UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Multi-dimensional Channel Parameter Estimation for mmWave Cylindrical Arrays

by Zhipeng Lin

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Prof. Ren Ping Liu, Prof. J. Andrew Zhang, and Prof. Tiejun Lv

University of Technology Sydney
Faculty of Engineering and Information Technology

March, 2021

Certificate of Original Authorship

I, Zhipeng Lin, declare that this thesis is submitted in fulfilment of the requirements for the award of PhD degree, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with Beijing University of Posts and Telecommunications.

This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

Signature: Zhipeng Lin

Date: 12/03/2021

ABSTRACT

Millimeter-wave (mmWave) large-scale antenna arrays, standardized for the fifth-generation (5G) communication networks, have the potential to estimate channel parameters with unprecedented accuracy, due to their high temporal resolution and excellent directivity. However, most existing techniques have very high complexities in hardware and software, and they cannot effectively exploit the properties of mmWave large-array systems for channel estimation. As a result, their application in 5G mmWave large array systems is limited in practice.

This thesis develops new and efficient solutions to channel parameter estimation using large-scale mmWave uniform cylindrical arrays (UCyAs). The key contributions of this thesis are on the following four aspects:

We first present a channel compression-based channel estimation method, which reduces the computational complexity substantially at a negligible cost of estimation accuracy. By capitalizing on the sparsity of mmWave channel, the method effectively filters out the useless signal components. As a result, the dimension of the element space of the received signals can be reduced.

Next, we extend the channel estimation to the hybrid UCyA case, and design new hybrid beamformers. By exploiting the convergence property of the Bessel function, the designed beamformers can preserve the recurrence relationship of the received signals with a small number of radio frequency (RF) chains.

We then arrange the received signals in a tensor form and propose a new tensor-based channel estimation algorithm. By suppressing the receiver noises in all dimensions (time, frequency, and space), the algorithm can achieve substantially higher estimation accuracy than existing matrix-based techniques.

Finally, to reduce cost and power consumption while maintaining a high network access capability, we develop a novel nested hybrid UCyA and present the corresponding parameter estimation algorithm based on the second-order channel

statistics. Simulation results show that by exploiting the sparse array technique to design the RF chain connection network, the angles of a large number of devices can be accurately estimated with much fewer RF chains than antennas.

Overall, this thesis presents several applicable UCyA design schemes and proposes the efficient channel parameter estimation algorithms. The presented new UCyAs can significantly reduce the hardware cost of the system with a marginal accuracy loss, and the proposed algorithms are capable of accurately estimating the channel parameters with low computational complexities. By employing the presented UCyAs and implementing the proposed novel algorithms cohesively, the different communication and deployment requirements of a variety of mmWave communication scenarios can be met.

KEYWORDS: Channel parameter estimation, millimeter-wave communications, large-scale antenna array, tensor processing, hybrid beamforming.

Dissertation directed by

Prof. Ren Ping Liu and Prof. J. Andrew Zhang

School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia

Prof. Tiejun Lv

School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China

Acknowledgements

I would like to extend my deep gratitude to all those who have offered me gracious help and support in the process of my thesis writing.

My deepest gratitude goes first and foremost to my supervisors, Prof. Ren Ping Liu, Prof. J. Andrew Zhang, and Prof. Tiejun Lv, for discovering my potential, constantly encouraging me to go further, and providing very valuable lessons on many different levels. Their meticulous academic attitudes and dedicated work styles have always been my model of learning. They have offered me numerous valuable suggestions with incomparable patience and encouraged me profoundly throughout my PhD study.

Also, I would like to express my sincere gratitude to Prof. Wei Ni for helping me review the thesis. The comments and the mutual discussions are very valuable for the thesis itself and for my work beyond it. Without his help, the completion of this thesis would have been impossible.

Special thanks goes to University Of Technology Sydney (UTS) and Beijing University of Posts and Telecommunications (BUPT) for providing wonderful working environments and useful study materials. I would also like to thank my former and current colleagues Xinyu Li, Jie Zeng, Yi Gong, Yuanyuan Ma, Shaoyang Wang, Yashuai Cao, Jintao Xing, Zhongyu Wang, Qixuan Zhang, Yun Zhang, Yuan Ren, Xuewei Zhang, and Qian Liu for countless technical discussions and personal talks, for sharing their knowledge and skills, for helping me out when help was dearly needed. I sincerely hope we can all stay in touch and I wish you all the best for your future.

Yet, none of all this would have been possible without the tremendous support of my parents. Finally, I am deeply indebted to my parent for their continuous support

and encouragement. Thank you for providing me safe and sorrow-free childhood, for discovering and supporting my abilities, and for being patient in times where work made me hard to get by. Their selfless love and silently supports give me the courage and determination to overcome all difficulties.

Zhipeng Lin

Beijing, China, 2021.

List of Publications

Journal Papers

J-1. **Z. Lin**, T. Lv, W. Ni, J. A. Zhang and R. P. Liu, "Nested Hybrid Cylindrical Array Design and DoA Estimation for Massive IoT Networks," *IEEE J. Sel. Areas Commun.*, Aug. 2020. (early access)

J-2. **Z. Lin**, T. Lv, and P. T. Mathiopoulos, "3-D indoor positioning for millimeter-Wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2472-2486, June 2018.

J-3. **Z. Lin**, T. Lv, W. Ni, J. A. Zhang and R. P. Liu, "Tensor-Based Multi-Dimensional Wideband Channel Estimation for mmWave Hybrid Cylindrical Arrays," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7608 - 7622, Dec. 2020.

J-4. **Z. Lin**, T. Lv, W. Ni, J. A. Zhang, J. Zeng and R. P. Liu, "Joint Estimation of Multipath Angles and Delays for Millimeter-Wave Cylindrical Arrays with Hybrid Front-ends," *IEEE Trans. Wireless Comm.*, Mar. 2021 (early access)

J-5. T. Lv, **Z. Lin**, P. Huang and J. Zeng, "Optimization of the Energy-Efficient Relay-Based Massive IoT Network," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 3043-3058, Jun. 2018.

J-6. Q. Liu, T. Lv and **Z. Lin**, "Energy-Efficient Transmission Design in Cooperative Relaying Systems Using NOMA," *IEEE Commun. Letters*, vol. 22, no. 3, pp. 594-597, Mar. 2018.

J-7. J. Zeng, T. Lv, **Z. Lin**, R. Liu, J. Mei, W. Ni and Y. Guo, "Achieving Ultrareliable and Low-Latency Communications in IoT by FD-SCMA," *IEEE Internet of Things J.*, vol. 7, no. 1, pp. 363-378, Jan. 2020.

J-8. Z. Wang, T. Lv, **Z. Lin**, J. Zeng and P. T. Mathiopoulos, "Outage Performance of URLLC NOMA Systems With Wireless Power Transfer," *IEEE Wireless*

Commun. Letters, vol. 9, no. 3, pp. 380-384, Mar. 2020.

J-9. Y. Cao, T. Lv, **Z. Lin**, P. Huang and F. Lin, "Complex ResNet Aided DoA Estimation for Near-Field MIMO Systems," *IEEE Trans. Vehic. Techno.*, vol. 69, no. 10, pp. 11139-11151, Oct. 2020.

J-10. X. Zhang, T. Lv, Y. Ren and **Z. Lin**, "Joint Content Push and Transmission in NOMA With SWIPT Caching Helper," *IEEE Commun. Letters*, vol. 24, no. 4, pp. 922-925, Apr. 2020.

J-11. S. Wang, T. Lv, X. Zhang, **Z. Lin**, and P. Huang, Learning-based multi-channel access in 5G and beyond networks with fast time-varying channels, *IEEE Trans. Vehic. Techno.*, vol. 69, no. 5, pp. 5203-5218, May 2020.

Conference Papers

C-1. **Z. Lin**, T. Lv, J. A. Zhang and R. P. Liu, 3D wideband mmWave localization for 5G massive MIMO systems, in *Proc. IEEE Int. Global Commun. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1-6.

C-2. **Z. Lin**, T. Lv, J. A. Zhang and R. P. Liu, Tensor-based high-accuracy position estimation for 5G mmWave massive MIMO systems, in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1-6.

C-3. Y. Zhang, **Z. Lin** and T. Lv, "Sum-Rate-Driven Energy Efficiency Optimization in Massive MIMO Relay Networks," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Beijing, China, Aug. 2018, pp. 158-162.

C-4. Q. Zhang, **Z. Lin** and T. Lv, "Fast Sparse Bayesian Channel Estimation for Wideband mmWave Systems," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Bologna, Sept. 2018, pp. 1-5.

C-5. Q. Cheng, **Z. Lin**, J. A. Zhang, D. N. Nguyen, X. Huang, A. Kekirigoda and K. Hui, "Multi-user MIMO with Jamming Suppression for Spectrum-Efficient Tactical Communications, in *Proc. Int. Con. Signal Process. Commun. Systems (ICSPCS)* (Accepted).

C-6. Q. Zhang, T. Lv and **Z. Lin**, "Variational Bayesian Channel Estimation for Wideband Multiuser mmWave Systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019, pp. 1-6.

C-7. A. Kekirigoda, K. Hui, Q. Cheng, **Z. Lin**, J. A. Zhang, D. N. Nguyen and X. Huang, "Massive MIMO for Tactical Ad-hoc Networks in RF Contested Environments, in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Norfolk, VA, USA, 2019, pp. 658-663.

Contents

List of Figures

Abbreviation

1-D - One-Dimensional

2-D - Two-Dimensional

3-D - Three-Dimensional

5G - Fifth-Generation

ADC - Analog-to-Digital Converter

AoA - Angle of Arrival

AoD - Angle of Departure

AP - Access Point

AWGN - Additive White Gaussian Noise

B5G - Beyond Fifth-Generation

BS - Base Station

CMOS - Complementary Metal Oxide Semiconductor

CP - CANDECOMP/PARAFAC

CRLB - Cramér-Rao Lower Bound

CS - Compressed Sensing

CWSSM - Coherent Wideband Signal-Subspace Method

DAC - Digital-to-Analog Converter

DFT - Discrete Fourier Transformation

DoA - Direction-of-Arrival

DoF - Degree of Freedom

ESPRIT - Estimation of Signal Parameter via Rotational Invariance Technique

EVD - Eigenvalue-Decomposition

GPS - Global Position System

HOSVD - Higher-Order Singular Value Decomposition

LoS - line-of-sight

IoT - Internet of Things

ISSP - Incoherent Signal-Subspace Processing

IWSSM - Incoherent Wideband Signal-Subspace Method

MHA - Minimum Hole Array

mIoT - Massive Internet of Things

ML - Maximum Likelihood

mmWave - Millimeter-Wave

MRA - Minimum Redundancy Array

MS - Mobile Station

MUSIC - Multiple Signal Classification

NLoS - Non-Line-of-Sight

OBA - Open Box Array

OFDM - Orthogonal Frequency Division Multiplexing

OMP - Orthogonal Matching Pursuit

Q-DFT - Quasi-Discrete Fourier Transform

RF - Radio Frequency

RMa - Rural Macro

RMSE - Root Mean Square Error

RSS - Received Signal Strength

SNR - Signal-to-Noise Ratio

TDoA - Time Difference of Arrival

TLS - Total-Least-Squares

UCA - Uniform Circular Array

UCAMI - Unitary Constrained Array Manifold Interpolation

UCyA - Uniform Cylindrical Array

ULA - Uniform Linear Array

URA - Uniform Rectangular Array

UMa - Urban Macro

UMi - Urban Micro

WSSM - Wideband Signal-Subspace Method

Nomenclature and Notation

a, $\mathbf{a}$, $\mathbf{A}$, and $\mathbb{A}$ stand for a scalar, a column vector, a matrix, and a set, respectively.

$\mathbf{I}_K$ denotes a $K \times K$ identity matrix.

$\mathbf{0}_{M \times K}$ denotes an $M \times K$ zero matrix.

$\|\mathbf{A}\|_{\mathrm{F}}$ denotes the Frobenius norm of $\mathbf{A}$.

invec$(\cdot)$ denotes the inverse algorithm of vectorization.

det$(\mathbf{A})$ is the determinant of $\mathbf{A}$.

Tr$(\mathbf{A})$ denotes the trace of $\mathbf{A}$.

vec$(\mathbf{A})$ is the vectorization of $\mathbf{A}$.

$\mathbf{A}^*$ denotes the conjugate of $\mathbf{A}$.

$\mathbf{A}^T$ denotes the transpose of $\mathbf{A}$.

$\mathbf{A}^H$ denotes the conjugate transpose of $\mathbf{A}$.

$\mathbf{A}^{-1}$ denotes the inverse of $\mathbf{A}$.

$\mathbf{A}^\dagger$ denotes the Moore-Penrose pseudo inverse of $\mathbf{A}$.

$\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker products of $\mathbf{A}$ and $\mathbf{B}$.

$\mathbf{A} \diamond \mathbf{B}$ is the Khatri-Rao product of $\mathbf{A}$ and $\mathbf{B}$.

$\mathrm{E}\{\cdot\}$ denotes the expectation of a random variable.

$\odot$ is the Hadamard product operator.

$\lceil \cdot \rceil$ and $\mathrm{mod}(\cdot)$ represent the ceiling function and the modulo operator, respectively.

$O(\cdot)$ denotes the computational complexity.

$\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ denotes an order-N tensor, whose elements (entries) are $a_{i_1,i_2,\cdots,i_N}$, $i_n = 1, 2, \ldots, I_n$, and the index of $\mathcal{A}$ in the n-th mode ranges from 1 to I_n.

$\mathcal{A}_{:,:,\cdots,:,i_n=k,:,\cdots,:}$ denotes a subtensor of $\mathcal{A}$, where the index of the mode-n is set to k $(0 \leq k \leq I_n)$.

$[\mathcal{A} \sqcup_n \mathcal{B}]$ denotes the tensor concatenation of $\mathcal{A}$ and $\mathcal{B}$ in mode-n.

$\mathbf{A}_{(n)} \in \mathbb{C}^{I_n \times (I_1 I_2 \cdots I_N / I_n)}$ denotes the mode-n unfolding (also known as matricization) of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$.

$\mathrm{Rank}_n(\mathcal{A})$ is the rank of the mode-n unfolding of tensor $\mathcal{A}$, i.e., n-rank of $\mathcal{A}$.

$\mathcal{C} = \mathcal{A} \times_n \mathbf{B} \in \mathbb{C}^{I_1 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N}$ is the n-mode product of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ and a matrix $\mathbf{B} \in \mathbb{C}^{J_n \times I_n}$. It can be written in the form of the mode-n matricized tensor: $\mathbf{C}_{(n)} = \mathbf{B}\mathbf{A}_{(n)}$.

$\mathcal{C} = \mathcal{A} \circ \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N \times J_1 \times J_2 \times \cdots \times J_M}$ is the outer product of two tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ and $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \times \cdots \times J_M}$, whose elements are $c_{i_1,i_2,\cdots,i_N,j_1,j_2,\cdots,j_M} = a_{i_1,i_2,\cdots,i_N} \cdot b_{j_1,j_2,\cdots,j_M}$.

Some important properties of tensor operations used in this paper are presented in Appendix A.

Chapter 1

Introduction

1.1 Research Background

Rapid development of wireless transmission and mobile networking techniques has resulted in tremendous demands for data traffic and network capacity [6,7]. The millimeter-wave (mmWave) communication, which is promising for achieving high data rate and low latency communication thanks to its enormous unlicensed bandwidth, has drawn great attention in both research community and industry [8, 9]. However, mmWave brings new challenges of increased outage probability due to high pathloss attenuation. As a key technique in fifth-generation (5G) wireless communication systems, large antenna array can improve the spectrum efficiency of the systems, and also effectively compensate for the increased propagation losses of the mmWave communications [10–13]. The use of mmWave, which has short wavelengths, in turn, also makes the placement of a relatively large number of antennas on a small area possible [14, 15].

As an essential component in mmWave large antenna systems, precise channel parameter estimation is indispensable for mobile device localization, beamforming design, and optimization of power division and/or allocation [16–18]. The large bandwidth of mmWave systems allows for high-accuracy delay estimation thanks to its high temporal resolution [19, 20]. The high-directivity large array technique is also an enabler for accurate angle estimation [21, 22]. Combining with the large array technique, wideband mmWave large array systems are promising for precisely estimating channel parameters [23–25]. However, most current researches only fo-

cus on linear and rectangular arrays. Compared to antenna arrays with linear or rectangular layouts, circular and cylindrical antenna arrays have many advantages and are particularly suitable for mmWave communication systems [26–28]. One prominent advantage is that circular and cylindrical arrays are more robust to the changes in surrounding environments, in which the changes would lead to so-called vibration effects [26] on the performance of the mmWave communication systems. Specifically, due to the very high frequency of the mmWave, blockage occurs frequently due to rough terrains, passing-by vehicles, mobile users, and so on. As a consequence, the performance of mmWave systems is sensitive to the surrounding environments. The axial symmetry of the circular arrays enables that their main lobes have negligible gain fluctuation in any azimuth angle. Such architecture makes them immune to the angle variations caused by the vibrations. Moreover, the circular arrays have wider beam widths than other popular layout arrays. With a wider main lobe (i.e., antenna beam), less gain loss is suffered under the equal beam misalignment condition. Therefore, for mmWave systems with circular arrays, the corresponding performance losses caused by beam misalignment are less than those using other array architectures [26].

The high computational complexity holds back the application of large-scale antenna arrays into practical mmWave systems, due to the use of a large number of antennas [23–25]. Thanks to the sparsity of mmWave multi-antenna channels, the channels are expected to be low-rank, which can be exploited to reduce the implementation complexity of estimation algorithms. However, most existing researches only focus on mathematics-based ways to reduce the computational complexity of algorithms rather than the properties of the propagation channels [29–31]. For example, by decomposing the received signal components into signal and noise subspaces for parameter estimation, subspace-based channel parameter estimation algorithms, such as estimation of signal parameter via rotational invariance technique (ESPRIT)

and multiple signal classification (MUSIC) [29,30], have much a lower computational complexity than the maximum likelihood (ML)-based estimation algorithms. Nevertheless, in the context of mmWave large-scale array systems, the computational complexity of subspace-based algorithms operating in the element space field remains still very high, since the dimension of element space in large-scale array systems is very large.

High hardware cost and power consumption are also major challengers [32], because it is unrealistic to provide a radio frequency (RF) chain for each antenna, as fully digital beamforming techniques would require [33]. Hybrid beamforming is an appropriate architecture in which a low-dimensional digital beamforming in the baseband and a high-dimensional analog beamforming at the RF front-end are used [34, 35], but conventional parameter estimation algorithms are inapplicable in mmWave hybrid arrays. The state-of-the-art spatial spectrum estimation algorithms, such as ML estimators [36] and subspace-based algorithms [29, 30, 37], were designed to estimate continuous channel parameters using fully digital arrays, where each baseband observation is directly sampled from the signal received at an antenna. The effective operations of these algorithms are based on some important structures existing in the received signals, such as the multiple-invariance structure. With hybrid front-end, the received signals of multiple antennas are combined via an RF phase-shifting network. As a result, the required important structures would be obscured or even lost, and the algorithms cannot directly apply. On the other hand, current channel estimation schemes for hybrid beamforming are typically designed with given channel information [35]. They apply RF networks to directly combine the received signals from multiple antennas, and would result in estimation accuracy losses [18, 38].

1.2 Research Objective

The above-mentioned challenges serve as the main motivation for the general train of thought of this thesis, which presents high-accuracy but low-complexity (both in hardware and software) channel parameter estimation methods for large-scale mmWave uniform cylindrical arrays (UCyAs). The methods complement the advantages of the large antenna array technique and the mmWave communications, and can be applicable to various mmWave communication scenarios, including indoor, urban micro (UMi), rural macro (RMa), and Internet of Things (IoT) application scenarios, as shown in Figs. 1.1 and 1.2. Apart from finding algebraic-based signal processing solutions, we exploit the peculiar properties of mmWave large arrays, such as sparse propagation, high angular resolvability and broad bandwidth, to improve the estimation accuracy and reduce the computational complexity. We also design effective beamforming architectures to address the problems in mmWave large antenna array systems, including high hardware cost, severe pathloss of mmWave, and beam squint (which is caused by wide bandwidths of mmWave signals [39]).

