analyzing various conflicts of shared resources between cores to improve accuracy will remain the mainstream research direction of static analysis. However, in the field of the soft real-time, methods such as dynamic methods and hybrid testing have gradually become the mainstream direction of the industry.

References

- [1] LYU M S, GUAN N, ZHANG Y, et al. A survey of WCET analysis of real-time operating systems [C]//Proc. International Conference on Embedded Software and Systems. IEEE, 2009: 65 – 72. DOI: 10.1109/ICISS.2009.24
- [2] LYU M S. Research on worst-case execution time analysis techniques of real-time systems [D]. Shenyang: Northeastern University, 2010
- [3] HECKMANN R, LANGENBACH M, THESING S, et al. The influence of processor architecture on the design and the results of WCET tools [J]. Proceedings of the IEEE, 2003, 91(7): 1038 – 1054. DOI: 10.1109/JPROC.2003.814618
- [4] LI X F, ROYCHOUDHURY A, MITRA T. Modeling out-of-order processors for WCET analysis [J]. Real-time systems, 2006, 34(3): 195 – 227. DOI: 10.1007/s11241-006-9205-5
- [5] YAN J, ZHANG W. WCET analysis for multi-core processors with shared L2 instruction caches [C]//Proc. IEEE Real-Time and Embedded Technology and Applications Symposium. IEEE, 2008: 80 – 89. DOI: 10.1109/RTAS.2008.6
- [6] LYU M, YI W, GUAN N, et al. Combining abstract interpretation with model checking for timing analysis of multicore software [C]//31st IEEE Real-Time Systems Symposium. IEEE, 2010: 339 – 349. doi: 10.1109/RTSS.2010.30.
- [7] ZHU Y A, SHI X C, YAO Y, et al. A WCET analysis method for multi-core processors with multi-tier coherence protocol [J]. Journal of computer research and development, 2023, 60(1): 30 – 42. DOI: 10.7544/jissn1000-1239.202111244
- [8] CHATTOPADHYAY S, ROYCHOUDHURY A, MITRA T. Modeling shared cache and bus in multi-cores for timing analysis [C]//Proc. 13th International Workshop on Software & Compilers for Embedded Systems. ACM, 2010: 1 – 10. DOI: 10.1145/1811212.1811220
- [9] KELTER T, FALK H, MARWEDEL P, et al. Bus-aware multicore WCET analysis through TDMA offset bounds [C]//Proc. 23rd Euromicro Conference on Real-Time Systems. IEEE, 2011: 3 – 12. DOI: 10.1109/ECRTS.2011.9
- [10] DEVERGE J F, PUAUT I. Safe measurement-based WCET estimation [C]//5th International Workshop on Worst-Case Execution Time Analysis (WCET'05). OASICS, 2007: 13 – 16. DOI: 10.4230/OASICS.WCET.2005.808
- [11] RIEDER B, WENZEL I, STEINHAMMER K, et al. Using a runtime measurement device with measurement-based WCET analysis [C]//Embedded System Design: Topics, Techniques and Trends, Proc. of IFIP TC10 Working Conference: International Embedded Systems Symposium (IESS). Springer, 2007: 15 – 26. DOI:10.1007/978-0-387-72258-0_2
- [12] WENZEL I, RIEDER B, KIRNER R, et al. Automatic timing model generation by CFG partitioning and model checking [C]//Design, Automation and Test in Europe. IEEE, 2005: 606-611. DOI: 10.1109/DATE.2005.76
- [13] DASARI D, NELIS V. An analysis of the impact of bus contention on the WCET in multicores [C]//Proc. IEEE 14th International Conference on High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems. IEEE, 2012: 1450 – 1457. DOI: 10.1109/HPCC.2012.212
- [14] CHATTOPADHYAY S, ROYCHOUDHURY A. Unified cache modeling for WCET analysis and layout optimizations [C]//Proc. 30th IEEE Real-Time Systems Symposium. IEEE, 2009: 47 – 56. DOI: 10.1109/RTSS.2009.20
- [15] PUSCHNER P, BURNS A. Guest editorial: a review of worst-case execution-time analysis [J]. Real-time systems, 2000, 18(2): 115 – 128. DOI: 10.1023/A:1008119029962
- [16] THEILING H, FERDINAND C, WILHELM R. Fast and precise WCET prediction by separated cache and path analyses [J]. Real-time systems, 2000, 18(2): 157 – 179. DOI: 10.1023/A:1008141130870
- [17] LI X F, LIANG Y, MITRA T, et al. Chronos: a timing analyzer for embedded software [J]. Science of computer programming, 2007, 69(1/2/3): 56 – 67
- [18] HUYNH B K, JU L, ROYCHOUDHURY A. Scope-aware data cache analysis for WCET estimation [C]//Proc. 17th IEEE Real-Time and Embedded Technology and Applications Symposium. IEEE, 2011: 203 – 212. DOI: 10.1109/RTAS.2011.27
- [19] LIANG Y, DING H P, MITRA T, et al. Timing analysis of concurrent programs running on shared cache multi-cores [J]. Real-time systems, 2012, 48(6): 638 – 680. DOI: 10.1007/s11241-012-9160-2
- [20] CHEN F Y, ZHANG D S, WANG Z Y. Static analysis of Run-time inter-thread interferences in shared cache multi-core architectures based on instruction fetching timing [C]//Proc. IEEE International Conference on Computer Science and Automation Engineering. IEEE, 2011: 208 – 212. DOI: 10.1109/CSAE.2011.5952455
- [21] CHEN F Y, ZHANG D S, WANG Z Y. Static analysis of run-time inter-core interferences for concurrent programs in shared cache multicore architectures [J]. Applied mechanics and materials, 2012, 198/199: 523 – 527. DOI: 10.4028/www.scientific.net/AMM.198-199.523
- [22] CHATTOPADHYAY S, CHONG L K, ROYCHOUDHURY A, et al. A unified WCET analysis framework for multicore platforms [J]. ACM transactions on embedded computing systems, 13(4s): 124. DOI: 10.1145/2584654
- [23] DAVIS R I, CUCU-GROSJEAN L. A survey of probabilistic timing analysis techniques for real-time systems [J]. Leibniz Transactions on Embedded Systems (LITES), 2019, 6(1): 03:1 – 03:60. DOI: 10.4230/LITES-v006-i001-a003
- [24] AbsInt. aiT [EB/OL]. [2023-01-17]. <https://www.absint.com/ait>
- [25] Rapita Systems. Rapitime [EB/OL]. [2023-01-17]. <https://www.rapitasystems.com/products/rapitime>
- [26] XMOS Ltd. XMOS [EB/OL]. [2023-01-17]. <https://www.xmos.ai/software-tools>
- [27] National University of Singapore. Chronos [EB/OL]. [2023-01-17]. <https://www.comp.nus.edu.sg/~rpembed/chronos>
- [28] Tidorum Ltd. Bound-T execution time analyzer [EB/OL]. [2023-01-17]. <http://www.bound-t.com>
- [29] IRISA team. Heptane [EB/OL]. [2023-01-17]. <https://team.nria.fr/pacaps/soft-ware/heptane>
- [30] MRTC. Swedish execution time analysis tool [EB/OL]. [2023-01-17]. <http://www.mrtc.mdh.se/projects/wcet/sweet.html>

Biographies

LUO Haoran received his bachelor's degree from Sichuan University, China. He is pursuing a master's degree at University of Electronic Science and Technology of China (UESTC). Driven by a deep passion for computer science, he actively engages in various academic activities and research projects.

HU Shuisong received his bachelor's degree from University of Electronic Science and Technology of China (UESTC). He is pursuing a master's degree at UESTC. He actively engages in research and exploration in various areas, including computer networks and operation system.

WANG Wenyong (wangwy@uestc.edu.cn) is a professor at the School of Computer Science, University of Electronic Science and Technology of China (UESTC). He is an IEEE member and a senior member of the China Computer Federation (CCF). He serves as a member of the CCF Information Network, Software Engineering, and High-Performance Computing committees, a council member of the China Internet Association, a member of the Expert Committee of China Education and Research Computer Network (CERNET), and a member of the Expert Committee of China Next Generation Internet (CNGI). He is also a member of the Ministry of Education's Next Generation Internet Technology Innovation Team and serves on the Technical Committee of the National Engineering Laboratory for Next Generation Internet Core Network at Tsinghua University, China. His research interests include computer networks and data sharing and exchange.

TANG Yuke is currently working at the tool support project team of the Department of Operating System Products, ZTE Corporation. He is mainly engaged in compilation, debugging, performance tuning and defect detection of operating systems.

ZHOU Junwei graduated from Southwest JiaoTong University, China, with a major in computer science and joined ZTE Corporation in 2022. He is currently working at the Department of Operation System Products and mainly responsible for the development of WCET tools.



Filter Design of Wireless Base Station Power Supply

LI Wei, GUO Wei, WANG Zhida
(ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTECOM.202401011

<https://kns.cnki.net/kcms/detail/34.1294.TN.202410307.1721.004.html>,
published online March 9, 2024

Manuscript received: 2023-04-18

Abstract: The design of electromagnetic interference (EMI) filters needs to fulfill the EMI standards. Designing a filter is a time-consuming process for new engineers as well as for those experienced engineers. This paper measures and compares the noise spectrum of the wireless base station power prototype with and without the original filter. The ideal insertion loss (IL) of the original filter is obtained by combining calculation and simulation. It is pointed out that the effect of the original filter is not good. Based on the improved insertion-loss method, the source impedance model of the prototype is established by combining measurement and theory. A procedure for designing EMI filters for switch power supply will be presented. The filter design procedure makes it possible to design filters quickly and easily. Finally, the proposed filter design method is proved to be effective by the EMI measurement of the wireless base station power supply prototype.

Keywords: electromagnetic compatibility; common mode noise; EMI filter

Citation (Format 1): LI W, GUO W, WANG Z D. Filter design of wireless base station power supply [J]. *ZTE Communications*, 2024, 22(1): 95 – 105. DOI: 10.12142/ZTECOM.202401011

Citation (Format 2): W. Li, W. Guo, and Z. D. Wang, “Filter design of wireless base station power supply,” *ZTE Communications*, vol. 22, no. 1, pp. 95 – 105, Mar. 2024. doi: 10.12142/ZTECOM.202401011.

1 Introduction

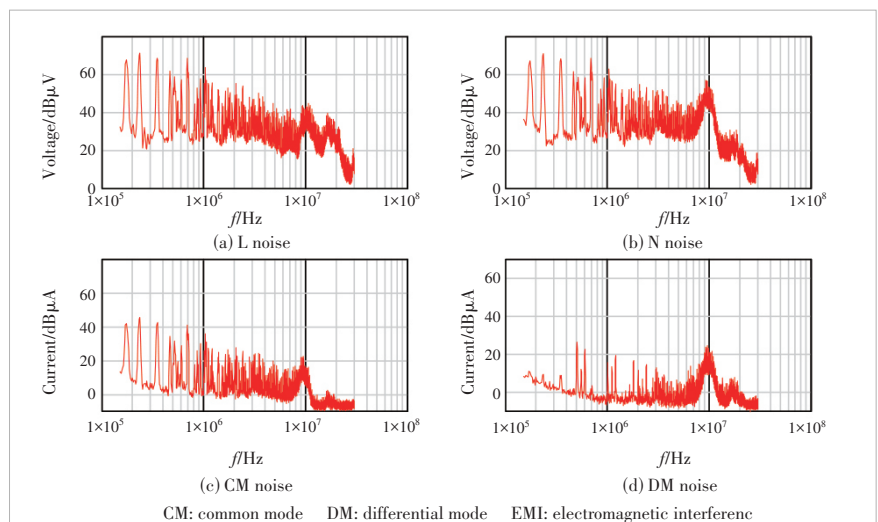
With the development of power electronic equipment, electromagnetic interference (EMI) has become more and more serious. According to the different coupling paths, EMI can be divided into conduction EMI and radiation EMI. The main solution to reducing these conducted emissions is based on the use of EMI filters^[1-2]. The EMI filters are made from coupled inductors combined with capacitors. The choice of filter topology depends on networks and impedances. Generally, common mode (CM) and differential mode (DM) filters are used for power converters. The conventional filter design flow is proposed by SHIH et al.^[3]

EMI filters are widely used with the advantages of simple structure and low cost^[4]. We first measure and compare the noise spectrum of the wireless base station power supply prototype with and without the original filter^[5]. The results show that the ground impedance is the key factor affecting the insertion loss (IL) of the EMI filter. After that, based on the improved insertion-loss method^[6], the electrical model of the source impedance of the prototype is established by combining measurement and theory, and

the accuracy of the insertion-loss method to extract the source impedance is verified by experiments^[7-9]. Finally, the filter is designed for the prototype, and the correctness of the EMI filter design is verified by simulation and measurement.

2 Prototype Raw EMI Noise

The EMI spectrum of the prototype without filters is obtained by the measurement shown in Fig. 1.



▲ Figure 1. Noise spectrum without EMI filter

The original filter is shown in Fig. 2. The noise spectrum of the prototype with the original EMI filter is shown in Fig. 3.

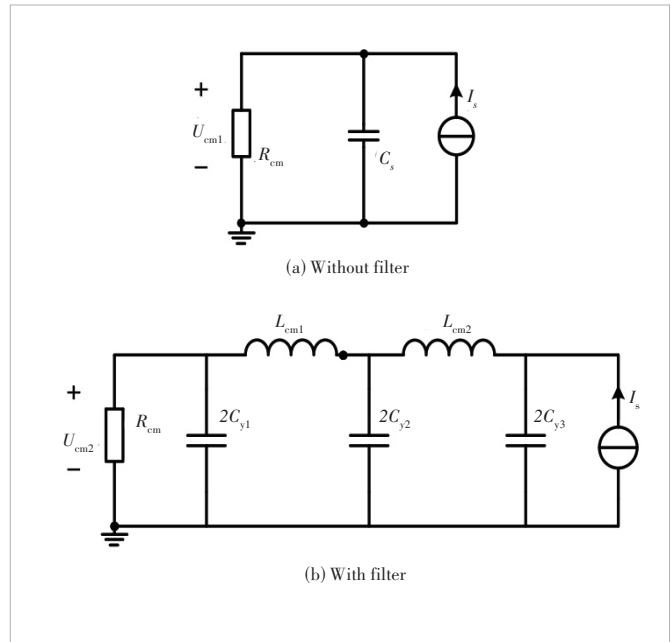
Comparing the CM noise spectrum in Fig. 1 with that in Fig. 3, we find that the CM insertion loss with the original filter is about 20 dB – 40 dB in the low and medium frequency bands.