1.3 Thesis Organization

We present four research topics. The research scenarios studied in Chapters 3, 4, and 5 are illustrated in Fig. 1.1, where frequencies are allocated according to the mmWave channel models constructed in [40–43]. Specifically, mmWave is used in short-range high-density (such as the indoor scenario and UMi) and long-distance low-density (such as RMa) wireless communication scenarios. Microwave (below 6 GHz) is applied to long-distance communications in complex environments e.g., UMa, where steady and reliable mmWave transmission usually cannot be guaranteed. This is because the mmWave propagation suffers from serious attenuation and frequent blockage [9, 41]. The typical scenarios in Fig. 1.1 are briefly described as follows:

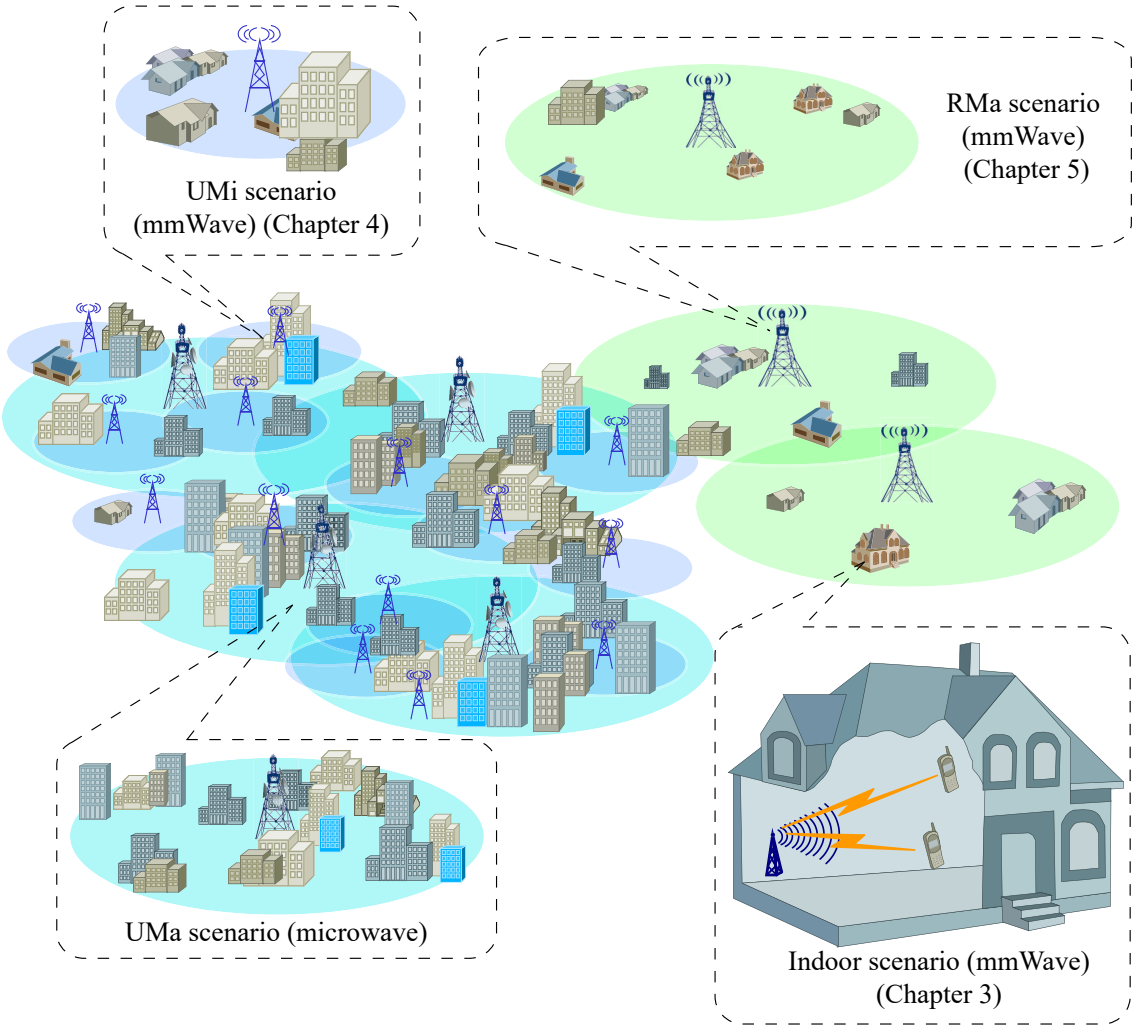


Figure 1.1 : The research scenarios studied in Chapters 3, 4, and 5.

- Indoor scenario (Chapter 3): The indoor scenarios generally include offices and shopping malls, whose dimensions range typically from 25 to 150 m. The typical offices are work spaces enclosed by doors and walls, and the BSs are mounted at a height of 2-3 m either on the ceilings or the walls. The shopping malls are typically more than 2 stories high, in which the BSs are installed at a height of approximately 3 m on the walls or ceilings.

- UMi scenario (Chapter 4): The typical UMi scenarios include street canyons and open squares, and they are usually with high user density. The lengths of the street canyons are often over 100 m. The main elements inside street canyons are roadside buildings, pedestrians, vehicles, and vegetations on both sides of the road. The open squares are propagation environments in the shape of squares, rectangles, or circles and usually surrounded by buildings. The cell radius for open squares are typically less than 100 m and they often contain pedestrians, vehicles, vegetation, and so on.

- RMa scenario (Chapter 5): The RMa scenarios focus on large and continuous coverage. The cell radii of RMa are typically above 200 m. Different from the UMa scenario, the key characteristics of the RMa scenario are flat terrains and continuous wide area coverage.

- Urban macro (UMa) scenario: The UMa refers to the a coverage scenario that has larger coverage than UMi and higher user density than RMa. In UMa, the LoS usually cannot be guaranteed and most part of the signal reaches terminals via diffraction or scattering.

The research scenario studied in Chapter 6 is the mIoT network, as shown in Fig. 1.2, which can connect hundreds to billions of IoT devices [44]. The goal of the mIoT network is to provide high connectivity and efficiently transmit large amounts of data with vast numbers of devices. As a technique which can significantly increase

Table 1.1 : The research scenarios and employed front-ends in different chapters.

Chapters	3	4	5	6
Research Scenarios	Indoor	UMi	RMa	mIoT
Front-end Structures	Digital Beamforming	3D Hybrid Beamforming	Hybrid Directional Beamforming	Nested Hybrid Beamforming

network capacity, the mmWave large antenna array is particularly useful in mIoT networks [44].

This subsection provides a brief introduction for the following four chapters and summarizes the major contributions of each individual chapters. A more detailed introduction on each chapter is presented at the beginning of each chapter separately. Table 1.1 shows the research scenarios and employed front-ends in different chapters.

1.3.1 Joint RSS-AoA Estimation and Localization for Digital UCyAs

Chapter 3 is devoted to the channel parameter estimation for the mmWave indoor scenario, where full digital UCyAs are deployed. As a widely used architecture of multi-antenna systems, digital beamforming can flexibly assign different powers and phases to different antennas. It also enables the received signal of every antenna to be available at the baseband, which prevents the resolution losses of the channel parameter estimation under consideration in this thesis. Thus, such architecture is appropriate for the applications which require flexible control of the antennas and high-accuracy channel parameter estimation.

However, the deployment of a large number of antennas leads to a very high computational complexity when estimating the channel parameters in large-scale

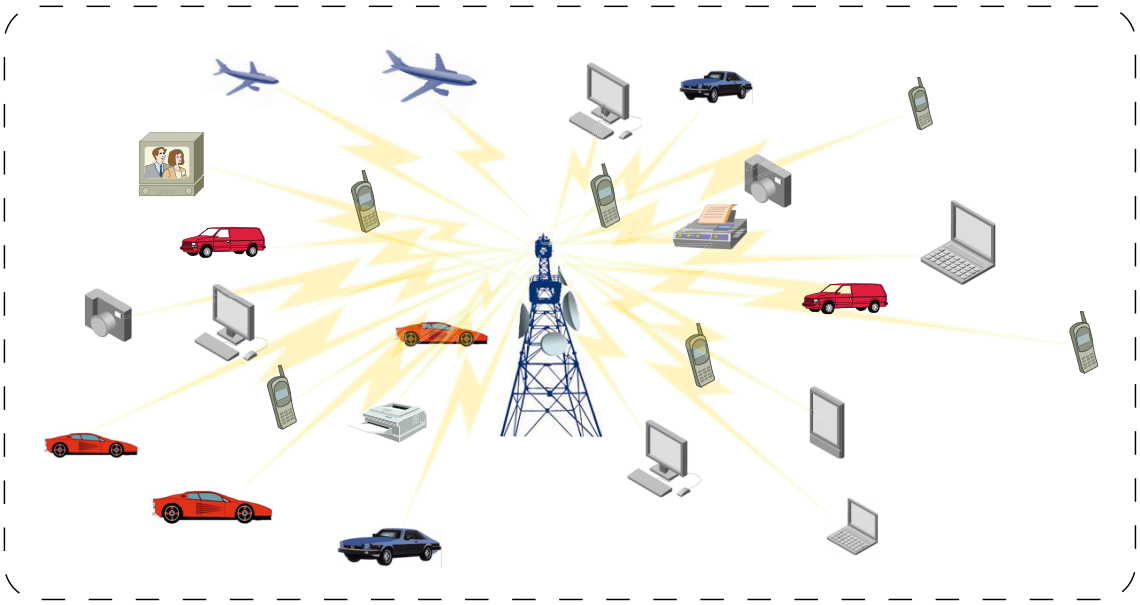


Figure 1.2 : The research scenario studied in Chapter 6.

mmWave array systems. To address this problem, in Chapter 3, we propose a new joint received signal strength (RSS)-angle of arrival (AoA) estimation method for large-scale, digital, mmWave arrays. The method employs both mathematics-based and mmWave property-based ways to reduce the computational complexity. We first provide the motivation and discuss the state of the art in Section 3.1 and introduce the system model in Section 3.2, including the channel model and the basic system operation assumptions. In Section 3.3, we propose a novel channel compression technique, which exploits the particular properties of the mmWave large array channel, i.e., the quasi-optical and sparse multipath propagation [45, 46], to reduce the dimension of the received signals. The technique is capable of reserving the principal components of the channel, i.e., the line-of-sight (LoS) and single-bounce scattering paths, while abandoning the secondary components, such as higher-order bounces and diffuse scattering paths.

A joint RSS-AoA estimation algorithm is described and analyzed in Section 3.4, where we transform the signals from the element space to a low-dimensional

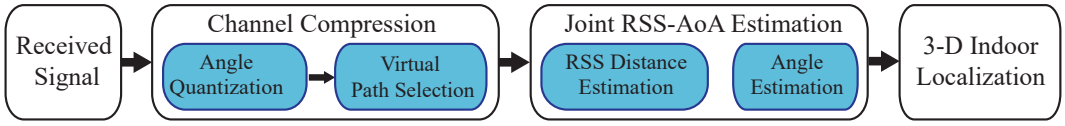


Figure 1.3 : The flow diagram of the channel parameter estimation and indoor localization approach in Chapter 3.

beamspace. As a result, the computational complexity of the proposed method can be further reduced. Based on the estimated distance and the orientation of the target mobile station (MS), we propose a new three-dimensional (3-D) indoor positioning method in Section 3.5 to locate the MS under LoS and non-LoS (NLoS) propagation conditions. By exploiting the quasi-optical propagation at mmWave frequencies with the knowledge on the environment, only a single base station (BS) is required to implement spatial 3-D localization. Performance evaluation results are provided in Section 3.6, which show that our proposed indoor positioning method can achieve significantly reduced computational complexity while maintaining high estimation accuracy.

The new steps developed in Chapter 3 are plotted in Fig. 1.3, and the contributions of the chapter are summarized, as follows.

- We propose channel compression as a preprocessing technique for the mmWave parameter estimation. This technique effectively filters and mitigates the received multiple reflection and diffuse scattering components. By capitalizing on the peculiar properties of mmWave channels, i.e., the quasi-optical and sparse multipath propagation [45, 46], the proposed algorithm is capable of acquiring the principal components of the channel, i.e., the LoS and single-bounce scattering paths, while suppressing the secondary components, such as higher-order bounces and diffuse scattering paths. Consequently, the dimension of the element space of the received signals is significantly reduced, while

maintaining the useful signals for obtaining the positioning information.

- A beamspace-based AoA estimation method is introduced. The advantage of the proposed approach is that it estimates angular parameters in the beamspace rather than in the element space, which has significantly higher dimensions due to the employment of a large number of antennas in large-scale antenna array systems. Because the number of the beamforming weight vectors is much smaller than that of antennas, after the beamspace transformation, the AoA estimation approach operates in a much lower-dimensional space. In this way, the computational complexity of positioning in large-scale mmWave array systems is further reduced.

- We design a novel mmWave indoor localization approach, which provides accurate estimates of both the distance and the orientation of the target MS under LoS and NLoS propagation conditions. By exploiting the quasi-optical propagation at mmWave frequencies and the a-priori knowledge of the environment (including the position and material of potential reflectors), only a single BS is required for the proposed approach to implement spatial 3-D localization. This is different from other previously published methods, for which each propagation area has to be covered by three or more BSs under the requirements of multilateration/triangulation algorithms [47]. Furthermore, since we consider the case where the LoS path may be blocked in practice, the proposed approach is applicable to both LoS and NLoS scenarios.

The proposed algorithm and its results are publications in IEEE Transactions on Communications, i.e., [J2].

1.3.2 Channel Parameter Estimation for 3-D Wideband Hybrid UCyAs

Chapter 4 contributes to the channel parameter estimation for the UMi scenarios with hybrid UCyAs. The range of the UMi scenario is typically larger than that of the indoor scenario, and thus large-scale antenna arrays would be deployed, which, however, would consume more power and incur higher costs due to a large number of RF chains. In this sense, hybrid beamforming is a more appropriate architecture than conventional digital beamforming in the UMi scenarios. It enables the number of RF chains to be much smaller than the numbers of the antennas. As a result, the system hardware cost and power consumption can be reduced.

In Chapter 4, a new 3-D wideband mmWave hybrid UCyA is presented for the UMi scenarios. The designed new array addresses the problem of beam squint and requires far fewer RF chains than antennas. We first provide the motivation and overview of channel estimation for hybrid beamforming in Section 4.1, and introduce the system model in Section 4.2. In Section 4.3, we design a novel two-step 3-D hybrid beamforming strategy, where we form a small number of vertical beams to pick up significant energy of received signals, and conduct quasi-discrete Fourier transform (Q-DFT) on the horizontal plane. We exploit the convergence of the Bessel function, so that the received signals can be converted to a smaller dimension. We show that this strategy can reduce the number of required RF chains while preserving the multiple-invariance structure in array response vectors. As a result, subspace-based algorithms remain effective for parameter estimation.

To suppress the beam squint, we reconstruct the output signals of the hybrid beamformer by generalizing linear interpolation into the 3-D space. We show in Section 4.3.3 that by using this method, we can coherently combine the wideband signals and achieve consistent array responses across the wideband. As a result, the high temporal resolution offered by wideband mmWave systems can be utilized to

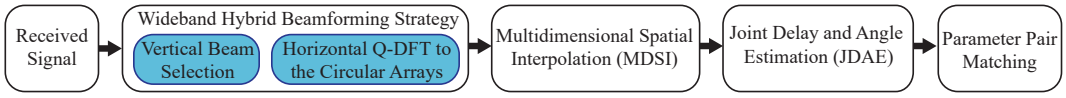


Figure 1.4 : The flow diagram of the proposed channel parameter estimation approach in Chapter 4.

improve the delay estimation accuracy. Given the multiple-invariance structure, we can jointly estimate the delay and AoAs of each path from the combined wideband signals, as described in Section 4.4. To pair the estimates for each path, we add perturbation matrices to the eigenvalue matrices to mitigate the mismatch of the estimated delays and angles caused by noises. In Section 4.5, simulation results are provided to illustrate the performance improvements of the proposed channel parameter estimation method. Finally, conclusions are drawn in Section 4.6.

We plot Fig. 1.4 to illustrate the new steps developed in Chapter 4, and the key contributions of the chapter are summarized, as follows.

- We propose a novel 3-D hybrid beamformer to reduce the number of required RF chains while preserving the multiple-invariance structure in array response vectors. As a result, subspace-based algorithms remain effective for parameter estimation. Specifically, we first form a small number of vertical beams to pick up significant energy of received signals. The Q-DFT is then conducted on the horizontal plane to convert the received signals to a smaller dimension by exploiting the convergence of the Bessel function.

- We generalize linear interpolation to the 3-D space, to reconstruct the output signals of the hybrid beamformer. By this means, we achieve consistent array responses across the wideband and suppress the beam squint effect. The wideband signals can be coherently combined, and the high temporal resolution offered by wideband mmWave systems can be utilized to improve the

delay estimation accuracy.

- We jointly estimate the delay and AoAs of each path, and match the estimated parameters for different paths. Specifically, the elevation AoAs and delays are estimated by utilizing ESPRIT to exploit the multiple-invariance structure, followed by the azimuth AoAs estimated by using MUSIC. Perturbation matrices are introduced to mitigate the mismatch between the estimated delays and angles in the presence of non-negligible noises. As a result, different paths can be correctly detected.

The proposed algorithm and its results are publications in IEEE Global Communications Conference, i.e., [C-1], and IEEE Transactions on wireless Communications, i.e., [J-4].

1.3.3 Tensor-based Parameter Estimation for Hybrid Directional UCyAs

Chapter 5 is devoted to high-accuracy channel parameter estimation algorithms in the RMa scenarios. Directional hybrid beamforming is employed, which can reduce the required number of RF chains and provide sufficient signal power to support long-distance transmission links. Based on tensor signal processing, Chapter 5 presents a novel multi-dimensional approach to channel parameter estimation with large-scale mmWave hybrid directional UCyAs, which are immune to mutual coupling, but also known to suffer from infinite-dimensional array responses and intractability.

Two salient steps for our new tensor-based parameter estimation are presented in Sections 5.3 and 5.4, respectively, where we design the hybrid beamformer and suppress the beam squinting effect in the received signals. Specifically, we first design a hybrid directional beamformer based on Q-DFT in Section 5.3. By employing sweeping directional beamforming and exploiting the convergence property of the

Bessel function, the hybrid directional beamformer can deal with the severe pathloss of mmWave links and maintain the angular resolution of the hybrid UCyA with a reduced number of RF chains. The second step, presented in Section 5.4, is a new low-complexity unitary constrained array manifold interpolation (UCAMI). By only optimizing the focusing matrices in the elevation angular domain, this step can suppress the beam squinting effect and enable coherent combining across the wideband, without quantization in the angular space. Hence, no quantization error occurs.

After the two preprocessing steps, we construct a truncated higher-order singular value decomposition (HOSVD) model of the signals and propose a new tensor-based subspace estimation algorithm in Section 5.5, which jointly estimates the delay, and the azimuth and elevation angles of each received signal by exploiting the important shift-invariance relations. The algorithm suppresses the receiver noises in all of the time, frequency, and space dimensions, and hence accurately estimates the high-dimensional channel parameters of multiple coherent or incoherent signal sources. To decorrelate coherent signals at the hybrid UCyA, we also introduce a new way to rearrange the measurement tensor of the received signals, i.e., spatial smoothing, in Section 5.5.3. Coherent signals can then be separated and can be estimated independently by using the proposed tensor subspace estimation algorithm. In Section 5.6, validated by the CRLB, simulation results show that the proposed tensor-based algorithm is able to achieve a much higher accuracy than state-of-the-art matrix-based techniques in the RMa scenario of 5G/B5G systems. The proposed tensor-based algorithm works well even when the signal-to-noise ratio (SNR) is low, credited to the effective noise suppression in all of the time, space, and frequency domains.

The steps of the proposed novel tensor-based wideband channel estimation approach are illustrated in Fig. 1.5, and will be elaborated in the Chapter 5. The key contributions of the chapter are summarized, as follows:

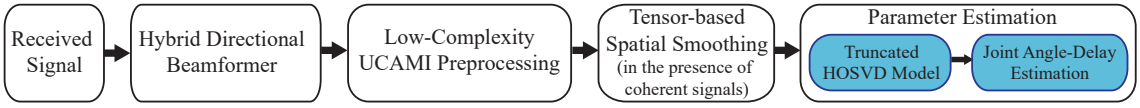


Figure 1.5 : The flow diagram of the proposed channel parameter estimation approach in Chapter 5.

- We design the hybrid beamformers by using Q-DFT to maintain the angular resolution of the hybrid UCyA with a reduced number of RF chains. Developing and applying a low-complexity UCAMI, we suppress the beam squinting effect and enable coherent combining across the wideband. These are two critical steps for our new tensor-based parameter estimation.

- We propose a new tensor-based subspace estimation algorithm to jointly estimate the delay, and the azimuth and elevation angles of each received signal by exploiting the important shift-invariance relations in the constructed truncated HOSVD model. The algorithm can suppress the receiver noises in all of the time, frequency, and space dimensions, and hence it accurately estimates the high-dimensional channel parameters of multiple coherent or incoherent signal sources.

- We introduce a new way to rearrange the measurement tensor of the received signals to decorrelate coherent signals at the hybrid UCyA, i.e., spatial smoothing. Coherent signals can be separated and estimated independently by using the proposed tensor subspace estimation algorithm.

The proposed algorithm and its results are publications in IEEE Transactions on communications, i.e., [J-3], and IEEE International Conference on Communications, i.e., [C-2].

1.3.4 Nested Hybrid UCyA Design and DoA Estimation

Chapter 6 focuses on the hybrid UCyA design and its corresponding 2-D DoA estimation algorithm in mIoT networks. The key physical-layer requirements of mIoT networks are to reduce cost and power consumption while maintaining a high network access capability. Deploying the hybrid arrays is a cost- and energy-efficient means to meet the requirement, but this method would penalize system DoF and channel estimation accuracy.

In Chapter 6, we exploit the sparse array technique and design a new nested massive hybrid UCyA in mIoT networks. The corresponding angle estimation algorithm is also proposed based on tensor processing. In Section 6.3.2, we first utilize the theory of phase-space transformation to transform the nonlinear phase of the UCyA steering vectors to be linear to the element locations. As a result, the horizontal symmetric structure of UCyA is preserved. Then, we flatten the 3-D RF-chain connection network into a 2-D plane and design the RF-chain connection network based on the "Configuration II" nested array [48], as presented in Section 6.3.3. By deploying the proposed sparse RF-chain connection network, we show that the proposed nested hybrid antenna array enables the BS to estimate the DoAs of a large number of devices with much fewer RF chains than antennas. As a result, the massive access requirement of mIoT can be met with significantly reduced hardware cost and network overhead.

To improve the accuracy of the channel parameter estimation, we formulate the received signals in the tensor form and propose a spatial smoothing-based method in Section 6.4 to enhance the n-rank of the constructed the second-order statistics of the signal tensor model. Here, n denotes the index of the tensor mode. A new tensor-based subspace 2-D DoA estimation algorithm for the designed nested hybrid UCyA is developed in Section 6.5, where the hardware and software complexities of

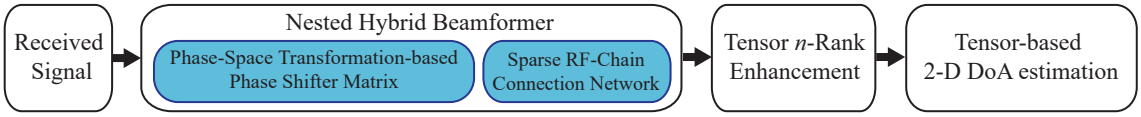


Figure 1.6 : The flow diagram of the proposed approach in Chapter 6.

the proposed estimation algorithm are also analyzed.

We plot Fig. 1.6 to illustrate the new steps developed in Chapter 6. The key contributions of the chapter are summarized as follows.

- We design a new nested hybrid UCyA, which reduces the required number of RF chains while preserving the inherently horizontal symmetric structure of the UCyA to maintain a good channel estimation accuracy. The theory of phase-space transformation is first used to transform the nonlinear phase of the UCyA steering vectors to be linear to the element locations. Then, we design the RF-chain connection network by exploiting the sparse array technique, and utilize its generated difference coarray for parameter estimation.

- We analyze the rank relationship between signal matrix and the signal tensor model in each dimension, and propose a tensor n-rank enhancement method which ensures that the signal and noise subspaces can be properly decomposed in all dimensions.

- We propose a new tensor-based 2-D DoA estimation algorithm, based on our hybrid array design. We combine the tensor tool with ESPRIT to estimate the elevation angles. Then, we substitute the estimates to derive the azimuth angles by using tensor MUSIC. Simulation results show that, by suppressing the noise components in all tensor modes, the proposed algorithm can significantly improve the estimation accuracy in the mIoT networks, as compared to the state of the art.

The proposed algorithm and its results are publications in IEEE Journal on Selected Areas in Communications, i.e., [J-1].

Chapter 2

Literature Review

The mmWave communication, with a large amount of available spectrum, has been considered as a promising approach for achieving high data rate and low latency communication in the 5G/beyond 5G (B5G) wireless networks [49–51]. Not only can the wide bandwidth of mmWave systems improve the accuracy of delay estimation by exploit its potentially high temporal resolution, but also reduces the effect of small scale fading to the level [19,20,52] where the knowledge of the path loss model can be utilized for accurate and reliable distance estimation. For instance, [53] used a weighted combination of RSS indicator measurements for 60 GHz indoor localization applications. It was shown that an accuracy of approximately 1 m can be achieved. On the other hand, thanks to the very small wavelengths of mmWave signals and the currently advanced complementary metal oxide semiconductor (CMOS) technology, nowadays, a massive number of antennas can be integrated into small-scale areas [14]. It enables the large antenna array, another leading 5G technology relying on the use of a large number of antennas, to be deployed even at terminals with restricted sizes and weights [10, 54]. Besides the advantages in cellular communications, including increased system capacity and high spectral efficiency [9], the large array technique is also an enabler for high-accuracy angle estimation thanks to its high directivity [21]. Combining with the large array technique, wideband mmWave systems are promising to precisely estimate channel parameters [23–25, 55].

The ML estimator [56] and the approximate-ML estimator [16] are widely used channel parameter estimation algorithms, which can achieve optimal or near optimal

estimation performance. However, the prohibitively high computational costs hinder their practical implementation in large-scale antenna systems, in particular, for multidimensional parameter estimations, in which complicated high-dimension searches need to be conducted. The covariance matching estimation algorithm (COMET) decomposes the highly dimensional search into multiple successive one-dimensional versions [57, 58]. As a result, the computational complexity caused by the high-dimension search can be reduced. Nevertheless, there is a drawback that such algorithm cannot be extended to the scenario with multiple sources [36]. By approximating the array covariance matrix, [59] proposed an iterative search-based multiple sources parameters estimation algorithm, which, however, needs the preliminary estimates of the source directions. Based on the decomposition of signal and noise subspaces, subspace estimation algorithms [15, 29–31, 37, 60], such as ESPRIT and MUSIC, not only have lower computational complexities than the ML-based estimation algorithms [36], but also obtain attractive parameter estimation performance. However, in mmWave large-scale array systems, due to the deployment of a large number of antennas, the computational complexity of the subspace-based algorithms is still too high, because the dimension of the element space in large-scale array systems is very large. For example, the typical ESPRIT has a cubic complexity with regards to the number of antennas [30], which means that the direct use of this algorithm in large-scale array systems is not appropriate.

Most existing researches, including COMET, ESPRIT and MUSIC, only focus on exploring mathematics-based ways to reduce the computational complexity of algorithms, rather than exploring the properties of the propagation channels [61–63]. Due to the sparsity of mmWave, mmWave multi-antenna channels are expected to be low-rank, and can be exploited to reduce the implementation complexity of estimation algorithms. A few sparse representation techniques were developed in [64,65] and [66] to exploit the sparsity of mmWave multi-antenna systems for channel esti-

mation. The techniques first perform DFT to formulate the channel representation as a sparse signal recovery problem, and then apply compressed sensing (CS) to solve the problem. However, since their techniques are based on the DFT of the array steering vectors, where linear recurrence relations exist between the array steering vectors, their techniques cannot be directly applied to arrays with circular layouts, e.g., uniform circular arrays (UCAs) and UCyAs, where the recurrence relations between the array steering vectors are nonlinear. In addition, only 1-D uniform linear arrays (ULAs) are considered in [64, 65] and [66] and it is not straightforward to extend their techniques to 2-D arrays, even URAs. This is because a 2-D sparse signal recovery problem would need to be formulated and solved. The problem would incur a high computational complexity, if solved with CS techniques.

High hardware cost and high power consumption are also obstacles for applying large-scale array techniques into actual mmWave systems. Conventional low-frequency multi-antenna systems mainly use digital beamforming architectures, in which the numbers of RF chains, digital-to-analog converters (DACs), and analog-to-digital converters (ADCs) are equal to the number of antennas [9]. However, it is unrealistic to deploy the same number of expensive RF chains and high power-consumed ADCs/DACs in large-scale arrays. As a result, it is clearly not feasible to install a separate RF chain and a data converter for every antenna. One simplest solution is analog beamforming, which, unfortunately, can support only the transmission of a single user or a single stream. Hybrid beamforming is a more appropriate architecture, in which a low-dimensional digital beamforming in the baseband and a high-dimensional analog beamforming (or beam selector) at the RF front-end are used to process each transmission signal, as illustrated in Figs. 2.1 (a) and (b), where analog beamforming is based on the phase shifters or the lens arrays. Channel parameter estimation techniques have been well studied in low-frequency multi-antenna systems, but little results are available for large-scale

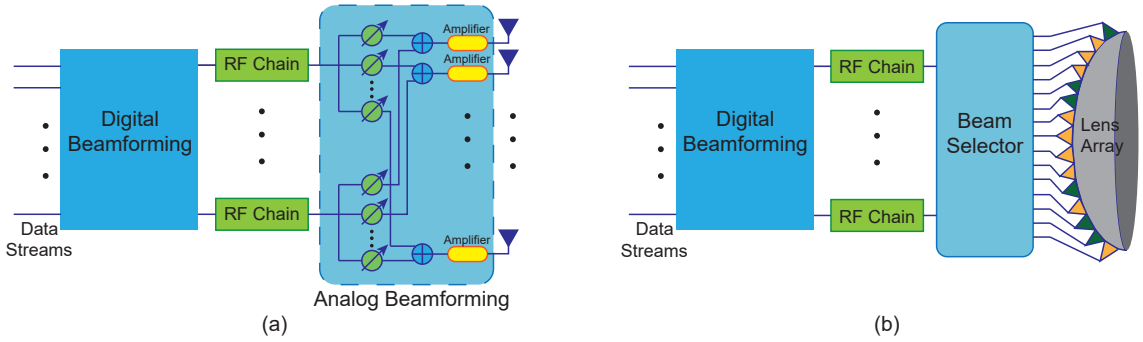


Figure 2.1 : Hybrid beamforming architectures: a) phase shifter-based analog beamforming; and b) lens-based analog beamforming.

mmWave antenna arrays using hybrid front-end [34,35]. Current hybrid beamforming schemes are typically based on CS. They need to discretize channel coefficients and would result in estimation accuracy losses [33,67]. The state-of-the-art spatial spectrum estimation algorithms, such as ML estimators [36] and subspace-based algorithms [29,30,37], were designed to estimate continuous channel parameters using fully digital arrays, where each baseband observation is directly sampled from the signal received at an antenna. With hybrid front-end, the received signals of multiple antennas are combined via an RF phase-shifting network. As a result, the multiple-invariance structure would be obscured or even lost, and the spatial spectrum estimation algorithms cannot directly apply [18,38].

Challenges also arise from beam squint [39], due to typically wide bandwidths of mmWave signals; in other words, the beam directions can change markedly over the different frequencies of a signal bandwidth. Large available bandwidth of mmWave systems makes accurate range measurements possible thanks to high temporal resolution, but the beam squint can lead to channel dispersion in a spatial angle across the bandwidth [39]. Most existing channel parameter estimation methods, e.g., [14,22,68], were designed for narrowband signals. They cannot deal with the beam squint. One existing solution which does support wideband oper-

ations is applying wideband signal-subspace methods (WSSMs) [65, 69] to preprocess the wideband signals. These methods can remove the frequency dependence of array steering vectors and suppress the beam squinting effect. Specifically, the incoherent WSSM (IWSSM) [14, 70] decomposes the received signals into multiple non-overlapping narrowbands, and estimates the parameters independently at each narrowband. This method requires extra steps to combine the results of all the narrow bands [71] and does not utilize the high temporal resolution offered by wideband mmWave systems. In [69, 72], the coherent WSSM (CWSSM) maps the frequency-dependent array steering matrices to a reference frequency by producing so-called focusing matrices. The generation of the focusing matrices in this method requires initial values, and the performance of the methods is susceptible to the initial values. A variation of CWSSM, named unitary constrained array manifold interpolation (UCAMI), was proposed in [73, 74]. It eliminates the need for initial estimates and avoids focusing loss[1]. However, the focusing matrices of UCAMI are obtained by solving multi-dimensional optimization problems. Since the dimension of the problems is equal to the number of frequency-dependent parameters, UCAMI is computationally expensive.

On the other hand, existing channel parameter estimation algorithms are typically matrix-based. By those matrix-based algorithms, the relations between different dimensions (i.e., domains) of the signal become obscure, because the received multi-dimensional (i.e., space, time and frequency) signals are stacked into two-dimensional matrices [1, 75]. Tensor-based channel parameter estimations have been demonstrated to be more powerful than conventional matrix-based techniques in [75–77]. By arranging and processing the received signals in a tensor form, the

[1]Focusing loss refers to the ratio between the array signal-to-noise ratios after and before focusing operations. Focusing loss can be avoided by constraining the focusing matrices that are unitary [69].

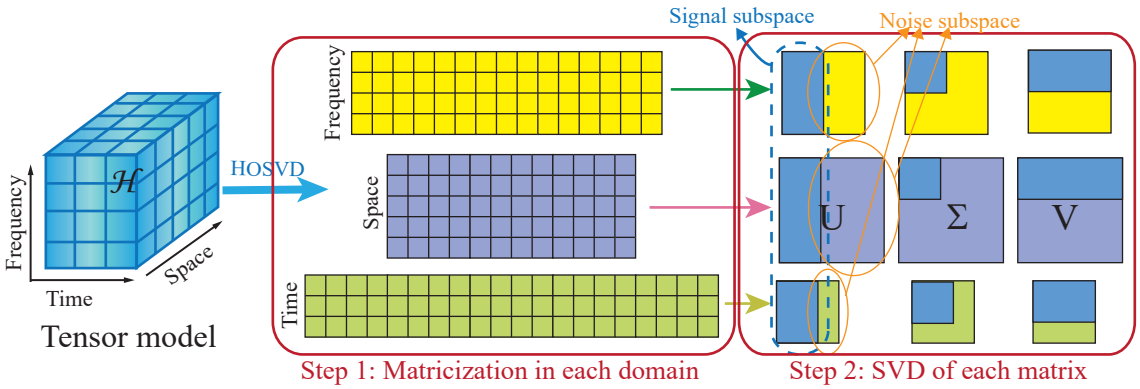


Figure 2.2 : The illustration of the tensor processing in [1–4].

relations between each dimension/domain of the received signal can be exploited. As a result, the multi-dimensional parameters can be estimated with super-high accuracy [1, 75]. The papers [1–4] presented tensor-based algorithms for multi-dimensional channel parameter estimation, which preserve the multi-dimensional structure of signals. They constructed channel tensor model and applied HOSVD to decompose the signal and noise subspaces in every domain (space, frequency, and time), as shown in Fig. 2.2. As a result, the accuracy of the channel parameter estimation can be improved in scatter-rich microwave-band channels. The authors of [78] and [79] exploited the sparsity of mmWave channels to further improve the estimation accuracy. However, their algorithms require an alternating-least-squares procedure with no guarantee of convergence. In addition, the algorithm in [79] is only suitable for narrowband systems with URAs.

As an emerging technology attracting significant attention, IoT promotes a high level of situational awareness and has been used in various areas, such as governments, industry, and academia [80–83]. It is envisioned that more than five billion devices will be connected in an IoT network by 2025 which will allow new ways of living and communicating [5]. As one major segment of the IoT network, mIoT refers to the applications that are capable of connecting a large number of IoT de-

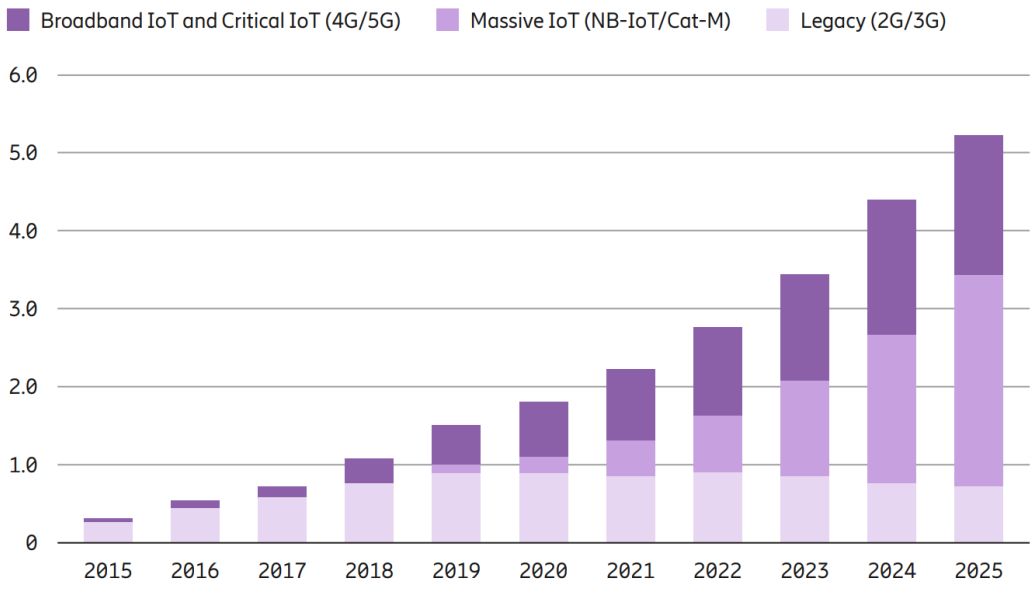


Figure 2.3 : Expected growth of IoT connections (billion) [5].

vices to an internet-enabled system [32, 84–86]. As such, the large network capacity is an urgent demand for mIoT networks. Combined with advanced multiple access techniques, large-scale mmWave array, which can significantly increase network capacity, can be potentially applied to mIoT networks [44]. However, it is unrealistic to provide an RF chain for each antenna, as digital beamforming techniques would require [33]. Hybrid beamforming is an appropriate architecture, but most conventional channel estimation schemes for hybrid beamforming were designed with given channel information [35]. And some of them apply RF networks to directly combine the received signals from multiple antennas, resulting in resolution losses of channel estimation accuracy [18, 38]. As a result, the system degree of freedom (DoF), referring to the number of targets which can be sensed and estimated at the BS [87], would decrease.

To increase the system DoF with a limited number of antennas, the concept of sparse array, such as minimum redundancy array (MRA) [88], minimum hole array (MHA) [89], nested array [87], and coprime array [90], has attracted considerable at-

tention. By exploiting the second-order statistics of impinging signals, these sparse arrays are capable of identifying $O(N^2)$ uncorrelated sources with only N physical antenna elements. However, existing sparse array techniques have been typically used to design linear or square arrays. Compared to square arrays, circular arrays have a much more compact size, less sensitivity to mutual coupling, and inherently more symmetric structure [27], and hence, they are more suitable for mIoT applications. Authors of [91] and [92] proposed nested sparse circular arrays for DoA estimation. However, they directly computed the autocorrelation of impinging signals, which unfortunately compromises the original symmetric structures of circular arrays and penalizes the channel estimation accuracy significantly. Channel estimation is also challenging for sparse arrays. There have been attempts to apply the MUSIC algorithm to networks equipped with sparse arrays [14, 48, 87, 90, 93]. However, the estimation accuracy of those algorithms is unsatisfactory, depending on the searching step and signal correlation. Tensor-based multi-dimensional MUSIC algorithms were proposed in [2, 94] for sparse arrays to improve estimation accuracy. However, since the MUSIC spectrum of their algorithms is a product of multiple separable second-order spectra, undesirable *cross-terms* [94] would arise, leading to incorrect spectral peak search results. To solve this problem, CANDECOMP/PARAFAC (CP)-based tensor channel estimation algorithms were proposed [78], but these algorithms have a very high computational complexity. Table 2.1 shows a comparison between the existing studies and our work.

Table 2.1 : A comparison between the existing studies and our work.

	Existing studies	Our work
Channel estimation algorithms	1) ML-based algorithms: [56] and [16]; 2) COMET: [36, 57, 58]; 3) Typical subspace estimation algorithms: [29–31, 37, 60].	Beamspace-based subspace algorithms
Signal processing methods	1) Matrix-based methods: [29–31, 37, 60–66]; 2) Tensor CP decomposition: [78] and [79]; 3) Tensor HOSVD (narrowband): [1–3, 94].	Tensor HOSVD (wideband)
Reducing computational complexities	1) Exploring mathematics-based ways: [61–63]; 2) Exploring the properties of the propagation channels: [64, 65] and [66].	Exploring both the mathematics-based ways and the sparsity of mmWave channels:
Beamforming structures	1) Digital beamforming: [29–31, 36, 37, 57, 58, 60]; 2) CS-based hybrid beamforming: [33–35].	Subspace-based hybrid beamforming.
Wideband operations	1) IWSSMs: [70]; 2) CWSSMs: [69, 72, 73].	Low-complexity UCAMI
Antenna arrays	1) ULAs: [64, 65] and [66]; URAs: [29, 30, 60, 63, 79].	UCyAs

Chapter 3

Joint RSS-AoA Estimation and Localization for Digital UCyAs

This chapter is devoted to the channel parameter estimation for UCyAs with digital beamforming, where the number of RF chains, DACs as well as ADCs are equal to that of antennas [9]. In this chapter, we consider the mmWave indoor scenario, as illustrated in Fig. 3.1. Digital beamforming can flexibly assign different powers and phases to different antennas, and it enables the received signal of every antenna to be available at the baseband, preventing the resolution losses of channel estimation and reducing the difficulties in beamforming control. Thus, this architecture has been widely used in multi-antenna systems.