Fig. 4 shows the schematic of the CM equivalent circuit of the second-order π -type EMI filter. In Fig. 4(b), Z_s is connected in parallel with the Y capacitor in the second-order π -type filter, and Z_s is much smaller than the Y capacitor in the low and middle frequency bands, so the effect of Z_s can be ignored in this analysis.

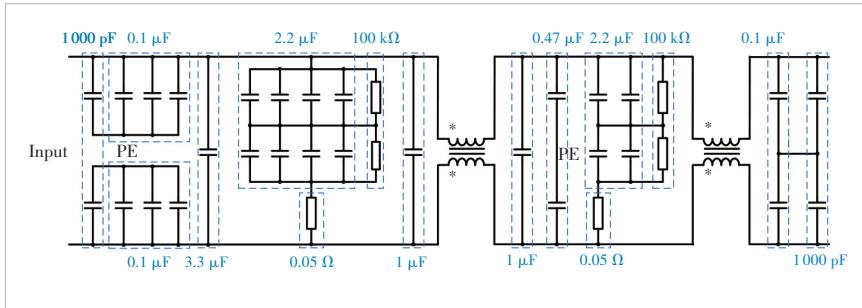
The insertion loss of the second-order π -type filter can be obtained as:

$$\begin{aligned}
 |H_{cm}| = 20 \log & \left[8L_{cm1}L_{cm2}C_{y1}C_{y2}C_{y3}R_{cm} \cdot s^5 + 4L_{cm1}L_{cm2}C_{y2}C_{y3}s^4 + \right. \\
 & \left. [4L_{cm1}R_{cm}C_{y1}(C_{y2} + C_{y3}) + 4L_{cm2}R_{cm}C_{y3}(C_{y1} + C_{y2})]s^3 + \right. \\
 & \left. 2(L_{cm1}C_{y2} + L_{cm2}C_{y3} + L_{cm1}C_{y3})s^2 + 2R_{cm}(C_{y1} + C_{y3})s + 1 \right]. \quad (1)
 \end{aligned}$$

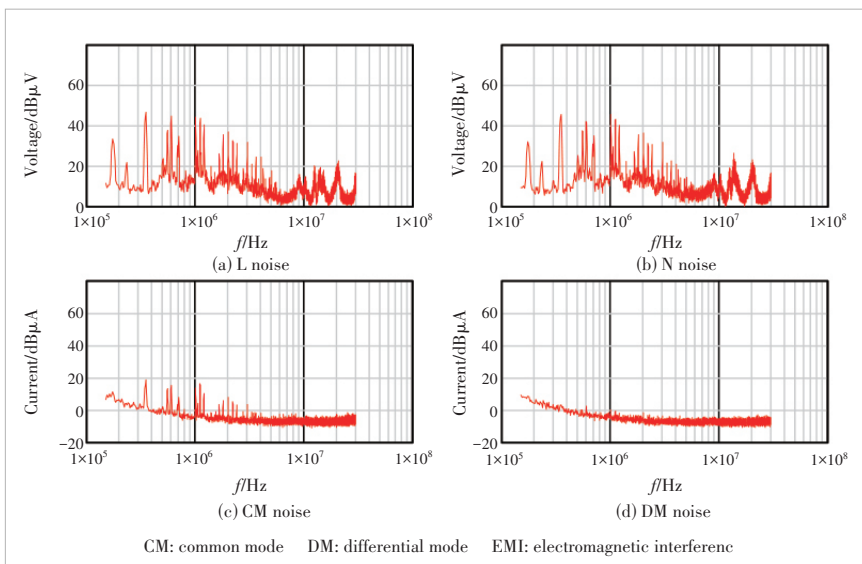
According to Eq. (1), the insertion loss of the original filter is about 200 dB – 300 dB in the low and medium frequency bands. The actual insertion loss of the filter does not reach the theoretical value.



▲ Figure 4. Common mode (CM) equivalent circuit of the second-order π -type electromagnetic interference (EMI) filter



▲ Figure 2. Prototype filter



▲ Figure 3. Raw noise with EMI filter

3 Filter Ground Impedance

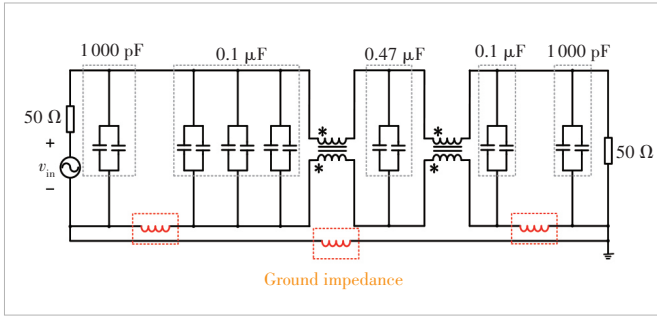
The original EMI filter has a large ground loop, which results in a large ground impedance. The effect of ground impedance on CM insertion loss will be verified by simulation. The simulation schematic is shown in Fig. 5. The CM insertion loss considering ground impedance is shown in Fig. 6(a), and Fig. 6(b) shows the insertion loss of the ideal filter.

Fig. 6(a) illustrates that the resonance of the insertion loss of the filter is advanced and the insertion loss is reduced after the addition of the ground impedance. Due to the large Y capacitor used in the original filter, even a small ground impedance will result in the resonance frequency decrease of the insertion loss. Therefore, in order to reduce EMI noise, the EMI filter electrical parameters need to be designed.

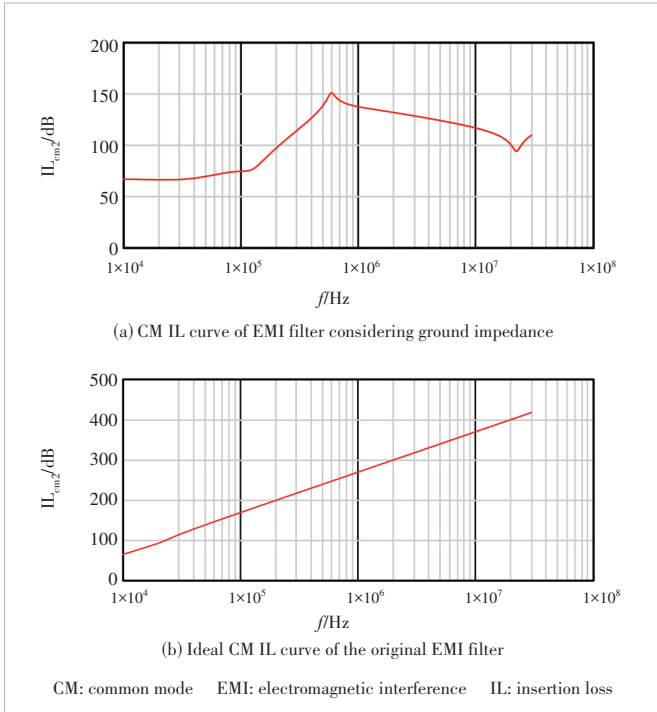
4 Noise Source Impedance Extraction

It is necessary to extract the source impedance of the prototype noise source before designing the filter. The equivalent model of CM noise transmission path without a filter is shown in Fig. 7.

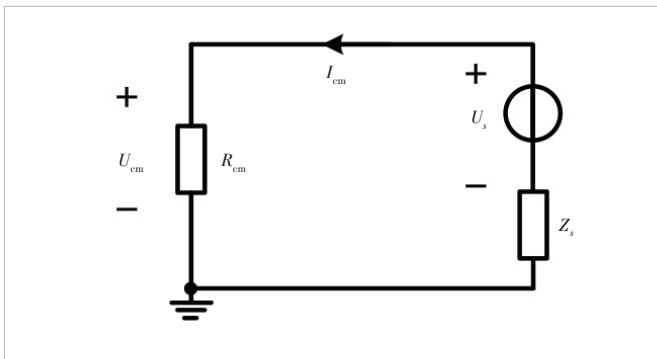
The equivalent CM noise current I_{cm} is:



▲ Figure 5. Ideal common mode (CM) insertion loss (IL) simulation schematic of original electromagnetic interference (EMI) filter



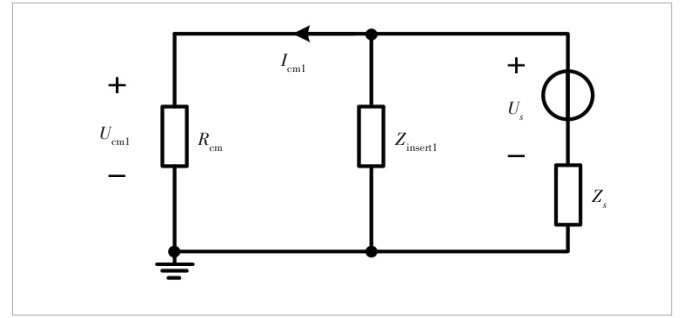
▲ Figure 6. Impact of ground impedance on filter insertion loss



▲ Figure 7. Common mode (CM) noise transmission model

$$I_{cm} = \frac{U_s}{R_{cm} + Z_s} \quad (2)$$

After the impedance $Z_{insert1}$ is inserted, its equivalent model is shown in Fig. 8.



▲ Figure 8. A common mode (CM) noise transmission model with $Z_{insert1}$

CM noise current I_{cm1} in Fig. 8 is:

$$I_{cm1} = U_s \cdot \frac{\frac{Z_{insert1}}{R_{cm} + Z_{insert1}}}{Z_s + \frac{R_{cm} \cdot Z_{insert1}}{R_{cm} + Z_{insert1}}} \quad (3)$$

According to the definition of insertion loss, the insertion loss can be obtained by combining Eqs. (2) and (3):

$$\begin{aligned} IL_1 &= 20 \log \left| \frac{U_{cm1}}{U_{cm}} \right| = 20 \log \left| \frac{I_{cm1} \cdot R_{cm}}{I_{cm} \cdot R_{cm}} \right| = \\ &= 20 \log \left| \frac{Z_s + R_{cm}}{Z_s + \frac{R_{cm} \cdot Z_{insert1}}{R_{cm} + Z_{insert1}}} \right| \end{aligned} \quad (4)$$

Similarly, the insertion loss with $Z_{insert2}$ and $Z_{insert3}$ is:

$$\begin{aligned} IL_2 &= 20 \log \left| \frac{U_{cm2}}{U_{cm}} \right| = 20 \log \left| \frac{I_{cm2} \cdot R_{cm}}{I_{cm} \cdot R_{cm}} \right| = \\ &= 20 \log \left| \frac{Z_s + R_{cm}}{Z_s + \frac{R_{cm} \cdot Z_{insert2}}{R_{cm} + Z_{insert2}}} \right| \end{aligned} \quad (5)$$

$$\begin{aligned} IL_3 &= 20 \log \left| \frac{U_{cm3}}{U_{cm}} \right| = 20 \log \left| \frac{I_{cm3} \cdot R_{cm}}{I_{cm} \cdot R_{cm}} \right| = \\ &= 20 \log \left| \frac{Z_s + R_{cm}}{Z_s + \frac{R_{cm} \cdot Z_{insert3}}{R_{cm} + Z_{insert3}}} \right| \end{aligned} \quad (6)$$

It can be obtained from Eqs. (4) and (5):

$$IL_1 - IL_2 = 20 \log \left| \frac{Z_s + \frac{R_{cm} \cdot Z_{insert2}}{R_{cm} + Z_{insert2}}}{Z_s + \frac{R_{cm} \cdot Z_{insert1}}{R_{cm} + Z_{insert1}}} \right| \quad (7)$$

It can be obtained from Eqs. (4) and (6):

$$IL_1 - IL_3 = 20 \log \left| \frac{Z_s + \frac{R_{cm} \cdot Z_{insert3}}{R_{cm} + Z_{insert3}}}{Z_s + \frac{R_{cm} \cdot Z_{insert1}}{R_{cm} + Z_{insert1}}} \right|. \quad (8)$$

In Eqs. (7) and (8), IL_1 , IL_2 and IL_3 are measured, and the value of inserting impedance Z_{insert} is known. Z_s is the source impedance of the noise source that needs to be extracted, which is divided into real parts and imaginary parts:

$$Z_s = a + j \cdot b. \quad (9)$$

A system of two quadratic equations is obtained:

$$\begin{cases} \frac{a + j \cdot b + \frac{R_{cm} \cdot Z_{insert2}}{R_{cm} + Z_{insert2}}}{a + j \cdot b + \frac{R_{cm} \cdot Z_{insert1}}{R_{cm} + Z_{insert1}}} = 10^{\frac{IL_1 - IL_2}{20}} \\ \frac{a + j \cdot b + \frac{R_{cm} \cdot Z_{insert3}}{R_{cm} + Z_{insert3}}}{a + j \cdot b + \frac{R_{cm} \cdot Z_{insert1}}{R_{cm} + Z_{insert1}}} = 10^{\frac{IL_1 - IL_3}{20}} \end{cases}. \quad (10)$$

For the convenience of the experiment, three sets of Y capacitors: 22 nF, 47 nF and 100 nF, were selected as the inserting impedance. Before the insertion-loss method experiment, the impedance characteristic curves of the three groups of Y capacitors were measured, and the measured impedance curves were shown in Fig. 9.

The high-frequency model of the capacitor is connected in series through the capacitor, parasitic inductance, and parasitic resistance. These three values can be calculated through numerical fitting.