In this chapter, a novel low-complexity joint RSS-AoA estimation method is proposed for mmWave large array systems. We first propose a novel channel compression technique by exploiting the sparsity and quasi-optical propagation property of mmWave. By properly quantizing and selecting the received signals, the technique reduces the dimension of the received signal space while maintaining the accuracy of the parameter estimation. After estimating the distance between the BS and the MS, we apply beamspace transformation to transform signal vectors in the element space to the low-dimensional beamspace. As a result, the computational complexity of the angle estimation is significantly reduced. We finally present a novel 3-D indoor positioning approach to estimate the 3-D coordinates of the MS. Simulation results show that the proposed indoor approach is capable of achieving high accuracy with significantly lower computational complexity as compared to other previously known indoor positioning approaches.

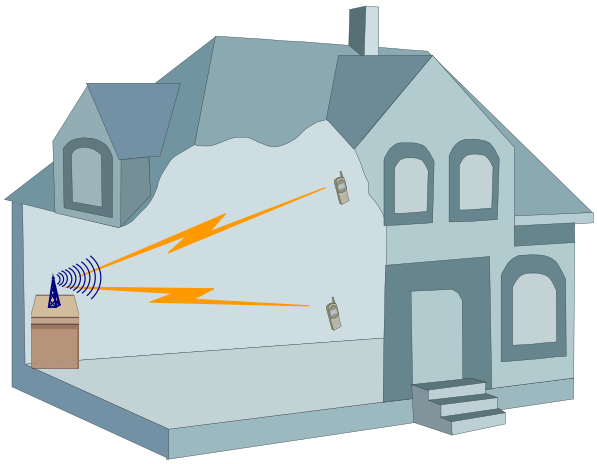


Figure 3.1 : The research scenario studied in Chapter 3.

3.1 Motivation and State of the Art

In mmWave large-scale array systems, the computational complexity of the most existing parameter estimation methods is very high, due to the deployment of a large number of antennas. To solve this problem, in this chapter, we propose a low-complexity joint RSS-AoA estimation method. We first propose a novel channel compression technique in Section 3.3 to reduce the dimension of the received signals to be processed. The quasi-optical property of mmWave propagation leads to the signal path components, which are incident on the surrounding walls, producing specular reflections and little diffuse scatterings. Due to the surface roughness of building materials, there are some useless components of diffuse scattering superimposed on the deterministic specular part of the received signal, which not only increases the dimension of the received signal space, but also reduces the parameter estimation accuracy. By exploiting the sparsity of the mmWave large array channel, the proposed technique remains the principal components of the channel, i.e., the LoS and single-bounce scattering paths, and abandons the secondary components, such as higher-order bounces and diffuse scattering paths. As a result, the estimation accuracy of the proposed estimation method can be improved while reducing

the computational complexity.

Relying on the aforementioned channel compression technique, a joint RSS-AoA estimation algorithm is introduced and analyzed in Section 3.4. In particular, we develop a RSS-based ranging measurement method, *Fredi* [95], which exploits the different transmission power of paths and can overcome the measuring error caused by the multipath effects, to estimate the distance between the BS and the MS. Then, we propose a low-complexity beamspace-based AoA estimation algorithm. We apply beamspace transformation to transform signal vectors in the element space to a beamspace, which has much lower dimensions than the original signal element space, and thus the computational complexity of the parameter estimation can be significantly reduced.

Based on the estimated parameters, Section 3.5 presents a novel 3-D NLoS indoor positioning approach. Most positioning algorithms proposed for mmWave large array systems, such as [22, 47, 50, 63, 68, 96, 97], assume that the LoS path exists between the MS and the BS, and they only focused on 2D scenarios. It is well-known that LoS signal reception is not always guaranteed in practical propagation environments. NLoS localization was considered in [21], but it is also a 2-D approach, and it needs to deploy at least four BSs, increasing the overall system complexity. By exploiting the quasi-optical propagation at mmWave frequencies, we calculate the angle of departure (AoD) of each path and employ the single-bounce specular reflection components to accurately locate the MS in 3-D mmWave indoor scenarios with only a single BS. This section also analyses the computational complexity of the proposed approach. Compared with the existing methods, the proposed approach requires significantly less computational complexity, i.e. reductions up to several orders of magnitude. We also show that the complexity performance improvement of the proposed approach becomes even higher as the the number of antennas increases.

Section 3.6 presents extensive simulation results to show that the proposed method achieves high precision for parameter estimation and indoor localization applications. We compare the performance of the proposed parameter estimation algorithm with that of the existing methods, and use the CRLB as a reference. Results show that the proposed parameter estimation method outperforms the existing methods in terms of AoA estimation when the number of the antennas is larger, and the proposed positioning scheme can achieve decimeter-level positioning accuracy even with low SNR values. In subsection 3.6.3, we also evaluate the sensitivity of the proposed method to the quantization error introduced by the channel compressing. We show that as long as the values of the resolution parameters are sufficiently large, the effect on the overall performance of neglecting the quantization error is minimal.

3.2 System Model

The mmWave large array channel model can be derived from standard large-scale array channel models used in lower frequency bands [35, 98]. Consider a large-scale array system with N_{t} transmit and N_{r} receive antennas. The received narrowband mmWave signal, $\mathbf{y}(t) \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$, is given by

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t), \tag{3.1}$$

where $\mathbf{x}(t) \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ is the transmitted signal vector, $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{r}} \times N_{\mathrm{t}}}$ represents the mmWave channel matrix between BS and MS, and $\mathbf{n}(t) \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$ is the Gaussian noise corrupting the received signal $\mathbf{y}(t)$. For 3-D channel models, $\mathbf{H}$ can be mathematically expressed as

$$\mathbf{H} = \sum_{l=0}^{N_{\mathrm{p}}} \beta_l \mathbf{a}^{\mathrm{r}}(\phi_l^{\mathrm{r}}, \theta_l^{\mathrm{r}}) \mathbf{a}^{\mathrm{t}\,H}(\phi_l^{\mathrm{t}}, \theta_l^{\mathrm{t}}), \tag{3.2}$$

where

$$\mathbf{a}^{\mathrm{r}}(\phi_l^{\mathrm{r}}, \theta_l^{\mathrm{r}}) = \frac{1}{\sqrt{N_{\mathrm{r}}}} \left[a_1^{\mathrm{r}}(\phi_l^{\mathrm{r}}, \theta_l^{\mathrm{r}}), a_2^{\mathrm{r}}(\phi_l^{\mathrm{r}}, \theta_l^{\mathrm{r}}), \ldots, a_{N_{\mathrm{r}}}^{\mathrm{r}}(\phi_l^{\mathrm{r}}, \theta_l^{\mathrm{r}}) \right]^T \tag{3.3}$$

and

$$\mathbf{a}^{\mathrm{t}}(\phi_l^{\mathrm{t}}, \theta_l^{\mathrm{t}}) = \frac{1}{\sqrt{N_{\mathrm{t}}}} \left[a_1^{\mathrm{t}}(\phi_l^{\mathrm{t}}, \theta_l^{\mathrm{t}}), a_2^{\mathrm{t}}(\phi_l^{\mathrm{t}}, \theta_l^{\mathrm{t}}), \ldots, a_{N_{\mathrm{t}}}^{\mathrm{t}}(\phi_l^{\mathrm{t}}, \theta_l^{\mathrm{t}})\right]^T \tag{3.4}$$

are the array steering vectors of receive and transmit antennas, respectively, with $\phi_l^{\mathrm{t}}/\theta_l^{\mathrm{t}}$ and $\phi_l^{\mathrm{r}}/\theta_l^{\mathrm{r}}$ being the azimuth/elevation of AoDs and AoAs of the l-th propagation path, N_{p} denotes the number of received paths[1] and β_l represents the complex amplitude of the l-th propagation path.

We apply the *close-in free space reference distance* (CI) model [99], which has the solid physical basis in both frequency f and distance d, to model the path loss of mmWave indoor scenarios. According to [99], the path loss, $\mathrm{PL}^{\mathrm{CI}}$, is given by

$$\mathrm{PL}^{\mathrm{CI}}(f, d)[dB] = \mathrm{FSPL}(f, d_0) + 10\alpha \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma^{\mathrm{CI}}, \tag{3.5}$$

where α is the path loss exponent; $\mathrm{FSPL}(f, d_0) = 20\log_{10}(4\pi f/c)$ is the free space loss at physically-based reference distance $d_0 = 1\,\mathrm{m}$; c denotes the speed of light; and X_σ^{CI} is a zero mean Gaussian random variable with standard deviation σ in dB, which describes the shadow fading term[2].

In mmWave large array channel, the highly directional antennas as well as the small wavelengths are sensitive to the existence of LoS paths. We use the LoS probability of mmWave channels to determine whether the signal from the LoS path can be received. Apart from LoS paths, when mmWave signals propagates in indoor enviroments, the BS can also receive a lot of diffuse scattering and high-order bounces from all directions. As a result, in a multipath mmWave indoor scenario, the received paths in (3.2) include LoS, single-bounce specular reflections, diffuse scatterings and higher-order bounces. Since LoS blockage between the BS and the MS causes high attenuation at mmWave frequencies, channel characteristics are usually modeled separately for the LoS and NLoS cases.

---

[1]Note that the LoS path in (3.2) corresponds to $l = 0$.

[2]According to [99], for 28GHz mmWave NLOS paths, the variance is about 10 dB.

The LoS probability is defined as the probability that the signal, which propagates from the BS to the MS along the geometrically shortest route, is not blocked by any objects available in the propagation environment [43]. According to [100], we use the following exponential LoS probability model for the indoor mmWave large array channel, as given by

$$P_{\mathrm{LoS}}(d) = (1 - P_{\infty}) \exp\left(-\left(d/\kappa_1\right)^{\kappa_2}\right) + P_{\infty}, \tag{3.6}$$

where d (in m) is the distance between the BS and the MS, $P_{\infty} = P_{\mathrm{LoS}|d\to\infty}$, κ_1 is the decay parameter and κ_2 is the exponent parameter. Obviously, if the distance between the BS and the MS can guarantee LoS propagation, $P_{\mathrm{LoS}}(d) = 1$.

In this chapter we assume that, when the distance between the BS and the MS can guarantee LoS, using the received signal from LoS path for localization is sufficient. Otherwise, the components of single-bounce specular reflection are used to ensure positioning accuracy.

3.3 Channel Compression

Before estimating parameters, we first propose a novel channel compression technique to filter out the unwanted received components while maintaining the useful signals required for the accurate parameter estimation.

In practical mmWave transmission scenarios, due to the surface roughness of building materials, there are some components of diffuse scattering superimposed on the deterministic specular part of the received signal[3]. If these components cannot be handled properly, the estimation accuracy would be harmed. This is because diffuse scattering, defined as the interaction of an impinging wave with a rough surface,

[3]As stated by the Rayleigh criterion, the roughness of the surface is directly proportional to the wavelength of the signal [101]. At mmWave frequencies, there are many objects whose roughness is in the order of the signal wavelength.

produces a spread of the energy in multiple direction as opposed to specular reflection from a smooth surface [43]. Taking advantage of the not so well behaving scattering nature of the mmWave channel, by filtering out the useless diffuse scatterings and multiple reflection components [4], the proposed channel compression method does not only yield to very accurate channel parameter information, but it also reduces the computational complexity of the proposed estimation algorithm.

The rest of the section is organized as follows. Firstly, angle quantization, as the preprocessing for achieving the channel compression, will be presented. Then, based on the quasi-optical and sparse multipath properties of mmWave, the method named virtual path selection will be designed to choose useful paths from the received signals.

3.3.1 Angle Quantization

The channel matrix $\mathbf{H}$ in (3.2) can be rewritten, in a more compact way, as

$$\mathbf{H} = \mathbf{A}^{\mathrm{r}} \mathbf{\Lambda}_{\beta} \mathbf{A}^{\mathrm{t}\,H}, \tag{3.7}$$

where the array steering matrices are given by

$$\mathbf{A}^{\mathrm{r}} = \left[\mathbf{a}^{\mathrm{r}}(\phi_1^{\mathrm{r}}, \theta_1^{\mathrm{r}}), \mathbf{a}^{\mathrm{r}}(\phi_2^{\mathrm{r}}, \theta_2^{\mathrm{r}}), \ldots, \mathbf{a}^{\mathrm{r}}(\phi_{N_{\mathrm{p}}}^{\mathrm{r}}, \theta_{N_{\mathrm{p}}}^{\mathrm{r}}) \right] \in \mathbb{C}^{N_{\mathrm{r}} \times N_{\mathrm{p}}} \tag{3.8}$$

and

$$\mathbf{A}^{\mathrm{t}} = \left[\mathbf{a}^{\mathrm{t}}(\phi_1^{\mathrm{t}}, \theta_1^{\mathrm{t}}), \mathbf{a}^{\mathrm{t}}(\phi_2^{\mathrm{t}}, \theta_2^{\mathrm{t}}), \ldots, \mathbf{a}^{\mathrm{t}}(\phi_{N_{\mathrm{p}}}^{\mathrm{t}}, \theta_{N_{\mathrm{p}}}^{\mathrm{t}}) \right] \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{p}}}, \tag{3.9}$$

respectively, and $\mathbf{\Lambda}_{\beta} = \mathrm{diag}(\beta)$, with $\beta = [\beta_1, \beta_2, \ldots, \beta_{N_{\mathrm{p}}}]$.

Assume that the $\phi_l^{\mathrm{r}}, \phi_l^{\mathrm{t}}$ and $\theta_l^{\mathrm{r}}, \theta_l^{\mathrm{t}}$ are taken from uniform grids of G_1 and G_2 points respectively, i.e., $\phi_l^{\mathrm{r}}, \phi_l^{\mathrm{t}} \in \{0, 2\pi/G_1, \ldots, 2\pi(G_1 - 1)/G_1\}$ and $\theta_l^{\mathrm{r}}, \theta_l^{\mathrm{t}} \in \{0, \pi/G_2, \ldots,$

[4] From now on, and unless otherwise noted, it will be assumed that the number of multiple signal reflections is at least 2.

$\pi(G_2-1)/G_2\}$, $l=0,1,\ldots,N_{\rm p}$, where $G=G_1G_2 \gg N_{\rm p}$ is the number of received virtual paths. The quantized array steering matrices, $\tilde{\mathbf{A}}^{\rm r}$ and $\tilde{\mathbf{A}}^{\rm t}$, are expressed as

$$\begin{aligned}\tilde{\mathbf{A}}^{\rm r} = \Big[&\mathbf{a}^{\rm r}(\tilde{\phi}_1^{\rm r},\tilde{\theta}_1^{\rm r}),\mathbf{a}^{\rm r}(\tilde{\phi}_1^{\rm r},\tilde{\theta}_2^{\rm r}),\ldots,\mathbf{a}^{\rm r}(\tilde{\phi}_1^{\rm r},\tilde{\theta}_{G_2}^{\rm r}),\\ &\mathbf{a}^{\rm r}(\tilde{\phi}_2^{\rm r},\tilde{\theta}_1^{\rm r}),\mathbf{a}^{\rm r}(\tilde{\phi}_2^{\rm r},\tilde{\theta}_2^{\rm r}),\ldots,\mathbf{a}^{\rm r}(\tilde{\phi}_2^{\rm r},\tilde{\theta}_{G_2}^{\rm r}),\\ &\quad\vdots\qquad\quad\vdots\qquad\vdots\qquad\vdots\\ &\mathbf{a}^{\rm r}(\tilde{\phi}_{G_1}^{\rm r},\tilde{\theta}_1^{\rm r}),\mathbf{a}^{\rm r}(\tilde{\phi}_{G_1}^{\rm r},\tilde{\theta}_2^{\rm r}),\ldots,\mathbf{a}^{\rm r}(\tilde{\phi}_{G_1}^{\rm r},\tilde{\theta}_{G_2}^{\rm r})\Big];\end{aligned} \tag{3.10}$$

$$\begin{aligned}\tilde{\mathbf{A}}^{\rm t} = \Big[&\mathbf{a}^{\rm t}(\tilde{\phi}_1^{\rm t},\tilde{\theta}_1^{\rm t}),\mathbf{a}^{\rm t}(\tilde{\phi}_1^{\rm t},\tilde{\theta}_2^{\rm t}),\ldots,\mathbf{a}^{\rm t}(\tilde{\phi}_1^{\rm t},\tilde{\theta}_{G_2}^{\rm t}),\\ &\mathbf{a}^{\rm t}(\tilde{\phi}_2^{\rm t},\tilde{\theta}_1^{\rm t}),\mathbf{a}^{\rm t}(\tilde{\phi}_2^{\rm t},\tilde{\theta}_2^{\rm t}),\ldots,\mathbf{a}^{\rm t}(\tilde{\phi}_2^{\rm t},\tilde{\theta}_{G_2}^{\rm t}),\\ &\quad\vdots\qquad\quad\vdots\qquad\vdots\qquad\vdots\\ &\mathbf{a}^{\rm t}(\tilde{\phi}_{G_1}^{\rm t},\tilde{\theta}_1^{\rm t}),\mathbf{a}^{\rm t}(\tilde{\phi}_{G_1}^{\rm t},\tilde{\theta}_2^{\rm t}),\ldots,\mathbf{a}^{\rm t}(\tilde{\phi}_{G_1}^{\rm t},\tilde{\theta}_{G_2}^{\rm t})\Big].\end{aligned} \tag{3.11}$$

Thus the virtual representation of the channel matrix based on the quantization can be expressed as

$$\tilde{\mathbf{H}} = \tilde{\mathbf{A}}^{\rm r}\tilde{\mathbf{\Lambda}}_\beta(\tilde{\mathbf{A}}^{\rm t})^H, \tag{3.12}$$

where $\tilde{\mathbf{\Lambda}}_\beta \in \mathbb{C}^{G\times G}$ includes the path gains of the corresponding quantized directions.

In this chapter, only the quantized AoAs/AoDs will be further considered. This is because, as will be explained in Section 3.6, numerical simulations have clearly shown that the impact of the quantization error on the performance of proposed algorithms is negligible [33,35]. Therefore, $\mathbf{H}$ is equivalent to its virtual representation in (3.12) , i.e.,

$$\mathbf{H} \approx \tilde{\mathbf{H}} = \tilde{\mathbf{A}}^{\rm r}\tilde{\mathbf{\Lambda}}_\beta(\tilde{\mathbf{A}}^{\rm t})^H. \tag{3.13}$$

3.3.2 Virtual Path Selection

Let us construct an $N_{\rm t}N_{\rm r}\times 1$ column channel vector $\mathbf{h}=\mathrm{vec}(\mathbf{H})$ with the following channel covariance matrix

$$\mathbf{R}_{\mathbf{h}} = \mathrm{E}\left\{\mathbf{h}\mathbf{h}^H\right\}, \tag{3.14}$$

which captures the second-order statistics of the mmWave large array channel. Using the property of Kronecker product, i.e., $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$, we obtain

$$\mathbf{R_h} = \mathrm{E}\left\{\mathbf{h}\mathbf{h}^H\right\} = ((\tilde{\mathbf{A}}^{\mathrm{t}})^* \otimes \tilde{\mathbf{A}}^{\mathrm{r}})\mathbf{R}_{\Lambda_\beta}((\tilde{\mathbf{A}}^{\mathrm{t}})^* \otimes \tilde{\mathbf{A}}^{\mathrm{r}})^H, \tag{3.15}$$

where $\mathbf{R}_{\Lambda_\beta} = \mathrm{E}\left\{\Lambda_\beta \Lambda_\beta^H\right\} \in \mathbb{C}^{G^2 \times G^2}$ is a diagonal matrix and $\Lambda_\beta = \mathrm{vec}(\tilde{\mathbf{\Lambda}}_\beta) \in \mathbb{C}^{G^2 \times 1}$. It is evident from (3.13) that the channel gain appears in $\tilde{\mathbf{\Lambda}}_\beta$, so that according to (3.15), the estimation of $\mathbf{R_h}$ becomes equivalent to estimating $\mathbf{R}_{\Lambda_\beta}$ [102].

In mmWave systems, the received power of the diffuse scattering and multiple reflection paths is much lower than that in LoS and single-bounce specular reflection paths, so the covariance matrix $\mathbf{R}_{\Lambda_\beta}$ of the virtual paths possesses a special sparsity structure. In order to filter out the received diffuse scatterings and multiple reflections which will interfere with parameter estimation, the paths of LoS and specular reflections need to be selected from the received paths.

Let $\sigma_v = \left[\mathbf{R}_{\Lambda_\beta}\right]_{v,v}$ with $v \in \mathcal{V} = \{1, 2, \ldots, G^2\}$, which is the diagonal element of the matrix $\mathbf{R}_{\Lambda_\beta}$, and $\mathcal{V}$ is the set of the sequence numbers of $\mathbf{R}_{\Lambda_\beta}$ diagonal entries. We define a selection-set $\mathcal{U}$ as

$$\mathcal{U} \triangleq \left\{\sigma_{\eta(1)}, \sigma_{\eta(2)}, \ldots, \sigma_{\eta(G^2)}\right\}, \tag{3.16}$$

where the index $\eta(u)$ of $\sigma_{\eta(u)}$ with $u = 1, 2, \ldots, G^2$, represents the sequence number in $\mathcal{V}$. The index $\eta(u)$ can be obtained by

$$\begin{cases} \eta(1) = \arg\max_{v \in \mathcal{V}} \sigma_v, \\ \eta(2) = \arg\max_{v \in \mathcal{V} \setminus \{\eta(1)\}} \sigma_v, \\ \quad \vdots \\ \eta(G^2) = \arg\max_{v \in \mathcal{V} \setminus \{\eta(1), \eta(2), \ldots, \eta(G^2-1)\}} \sigma_v. \end{cases} \tag{3.17}$$

Assume that the total number of LoS path and specular reflection paths is N'_{p}, with $N'_{\mathrm{p}} \ll N_{\mathrm{p}} \ll G$. The following criterion is proposed for selecting the strongest

N'_{p} components from $\mathcal{U}$:

$$\sum_{u=1}^{N'_{\mathrm{p}}} \sigma_{\eta(u)} \geq \gamma \mathrm{tr}(\mathbf{R}_{\Lambda_\beta}), \tag{3.18}$$

where γ is a threshold, which should be close to 1, e.g., 0.9[5]. Then, the following selection procedure is undertaken. Firstly, the strongest component is selected. If the value of $\sigma_{\eta(k)}$ of the strongest component does not satisfy (3.18), the second strongest component will be selected, and this procedure will continue until the sum of $\sigma_{\eta(k)}$ of the strongest N'_{p} components can satisfy the criterion (3.18).

Assuming that an appropriate threshold γ has been selected, the received signal used for channel parameter purposes will come from the LoS and single-bounce specular reflection paths[6]. Because $\tilde{\mathbf{\Lambda}}_\beta$ contains the path gains of the corresponding quantized directions, and the index $\eta(u)$ of $\sigma_{\eta(u)}$ in (3.18) represents the sequence number of the selected paths, the quantized directions corresponding to the selected paths can be obtained after channel compression. Given this, the channel matrix $\dot{\mathbf{H}}$ can be mathematically expressed as

$$\dot{\mathbf{H}} = \dot{\mathbf{A}}^{\mathrm{r}} \dot{\mathbf{\Lambda}}_\beta (\dot{\mathbf{A}}^{\mathrm{t}})^H, \tag{3.19}$$

[5]It has been shown in [45] that for indoor scenarios the contributions of the multiple reflections, i.e., the reflective times are more than two, and that of the diffuse scattering components compared to the total received energy are very weak, only accounting for about 10% of the total energy. Thus for our positioning scheme, we have chosen the threshold $\gamma = 0.9$ to filter out the received diffuse scattering or multiple reflection components, and then use the LoS path or single-bounce specular scattering paths to locate the MS.

[6]It should be noted that, as already mentioned in Section 3.2, when the path between the BS and the MS can guarantee LoS, e.g., the scenario is free of blocking objects or the distance between the BS and the MS is very short, using received signal from LoS path, i.e., $l = 0$, is sufficient for channel parameter estimation purposes. In this case, only the strongest LoS component needs to be selected.

where

$$\dot{\mathbf{A}}^{\mathrm{r}} = \Big[\mathbf{a}^{\mathrm{r}}\left(\tilde{\phi}^{\mathrm{r}}_{\lceil((\eta(1)-1)\mathrm{mod}G+1)/G_2\rceil}, \tilde{\theta}^{\mathrm{r}}_{((\eta(1)-1)\mathrm{mod}G)\mathrm{mod}G_2+1}\right), \\ \mathbf{a}^{\mathrm{r}}\left(\tilde{\phi}^{\mathrm{r}}_{\lceil((\eta(2)-1)\mathrm{mod}G+1)/G_2\rceil}, \tilde{\theta}^{\mathrm{r}}_{((\eta(2)-1)\mathrm{mod}G)\mathrm{mod}G_2+1}\right), \ldots, \\ \mathbf{a}^{\mathrm{r}}\left(\tilde{\phi}^{\mathrm{r}}_{\lceil((\eta(N'_{\mathrm{p}})-1)\mathrm{mod}G+1)/G_2\rceil}, \tilde{\theta}^{\mathrm{r}}_{((\eta(N'_{\mathrm{p}})-1)\mathrm{mod}G)\mathrm{mod}G_2+1}\right)\Big]; \tag{3.20}$$

$$\dot{\mathbf{A}}^{\mathrm{t}} = \Big[\mathbf{a}^{\mathrm{t}}\left(\tilde{\phi}^{\mathrm{t}}_{\lceil\lceil\eta(1)/G\rceil/G_2\rceil}, \tilde{\theta}^{\mathrm{t}}_{(\lceil\eta(1)/G\rceil-1)\mathrm{mod}G_2+1}\right), \\ \mathbf{a}^{\mathrm{t}}\left(\tilde{\phi}^{\mathrm{t}}_{\lceil\lceil\eta(2)/G\rceil/G_2\rceil}, \tilde{\theta}^{\mathrm{t}}_{(\lceil\eta(2)/G\rceil-1)\mathrm{mod}G_2+1}\right), \ldots, \\ \mathbf{a}^{\mathrm{t}}\left(\tilde{\phi}^{\mathrm{t}}_{\lceil\lceil\eta(N'_{\mathrm{p}})/G\rceil/G_2\rceil}, \tilde{\theta}^{\mathrm{t}}_{(\lceil\eta(N'_{\mathrm{p}})/G\rceil-1)\mathrm{mod}G_2+1}\right)\Big] \tag{3.21}$$

contain the selected array steering vectors, and $\dot{\mathbf{\Lambda}}_{\beta} = \mathrm{diag}(\tilde{\beta}') \in \mathbb{C}^{N'_{\mathrm{p}}\times N'_{\mathrm{p}}}$ with $\tilde{\beta}' = [\tilde{\beta}_{\eta(1)}, \tilde{\beta}_{\eta(2)}, \ldots, \tilde{\beta}_{\eta(N'_{\mathrm{p}})}]$ carries the path gains of the corresponding selected paths. At the t-th time instant, the received signal vector at the BS after the channel compression is rewritten as

$$\dot{\mathbf{y}}(t) = \dot{\mathbf{H}}\mathbf{x}(t) + \dot{\mathbf{n}}(t). \tag{3.22}$$

3.4 Joint RSS-AoA Estimation

In this section, we propose a new joint RSS-AoA estimation algorithm to estimate the distance between the BS and the MS and the arrival angle of each paths. We first estimate the RSS by exploiting the transmission power of each path, and then, propose a low-complexity beamspace-based AoA estimation algorithm. The AoDs of the paths are calculated in the next section by exploiting the quasi-optical property of mmWave.

3.4.1 RSS Distance Estimation

In multipath propagation environments, the propagated waves of the different paths will be combined either constructively or destructively depending upon their

received phases. As these phases depend on the signal frequency, the electromagnetic waves of the combined paths vary according to the signal frequencies, resulting in different RSS values.

In this chapter, we employ a RSS-based ranging measurement method, *Fredi* [95], which can overcome the measuring error caused by the multipath effects, to estimate the distance between the BS and the MS. Considering a simple sine radio wave, according to (3.5), the energy field strength of each propagation path can be modeled as

$$E(d,\lambda,\Gamma,t)=\Gamma\frac{G_{\mathrm{t}}G_{\mathrm{r}}P_{\mathrm{t}}}{\sqrt{\mathrm{PL}^{\mathrm{CI}}}}\sin\left(\frac{2\pi c}{\lambda}t+2\pi\frac{d}{\lambda}\right)=\frac{\Gamma\lambda}{d'}\sqrt{\frac{C}{X_{\sigma}^{\mathrm{CI}}}}\sin\left(\frac{2\pi c}{\lambda}t+2\pi\frac{d}{\lambda}\right), \tag{3.23}$$

where $d'=(d/d_0)^{\frac{\alpha}{2}}$, and $\Gamma\in(0,1]$ is a reflection coefficient[7]. $C=G_{\mathrm{t}}G_{\mathrm{r}}P_{\mathrm{r}}/(4\pi)$ is a hardware-dependent constant, where G_{t} is the gain of the transmitter, G_{r} is the receiver gain, and P_{t} is the transmission power. Assume that there are N_{f} different frequencies. The N_{f} RSS measurements, denoted as $\hat{s}_{n_{\mathrm{f}}}$, where $n_{\mathrm{f}}=1,2,\ldots,N_{\mathrm{f}}$, will be obtained. In order to eliminate the uncertainty introduced by the shadowing fading, each RSS value is averaged by performing N_{t} runs. If M different paths exist, the RSS at the receiver is the averaged power of the received signal, which can be expressed as

$$\begin{aligned}\hat{s}_{n_{\mathrm{f}}}(d_1,\ldots,d_M,\Gamma_1,\ldots,\Gamma_M,\lambda_{n_{\mathrm{f}}})&=\frac{1}{N_{\mathrm{t}}}\sum_{t=1}^{N_{\mathrm{t}}}\left(\sum_{m=1}^{M}E_m(t)\right)^2\\&=C\lambda_{n_{\mathrm{f}}}^2\left(\sum_{m=1}^{M}\frac{\Gamma_m^2}{2{d'_m}^2}+\sum_{m\neq m'}^{M}\frac{\Gamma_m\Gamma_{m'}}{d'_m d'_{m'}}\cos\left(\frac{d'_m-d'_{m'}}{\lambda_{n_{\mathrm{f}}}}\right)\right).\end{aligned} \tag{3.24}$$

For a sequence of N_{f} discrete values, we use Discrete Fourier Transformation (DFT) to transform them into another sequence of N_{f} numbers, as given by

$$P_{n_{\mathrm{k}}}=\sum_{n_{\mathrm{f}}=1}^{N_{\mathrm{f}}}\frac{\hat{s}_{n_{\mathrm{f}}}}{C\lambda_{n_{\mathrm{f}}}^2}\exp(-j2\pi\frac{n_{\mathrm{k}}}{N_{\mathrm{f}}}n_{\mathrm{f}}),n_{\mathrm{k}}=1,2,\ldots,N_{\mathrm{f}}. \tag{3.25}$$

[7]For the LoS path, $\Gamma=1$.

Note that we only consider the positive part of $P(n_{\mathrm{k}})$ in this chapter. Substituting (3.24) into (3.25), we have

$$P_{d'_m - d'_{m'}} = \frac{\Gamma_m \Gamma_{m'}}{d'_m d'_{m'}}, \tag{3.26}$$

where $d'_m - d'_{m'}$ is the index of P. In our system, the antenna deployed in the MS is vertically polarizated, thus the electric field vector, E, is parallel to the plane of walls, so that $\Gamma_1 = \Gamma_2 = \ldots = \Gamma_M$. Due to the existance of M distinct propagation paths, there are, in total, $M(M-1)/2$ different results for $P_{d'_m - d'_{m'}}$, which means that there are $M(M-1)/2$ equation. Then by solving these equations, the distance estimator, $\hat{d}$, can be figured out.

3.4.2 Angle Estimation

As explained in Section 3.3, the path components incident on the smooth surfaces of the surrounding walls mainly produce specular reflections. Assuming that, after channel compression, the diffuse scattering and multiple reflection components have been filtered out. Then, the remaining received paths include the LoS and the specular reflection paths from the surrounding walls. Since the LoS can be dealt with as a special case of NLoS, the general case of specular reflected signal paths will be further considered. Such specular reflection paths, which reflect from the four distinct surrounding walls, can be seen as transmitted paths by four virtual MSs, and in this sense they can be considered separately. As previously explained, when considering the single-bounce specular reflection path, there are deterministic geometrical relationships between AoAs and AoDs, so that by evaluating azimuth and elevation AoAs the relevant parameters for AoDs can be obtained.

The received signal vector in (3.22) can be mathematically expressed as

$$\dot{\mathbf{y}}(t) = \sum_{l'=1}^{N'_{\mathrm{p}}} \sum_{\xi=1}^{N_{\mathrm{s}}} \beta_{l',\xi} \mathbf{a}^{\mathrm{r}}(\tilde{\phi}^{\mathrm{r}}_{l',\xi}, \tilde{\theta}^{\mathrm{r}}_{l',\xi}) M_{\xi}(t) + \dot{\mathbf{n}}(t), \tag{3.27}$$

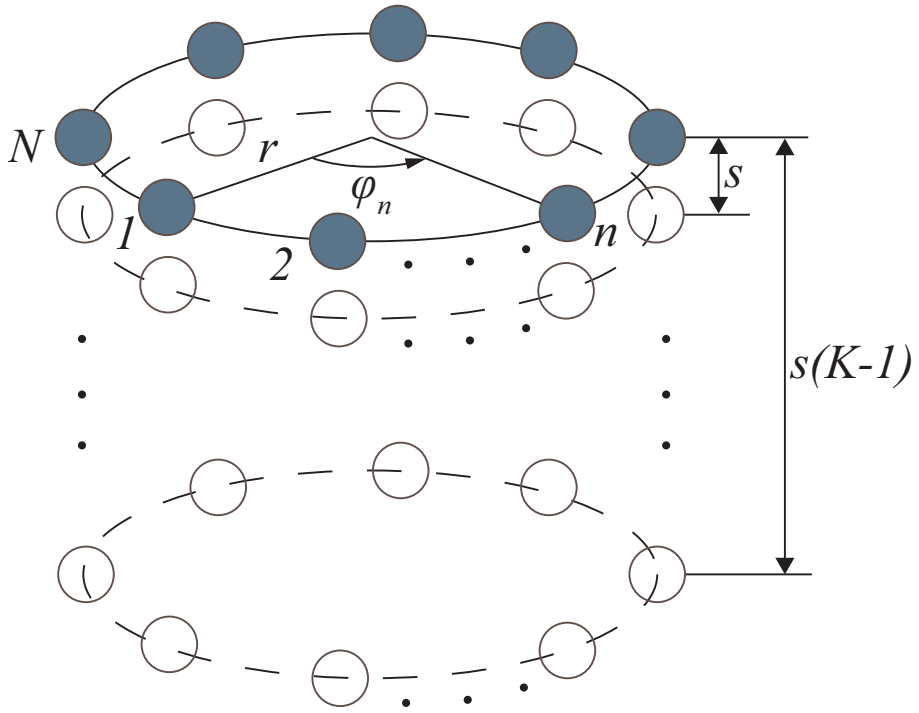


Figure 3.2 : The considered geometric model of the UCyA of the BS consisting of K vertically arranged and coaxially aligned UCAs, with uniform distance, s, between them, and each of them is composed of N antennas.

where

$$M_{\xi}(t) = \mathbf{a}^{\mathrm{t}H}(\tilde{\phi}^{\mathrm{t}}_{l',\xi}, \tilde{\theta}^{\mathrm{t}}_{l',\xi})\mathbf{x}(t) \tag{3.28}$$

$\xi = 1, 2, \ldots, N_{\mathrm{s}}$ represents the number of the specular reflection paths from the four distinct surrounding walls, and l' is the sequence number of selected propagation paths after channel compression[8].

Our system employs an UCyA at the BS [15, 29]. The geometric model of the UCyA is shown in Fig. 3.2. The array consists of K vertically arranged and coaxially aligned UCAs, each of which has a radius r, and they have equal vertical distances, s. Each UCA consists of N antennas, so that the total number of antennas in the UCyA is $N_{\mathrm{r}} = NK$, and the height of this UCyA is $s(K-1)$. As illustrated in Fig. 3.2, the N antennas of any UCA are uniformly distributed over the circumference of a circle, and thus the i-th element of array response vector in the k-th UCA is given by

$$a^{\mathrm{r}}_i(\tilde{\phi}^{\mathrm{r}}_{l',\xi}, \tilde{\theta}^{\mathrm{r}}_{l',\xi}) = \exp\left(j\frac{2\pi}{\lambda}\left[r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\cos(\tilde{\phi}^{\mathrm{r}}_{l',\xi} - \varphi_n) - h(k-1)\cos(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right]\right), \tag{3.29}$$

where $i = N(k-1)+n$, $k = 1, 2, \ldots, K$, $n = 1, 2, \ldots, N$, and $\varphi_n = 2\pi(n-1)/N$ is the central angle difference between the n-th antenna and the first antenna, which is measured counterclockwise.

In the next subsection, we first discuss the beamspace transformation process, and then we present the estimation of azimuth angle estimation procedure.

Beamspace Transformation

In order to reduce the large dimension of received signal vectors, we first use a phase-specific beamforming vector to transform the element space into the beamspace.

[8]Note that $\mathbf{a}^{\mathrm{r}}(\tilde{\phi}^{\mathrm{r}}_{l',\xi}, \tilde{\theta}^{\mathrm{r}}_{l',\xi})$ or $\mathbf{a}^{\mathrm{t}H}(\tilde{\phi}^{\mathrm{t}}_{l',\xi}, \tilde{\theta}^{\mathrm{t}}_{l',\xi})$ in (3.27) is always different from that in (3.10) or (3.11), because l' is the sequence number of paths in each of the N'_p propagation paths obtained after channel compression.

Let us rewrite (3.29) in a more compact form as

$$a_i^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}) = \left[\mathbf{a}_k^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}})\right]_{n,1}, \tag{3.30}$$

where

$$\mathbf{a}_k^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}) = \frac{1}{\sqrt{N}}[a_{N(k-1)+1}^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}), \ldots, a_{Nk}^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}})]^T. \tag{3.31}$$

Define $p \triangleq -P, -P+1, \ldots, P$ as the phase mode. We express the beamforming weight vector, $\mathbf{w}_p$, as

$$\mathbf{w}_p = \frac{1}{N}\left[e^{jp\varphi_1}, e^{jp\varphi_2}, \ldots, e^{jp\varphi_N}\right]^H. \tag{3.32}$$

The array pattern $f_{p,k}(\tilde{\phi}_{l'}^{\mathrm{r}}, \tilde{\theta}_{l'}^{\mathrm{r}})$ can be mathematically expressed as

$$f_{p,k}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}) = \mathbf{w}_p^H \mathbf{a}_k^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}). \tag{3.33}$$

Transforming the received signal vector of the k-th UCA from the element space into the beamspace, we have

$$\tilde{\mathbf{y}}_k(t) = \mathbf{F}^H \dot{\mathbf{y}}_k(t) = \sum_{l'=1}^{N_p'} \beta_{l'} \tilde{\mathbf{a}}_k^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}) M(t) + \tilde{\mathbf{n}}_k(t), \tag{3.34}$$

where

$$\mathbf{F} = \sqrt{N}[j^{-P}\mathbf{w}_{-P}, j^{-P+1}\mathbf{w}_{-P+1}, \ldots, j^P \mathbf{w}_P],$$

is the beamforming matrix [103], $\tilde{\mathbf{n}}_k(t) = \mathbf{F}^H \mathbf{n}_k(t)$ and $\tilde{\mathbf{a}}_k^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}})$ are the beamspace transformed noise and array response vectors, respectively. The beamspace transformed signal vector can be expressed as

$$\tilde{\mathbf{y}}(t) = [\tilde{\mathbf{y}}_1^T(t), \tilde{\mathbf{y}}_2^T(t), \ldots, \tilde{\mathbf{y}}_K^T(t)]^T. \tag{3.35}$$

If the phase mode satisfies $|p| < N$, the array pattern in (3.33) can be rewritten as [103]

$$\begin{aligned} f_{p,k}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}) = \Bigg[& j^p J_p\left(\frac{2\pi}{\lambda} r \sin(\tilde{\theta}_{l',\xi}^{\mathrm{r}})\right) e^{jp\tilde{\phi}_{l',\xi}^{\mathrm{r}}} \\ & + \varepsilon_p\left(\frac{2\pi}{\lambda} r \sin(\tilde{\theta}_{l',\xi}^{\mathrm{r}}), \tilde{\phi}_{l',\xi}^{\mathrm{r}}\right)\Bigg] e^{-j\frac{2\pi}{\lambda}h(k-1)\cos(\tilde{\theta}_{l',\xi}^{\mathrm{r}})}, \end{aligned} \tag{3.36}$$

where

$$\varepsilon_p\left(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi}),\tilde{\phi}^{\mathrm{r}}_{l',\xi}\right)=\sum_{m=1}^{\infty}\left[j^g J_g\left(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right)\times e^{jg\tilde{\phi}^{\mathrm{r}}_{l',\xi}}\right.$$
$$\left.+j^{g'}J_h\left(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right)e^{jg'\tilde{\phi}^{\mathrm{r}}_{l',\xi}}\right], \tag{3.37}$$

with $g=Nm-p$, $g'=Nm+p$, and $J_p(\cdot)$ is the Bessel function of the first kind of order p. $j^pJ_p\left(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right)e^{jp\tilde{\phi}^{\mathrm{r}}_{l',\xi}}$ and $\varepsilon_p(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi}),\tilde{\phi}^{\mathrm{r}}_{l',\xi})$ are the principle and residual terms of the array pattern, respectively. Since in mmWave large array systems, the number of antennas, N, at each UCA is typically very large, the residual term is much smaller than the principle term for any azimuth and elevation AoAs[9]. Furthermore, it is shown in Appendix B.1 that when $N\to\infty$, the residual error tends to zero. As a result, in this case, the residual term can be neglected, and the space transforming error in the beamspace transformation is very small. The array pattern can be approximated as

$$f_{p,k}(\tilde{\phi}^{\mathrm{r}}_{l',\xi},\tilde{\theta}^{\mathrm{r}}_{l',\xi})\approx j^pJ_p\left(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right)e^{j\left[p\tilde{\phi}^{\mathrm{r}}_{l',\xi}-\frac{2\pi}{\lambda}h(k-1)\cos(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right]}. \tag{3.38}$$