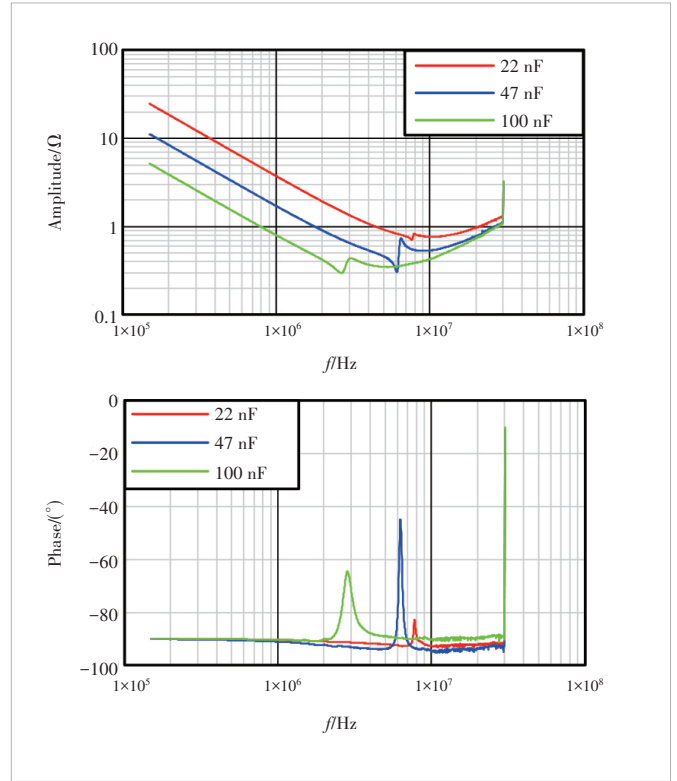
When two Y capacitors are inserted between the device under test (DUT) and LISN, there are two parallel C_y in the CM loop, and their equivalent impedance $Z_c = 1/(j2\omega C_y)$. The equivalent circuit is shown in Fig. 10.

CM noise current I_{cm} in Fig. 10 can be obtained:

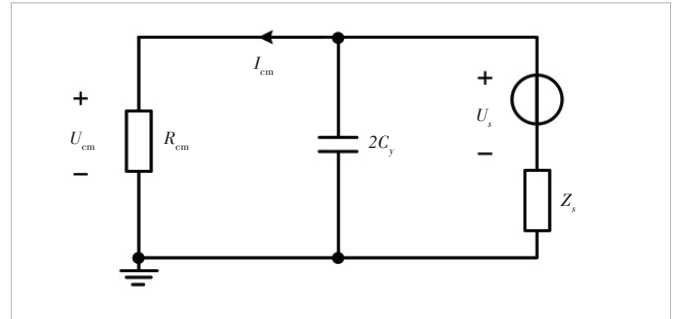
$$I_{cm} = \frac{U_{cm}}{R_{cm}} = \frac{\frac{R_{cm} \cdot Z_c}{R_{cm} + Z_c}}{Z_s + \frac{R_{cm} \cdot Z_c}{R_{cm} + Z_c}} \cdot \frac{U_s}{R_{cm}}. \quad (11)$$

Three sets of Y capacitors were connected to the prototype, and the CM noise spectrum was measured to obtain the insertion loss, as shown in Fig. 11.

The conducted CM noise current spectrum obtained by measurement needs to be converted into amplitudes, and the conversion relationship is as follows:



▲ Figure 9. Measured amplitude-frequency characteristic



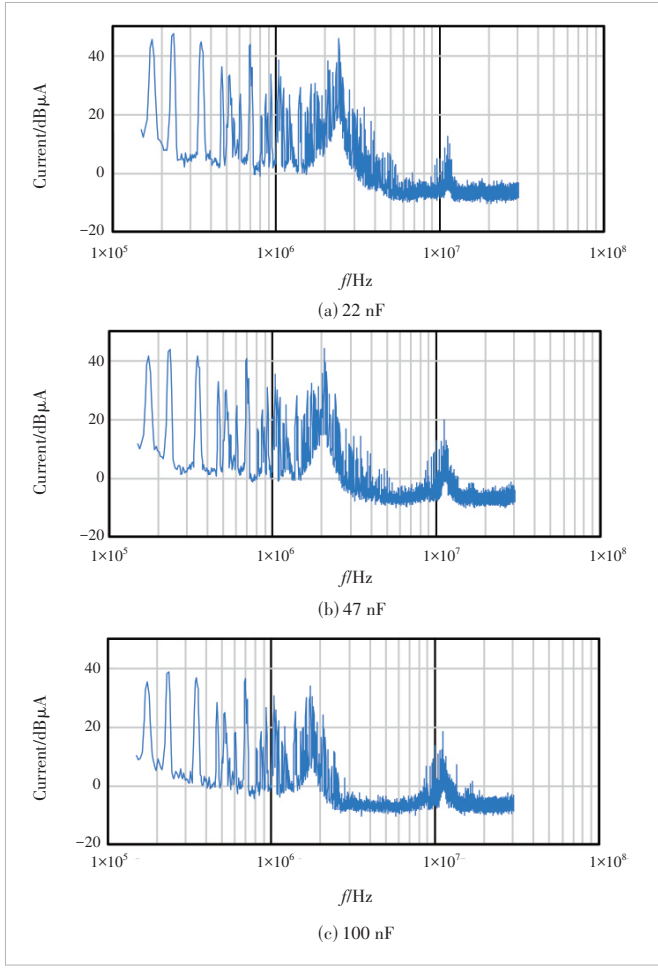
▲ Figure 10. Equivalent circuit with C_y

$$I_{CM_LISN} = 20 \log(i_L + i_N), \quad (12)$$

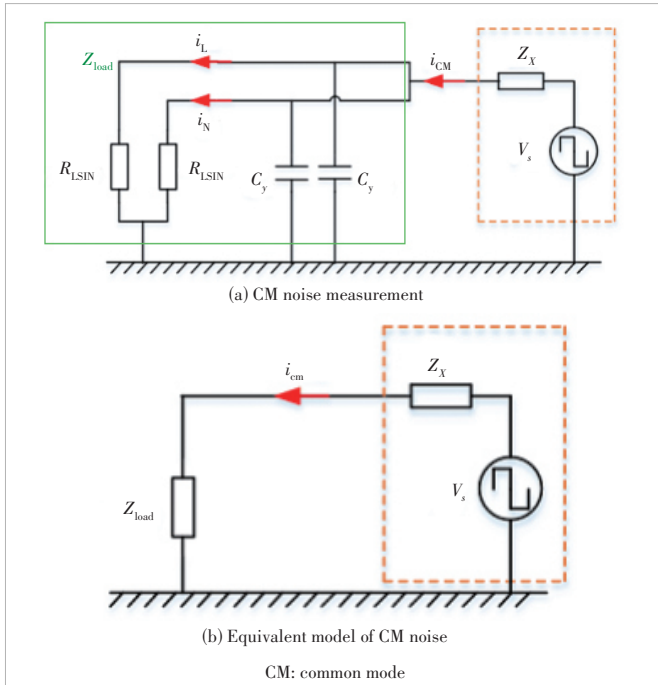
where I_{CM_LISN} represents the amplitude of the current through LISN, and i_L and i_N represent the actual noise current through L-line and N-line flowing through resistance R_{LISN} respectively. The CM noise current flowing through the source impedance Z_s of the noise source is defined as i_{cm} , as shown in Fig. 12(a).

$$|i_{cm}| = \left| (i_L + i_N) \cdot \left(1 + \frac{25}{Z_{C_y}} \right) \right|. \quad (13)$$

By connecting Eq. (12) with Eq. (13), the actual CM noise current I_{cm} is obtained. In each group of noise measurements, R_{LISN} and C_y are known values, which are defined as Z_{load} . The Z_{load} in Fig. 12(b) is:



▲ Figure 11. Common mode (CM) spectrum of CM noise



▲ Figure 12. CM noise measurement and equivalent model of CM noise

$$Z_{\text{load}} = \frac{1}{\frac{1}{R_{\text{cm}}} + \frac{1}{Z_C}}, \quad (14)$$

where R_{CM} stands for the parallel R_{LISN} . The amplitude of i_{CM} is expressed as follows:

$$|I_{\text{cm}}| = \left| \frac{U_s}{Z_{\text{load}} + Z_s} \right|, \quad (15)$$

where the relationship between $|I_{\text{cm}}|$ and the source impedance Z_s is established. The CM noise spectrum measured by connecting the three groups can be obtained by corresponding three groups: $|I_{\text{cm}1}|$, $|I_{\text{cm}2}|$, and $|I_{\text{cm}3}|$, which can be combined by:

$$\frac{|I_{\text{cm}1}|}{|I_{\text{cm}2}|} = \left| \frac{Z_s + Z_{\text{load}2}}{Z_s + Z_{\text{load}1}} \right| = \left| \frac{Z_{s_re} + j \cdot Z_{s_im} + Z_{\text{load}2_re} + j \cdot Z_{\text{load}2_im}}{Z_{s_re} + j \cdot Z_{s_im} + Z_{\text{load}1_re} + j \cdot Z_{\text{load}1_im}} \right|, \quad (16)$$

$$\frac{|I_{\text{cm}1}|}{|I_{\text{cm}2}|} = \left| \frac{Z_s + Z_{\text{load}3}}{Z_s + Z_{\text{load}1}} \right| = \left| \frac{Z_{s_re} + j \cdot Z_{s_im} + Z_{\text{load}3_re} + j \cdot Z_{\text{load}3_im}}{Z_{s_re} + j \cdot Z_{s_im} + Z_{\text{load}1_re} + j \cdot Z_{\text{load}1_im}} \right|. \quad (17)$$

$Z_{\text{load}2_re}$ and $Z_{\text{load}2_im}$ represent the real and imaginary parts of the equivalent impedance, respectively. Z_{s_re} and Z_{s_im} represent the real and imaginary parts of the source impedance respectively. According to Eqs. (12) – (17), the source impedance constraint equation of the noise source can be further obtained by:

$$Z_{s_im} = -\frac{Z_{\text{load}1_im} \cdot I_{\text{cm}1}^2 - Z_{\text{load}2_im} \cdot I_{\text{cm}2}^2}{I_{\text{cm}1}^2 - I_{\text{cm}2}^2} \pm \sqrt{\left(\frac{Z_{\text{load}1_im} \cdot I_{\text{cm}1}^2 - Z_{\text{load}2_im} \cdot I_{\text{cm}2}^2}{I_{\text{cm}1}^2 - I_{\text{cm}2}^2} \right)^2 + \frac{I_{\text{cm}2}^2 (Z_{s_re} + Z_{\text{load}2_re})^2}{I_{\text{cm}1}^2 - I_{\text{cm}2}^2} - \frac{I_{\text{cm}1}^2 (Z_{s_re} + Z_{\text{load}1_re})^2}{I_{\text{cm}1}^2 - I_{\text{cm}2}^2} - \frac{Z_{\text{load}1_im} \cdot I_{\text{cm}1}^2 - Z_{\text{load}2_im} \cdot I_{\text{cm}2}^2}{I_{\text{cm}1}^2 - I_{\text{cm}2}^2}}, \quad (18)$$

$$Z_{s_im} = -\frac{Z_{\text{load}1_im} \cdot I_{\text{cm}1}^2 - Z_{\text{load}3_im} \cdot I_{\text{cm}3}^2}{I_{\text{cm}1}^2 - I_{\text{cm}3}^2} \pm \sqrt{\left(\frac{Z_{\text{load}1_im} \cdot I_{\text{cm}1}^2 - Z_{\text{load}3_im} \cdot I_{\text{cm}3}^2}{I_{\text{cm}1}^2 - I_{\text{cm}3}^2} \right)^2 + \frac{I_{\text{cm}3}^2 (Z_{s_re} + Z_{\text{load}3_re})^2}{I_{\text{cm}1}^2 - I_{\text{cm}3}^2} - \frac{I_{\text{cm}1}^2 (Z_{s_re} + Z_{\text{load}1_re})^2}{I_{\text{cm}1}^2 - I_{\text{cm}3}^2} - \frac{Z_{\text{load}1_im} \cdot I_{\text{cm}1}^2 - Z_{\text{load}3_im} \cdot I_{\text{cm}3}^2}{I_{\text{cm}1}^2 - I_{\text{cm}3}^2}}. \quad (19)$$

Since there are plus or minus signs in Eqs. (18) and (19), both Eqs. (18) and (19) can represent two constraints (with four constraints in total) to form two quadratic equations. The trajectories of the solution sets are circles, and the intersec-

tion point of the trace of the two solution sets is the solution to the two equations.

According to the calculation, combined with the program written by Mathcad, the source impedance of the noise source is solved, and the results are shown in Fig. 13.

As shown in Fig. 13, the source impedance of the prototype is capacitive in the middle and low frequency bands of the conduction band and sensitive in the high frequency band. Therefore, the high-frequency equivalent capacitance model can be used to characterize the internal impedance characteristics of the noise source.

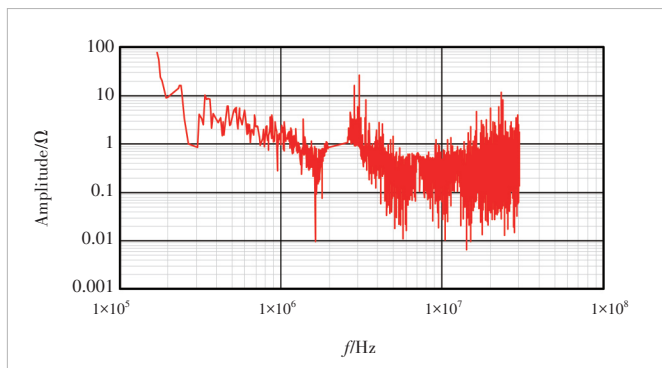
The corresponding source impedance expression is shown as follows:

$$Z_s = R_s + \frac{1}{j\omega C_s} + j\omega L_s \quad (20)$$

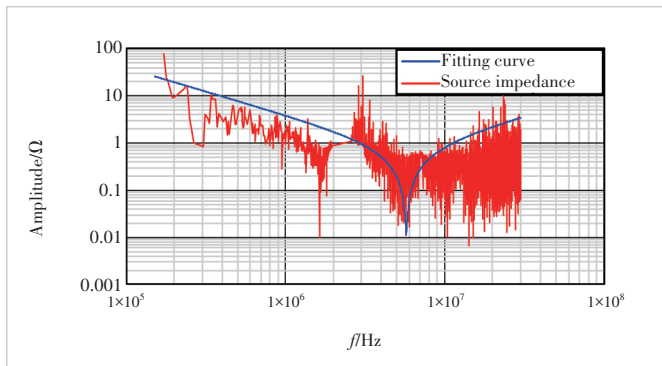
The circuit parameters are obtained by mathematical fitting: $R_s=0.011 \Omega$, $C_s=41.6 \text{ nF}$, and $L_s=18.7 \text{ nH}$. The fitting results for the circuit parameters are shown in Fig. 14.

According to Fig. 14, source impedance is approximately capacitive in the range of 150 kHz to 5.7 MHz. In the range of 5.7 MHz to 30 MHz, the internal impedance of the CM noise source is approximately inductive.

In order to verify the accuracy of the insertion-loss method to extract the internal impedance of the noise source, the Y capacitors of 220 nF and 470 nF were connected to the proto-



▲ Figure 13. Source impedance of common mode (CM) noise source



▲ Figure 14. Fitting results for the circuit parameters

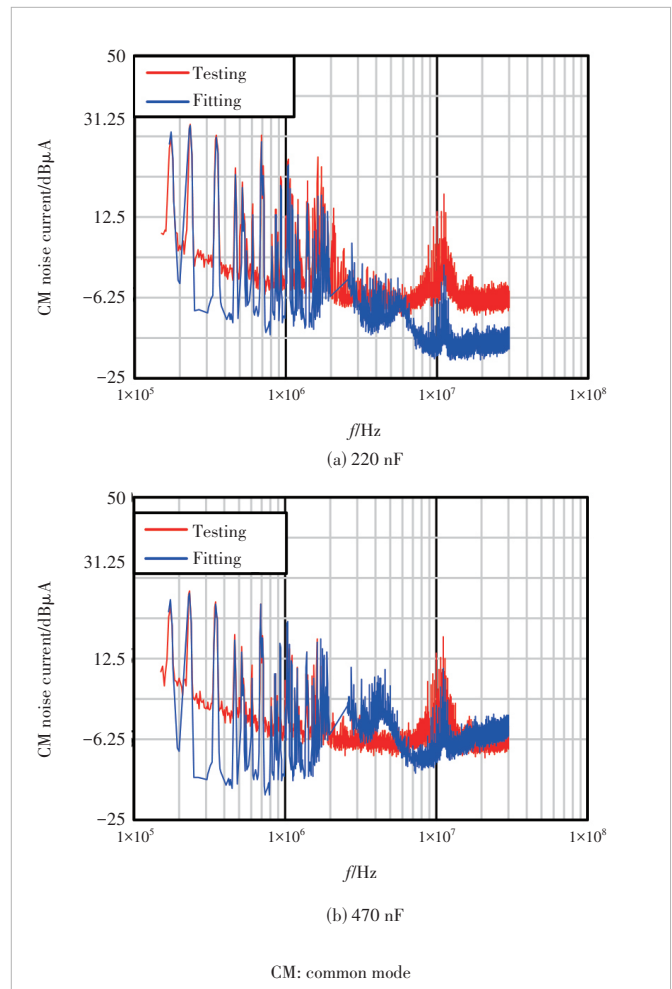
type respectively, and the corresponding CM noise spectrum was measured experimentally. The result is shown in Fig. 15.

In Fig. 15, the red curve is the CM noise spectrum measured when the Y capacitor is connected, and the blue curve is the CM noise spectrum calculated by the insertion-loss method. This experiment verifies the accuracy of insertion-loss method in extracting the source impedance parameters of the noise source.

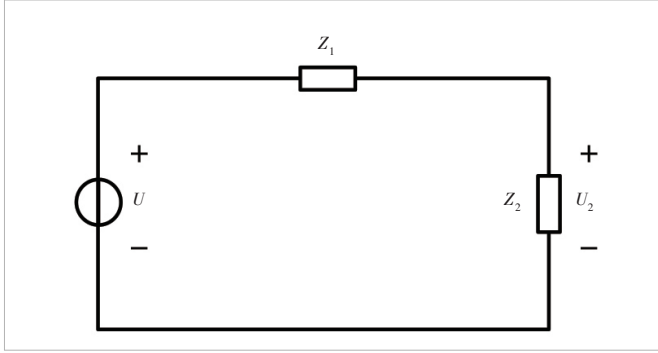
5 Impedance Mismatch Principle

When impedance is matched, the signal can be transmitted efficiently and completely. For filters, impedance mismatch must be made. The transmission efficiency of the interference signal is the lowest when impedance is mismatched. Taking Fig. 16 for example, mismatch is achieved by adjusting the values of Z_1 and Z_2 to adjust voltage distribution. The voltage U_2 assigned to Z_2 is:

$$U_2 = U \frac{Z_2}{Z_1 + Z_2} \quad (21)$$



▲ Figure 15. Verification of insertion-loss method



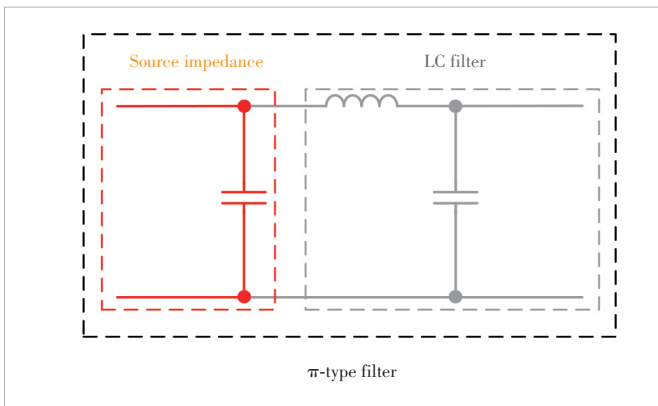
▲ Figure 16. Impedance mismatch

If U is the noise source, Z_1 is the filter impedance, and Z_2 is the load impedance, it is hoped that the filter impedance can withstand more interference voltage. Usually, the load impedance cannot be changed. In this case, the value of Z_1 should be increased, so that Z_1 is much larger than Z_2 , to ensure that the load is less disturbed.

If Z_1 represents the source impedance and Z_2 represents the filter impedance, in this case it is necessary to ensure that the noise that enters the filter is as little as possible, and the source impedance cannot be changed, and then Z_2 should be much less than Z_1 .

The conclusion can be drawn from the above. On the load side, the impedance of the filter should be higher than that of the load; on the noise source side, the impedance of the filter should be lower than the load impedance. Usually in the EMI filter, the inductor in series indicates high resistance, and the capacitor in parallel indicates low resistance.

The source impedance of the prototype noise source is capacitive in the low frequency band, corresponding to low resistance, so the LC-type (one inductor and one capacitor) filter should be selected. The source impedance of the noise source can be regarded as a part of the filter, that is, the noise source impedance side of the LC filter is parallel to an internal impedance capacitor, which can be equivalent to a π -type filter, as shown in Fig. 17. Therefore, in order to better filter effect, you can choose a π -type filter.



▲ Figure 17. Filter diagram

6 Filter Design

EMI filters play a role in the conductive EMI frequency range of the lower frequency band (from 150 kHz to several megahertz), so the filter design is mainly for this frequency band design. In the design of the EMI filter, if the input of the device is an alternating current (AC) source, it is necessary to select the appropriate Y capacitor according to the AC leakage current limit of the corresponding device. In this paper, the input of the wireless base station power supply prototype is direct current (DC) input, so the AC leakage current limit is not considered. In the design of the EMI filter, the capacitance of the Y capacitor can be used to reduce the inductance of the required CM inductor, reduce the volume of the CM choke, and achieve the compact design of the filter.

The filter of the prototype adopts a two-stage and two-channel π -type filter, and the CM choke adopts double-wire and winding, with the line diameter of 1.1 mm. The single-channel and two-channel co-currents are 26 A and 52 A, and the current density is about 16.522 A/mm². The circuit of the original filter is shown in Fig. 2.

Fig. 18(a) represents the CM noise equivalent circuit with a first-order π -type filter, C_1 and C_2 represent the Y capacitors, L_{cm} represents the CM inductance component of the CM choke, and R_{cm} represents the LISN equivalent resistance. C_s indicates the impedance within the noise source. The noise source is transformed with the source impedance part to make the power supply equivalent, and the voltage source is equated with the current source, and the equivalent circuit diagram is obtained as shown in Fig. 18(b).

In Fig. 19, the noise voltage U_{cm1} across R_{cm} in the absence of the filter is defined as:

$$U_{cm1} = \frac{R_{cm} \cdot I_s}{j\omega C_s R_{cm} + 1}. \quad (22)$$

The parallel connection of $2C_2$ and C_s is set to C_3 , and the noise voltage U_{cm2} across R_{cm} after accessing the π -type filter is:

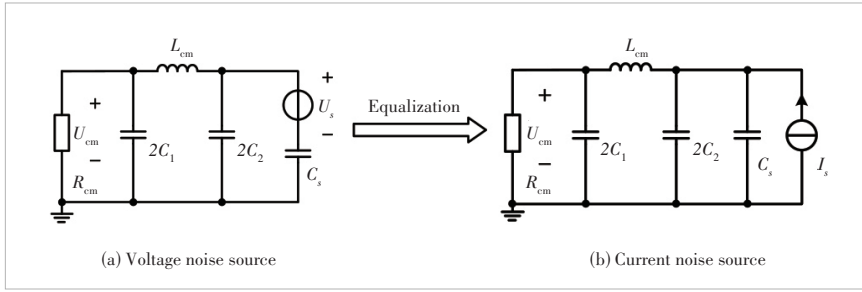
$$U_{cm2} = \frac{R_{cm} \cdot I_s}{1 - \omega^2 L_{cm} (C_3) + j\omega R_{cm} (2C_1 + C_3) - j2\omega^3 R_{cm} L_{cm} C_1 C_3}. \quad (23)$$

Then the CM IL of the first order π -type filter is:

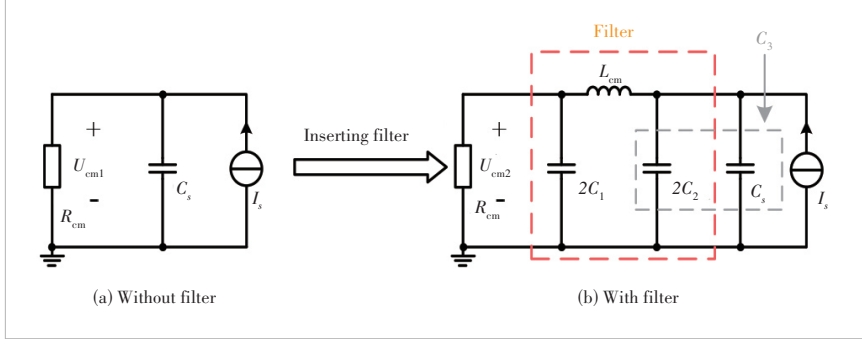
$$IL_{cm} = 20 \log \left| \frac{1 - \omega^2 L_{cm} C_3 + j\omega R_{cm} (2C_1 + C_3) - j2\omega^3 R_{cm} L_{cm} C_1 C_3}{j\omega R_{cm} C_s + 1} \right|. \quad (24)$$

Let $s=j\omega$ and substitute it into Eq. (24) to obtain the IL expression for the first-order π -type filter, we can have:

$$IL_{cm}(s) = 20 \log \left| \frac{2R_{cm} L_{cm} C_1 C_3 \cdot s^3 + L_{cm} C_3 \cdot s^2 + R_{cm} (2C_1 + C_3) \cdot s + 1}{R_{cm} C_s \cdot s + 1} \right|. \quad (25)$$



▲ Figure 18. Common mode (CM) noise equivalent model of a first-order π -type filter



▲ Figure 19. Equivalent model of common mode (CM) noise before and after connecting to the filter

We extract the items containing “s” in Eq.(25) and define them as follows:

$$IL_{cm1} = 20 \log \left| \frac{2R_{cm} L_{cm} C_1 C_3 \cdot s^3}{R_{cm} C_s \cdot s + 1} \right|, \quad (26)$$

$$IL_{cm2} = 20 \log \left| \frac{L_{cm} C_3 \cdot s^2}{R_{cm} C_s \cdot s + 1} \right|, \quad (27)$$

$$IL_{cm3} = 20 \log \left| \frac{R_{cm} (2C_1 + C_3) \cdot s + 1}{R_{cm} C_s \cdot s + 1} \right|. \quad (28)$$

Eqs. (26) – (28) represent the IL of the filter in the high, medium and low frequency bands, respectively. According to Eq. (27), the IL of the first-order π -type filter in the conduction band is mainly determined by Eq. (26) with a slope of 40 dB/dec.