Estimation of Elevation Angle

In this subsection, we propose a beamspace ESPRIT approach to estimate the elevation AoAs. In particular, by using the Taylor series expansion, the beamspace array response vector $\tilde{\mathbf{a}}^{\mathrm{r}}(\tilde{\phi}^{\mathrm{r}}_{l'},\tilde{\theta}^{\mathrm{r}}_{l'})$ of the UCyA can be expressed as

$$\tilde{\mathbf{a}}^{\mathrm{r}}(\tilde{\phi}^{\mathrm{r}}_{l',\xi},\tilde{\theta}^{\mathrm{r}}_{l',\xi})=\tilde{\mathbf{a}}^{\mathrm{r}}(\bar{\phi}^{\mathrm{r}}_{\xi},\bar{\theta}^{\mathrm{r}}_{\xi})+\frac{\partial\tilde{\mathbf{a}}^{\mathrm{r}}(\bar{\phi}^{\mathrm{r}}_{\xi},\bar{\theta}^{\mathrm{r}}_{\xi})}{\partial\bar{\phi}^{\mathrm{r}}_{\xi}}\Delta_{\tilde{\phi}^{\mathrm{r}}_{l',\xi}}+\frac{\partial\tilde{\mathbf{a}}^{\mathrm{r}}(\bar{\phi}^{\mathrm{r}}_{\xi},\bar{\theta}^{\mathrm{r}}_{\xi})}{\partial\bar{\theta}^{\mathrm{r}}_{\xi}}\Delta_{\tilde{\theta}^{\mathrm{r}}_{l',\xi}}+\epsilon_{l'}, \tag{3.39}$$

where $\bar{\phi}^{\mathrm{r}}_{\xi}/\bar{\theta}^{\mathrm{r}}_{\xi}$ is the mean of $\tilde{\phi}^{\mathrm{r}}_{l',\xi}/\tilde{\theta}^{\mathrm{r}}_{l',\xi}$, $\Delta_{\tilde{\phi}^{\mathrm{r}}_{l',\xi}}/\Delta_{\tilde{\theta}^{\mathrm{r}}_{l',\xi}}$ is angular deviations of $\tilde{\phi}^{\mathrm{r}}_{l',\xi}/\tilde{\theta}^{\mathrm{r}}_{l',\xi}$, and $\epsilon_{l'}$ is the least significant term, which can be neglected for small angular deviations. Then, the beamspace received signal vector, $\tilde{\mathbf{y}}(t)$, given in (3.35) can be

[9]As discussed in [15], the number of UCA antennas, N, should satisfy $N>2P$ and $P>2\pi r/\lambda$. Clearly, this is the case for mmWave large array systems.

expressed as

$$\tilde{\mathbf{y}}(t) = \mathbf{A}\mathbf{b}(t) + \tilde{\mathbf{n}}(t), \tag{3.40}$$

where

$$\begin{aligned}\mathbf{b}(t) = [&b_{1,1}(t), b_{2,1}(t), \ldots, b_{N_\mathrm{s},1}(t),\\ &b_{1,2}(t), \ldots, b_{N_\mathrm{s},2}(t), b_{1,3}(t), \ldots b_{N_\mathrm{s},3}(t)],\end{aligned} \tag{3.41}$$

$$\begin{aligned}\mathbf{A} = \Bigg[&\tilde{\mathbf{a}}^\mathrm{r}(\bar{\phi}_1^\mathrm{r}, \bar{\theta}_1^\mathrm{r}), \ldots, \tilde{\mathbf{a}}^\mathrm{r}(\bar{\phi}_4^\mathrm{r}, \bar{\theta}_4^\mathrm{r}), \frac{\partial \tilde{\mathbf{a}}^\mathrm{r}(\bar{\phi}_1^\mathrm{r}, \bar{\theta}_1^\mathrm{r})}{\partial \bar{\phi}_1^\mathrm{r}}, \ldots,\\ &\frac{\partial \tilde{\mathbf{a}}^\mathrm{r}(\bar{\phi}_{N_\mathrm{s}}^\mathrm{r}, \bar{\theta}_{N_\mathrm{s}}^\mathrm{r})}{\partial \bar{\phi}_{N_\mathrm{s}}^\mathrm{r}}, \frac{\partial \tilde{\mathbf{a}}^\mathrm{r}(\bar{\phi}_1^\mathrm{r}, \bar{\theta}_1^\mathrm{r})}{\partial \bar{\theta}_1^\mathrm{r}}, \ldots, \frac{\partial \tilde{\mathbf{a}}^\mathrm{r}(\bar{\phi}_{N_\mathrm{s}}^\mathrm{r}, \bar{\theta}_{N_\mathrm{s}}^\mathrm{r})}{\partial \bar{\theta}_{N_\mathrm{s}}^\mathrm{r}}\Bigg],\end{aligned} \tag{3.42}$$

and $b_{\xi,1}(t) = \sum_{l'=1}^{N'_\mathrm{p}} \beta_{l',\xi} M_\xi(t)$, $b_{\xi,2}(t) = \sum_{l'=1}^{N'_\mathrm{p}} \beta_{l',\xi} M_\xi(t) \Delta_{\tilde{\phi}^\mathrm{r}_{l',\xi}}$, $b_{\xi,3}(t) = \sum_{l'=1}^{N'_\mathrm{p}} \beta_{l',\xi} \times M_\xi(t) \Delta_{\tilde{\theta}^\mathrm{r}_{l',\xi}}$.

To remove the randomness in the received signal vector, the covariance matrix of the beamspace received signal vector is calculated as

$$\mathbf{R}_{\tilde{\mathbf{y}}} = \mathrm{E}\left\{\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)\right\} = \mathbf{A}\boldsymbol{\Lambda}_b\mathbf{A}^H + \sigma_\mathrm{n}^2\mathbf{I}_{N'_\mathrm{r}}, \tag{3.43}$$

where $\boldsymbol{\Lambda}_b$ is a diagonal matrix and $\mathbf{R}_{\tilde{\mathbf{y}}}$ is a normal matrix according to (3.43). The eigenvalue-decomposition (EVD) of $\mathbf{R}_{\tilde{y}}$ is obtained by

$$\begin{aligned}\mathbf{R}_{\tilde{y}} &= [\mathbf{E}_\mathrm{s}, \mathbf{E}_\mathrm{n}] \begin{bmatrix} \boldsymbol{\Sigma}_\mathrm{s} & \mathbf{0}_{3N_\mathrm{s}\times(N'_\mathrm{r}-3N_\mathrm{s})} \\ \mathbf{0}_{(N'_\mathrm{r}-3N_\mathrm{s})\times 3N_\mathrm{s}} & \sigma_\mathrm{n}^2\mathbf{I}_{N'_\mathrm{r}-3N_\mathrm{s}} \end{bmatrix} [\mathbf{E}_\mathrm{s}, \mathbf{E}_\mathrm{n}]^H\\ &= \mathbf{E}_\mathrm{s}\boldsymbol{\Sigma}_\mathrm{s}\mathbf{E}_\mathrm{s}^H + \sigma_\mathrm{n}^2\mathbf{E}_\mathrm{n}\mathbf{E}_\mathrm{n}^H,\end{aligned} \tag{3.44}$$

where $\mathbf{E}_\mathrm{s} \in \mathbb{C}^{N'_\mathrm{r}\times 3N_\mathrm{s}}$ and $\mathbf{E}_\mathrm{n} \in \mathbb{C}^{N'_\mathrm{r}\times(N'_\mathrm{r}-3N_\mathrm{s})}$ correspond to the signal subspace and noise subspace of the UCyA, respectively. As each path contains three parts due to the use of the Taylor series expansion in (3.39), there are, in total, $3N_\mathrm{s}$ elements in signal subspace. This means that $\boldsymbol{\Sigma}_\mathrm{s} \in \mathbb{R}^{3N_\mathrm{s}\times 3N_\mathrm{s}}$ is a diagonal matrix whose elements are the largest $3N_\mathrm{s}$ eigenvalues of $\mathbf{R}_{\tilde{\mathbf{y}}}$. Based on $\mathbf{E}_\mathrm{n}\mathbf{E}_\mathrm{n}^H + \mathbf{E}_\mathrm{s}\mathbf{E}_\mathrm{s}^H = \mathbf{I}_{N'_\mathrm{r}}$, (3.44) is rewritten as

$$\mathbf{R}_{\tilde{\mathbf{y}}} = \mathbf{E}_\mathrm{s}(\boldsymbol{\Sigma}_\mathrm{s} - \sigma_\mathrm{n}^2\mathbf{I}_{12})\mathbf{E}_\mathrm{s}^H + \sigma_\mathrm{n}^2\mathbf{I}_{N'_\mathrm{r}}. \tag{3.45}$$

According to (3.43), (3.45), and because the diagonal entries of $\mathbf{\Sigma}_{\mathrm{s}} - \sigma_{\mathrm{n}}^{\mathbf{2}}\mathbf{I}_{3N_{\mathrm{s}}}$ and $\mathbf{\Lambda}_b$ have nonzero values, we have

$$\mathbf{E}_{\mathrm{s}} = \mathbf{AT}, \tag{3.46}$$

where $\mathbf{T} \in \mathbb{C}^{3N_{\mathrm{s}} \times 3N_{\mathrm{s}}}$ is a full rank matrix. According to [15], when $2\pi r/\lambda < P < N/2$, the linear recurrence relation between the beamspace array response matrices of each pair of UCAs can be formulated as

$$\mathbf{A}_{k'+1} = \mathbf{A}_{k'}\mathbf{\Theta}. \tag{3.47}$$

In the above equation,

$$\mathbf{\Theta} = \begin{bmatrix} \mathbf{\Theta}_0 & \mathbf{0}_{N_{\mathrm{s}} \times N_{\mathrm{s}}} & \mathbf{\Theta}_1 \\ \mathbf{0}_{N_{\mathrm{s}} \times N_{\mathrm{s}}} & \mathbf{\Theta}_0 & \mathbf{0}_{N_{\mathrm{s}} \times N_{\mathrm{s}}} \\ \mathbf{0}_{N_{\mathrm{s}} \times N_{\mathrm{s}}} & \mathbf{0}_{N_{\mathrm{s}} \times N_{\mathrm{s}}} & \mathbf{\Theta}_0 \end{bmatrix}, \tag{3.48}$$

$$\mathbf{\Theta}_0 = \mathrm{diag}\left(e^{-j\frac{2\pi}{\lambda}h\cos(\bar{\theta}_1^{\mathrm{r}})}, \ldots, e^{-j\frac{2\pi}{\lambda}h\cos(\bar{\theta}_{N_{\mathrm{s}}}^{\mathrm{r}})} \right), \tag{3.49}$$

$$\mathbf{\Theta}_1 = \mathrm{diag}\left(j\frac{2\pi}{\lambda}h\sin(\bar{\theta}_1^{\mathrm{r}})e^{-j\frac{2\pi}{\lambda}h\cos(\bar{\theta}_1^{\mathrm{r}})}, \ldots, j\frac{2\pi}{\lambda}h\sin(\bar{\theta}_{N_{\mathrm{s}}}^{\mathrm{r}})e^{-j\frac{2\pi}{\lambda}h\cos(\bar{\theta}_{N_{\mathrm{s}}}^{\mathrm{r}})} \right), \tag{3.50}$$

where $k' = 1, 2, \ldots K-1$ and $\mathbf{A}_k = \mathbf{S}_k\mathbf{A} \in \mathbb{C}^{P' \times P'}$ is a submatrix of $\mathbf{A}$, while the selection matrix $\mathbf{S}_k = \left[\mathbf{0}_{p' \times p'(l-1)}, \mathbf{I}_{p'}, \mathbf{0}_{p' \times p'(Q-1)}\right] \in \mathbb{R}^{P' \times N_{\mathrm{r}}'}$. Therefore, the submatrix of the signal subspace matrix can be expressed as

$$\mathbf{E}_k = \mathbf{S}_k\mathbf{E}_{\mathrm{s}} = \mathbf{A}_k\mathbf{T} \in \mathbb{C}^{P' \times 12}. \tag{3.51}$$

Substituting (3.47) into (3.51), one obtains

$$\mathbf{E}_{k'+1} = \mathbf{A}_{k'}\mathbf{\Theta}\mathbf{T} = \mathbf{E}_{k'}\mathbf{T}^{-1}\mathbf{\Theta}\mathbf{T} = \mathbf{E}_{k'}\mathbf{\Psi}. \tag{3.52}$$

By using the total least-squares criterion, $\mathbf{\Psi}$ can be estimated as $\hat{\mathbf{\Psi}}_{k'}$, and each $\hat{\mathbf{\Psi}}_{k'}$ has 12 sorted eigenvalues, i.e., $\lambda_{k',\xi'}$ with $\xi' = 1, 2, \ldots, 3N_{\mathrm{s}}$. Because the eigenvalues of an upper triangular matrix are also the diagonal elements of this matrix,

the total $3(K-1)$ different estimates for each $\mathbf{\Theta}$ in (3.47) can be obtained. Thus, the elevation AoA from the ξ-th surrounding wall is estimated as

$$\hat{\theta}_{\xi}^{\mathrm{r}} = \frac{1}{3(K-1)} \sum_{\xi'=3\xi-2}^{3\xi} \sum_{k'}^{K-1} \left[\arccos\left(\frac{j \ln(\lambda_{k',\xi'})}{\mu s} \right) \right]. \tag{3.53}$$

Estimation of Azimuth Angle

Since the azimuth AoAs does not have the linear relation shown in (3.51), the ESPRIT is not applicable for our scheme. Instead, we propose the beamspace MUSIC to estimate the azimuth AoAs, which also achieves relatively good performance. Without loss of the generality, we consider the beamspace received signal vector of the k-th UCA, which has the following covariance matrix:

$$\mathbf{R}_{\tilde{\mathbf{y}}_k} = \mathbf{E}_{\mathrm{s}_k} \mathbf{\Sigma}_{\mathrm{s}_k} \mathbf{E}_{\mathrm{s}_k}^H + \sigma_{\mathrm{n}_k}^2 \mathbf{E}_{\mathrm{n}_k} \mathbf{E}_{\mathrm{n}_k}^H, \tag{3.54}$$

where $\mathbf{E}_{\mathrm{s}_k} \in \mathbb{C}^{P' \times 3}$ and $\mathbf{E}_{\mathrm{n}_k} \in \mathbb{C}^{P' \times (P'-3)}$ correspond to the signal subspace and noise subspace of the k-th UCA, respectively. Similar to (3.51), we can obtain

$$\mathbf{E}_{\mathrm{s}_k} = \mathbf{A}_k \mathbf{T}_k. \tag{3.55}$$

Based on the orthogonality between the columns of $\mathbf{E}_{\mathrm{s}_k}$ and that of $\mathbf{E}_{\mathrm{n}_k}$, the beamspace MUSIC spectrum is formulated as

$$Q(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}}) = \frac{1}{\left\| \mathbf{E}_{\mathrm{n}_k}^H \bar{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}}) \right\|_{\mathrm{F}}^2}, \tag{3.56}$$

where

$$\bar{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}}) = \left[\tilde{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}}), \frac{\partial \tilde{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}})}{\partial \Phi^{\mathrm{r}}}, \frac{\partial \tilde{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}})}{\partial \hat{\theta}_{\xi}^{\mathrm{r}}} \right] \tag{3.57}$$

In the above equation, $\tilde{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}})$, $\partial \tilde{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}}) / \partial \Phi^{\mathrm{r}}$ and $\partial \tilde{\mathbf{a}}_k^{\mathrm{r}}(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}}) / \partial \hat{\theta}_{\xi}^{\mathrm{r}}$ can be calculated by substituting the array pattern in (3.38) and its partial derivatives into (3.39) respectively, and Φ^{r} is the azimuth of the AoA. The Φ^{r} can be estimated by 1-D search, and the estimator of the azimuth of the AoA from the ξ-th surrounding wall is given by

$$\hat{\phi}_{\xi}^{\mathrm{r}} = \arg \max_{\Phi^{\mathrm{r}}} Q(\Phi^{\mathrm{r}}, \hat{\theta}_{\xi}^{\mathrm{r}}). \tag{3.58}$$

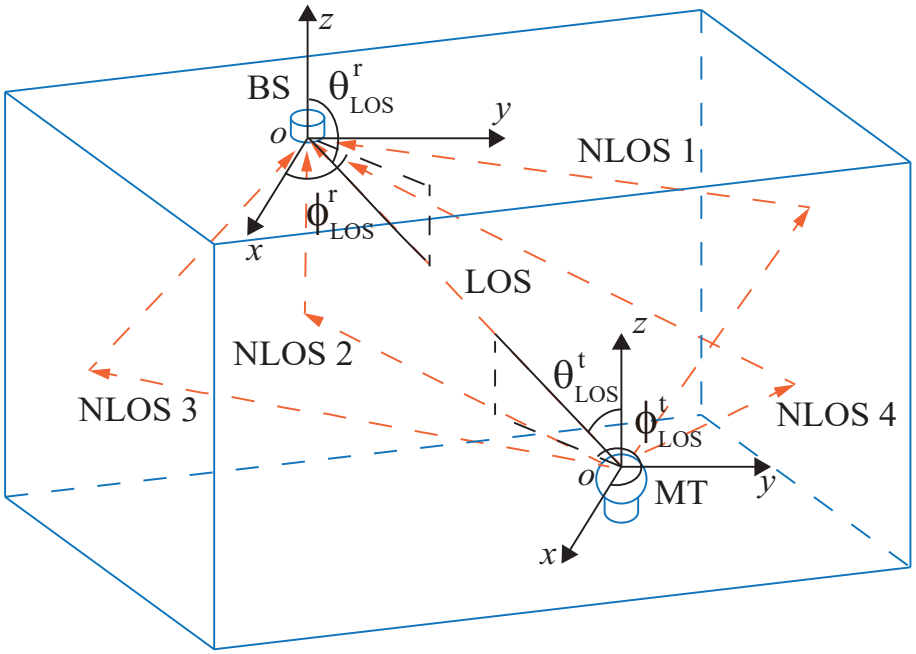


Figure 3.3 : The 3-D indoor positioning system under consideration illustrating the signal propagation in which a single BS is installed on the ceiling at known location and an unknown MS moving within the room.

3.5 3-D Localization and Complexity Analysis

In this section, we first introduce the considered indoor localization configuration. We then present an overview of the proposed localization approach and calculate the AoDs of all the paths by exploiting the quasi-optical propagation property of mmWave. The computational complexity of the proposed approach is also analyzed.

3.5.1 Indoor Configuration

As illustrated in Fig. 3.3, let us consider a 3-D indoor scenario, consisting of a BS as the receiver, which employs one large UCyA, located at a known position on the ceiling of a room, and an MS as the transmitter employing a single antenna whose location in the room is unknown and needs to be estimated.

We then exploit the the quasi-optical propagation property of mmWave to esti-

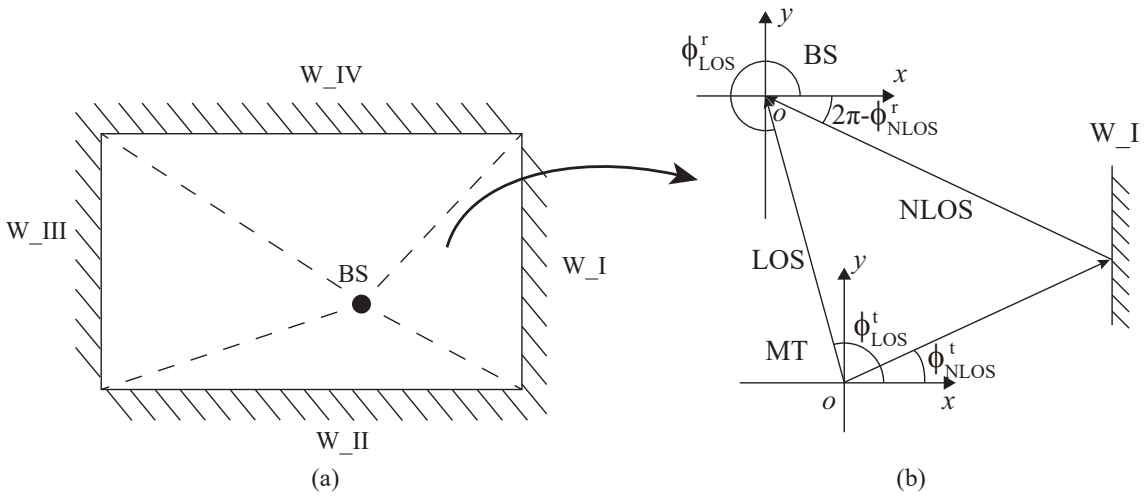


Figure 3.4 : The considered indoor positioning system showing the BS, the MS and the four Walls (W_I-W_IV). (a) Top view; (b) Azimuth angular relationships between AoAs of the BS and AoDs of the MS. Note that the NLoS path is reflected from the right side wall (W_I).