After the IL of the first-order π -type filter is determined, the IL required to pass the EMI test is obtained by comparing the original noise spectrum with the EMI standard limits, as shown in Eq. (29):

$$IL = \text{RAW NOISE} - \text{LIMIT} + \text{MARGIN}. \quad (29)$$

Fig. 20(a) shows the original CM noise spectrum measured by the current method, and the unit used in the EN55032 standard is $\text{dB}\mu\text{V}$. The current is thus converted into a voltage quantity, and the conversion relationship is as follows:

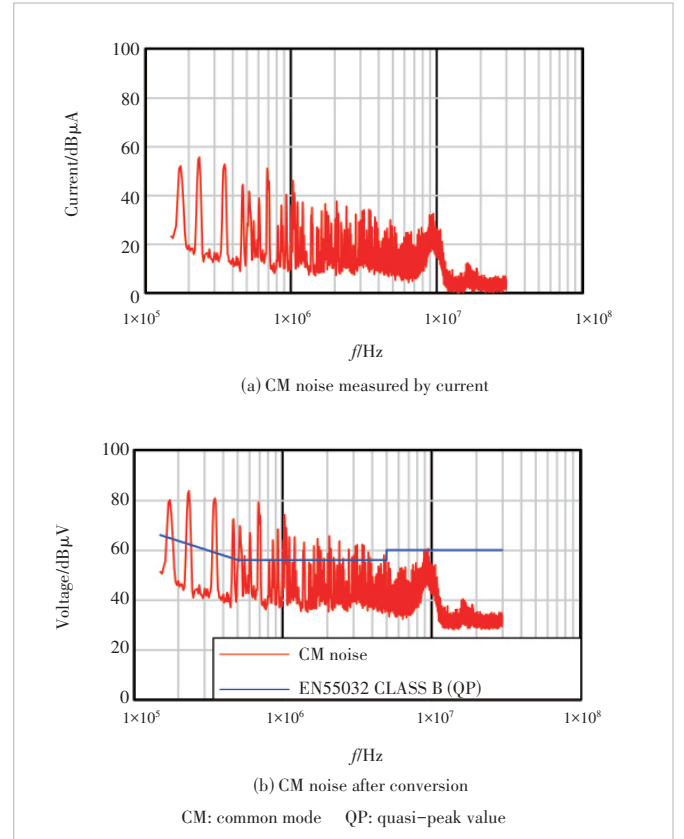
$$20 \log |U_{cm}| = 20 \log |I_{cm} \cdot R_{cm}| \approx 20 \log |I_{cm}| + 28 \text{ dB}. \quad (30)$$

The CM noise voltage spectrum after conversion is obtained according to Eq. (30), as shown in Fig. 20(b).

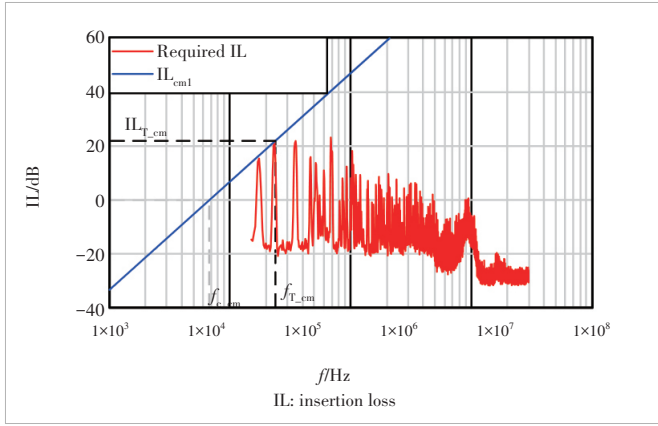
According to EN55032 CLASS B, the bare noise of the prototype is subtracted from the EMI standard limit, and the difference spectrum between the noise and the limit is obtained as shown in Fig. 21, and the difference is the minimum IL required.

In Fig. 21, $f_{T_{cm}}$ indicates the frequency of IL_{cm1} and the tangent point of the required IL, and $IL_{T_{cm}}$ corresponding to the tangent point indicates the minimum CM IL required by the prototype. According to Fig. 21, the minimum CM IL required by the prototype can be determined to be about 21.798 dB, while the IL of the actual filter design needs to be greater than this value to ensure the filter margin of the actual filter. To ensure that the filter has sufficient filtering effect, the filter insertion loss IL_{cm} is 40 dB. Based on the slope, the turning frequency of the curve $f_{c_{cm}}$ is obtained as follows:

$$f_{c_{cm}} = 10^{\frac{-IL_{cm}}{40}} \cdot f_{T_{cm}} \approx 23.4 \text{ kHz}. \quad (31)$$



▲ Figure 20. Measured CM noise spectrum



▲ Figure 21. Required common mode (CM) minimum IL spectrogram

For C_1 and C_2 , we take 100 nF. According to Eq. (32), the inductance of the CM filter L_{cm} is about 300.7 μ H. The actual CM inductance is set to 400 μ H to ensure sufficient filtering effect.

$$L_{cm} = \frac{2\pi R_{cm} C_s f_{c_cm} + 1}{16\pi^3 R_{cm} C_1 C_3 f_{c_cm}^3}. \quad (32)$$

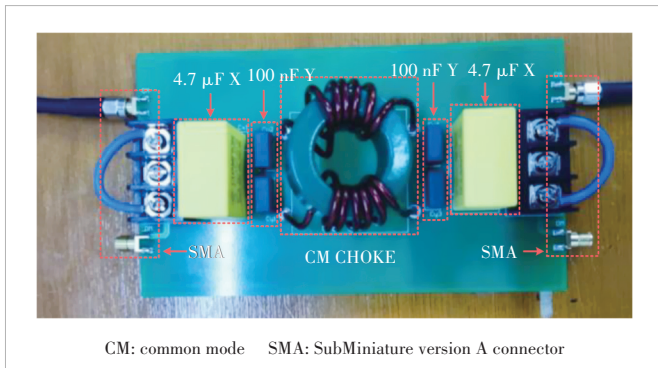
7 Experiment

Based on the calculation and analysis, the first-order π -type filter of the prototype is designed, as shown in Fig. 22.

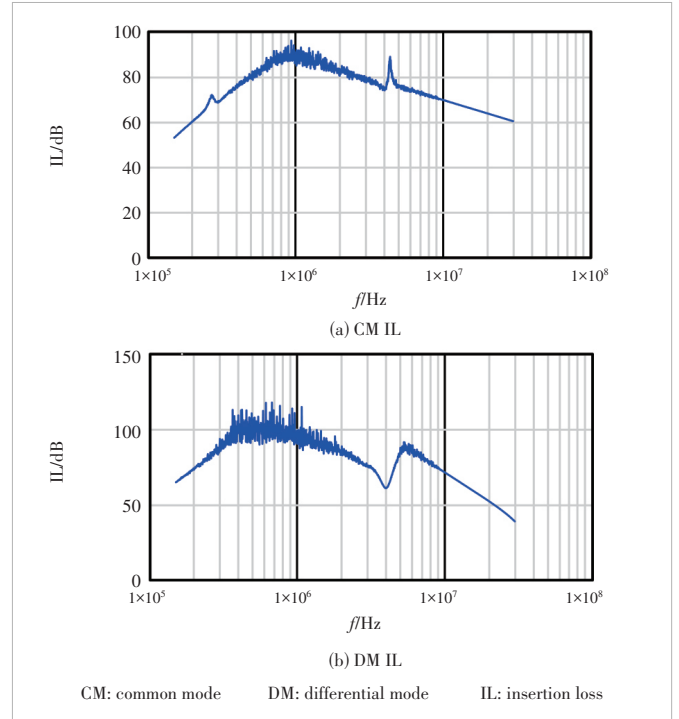
The IL of the new filter is measured using a network analyzer, KEYSIGHT model E5072A. The CM and DM IL of the filter are measured separately and the IL spectrum is shown in Fig. 23.

The EMI filter is designed to meet the design requirements for low and medium frequency bands.

The experimental environment in the filter effect verification experiment is different from the one in the previous paper, so the bare noise of the prototype needs to be measured again in the new environment. The receiver model used in the new experimental environment is R&S ESCI3, the LISN is Schwarzbeck NNLK 8129, and the current probe is FCC F-35A. The noise spectrum of the wireless base station power supply prototype connected to the original filter is first mea-



▲ Figure 22. First-order π -type filter



▲ Figure 23. Measured IL of first-order π -type filter

sured as shown in Fig. 24.

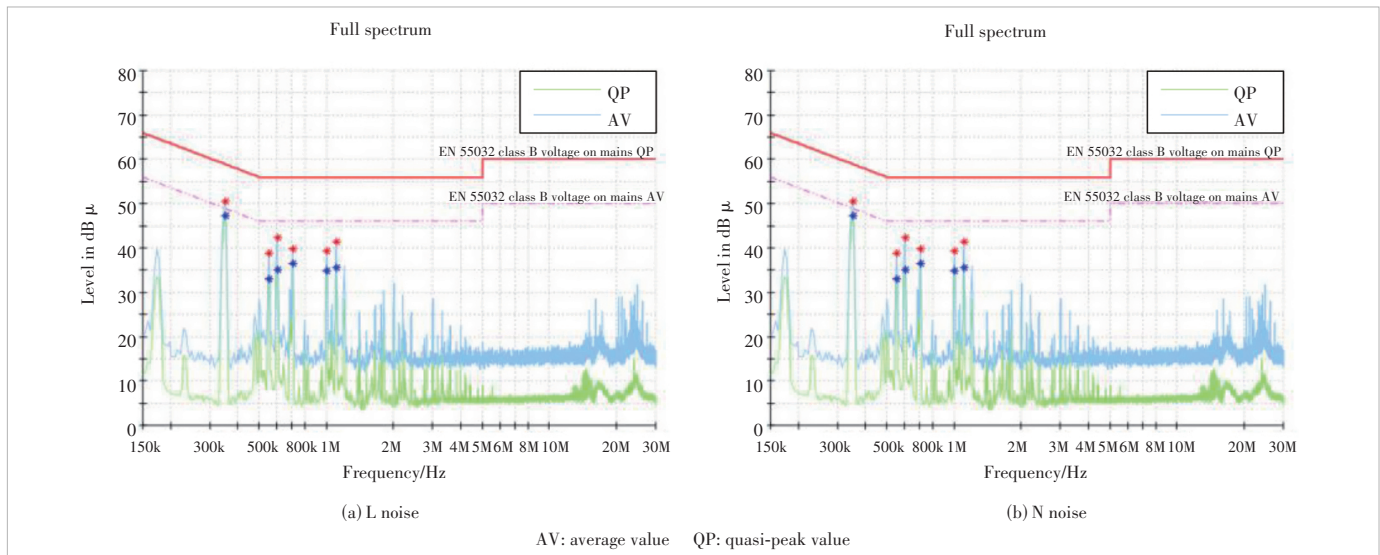
The noise spectrum of the prototype without the original filter is shown in Fig. 25. After the designed filter is connected, the noise spectrum of the wireless base station power supply prototype is shown in Fig. 26.

According to the spectrum, which can be seen in the CM noise spectrum and DM noise spectrum, 150 kHz - 10 MHz noise has been basically reduced to the bottom of the EMI standard line. There is a small spike at 10 MHz, but it does not exceed the limits of EN 55032 Class B, and the conductive EMI filter noise suppression effect has been verified. The designed filter, without degrading the noise suppression effect, reduces the board area by 60% compared with the original filter, significantly reducing the size of the filter and production costs.

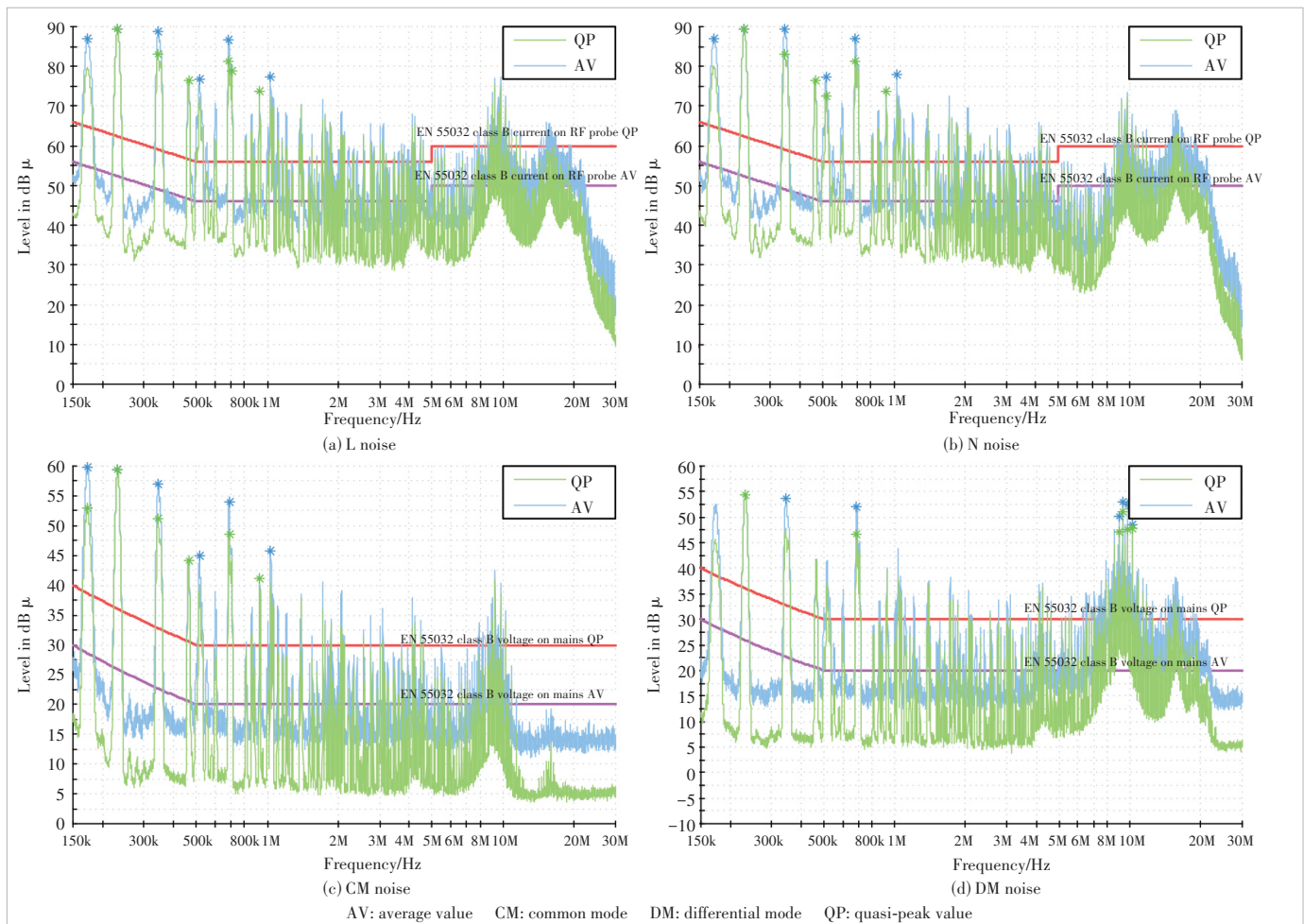
8 Conclusions

A practical CM EMI filter design procedure for a wireless base station power supply is presented. The filtering effect of the original filter is determined by comparing the noise spectrum before and after accessing the original filter, and the unsatisfactory filtering effect of the original filter is pointed out by combining the IL of the original filter with the theoretical calculation. The effect of the ground impedance in the original filter on the IL is verified by theoretical analysis and simulation.