Table 3.1 : Azimuth angle relationships between the BS and the MS for the specular reflection paths.

Walls	Azimuth Angle's Relationships ($l = 1, 2, \ldots, N_p$)
W_I	$\phi_l^{\mathrm{r}} + \phi_l^{\mathrm{t}} = 2\pi$
W_II	$\phi_l^{\mathrm{r}} + \phi_l^{\mathrm{t}} = 3\pi$
W_III	$\phi_l^{\mathrm{r}} + \phi_l^{\mathrm{t}} = 2\pi$
W_IV	$\phi_l^{\mathrm{r}} + \phi_l^{\mathrm{t}} = \pi$

mate the AoDs of all the paths. As was illustrated in Fig. 3.3, we regard the received signal angles at the origin of the coordinates as the reference, and the AoAs and AoDs are measured on an absolute frame of reference. In Fig. 3.4 the localization system under consideration is further detailed.

For the LoS path, it is clear that the sum of elevation AoAs and AoDs is π [10], and the difference between AoAs and AoDs for LoS path always equals to π, as shown in Fig. 3.4(b). For the single-bounce specular reflection paths, the relationships of azimuth angle between BS and MS are complex due to the reflections from the four different directional surrounding walls (W_I-W_IV). We assume that the LoS path between the BS and the MS is not blocked and the components of diffuse scattering and multiple reflection have been filtered out. As illustrated in Fig. 3.4(b), the BS receives the signal not only from the LoS path, but also from the four specular reflection paths, one of which has experienced the reflection from the right side wall (W_I). In order to locate the MS through AoAs, the specular reflection paths from the four different surrounding walls are considered separately, as shown in Fig. 3.4(a), and the relationships of the azimuth angle between the BS and the MS are given in Table 3.1.

In our system, when the LoS cannot be guaranteed, the single-bounce specular reflection components is employed to ensure the positioning accuracy. Once the distance, $\hat{d}$, between BS as well as MS and the estimated azimuth/elevation AoAs, $\hat{\phi}_{\xi}^{\mathrm{r}}/\hat{\theta}_{\xi}^{\mathrm{r}}$, have been obtained by the above mentioned approaches, the accurate position of the MS can be easily determined. It is explained in Appendix B.2 that each reflected path will identify two likely positions. For the case where the specular reflection path experiences the reflection from the right side wall (W_I), 3-D

[10]For the convenience of the geometrical illustration, the angle relationships between the BS and the MS discussed here do not consider the received paths after channel compression.

coordinates of the two likely positions of the MS are calculated as

$$\begin{cases} \hat{x}_{\mathrm{WI}} = \frac{\pm 2D_1 \tan^2(\hat{\phi}_1^{\mathrm{r}}) + \sqrt{\hat{d}^2 \sin^2(\hat{\theta}_1^{\mathrm{r}})(\tan^2(\hat{\phi}_1^{\mathrm{r}})+1) - 4D_1^2 \tan^2(\hat{\phi}_1^{\mathrm{r}})}}{tan^2(\hat{\phi}_1^{\mathrm{r}})+1} \\ \hat{y}_{\mathrm{WI}} = \frac{\left(2D_1 \mp \sqrt{\hat{d}^2 \sin^2(\hat{\theta}_1^{\mathrm{r}})(\tan^2(\hat{\phi}_1^{\mathrm{r}})+1) - 4D_1^2 \tan^2(\hat{\phi}_1^{\mathrm{r}})}\right) \tan(\hat{\phi}_1^{\mathrm{r}})}{\tan^2(\hat{\phi}_1^{\mathrm{r}})+1} \\ \hat{z}_{\mathrm{WI}} = -\hat{d}\cos(\hat{\theta}_1^{\mathrm{r}}) \end{cases}, \tag{3.59}$$

where D_1 is the distance between the W_I and the BS. Note that for our system model this distance is assumed to be known *a priori*, and that the center of the UCyA of the BS is taken as the origin of the coordinates (see Fig. 3.3). Similarly, by using the specular reflection paths from other walls , i.e., W_II, W_III and W_IV, the most probable position of the MS can be obtained from each of these paths. Since all four estimated positions are expected to be close to each other, they can be easily selected from two likely positions for the same specular reflection path through clustering algorithms [104]. By calculating the center of gravity of the tetrahedron with vertex $(\hat{x}_\xi, \hat{y}_\xi, \hat{z}_\xi)$ as the estimator of the MS's position, we have

$$\begin{cases} \hat{x} = \frac{1}{4}\sum_{\xi=1}^{4} \hat{x}_\xi \\ \hat{y} = \frac{1}{4}\sum_{\xi=1}^{4} \hat{y}_\xi \\ \hat{z} = \frac{1}{4}\sum_{\xi=1}^{4} \hat{z}_\xi \end{cases}. \tag{3.60}$$

3.5.2 Complexity Analysis

In this subsection, the computational complexity of the proposed localization scheme is discussed. Firstly, for the mmWave indoor scenario, the diffuse scattering accounts for approximately 10% of the total received power [43, 45]. Thus the threshold γ in (3.18) was set to 0.9 to filter out the received diffuse scattering and multiple reflection components from all directions. Since the number of received paths, N_p, is several orders of magnitude larger than that of the selected paths, N_p', (i.e., $N_p \gg N_p'$), the channel compression, which has been considered as the preprocessing of the proposed localization scheme, significantly reduces the computational

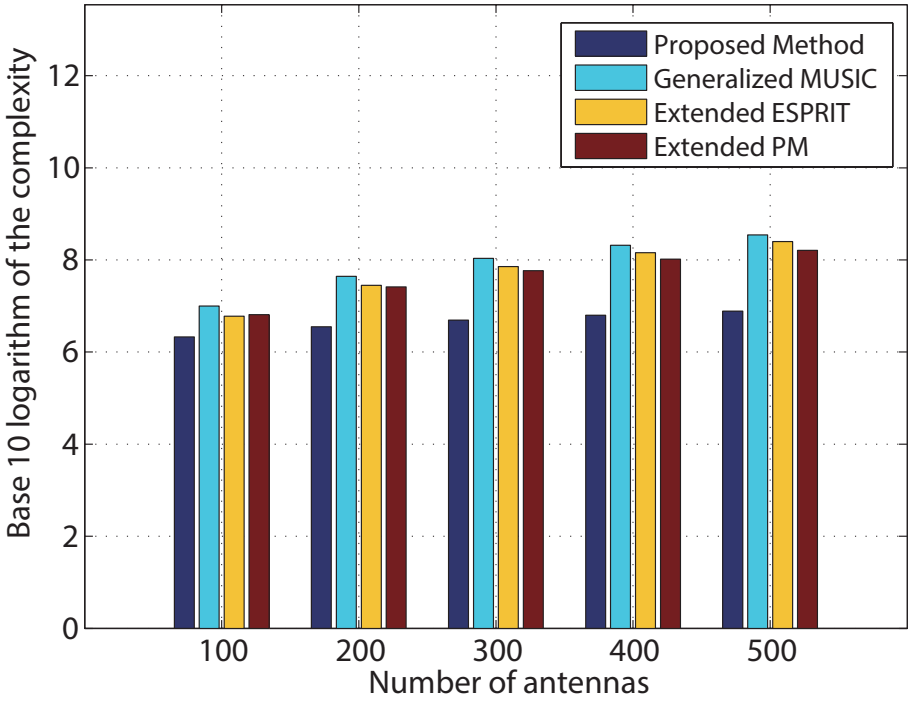


Figure 3.5 : Comparison of the computational complexity vs. number of antennas. Note that the y-axis uses a base 10 logarithmic scale.

complexity.

Secondly, the computational complexity of the proposed beamspace-based angle estimation is compared with that of the generalized MUSIC, the extended ESPRIT, and the extended propagator method (PM) presented in [29], [31], and [105], respectively. Because the number of the beamforming weight vector $\mathbf{w}_p$ in (3.32) is much smaller than the number of the antennas deployed in large-scale array systems, our proposed method has a significantly reduced computational complexity as compared to these three methods, all of which estimate angular parameters in the element space. Note that, by transforming the signal vectors from the element space to the beamspace, $\tilde{\mathbf{y}}_k(t)$ is obtained by using (3.34) and (3.35). The computational complexity of this step is equal to $O(N_{\mathrm{r}}P'T_s)$, where T_s is the number of snapshots. Since in practice, the covariance matrix can be estimated as $\mathbf{R}_{\tilde{\mathbf{y}}} = \frac{1}{T_s}\sum_{t=1}^{T_s}\tilde{\mathbf{y}}(t)\tilde{\mathbf{y}}^H(t)$, the

computational complexity for calculating the sample covariance matrix in (3.43) is $O(P'^2T_s)$. When performing the EVD on $\mathbf{R}_{\tilde{\mathbf{y}}_k}$ according to (3.44), and estimating $\hat{\phi}_\xi^{\mathrm{r}}$ with 1-D search using (3.58), the computational complexity is $O(P'^3 + P'^2D)$, where D is the search dimension of estimating $\hat{\phi}_\xi^{\mathrm{r}}$. The complexity of remaining steps, including performing EVD on $\mathbf{R}_{\tilde{\mathbf{y}}}$ in (3.43) and estimating $\hat{\mathbf{\Psi}}_{k'}$ and $\hat{\theta}_\xi^{\mathrm{r}}$, is $O(P'^3 + K^3)$. Thus, the overall computational complexity of the proposed angle estimation is $O(N_{\mathrm{r}}P'T_s + P'^2T_s + P'^3 + P'^2D)$. Note that, when the number of UCyA antennas N_{r} is large enough, by using the beamspace transformation, the computational complexity of proposed angle estimation is approximately equal to $O(N_{\mathrm{r}}P'T_s)$, while that of the approach used in [29] and [31] is $O(N_{\mathrm{r}}^3)$, and that of the approach in [105] is $O(N_{\mathrm{r}}^2(T_s + 4))$.

We evaluate the computational complexity performance of all four methods as a function of the number of antennas at the BS, where $T_s = 500$ and $P' = 28$, and the obtained results are provided in Fig. 3.5. For a fair comparison with the other three methods, the number of snapshots is fixed in our simulations. As the obtained results clearly show, compared with the existing methods, the proposed approach requires significantly less computational complexity, i.e. reductions up to several orders of magnitude. In addition, it is underlined that the complexity performance improvement becomes even higher as the the number of antennas increases.

3.6 Performance Evaluation and Discussion

This section presents extensive simulation results as an evidence that the proposed method achieves high precision for parameter estimation and indoor localization applications. We compare the performance of the proposed parameter estimation algorithm with that of the generalized MUSIC [29], the extended ESPRIT [31] and the extended PM [105]. We also use the CRLB as a reference to evaluate the performance of the proposed estimation algorithm. The derivation of the CRLB for

the system under consideration can be found in Appendix B.3. The performance of the proposed 3-D positioning approach for both the indoor office and shopping mall scenarios is evaluated. Additionally, it is shown that the quantization error of the channel compression method has a negligible effect on the parameter estimation performance of the proposed method.

As far as the simulation methodology is concerned, various experiments have been carried out in two different scenarios over the $f = 28$ GHz mmWave indoor channel[11], namely indoor office and shopping mall with sizes $15 \times 20 \times 4$ m^3 and $20 \times 50 \times 20$ m^3, respectively. The locations of the BS for two scenarios are set at (10 m, 9 m, 4 m) and (10 m, 20 m, 20 m). The distance, s, between adjacent UCAs and the radius, r, of each UCA are 0.5λ and 2λ, where λ is the wavelength of the mmWave carrier frequency.

3.6.1 Beamspace Angle Estimation

The measurements of azimuth/elevation AoAs of the mmWave signals are generated in the simulation scenarios by adding the AoA spread, $\sigma_\phi/\sigma_\theta$, whose values are set according to [106]. The path losses for different indoor wireless channels have been calculated by using (3.5) and the threshold, γ, in the channel compression was set to 0.9. The performance of angle estimations is evaluated by using the root mean square error (RMSE) criterion averaged over all the trials.

In Figs. 3.6 and 3.7, the RMSEs of the estimated azimuth and elevation AoAs versus the number of received antennas are shown for the indoor office and shopping mall scenarios, respectively, by considering SNR = 10 dB for both cases. It can be seen that the RMSEs of the AoA parameters get close to the CRLB as the number of received antennas at the BS increases. However, it is also noted that

[11] In this system, for a UCyA with 400 antennas (25×16), its dimension is about 40×80 mm^2, which is much smaller than the communication range. Therefore, the far-field condition is fulfilled.

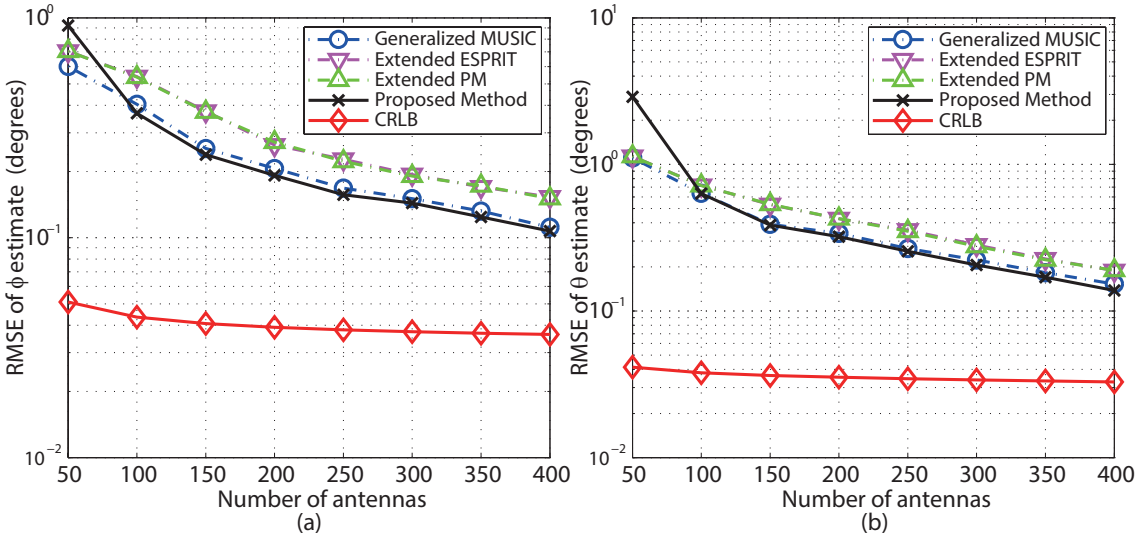


Figure 3.6 : Comparison of the RMSE performance vs. number of antennas for the estimation of different angular parameters by the different estimation methods for the indoor office scenario. (a) Azimuth AoAs; and (b) Elevation AoAs.

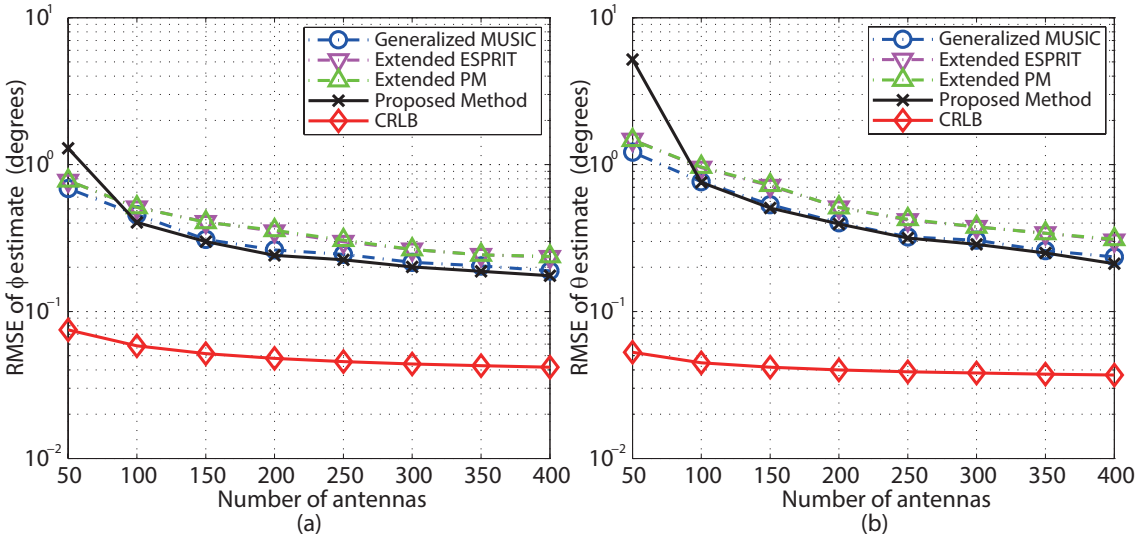


Figure 3.7 : Comparison of the RMSE performance vs. number of antennas for the estimation of different angular parameters by the different estimation methods for the shopping mall scenario. (a) Azimuth AoAs; and (b) Elevation AoAs.

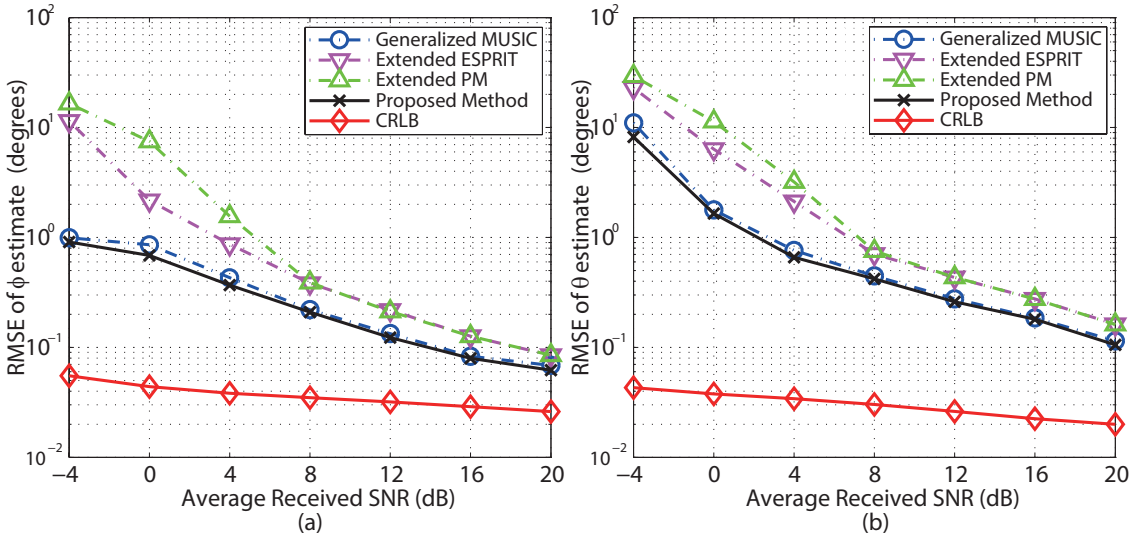


Figure 3.8 : Comparison of the RMSE performance vs. average received SNR for the estimation of different angular parameters by the different estimation methods for the indoor office scenario. (a) Azimuth AoAs; and (b) Elevation AoAs.