For the design of the EMI filter, the IL method is used to extract the source impedance of the noise source for a wireless base station power supply prototype, and the effectiveness of the source impedance extraction is verified by experiments. The circuit parameters of the EMI filter are designed with a



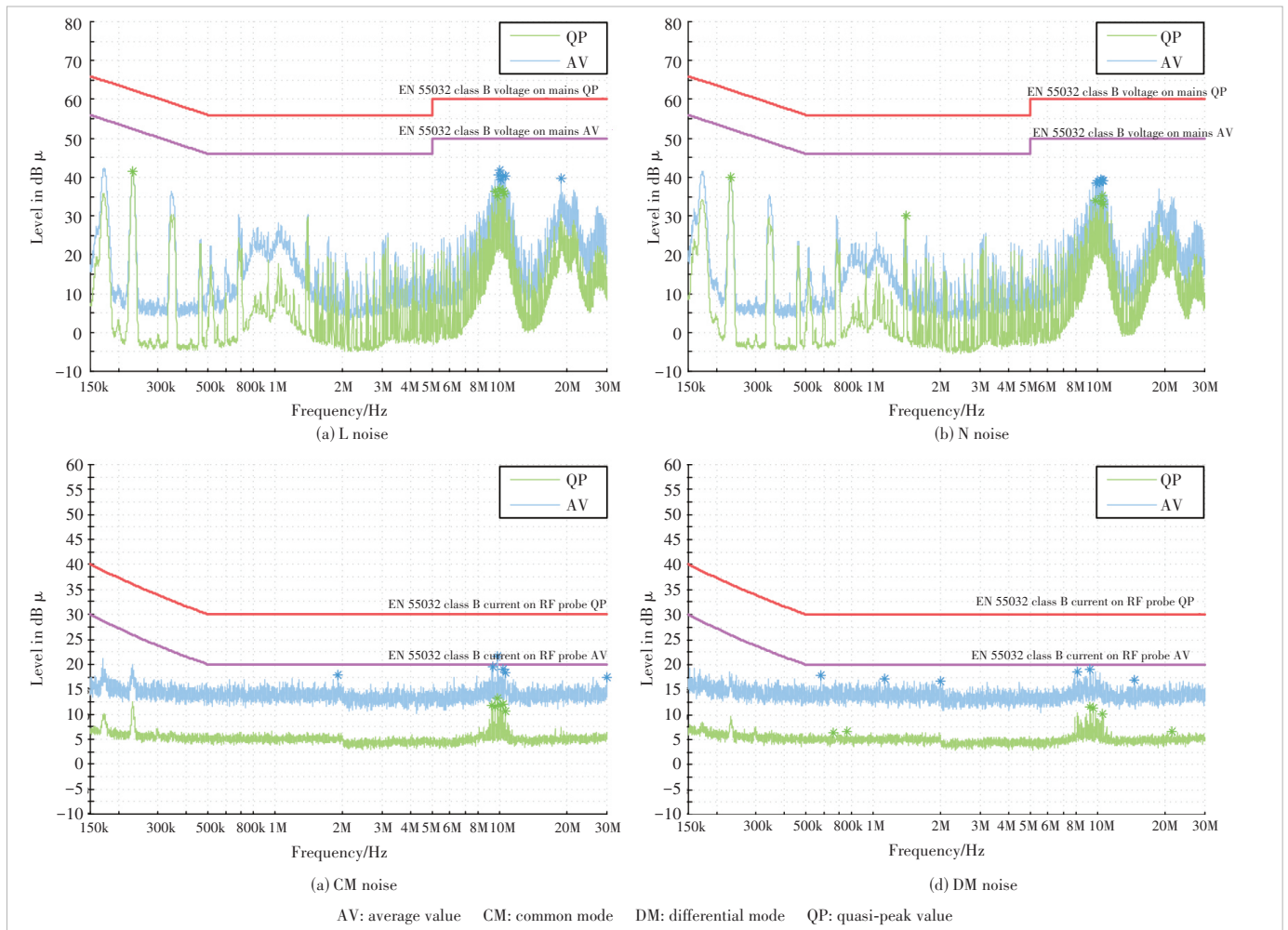
▲ Figure 24. Noise spectrum of the prototype with original filter



▲ Figure 25. Prototype removes raw noise from the filter

wireless base station power supply prototype, and the effectiveness of the filter design is verified by simulation and experimental measurements. Experimental results show that the de-

signed filter improves the noise rejection effect in the low and medium frequency bands by about 12 dB compared with the original filter, and the occupied board area is reduced by 60%.



▲ Figure 26. Noise spectrum of the prototype access filter

Reference

- [1] ZHANG D B, CHEN D Y, SABLE D. A new method to characterize EMI filters [C]//The thirteenth Annual Applied Power Electronics Conference and Exposition. IEEE, 2002: 929 - 933. DOI: 10.1109/APEC.1998.654009
- [2] CHEN H L, QIAN Z M, ZENG Z H, et al. Modeling of parasitic inductive couplings in a PI-shaped common mode EMI filter [J]. IEEE transactions on electromagnetic compatibility, 2008, 50(1): 71 - 79. DOI: 10.1109/TEMC.2007.913214
- [3] SHIH F Y, CHEN D Y, WU Y P, et al. A procedure for designing EMI filters for AC line applications [J]. IEEE transactions on power electronics, 1996, 11(1): 170 - 181. DOI: 10.1109/63.484430
- [4] KOTNY J L, DUQUESNE T, IDIR N. Design of EMI filters for DC-DC converter [C]//Proceedings of IEEE Vehicle Power and Propulsion Conference. IEEE, 2010: 1 - 6. DOI: 10.1109/VPPC.2010.5729047
- [5] HSIEH H I. A procedure including mix-mode noise for designing EMI filters for off-line applications [C]//Proceedings of IEEE Vehicle Power and Propulsion Conference. IEEE, 2008: 1 - 6. DOI: 10.1109/VPPC.2008.4677781
- [6] KOTNY J L, MARGUERON X, IDIR N. High frequency modeling method of EMI filters [C]//Proceedings of IEEE Energy Conversion Congress and Exposition. IEEE, 2009: 1671 - 1678. DOI: 10.1109/ECCE.2009.5316241
- [7] WANG S, LEE F C, ODENDAAL W G, et al. Improvement of EMI filter performance with parasitic coupling cancellation [J]. IEEE transactions on power electronics, 2005, 20(5): 1221 - 1228. DOI: 10.1109/TPEL.2005.854069
- [8] SCHNEIDER L M. Noise source equivalent circuit model for off-line converters and its use in input filter design [C]//Proceedings of IEEE International Symposium on Electromagnetic Compatibility. IEEE, 1983: 1 - 9. DOI: 10.1109/ISEMC.1983.7567390
- [9] WANG S, LEE F C, ODENDAAL W G. Characterization and parasitic extraction of EMI filters using scattering parameters [J]. IEEE transactions on power electronics, 2005, 20(2): 502 - 510. DOI: 10.1109/TPEL.2004.842949

Biographies

LI Wei (li.wei27@zte.com.cn) received his MS degree in mechatronic engineering from Southeast University, China in 2017. Currently he is working as an EMC engineer in ZTE Corporation. His research interests include switching power supply of EMC and lightning protection.

GUO Wei received his master's degree from Xidian University, China. He currently serves as the director of electromagnetic compatibility for wireless products with ZTE Corporation.

WANG Zhida graduated with a master's degree from South China University of Technology. He currently serves as the head of the Wireless Product Reliability Department, ZTE Corporation.



Real-Time 4-Mode MDM Transmission Using Commercial 400G OTN Transceivers and All-Fiber Mode Multiplexers

REN Fang¹, LI Yidan¹, YE Bing², LIU Jianguo²,
CHEN Weizhang²

(1. School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China;
2. Department of WDM System Design, ZTE Corporation, Beijing 100029, China)

DOI: 10.12142/ZTECOM.202401012

<http://kns.cnki.net/kcms/detail/34.1294.TN.20240226.1342.004.html>,
published online February 26, 2024

Manuscript received: 2023-06-15

Abstract: Weakly-coupled mode division multiplexing (MDM) technique is considered a promising candidate to enhance the capacity of an optical transmission system, in which mode multiplexers/demultiplexers (MMUX/MDEMUX) with low insertion loss and modal crosstalk are the key components. In this paper, a low-modal-crosstalk 4-mode MMUX/MDEMUX for the weakly-coupled triple-ring-core few-mode fiber (TRC-FMF) is designed and fabricated with side-polishing processing. The measurement results show that a pair of MMUX/MDEMUX and 25 km weakly-coupled TRC-FMF MDM link achieve low modal crosstalk of lower than -17.5 dB and insertion loss of lower than 11.56 dB for all the four modes. Based on the TRC-FMF and all-fiber MMUX/MDEMUX, an experiment for 25 km real-time 4-mode $3-\lambda$ wavelength division multiplexing (WDM)-MDM transmission is conducted using commercial 400G optical transport network (OTN) transceivers. The experimental results prove weakly-coupled MDM techniques facilitate a smooth upgrade of the optical transmission system.

Keywords: optical fiber couplers; mode (de)multiplexers; mode division multiplexing transmission

Citation (Format 1): REN F, LI Y D, YE B, et al. Real-time 4-mode MDM transmission using commercial 400G OTN transceivers and all-fiber mode multiplexers [J]. *ZTE Communications*, 2024, 22(1): 106 – 110. DOI: 10.12142/ZTECOM.202401012

Citation (Format 2): F. Ren, Y. D. Li, B. Ye, et al., “Real-time 4-mode MDM transmission using commercial 400G OTN transceivers and all-fiber mode multiplexers,” *ZTE Communications*, vol. 22, no. 1, pp. 106 – 110, Mar. 2024. doi: 10.12142/ZTECOM.202401012.

1 Introduction

The explosive increase of network users and various communication services such as the Internet of Things (IoT) and cloud computing has constantly requested the capacity enhancement of optical transmission systems and networks. However, recent studies have shown that the capacity usage of single-mode fiber (SMF) is tending to its nonlinear Shannon limit by combining various advanced transmission techniques^[1]. Mode division multiplexing (MDM) technique utilizing linearly-polarized (LP) modes in few-mode fibers (FMFs) has attracted great attention as a promising approach to enhancing the capacity of optical fiber communication systems^[2-3], but the inevitable modal crosstalk induced by the mode multiplexer/demultiplexer (MMUX/MDEMUX) and optical fiber transmission is hard to be suppressed. Two technical routes have been widely investigated to solve this problem. In a strongly-coupled MDM strategy without modal crosstalk

suppression^[4-6], coherent detection and massive $N \times N$ ($N > 2$) multiple-input multiple-output (MIMO) digital signal processing (DSP) have to be utilized to equalize each MDM channel at the receiver for achieving long-distance high-speed MDM transmission. However, the computational complexity and cost for MIMO-DSP can increase dramatically with longer FMF links and more LP modes, which is not compatible with commercial coherent optical modules utilizing polarization division multiplexing (PDM) and only 2×2 MIMO-DSP^[7]. Another promising MDM strategy is weakly-coupled transmission approaches that suppress the modal crosstalk from MMUX/MDEMUX and FMFs as much as possible and are compatible with conventional intensity modulation/direct detection (IM/DD) transceivers and commercial high-speed coherent optical modules^[8-9]. In this route, multiple-ring-core (MRC) weakly-coupled FMFs are designed to effectively suppress the distributed modal crosstalk (DMC) among different LP modes by enlarging effective index difference (Δn_{eff}) between different LP modes^[10]. Moreover, a 3-circular-mode real-time MDM transmission utilizing 200G quadrature phase shift keying (QPSK) optical transport network (OTN) transceiver-

This work was supported in part by the ZTE Industry-University-Institute Cooperation Funds.

ers over 300 km weakly-coupled MRC-FMF has been experimentally demonstrated^[9].

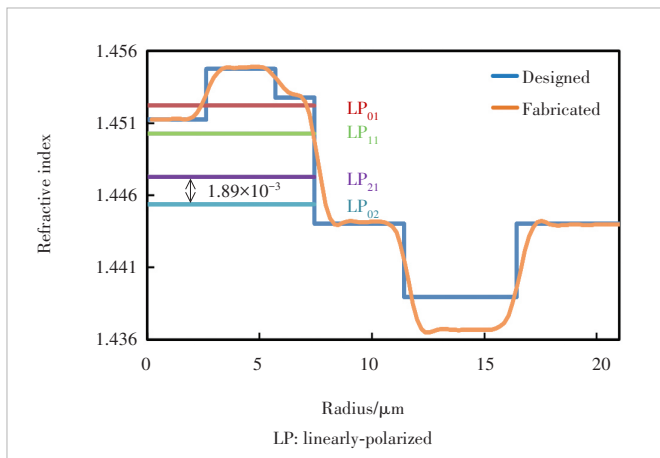
In addition to weakly-coupled FMF, the low-modal-crosstalk MMUX/MDEMUX with low insertion loss (IL) and low-modal crosstalk is also a key component in the MDM transmission link^[11]. Among various types of MMUXs and MDEMUXs, the one consisting of cascaded mode-selective couplers (MSCs) is promising for weakly-coupled MDM systems for its high conversion purity, high flexibility, and low coupling crosstalk with transmission FMFs^[12-15]. MSCs with side-polishing processing have high modal selectivity and low IL because their waveguide structure is not significantly damaged during the fabrication process. Real-time transmissions utilizing the side-polishing MMUX/MDEMUX and small form-factor pluggable plus (SFP+) optical modules have been experimentally demonstrated^[16-17].

In this paper, we adopt the side-polishing processing to design and fabricate MSCs for a weakly-coupled triple-ring-core (TRC) FMF. A pair of all-fiber low-modal-crosstalk 4-mode MMUX/MDEMUX are completed by cascading side-polished MSCs. The measurement results show that a pair of MMUX/MDEMUX and 25 km weakly-coupled TRC-FMF MDM link achieve low modal crosstalk of lower than -17.5 dB and insertion loss of lower than 11.56 dB for all the four modes. Based on the TRC-FMF and all-fiber MMUX/MDEMUX, a real-time 4-mode $3-\lambda$ 25 km wavelength division multiplexing (WDM)-MDM transmission is experimentally demonstrated using commercial 400G optical transport network (OTN) transceivers.

2 TRC-FMF and MMUX/DEMUX

2.1 Weakly-Coupled TRC-FMF

Fig. 1 shows the designed (blue line) and fabricated (orange line) index profile of the weakly-coupled TRC-FMF and n_{eff} of the supported four LP modes. This FMF is designed with a normalized frequency V of 4.8, and the refractive index differ-



▲ Figure 1. Refractive index profile of the designed triple-ring-core few-mode fiber (TRC-FMF) and n_{eff} of the supported four LP modes

ence between the core and cladding (Δn) is 0.6%. Three rings of index perturbations are applied into the fiber core to separate n_{eff} of different LP modes to suppress the distributed modal crosstalk among different LP modes in fiber transmission. To reduce the bending sensitivity of high-order LP modes, a depressed-index fluorine-doped trench is applied in the cladding. The designed weakly-coupled TRC-FMF is fabricated using the plasma chemical vapor deposition (PCVD) technique, which is widely used in commercial SMF fabrication. Table 1 lists the measured characteristics of the four LP modes at 1550 nm. A $\min|\Delta n_{\text{eff}}|$ up to 1.89×10^{-3} between different LP modes is achieved, which can effectively suppress the modal crosstalk among them. The propagation loss (PL) of each LP mode is all lower than 0.227 dB/km thanks to the depressed-index trench.

2.2 Weakly-Coupled Side-Polishing MMUX/MDEMUX

MMUX and MDEMUX in this work are completed by cascading all-fiber side-polishing MSCs. The FMF of each MSC is the same as the transmission fiber, which can reduce the extra IL and modal crosstalk induced by mode field mismatching at the connection point between the MMUX/MDEMUX and FMF. For phase-matching with high-index LP_{01} and LP_{11} modes of the TRC-FMF, a customized step-index SMF with a core/cladding radius of 2.48/62.5 μm and Δn of 1.23% at 1550 nm is adopted. The designed (blue line) and measured (orange line) index profiles are shown in Fig. 2,

▼ Table 1. Characteristics of the fabricated weakly-coupled TRC-FMF

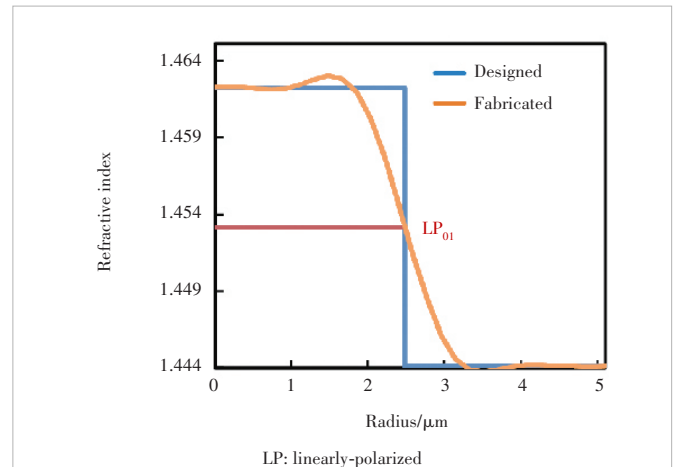
Characteristics	Four LP Modes			
	LP_{01}	LP_{11}	LP_{21}	LP_{02}
n_{eff}	1.452 215	1.450 277	1.447 285	1.445 394
$\Delta n_{\text{eff}} (\times 10^{-3})$	1.94	2.99	1.89	-
PL/(dB/km)	0.215	0.221	0.227	0.223
DC/(ps·nm ⁻¹ ·km ⁻¹)	20	18	23.4	24

DC: dispersion coefficient

PL: propagation loss

LP: linearly-polarized

TRC-FMF: triple-ring-core few-mode fiber

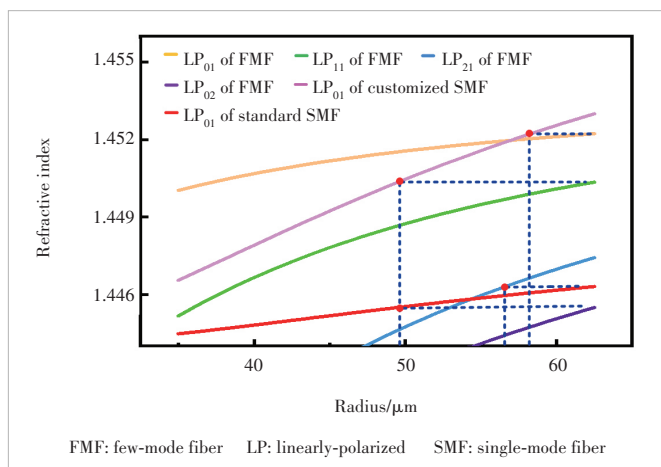


▲ Figure 2. Refractive index profile of the designed single-mode fiber (SMF) and n_{eff} of the LP_{01} mode

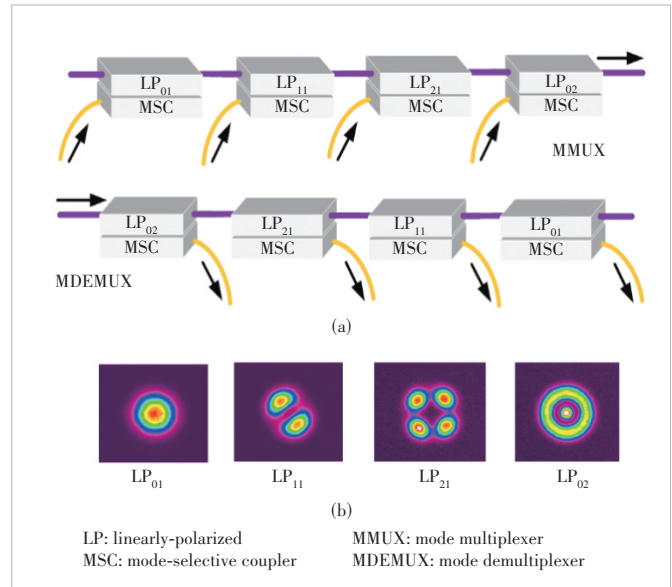
where n_{eff} of the LP_{01} mode of the customized SMF is 1.453 at 1550 nm. And a standard SMF with a core/cladding radius of 4.1/62.5 μm and Δn of 0.36% at 1550 nm is adopted to phase-matching with LP_{21} and LP_{02} modes of the TRC-FMF.

To satisfy the phase-matching condition, FMFs or SMFs are firstly pre-tapered on a fused biconical taper station. Fig. 3 shows n_{eff} of the four LP modes in the TRC-FMF and LP_{01} modes in the customized SMF and standard SMF as functions of the tapered radius, which are calculated by COMSOL multiphysics and Matlab based on the fabricated index profiles. The cross points of the dotted lines indicate the proper tapering radii for fabricating each MSC. We can see that the radius of the customized SMF needs to be pre-tapered to 58.5/49.2 μm respectively for the LP_{01} and LP_{11} MSCs, while the radius of the FMF is supposed to be pre-tapered to 57 μm for the LP_{21} MSC. For phase-matching with LP_{02} modes, the radius of the standard SMF should be pre-tapered to 49 μm . Then the FMFs and SMFs are respectively embedded into quartz blocks and polished on the grinding platform (SW-22C, SAMWELL). After polishing, the two half couplers are mated together to form the MSC.

The LP_{01} , LP_{11} , LP_{21} and LP_{02} MSCs are then cascaded orderly by fusion splicing to form a 4-mode MMUX, and the LP_{02} , LP_{21} , LP_{11} and LP_{01} MSCs are cascaded orderly by fusion splicing to form the 4-mode DEMMUX, which is shown in Fig. 4(a). It should be noted that since a single MSC can only demultiplex one spatial component of each pair of degenerate modes, a mode rotator is utilized before demultiplexing each degenerate mode to change its spatial orientation and achieve efficient demultiplexing. The output mode fields of the MMUX is presented in Fig. 4(b) and Table 2 shows IL and the modal crosstalk matrix of a pair of MMUX/MDMUX and 25 km weakly-coupled TRC-FMF. We can find that the maximum IL of the 25 km MDM link is 11.56 dB (LP_{21} mode) and the largest relative modal crosstalk, which is defined as the power ra-



▲ Figure 3. Effective index of the four LP modes in a triple-ring-core few-mode fiber (TRC-FMF) and the LP_{01} mode in two kinds of SMFs versus different tapered radii



▲ Figure 4. Diagrams of (a) MMUX/MDMUX and (b) output mode fields of MMUX

▼ Table 2. IL and modal crosstalk matrix of 25 km MDM link at 1550 nm

Input Mode	Output Power/dB			
	LP_{01}	LP_{11}	LP_{21}	LP_{02}
LP_{01}	-6.98	-27.18	-36.05	-27.85
LP_{11}	-30.52	-9.58	-32.06	-30.17
LP_{21}	-35.91	-31.78	-11.56	-33.49
LP_{02}	-26.89	-28.73	-31.87	-10.34

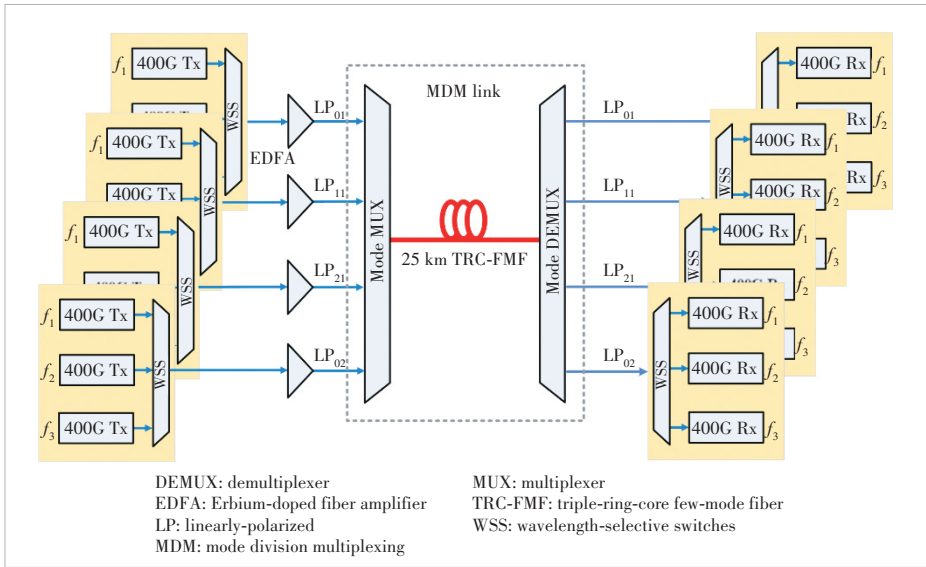
IL: insertion loss LP: linearly-polarized MDM: mode division multiplexing

tio between the crosstalk and signal for any two LP modes, is -17.5 dB (for the case of LP_{01} input and LP_{02} output).

3 Experimental Setup and Results

3.1 Experimental Setup

Real-time MDM-WDM transmission utilizing commercial 400G OTN transceivers and all-fiber side-polishing MMUX/MDMUX was experimentally conducted with the experimental setup shown in Fig. 5. The 400G OTN transceivers used dual-polarization (DP) 16 Quadrature Amplitude Modulation (QAM) with a baud rate of 69 Gbit/s (27% overhead). Twelve sets of 400G transceivers were configured with three wavelengths for each LP mode. The three central wavelengths for each mode were chosen to be 1529.528 nm, 1550.892 nm and 1564.649 nm to evaluate the characteristics of our transmission system over the C-band respectively. At the transmitter, four wavelength-selective switches (WSS) are utilized as wavelength multiplexers (WMUX) for each LP mode. Four Erbium-doped fiber amplifiers (EDFA) are followed to adjust the input power for balancing the span loss of the four LP modes in the MDM link. Then the wavelength-multiplexed signals are multiplexed by the 4-mode MMUX. After 25 km



▲ Figure 5. Experimental setup of the $3\text{-}\lambda \times 4\text{-mode} \times 400\text{ Gb/s}$ transmission

transmission, the received signal is demultiplexed by the 4-mode MDEMUX first. Another four WSSes are then utilized to demultiplex the received wavelength-multiplexed signals. Finally, each demultiplexed signal is detected by the 400G OTN receivers for real-time bit-error rate (BER) calculation.

3.2 Transmission Results and Discussion

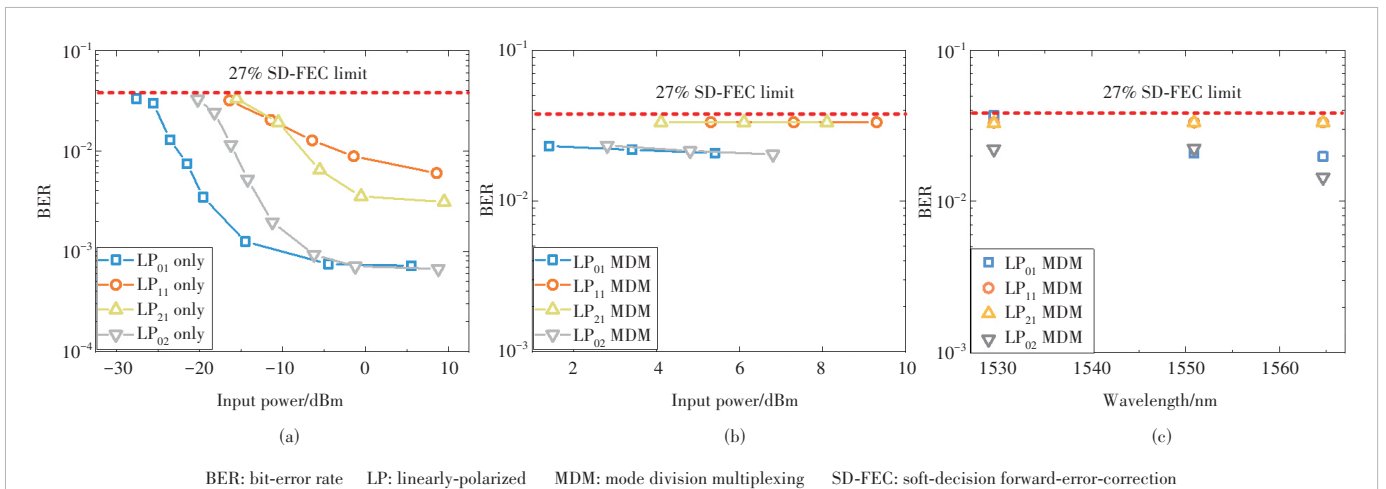
We input each LP mode one by one and measure the BER at different input power. The results are shown in Fig. 6(a). We can see that two non-degenerate modes (LP_{01} and LP_{02}) have low BERs while the BERs of two degenerate modes (LP_{11} and LP_{21}) are relatively large. This is because the degenerate modes have 4-fold degeneracy and suffer from large intra-LP-mode dispersion (ILMD), which is also the main reason that limits the transmission distance in this experiment^[18]. This part of the damage can be mitigated by optimizing the optical

fiber processing methods, which will achieve a longer transmission distance. Then the four LP modes are launched simultaneously and the curves of BER for all the four modes versus input power are depicted in Fig. 6(b). We can see that the BERs of the degenerate modes are still larger than those of the non-degenerate modes but the BERs are all below the 27%-overhead soft-decision forward-error-correction (SD-FEC) limit (3.7×10^{-2}) of the OTN transceivers. Fig. 6(c) shows the BERs of $3\text{-}\lambda \times 4\text{-mode} \times 400\text{G}$ simultaneous transmission. The input power of LP_{01} , LP_{11} , LP_{21} , and LP_{02} modes to the MDM link is 5.4 dBm, 9.3 dBm, 8.1 dBm, and 6.8 dBm for power balancing. We can find that the BERs of

the four LP modes at the three wavelengths across the C-band are all under the SD-FEC limit. These results indicate that our MDM transmission system has a wide operating bandwidth and can realize a 4-mode real-time transmission of more than 40 wavelengths with 100 GHz channel spacing over the C-band. By further optimizing the bandwidth of the MMUX/MDEMUX, we can achieve the 4-mode real-time transmission with more than 80 wavelengths using the same experimental setup.

4 Conclusions

In this paper, a low-modal-crosstalk 4-mode MMUX and MDEMUX for weakly-coupled TRC-FMF is designed and fabricated with side-polishing processing. The measurement results show that a pair of MMUX/MDEMUX and 25 km weakly-coupled TRC-FMF MDM link achieve low modal crosstalk of



▲ Figure 6. Transmission results: (a) BER curves of four LP modes in only by one transmission, (b) BER curves of 4 LP modes in MDM transmission, and (c) BERs of $3\text{-}\lambda \times 4\text{-mode} \times 400\text{G}$ simultaneous transmission

lower than -17.5 dB and insertion loss of lower than 11.56 dB for all the four modes. Based on the TRC-FMF and all-fiber MMUX/MDEMUX, a real-time 4-mode $3-\lambda$ 25 km MDM-WDM transmission is experimentally demonstrated using commercial 400G OTN transceivers, which proves the smooth upgrade of high-speed coherent transmission scenarios is possible by using weakly-coupled MDM techniques.

Acknowledgement:

The authors wish to thank Dr. CUI Jian of Peking University, China. Part of this work was published in ACP 2022.

References

- [1] RICHARDSON D J. Filling the light pipe [J]. *Science*, 2010, 330(6002): 327 – 328. DOI: 10.1126/science.1191708
- [2] RICHARDSON D J, FINI J M, NELSON L E. Space-division multiplexing in optical fibres [J]. *Nature photonics*, 2013, 7: 354 – 362. DOI: 10.1038/nphoton.2013.94
- [3] LI G F, BAI N, ZHAO N B, et al. Space-division multiplexing: the next frontier in optical communication [J]. *Advances in optics and photonics*, 2014, 6(4): 413 – 487
- [4] RYF R, RANDEL S, GNAUCK A H, et al. Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6 MIMO processing [J]. *Journal of lightwave technology*, 2012, 30(4): 521 – 531. DOI: 10.1109/jlt.2011.2174336
- [5] WANG Y Z, GAO T Y, LIU Y P, et al. Novel mirror-flipped mode permutation technique for long-haul mode-division multiplexing transmissions [C]/Proc. Optical Fiber Communication Conference (OFC). Optica Publishing Group, 2022. DOI: 10.1364/ofc.2022.m4b.5
- [6] RADEMACHER G, LUIS R S, PUTTNAM B J, et al. 1.53 Peta-bit/s C-band transmission in a 55-mode fiber [C]/European Conference on Optical Communication (ECOC). IEEE, 2022
- [7] SHEN L, GE D W, SHEN S K, et al. 16-Tb/s real-time demonstration of 100-km MDM transmission using commercial 200G OTN system [C]/Proc. Optical Fiber Communications Conference and Exhibition (OFC). OSA, 2021. DOI: 10.1364/OFC.2021.W11.2
- [8] GAO Y Y, GE D W, SHEN L, et al. Prototype of DSP-free IM/DD MDM transceiver for datacenter interconnection [J]. *Journal of lightwave technology*, 2022, 40(5): 1283 – 1295. DOI: 10.1109/jlt.2021.3123508
- [9] ZUO M Q, GE D W, GAO Y Y, et al. 3-mode real-time MDM transmission using single-mode OTN transceivers over 300 km weakly-coupled FMF [C]/Proc. Optical Fiber Communications Conference (OFC). OSA, 2022. DOI: 10.1364/OFC.2022.M4B.4
- [10] GE D W, GAO Y Y, YANG Y, et al. A 6-LP-mode ultralow-modal-crosstalk double-ring-core FMF for weakly-coupled MDM transmission [J]. *Optics communications*, 2019, 451: 97 – 103. DOI: 10.1016/j.optcom.2019.06.015
- [11] GAO Y Y, CUI J, GE D W, et al. A degenerate-mode-selective coupler for stable DSP-free MDM transmission [J]. *Journal of lightwave technology*, 2019, 37(17): 4410 – 4420. DOI: 10.1109/jlt.2019.2925116
- [12] SONG K Y, HWANG I K, YUN S H, et al. High performance fused-type mode-selective coupler using elliptical core two-mode fiber at 1550 nm [J]. *IEEE photonics technology letters*, 2002, 14(4): 501 – 503. DOI: 10.1109/68.992591
- [13] CHANG S H, CHUNG H S, RYF R, et al. Mode- and wavelength-division multiplexed transmission using all-fiber mode multiplexer based on mode selective couplers [J]. *Optics express*, 2015, 23(6): 7164 – 7172. DOI: 10.1364/OE.23.007164
- [14] IGARASHI K, PARK K J, TSURITANI T, et al. All-fiber-based selective mode multiplexer and demultiplexer for weakly-coupled mode-division multiplexed systems [J]. *Optics communications*, 2018, 408: 58 – 62. DOI: 10.1016/j.optcom.2017.08.049
- [15] CUI J, HUANG S L, GAO Y Y, et al. Real-time 4-mode MDM transmission Using triple-ring-core FMF and all-fiber mode multiplexers [C]/2022 Asia Communications and Photonics Conference (ACP). IEEE, 2022: 364 – 366. DOI: 10.1109/ACP55869.2022.10088884
- [16] CUI J, GAO Y Y, HUANG S L, et al. Low-modal-crosstalk orthogonal combine reception for degenerate modes in IM/DD MDM transmission [J]. *Optics express*, 2023, 31(5): 8586. DOI: 10.1364/oe.482879
- [17] CUI J, GAO Y Y, HUANG S L, et al. Five-LP-mode IM/DD MDM transmission based on degenerate-mode-selective couplers with side-polishing processing [J]. *Journal of lightwave technology*, 2023, 41(10): 2991 – 2998. DOI: 10.1109/JLT.2023.3240877
- [18] GE D W, ZUO M Q, ZHU J L, et al. Analysis and measurement of intra-LP-mode dispersion for weakly-coupled FMF [J]. *Journal of lightwave technology*, 2021, 39(22): 7238 – 7245. DOI: 10.1109/jlt.2021.3110821

Biographies

REN Fang (renfang@ustb.edu.cn) received her bachelor's and master's degrees in electronic science and technology from Tianjin University, China in 2008 and 2010, respectively. She received her PhD degree in information electronics from Hokkaido University, Japan in 2014. She is currently an associate professor with the School of Computer and Communication Engineering, University of Science and Technology Beijing, China. Her research interests include optical communication systems, optical devices, deep learning, and simulation with artificial intelligence.

LI Yidan received her bachelor's degree in communication engineering from Hebei University of Geosciences, China in 2022 and is currently a graduate student at University of Science and Technology Beijing, China. Her current research interests include optical communications.

YE Bing received his master's degree in electronic engineering from Tsinghua University, China. He currently works with the Department of WDM System Design, ZTE Corporation. His research interests include optical networks and systems and next-generation WDM networks.

LIU Jianguo received his master's degree in communication and information system from Huazhong University of Science and Technology, China. He currently works with the Department of WDM System Design, ZTE Corporation. His research interests include intelligent optical networks and next-generation WDM networks.

CHEN Weizhang received his master's degree in communication and information system from Huazhong University of Science and Technology, China. He currently works with the Department of WDM System Design, ZTE Corporation. His research interests include next-generation WDM networks and high-speed integrated photon and optical modules.

The 1st Youth Expert Committee

for Promoting Industry-University-Institute Cooperation

Director CHEN Wei, Beijing Jiaotong University
Deputy Director QIN Xiaoqi, Beijing University of Posts and Telecommunications
LU Dan, ZTE Corporation

Members (Surname in Alphabetical Order)

CAO Jin Xidian University
CHEN Li University of Science and Technology of China
CHEN Qimei Wuhan University
CHEN Shuyi Harbin Institute of Technology
CHEN Wei Beijing Jiaotong University
GUAN Ke Beijing Jiaotong University
HAN Kaifeng China Academy of Information and Communications Technology
HE Zi Nanjing University of Science and Technology
HU Jie University of Electronic Science and Technology of China
HUANG Chen Purple Mountain Laboratories
LI Ang Xi'an Jiaotong University
LIU Chunsen Fudan University
LIU Fan Southern University of Science and Technology
LIU Junyu Xidian University
LU Dan ZTE Corporation
LU Youyou Tsinghua University
NING Zhaolong Chongqing University of Posts and Telecommunications
QI Liang Shanghai Jiao Tong University
QIN Xiaoqi Beijing University of Posts and Telecommunications
QIN Zhijin Tsinghua University
SHI Yinghuan Nanjing University
WANG Jingjing Beihang University
WANG Xinggong Huazhong University of Science and Technology
WANG Yongqiang Tianjin University
WEN Miaowen South China University of Technology
WU Yongpeng Shanghai Jiao Tong University
XIA Wenchao Nanjing University of Posts and Telecommunications
XU Mengwei Beijing University of Posts and Telecommunications
XU Tianheng Shanghai Advanced Research Institute, Chinese Academy of Sciences
YANG Chuanchuan Peking University
YIN Haifan Huazhong University of Science and Technology
YU Jihong Beijing Institute of Technology
ZHANG Jiao Beijing University of Posts and Telecommunications
ZHANG Yuchao Beijing University of Posts and Telecommunications
ZHANG Jiayi Beijing Jiaotong University
ZHAO Yuda Zhejiang University
ZHOU Yi Southwest Jiaotong University
ZHU Bingcheng Southeast University

ZTE COMMUNICATIONS

中兴通讯技术(英文版)

ZTE Communications has been indexed in the following databases:

- Abstract Journal
- Cambridge Scientific Abstracts (CSA)
- China Science and Technology Journal Database
- Chinese Journal Fulltext Databases
- Index Copernicus
- Ulrich's Periodicals Directory
- Wanfang Data
- WJCI 2023

Industry Consultants:

DUAN Xiangyang, GAO Yin, HU Liujun, HUA Xinhai, LIU Xinyang, LU Ping, SHI Weiqiang, TU Yaofeng, WANG Huitao, XIONG Xiankui, ZHAO Yajun, ZHAO Zhiyong, ZHU Xiaoguang

ZTE COMMUNICATIONS

Vol. 22 No. 1 (Issue 86)

Quarterly

First Issue Published in 2003

Supervised by:

Anhui Publishing Group

Sponsored by:

Time Publishing and Media Co., Ltd.

Shenzhen Guangyu Aerospace Industry Co., Ltd.

Published by:

Anhui Science & Technology Publishing House

Edited and Circulated (Home and Abroad) by:

Magazine House of ZTE Communications

Staff Members:

General Editor: WANG Xiyu

Editor-in-Chief: WANG Li

Executive Editor-in-Chief: HUANG Xinming

Editorial Director: LU Dan

Editor-in-Charge: ZHU Li

Editors: REN Xixi, XU Ye, YANG Guangxi

Producer: XU Ying

Circulation Executive: WANG Pingping

Assistant: WANG Kun

Editorial Correspondence:

Add: 12F Kaixuan Building, 329 Jinzhai Road,

Hefei 230061, P. R. China

Tel: +86-551-65533356

Email: magazine@zte.com.cn

Website: <http://zte.magtechjournal.com>

Annual Subscription: RMB 120

Printed by:

Hefei Tiancai Color Printing Company

Publication Date: March 25, 2024

China Standard Serial Number: ISSN 1673-5188
CN 34-1294/TN