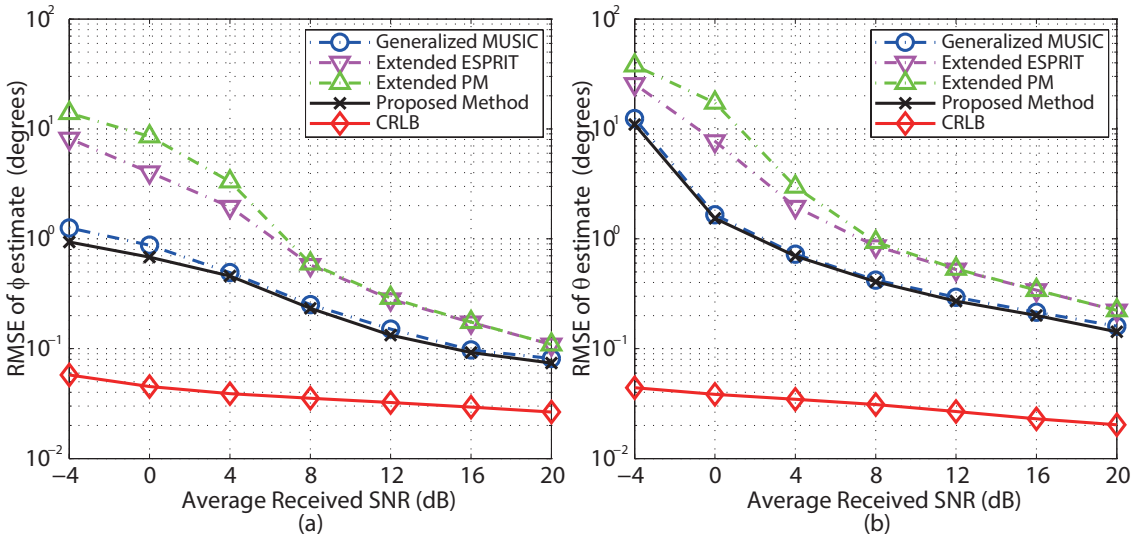


Figure 3.9 : Comparison of the RMSE performance vs. average received SNR for the estimation of different angular parameters by the different estimation methods for the shopping mall scenario. (a) Azimuth AoAs; and (b) Elevation AoAs.

when the number of UCyA antennas is not large, e.g. less then 100, the RMSE performance of the proposed approach is worse than that of the other three methods in [29], [31], and [105]. This can be explained because the residual term in (3.36) cannot be neglected when the number of antennas is small, since in this case the approximations made for the array pattern in (3.38) are not accurate. However, when the number of antennas increases, the RMSE performance of our method decreases faster than that of the other three methods, leading to a better accuracy of the azimuth angles estimated by the new method. Moreover, the performance results shown in Figs. 3.6 and 3.7 reveal that the localization precision is higher for the indoor office as compared to the shopping mall scenario. This is due to the fact that for the shopping mall scenario, the path losses and angle spreads are especially high in mmWave frequency bands, which causes this degradation in the estimation process.

Figs. 3.8 and 3.9 illustrate the performance of the RMSE for the estimated azimuth and elevation AoAs versus the average received SNR, by setting the number of receiving antennas to 300. These performance results show that the accuracy of angle estimation get close to the CRLB as the average received SNR increases and that the proposed method provides better accuracy than the previously known methods. Similarly with the previous set of evaluation results, it is also here observed that the positioning performance of indoor office outperforms that of the shopping mall scenario.

3.6.2 Position Estimation

In this subsection, we evaluate the performance of the proposed 3-D positioning method for both indoor office and shopping mall propagation scenarios. For each of these two scenarios, the simulations of the MS at three locations have been carried out under two SNR conditions, i.e. 10 and 0 dB. The number of received

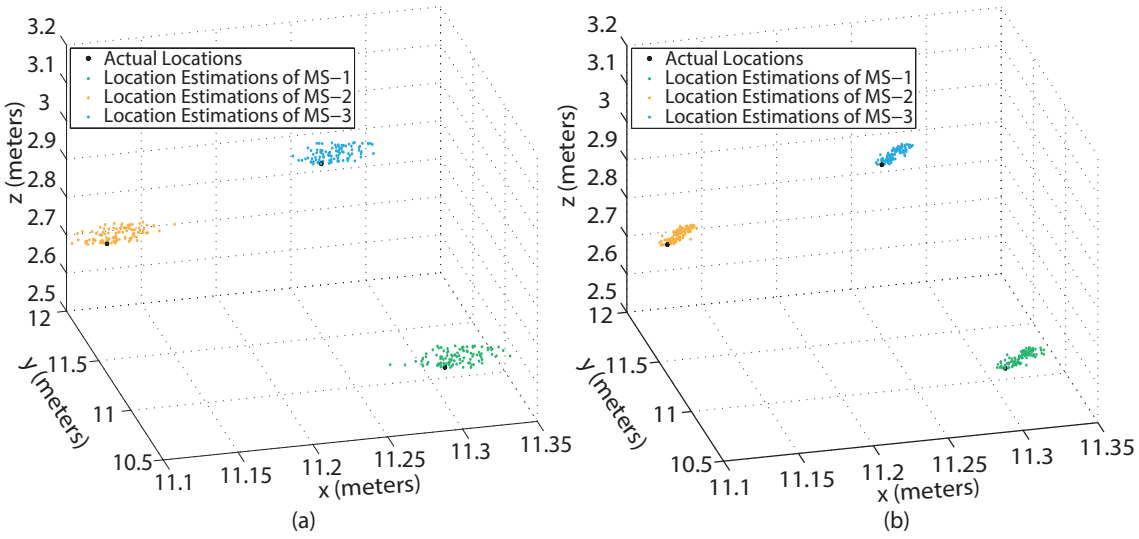


Figure 3.10 : Positioning performance for the indoor office scenario. (a) SNR=0 dB; and (b) SNR=10 dB.

antennas at the BS is 200, and the mmWave band (i.e., 28-29 GHz) is divided into 20 channels for RSS measurements. The actual locations of MS-1, MS-2 and MS-3 for the indoor office scenario are set to $(x1, y1, z1)$=(11.299 m, 10.750 m, 2.598 m), $(x2, y2, z2)$=(11.125 m, 11.948 m, 2.681 m) and $(x3, y3, z3)$=(11.237 m, 11.24 m, 3.031 m), respectively, while for shopping mall propagation scenario, they are $(x1, y1, z1)$=(8.66 m, 25.1 m, 17.321 m), $(x2, y2, z2)$=(7.071 m, 22.247 m, 16.853 m) and $(x3, y3, z3)$=(6.718 m, 26.7 m, 16.455 m). For each of the considered MS locations, at least 100 localizations trials have been carried out. The three MSs are sequentially localized. As it can be clearly seen from the results presented in Figs. 3.10 and 3.11, the proposed positioning scheme achieves decimeter-level positioning accuracy even with low SNR values, such as 0 dB. For normal operating SNR values, e.g. 10 dB, this accuracy improves even further.

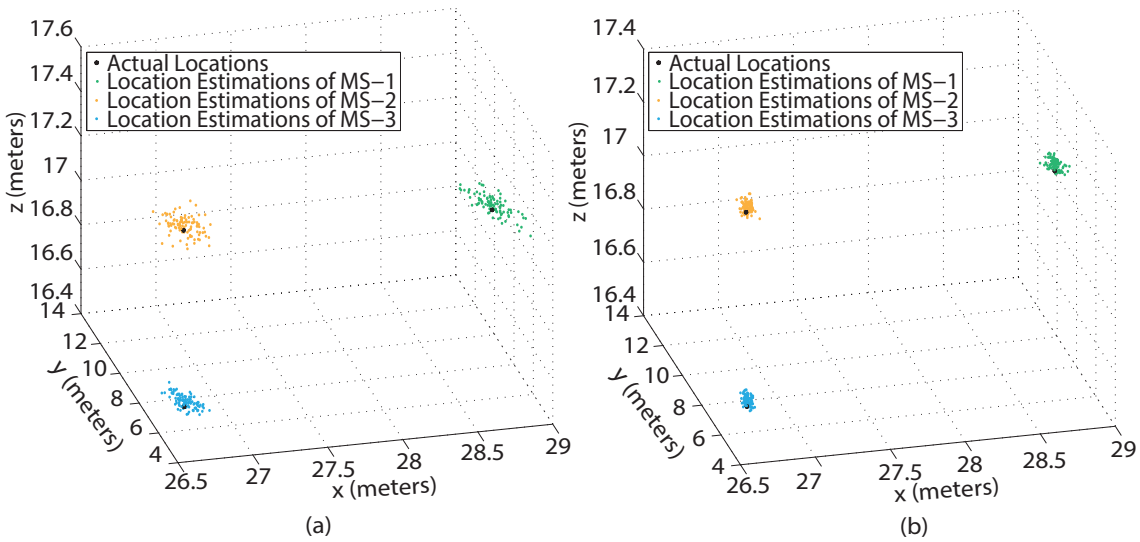


Figure 3.11 : Positioning performance for the shopping mall scenario. (a) SNR=0 dB; and (b) SNR=10 dB.

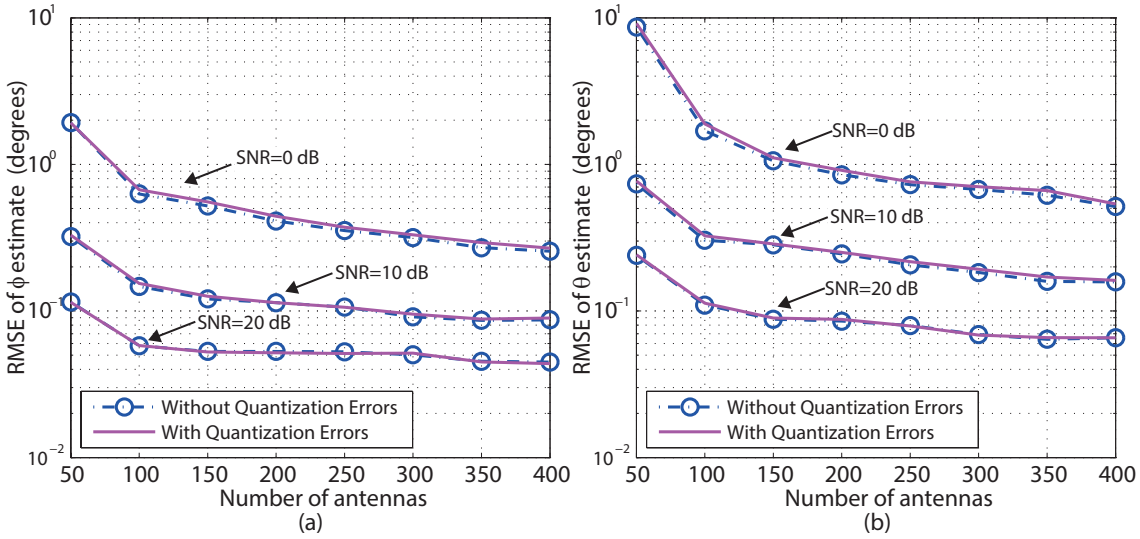


Figure 3.12 : The performance error due to the quantization assumption is evaluated under different conditions of received SNR for the indoor office scenario.

3.6.3 Effect of Quantization Error

In this subsection, the effect of the quantization on the estimation accuracy, which is introduced by the channel compressing method is evaluated. In order to validate the sensitivity of the proposed method to the quantization error, the following cases have been considered: i) The AoAs/AoDs are continuous, i.e., without quantization error; ii) The AoAs/AoDs are quantized, i.e., with quatization error. The resolution parameters G_1 and G_2 are set to 600 and 180 for azimuth and elevation angles, respectively. Note that for making a fair comparison, the virtual path selection in channel compression is not employed after quantization for the second case, and the performance simulations have been only performed for the indoor office scenario. Fig. 3.12 presents the RMSE performance evaluation results under various operating conditions with and without quantization. These results clearly show that, as long as the values of the resolution parameters G_1 and G_2 are sufficiently large, then the effect on the overall performance of neglecting the quantization error is minimal, if not non existent.

3.7 Summary

In this chapter, we have proposed a channel compression-based joint RSS-AoA estimation method for mmWave digital UCyAs. A new channel compression technique has been first designed to overcome the high computational complexity caused by the large number of antennas used at the BS. The technique filters out the received multiple reflection and diffuse scattering components, and hence, the accuracy of the parameter estimation method can also be significantly increased. We have also shown a beamspace-based channel parameter estimation approach in this chapter, which can transform the received signal vectors into the low-dimensional beamspace. As a result, the dimensions of the received signal vectors are reduced. Based on the estimated parameters, a novel mmWave indoor localization method has been pre-

sented, which only needs a single BS to achieve 3-D NLoS localization. As we have shown via simulations, in addition to the advantage of its low-complexity, the proposed indoor mmWave localization method is capable of obtaining a decimeter-level positioning accuracy even for very low SNR values. In the future, we will research the channel estimation methods in some realistic scenarios and not be narrowed to the cuboid scenario considered in this chapter.

Chapter 4

Channel Parameter Estimation for 3-D Wideband Hybrid UCyAs

The following chapters of the thesis contribute to the channel parameter estimation for hybrid UCyAs. Compared to conventional digital beamforming, the hybrid beamforming is a more appropriate architecture for mmWave large array systems, in which a low-dimensional digital precoder/combiner at baseband and a high-dimensional analog precoder/combiner at RF front-end are used to process each transmission signal. The number of expensive RF chains deployed in hybrid beamforming is typically much smaller than the numbers of the antennas and the ADCs/DACs. Hence, two major obstacles for applying large antenna arrays into actual mmWave systems, i.e., large hardware cost and high power consumption, can be effectively overcome.

In this chapter, we propose a novel channel estimation method using a wideband mmWave fully-connected hybrid UCyA. We consider the UMi scenario in 5G/B5G systems, as illustrated in Fig. 4.1. We first design a new hybrid beamformer to

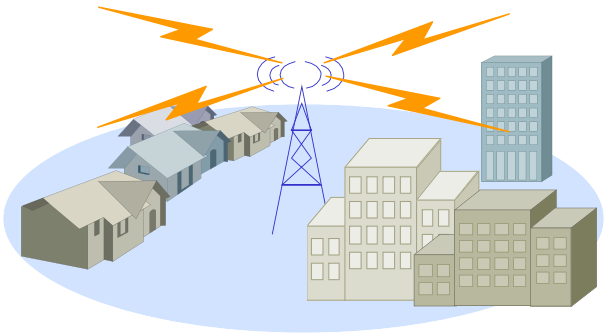


Figure 4.1 : The research scenario studied in Chapter 4.

reduce the dimension of received signals on the horizontal plane by exploiting the convergence of the Bessel function, and to reduce the active beams in the vertical direction through preselection. The important recurrence relationship of the received signals needed for subspace-based angle and delay estimation is preserved, even with substantially fewer RF chains than antennas. Then, linear interpolation is generalized to reconstruct the received signals of the hybrid beamformer, so that the signals can be coherently combined across the whole band and beam squint is suppressed. As a result, the subspace-based algorithms can be applied to estimate the angles and delays of the multi-paths. Simulations show that in the UMi scenario of future 5G/B5G systems, the proposed method can approach the CRLB of the estimation with a significantly lower computational complexity than existing techniques.

4.1 Motivation and Overview

Current hybrid beamforming schemes are typically based on CS. They exploit the angular sparsity of mmWave channels to reduce the number of RF chains. However, they need to discretize channel coefficients and would result in estimation accuracy losses [33, 67]. In this chapter, we propose a novel channel estimation method for wideband mmWave hybrid UCyAs. We address the problem of beam squint and the designed hybrid UCyA only requires a significantly small number of RF chains to accurately estimate the delay and the azimuth and elevation AoAs of the received paths. In Section 4.1, we first provide a motivation, the channel parameter estimation for wideband mmWave hybrid arrays, and a review of the state of the art. The system model is presented in Section 4.2, in which we provide an illustration of considered hybrid front-end architecture.

In Section 4.3, we propose a new two-step wideband hybrid beamforming strategy, which can reduce the number of required RF chains while preserving the multiple-invariance structure in array response vectors. At the first step, we se-

lect the needed beams vertically for angle and delay estimation by exploiting the sparsity (or low rank) nature of mmWave multi-antenna channels. By this means, we can estimate the number of paths and determine the number of vertical beams needed for the parameter estimation in Section 4.4. We then design a new hybrid beamformer to transform the high-dimensional signals of each UCA to a low dimension requiring much fewer RF chains than array elements. This is achieved by first applying Q-DFT to the signals and then exploiting the convergence property of the Bessel function to remove insignificant dimensions. Section 4.3.3 also presents a new spatial interpolation method. By applying the linear interpolation in both the vertical and horizontal spatial domains, the method combines the signals across the whole band, so that the subspace-based algorithms can be applied to estimate the angles and delays of the multi-paths.

In Section 4.4, we estimate the channel parameters by using the processed signals in Sections 4.3 and 4.3.3. Since all the operations conduced by the beamformer are linear transforms, the critical invariance structure for the validity of ESPRIT for the angle and delay estimation, can be recovered without losses between respective submatrices of the space-time response matrix for the subsequent angle and delay estimation. By exploiting the recurrence relations in the multiple-invariance structure, the delay and elevation angle of each path are estimated using ESPRIT. For the azimuth angles, since the expression for the horizontal array response vectors (4.23) does not exhibit any recurrence, we use MUSIC to estimate them based on the obtained corresponding elevation angles. Considering that the estimated parameters of each path cannot be matched automatically because of noises, a low-complexity multipath parameter matching is presented in Section 4.4.3. The method adds perturbation matrices to mitigate the mismatch of the estimated delays and angles caused by additive noises, and thus, the conventional high-complexity exhaustive search can be avoided. The hardware and software complexities of the proposed

estimation method are analyzed in the end.

Section 4.5 presents simulation results to demonstrate the performance of the proposed parameter estimation method in the UMi scenario of 5G/B5G systems. Compared with existing methods, we see that when the number of receive antennas is large, our proposed approach is better than other methods in terms of angle estimation accuracy. We also obverse that our approach achieves the best delay estimation accuracy. This is because, by using the proposed multidimensional spatial interpolation, the high temporal resolution offered by wideband mmWave signals is exploited. In this section, we also plot the RMSE of the angle estimation versus the value of the highest order under different numbers of horizontal array response vectors. We see that the number of phase-mode vectors needed in our approach does not depend on the number of array response vectors, which is important for complexity reduction, as discussed in Section 4.4.4.

4.2 System Model

We consider a mmWave multi-antenna orthogonal frequency division multiplexing (OFDM) system, where a BS with N_{R} antennas receives signals from a MS[1]. We assume that the directions and delays of the paths remain unchanged during parameter estimation. The received signal at subcarrier m ($m = 0, 1, \ldots, M-1$) is given by [64]

$$\mathbf{r}_m = \mathbf{H}_m x_m + \mathbf{n}_m, \tag{4.1}$$

[1]An omnidirectional antenna is deployed at the MSs to maintain connectivity irrespective of the orientation and posture of the MSs. One of the antenna elements at the BS is set to be the reference, so that the estimation of the MS would not rotate with respect to the BS. In the case where a directional antenna is installed at the MSs, the received signal-to-noise ratio (SNR) at BS could increase if the BS is inside the mainlobes of the MSs, or decrease otherwise. This could affect the accuracy of the proposed method in either way, while the operation of the method is unchanged.

where $\mathbf{H}_m \in \mathbb{C}^{N_\mathrm{R}\times 1}$, $\mathbf{n}_m \in \mathbb{C}^{N_\mathrm{R}\times 1}$, and x_m denote the channel matrix, the Gaussian noise, and the transmitted signal for subcarrier m, respectively; and N_p is the number of paths. The channel matrix, $\mathbf{H}_m$, can be expressed as

$$\mathbf{H}_m = \sum_{l=1}^{N_\mathrm{p}} \beta_l e^{-j2\pi f_m \tau_l} \mathbf{a}_m(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l}), \tag{4.2}$$

where β_l is the complex amplitude of the l-th path; $\mathbf{a}_m(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})$ is the array response vector with $\phi_{\mathrm{R},l}$ and $\theta_{\mathrm{R},l}$ being the azimuth and elevation of the AoAs of the l-th path. τ_l is the time delay of the l-th path. f_m is the frequency at the m-th subcarrier. $f_m = f_0 + m\Delta_\mathrm{F}$, where f_0 is the carrier frequency at the lower end of the band and Δ_F is the subcarrier spacing. If the signal bandwidth is much smaller than the carrier frequency, then $f_m \approx f_0$ and (4.2) reverts to the standard narrowband channel model.

The BS uses a hybrid UCyA antenna array. It consists of N_V horizontal layers of UCAs, each having N_H antennas, i.e., $N_\mathrm{R} = N_\mathrm{V} N_\mathrm{H}$. The radius of each UCA is r. The vertical distance between any two adjacent UCAs is h. Therefore, the array response vector is given by

$$\mathbf{a}_m(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l}) = \mathbf{a}_{\mathrm{V},m}(\theta_{\mathrm{R},l}) \otimes \mathbf{a}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l}), \tag{4.3}$$

where

$$[\mathbf{a}_{\mathrm{V},m}(\theta_{\mathrm{R},l})]_{n_\mathrm{V},1} = \frac{1}{\sqrt{N_\mathrm{V}}} \exp\left(-j\frac{2\pi}{c} f_m h (n_\mathrm{V} - \frac{N_\mathrm{V}+1}{2}) \cos(\theta_{\mathrm{R},l})\right) \tag{4.4}$$

and

$$[\mathbf{a}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})]_{n_\mathrm{H},1} = \frac{1}{\sqrt{N_\mathrm{H}}} \exp\left(j\frac{2\pi}{c} f_m r \sin(\theta_{\mathrm{R},l}) \cos(\phi_{\mathrm{R},l} - \varphi_{n_\mathrm{H}})\right) \tag{4.5}$$

are the array response vectors on the vertical and horizontal planes, respectively, with $n_\mathrm{V} = 1, 2, \ldots, N_\mathrm{V}$ and $n_\mathrm{H} = 1, 2, \ldots, N_\mathrm{H}$. c is the speed of light. Here, $\varphi_{n_\mathrm{H}} = 2\pi(n_\mathrm{H} - 1)/N_\mathrm{H}$ is the difference between the central angles of the n_H-th antenna and the first antenna of each UCA.

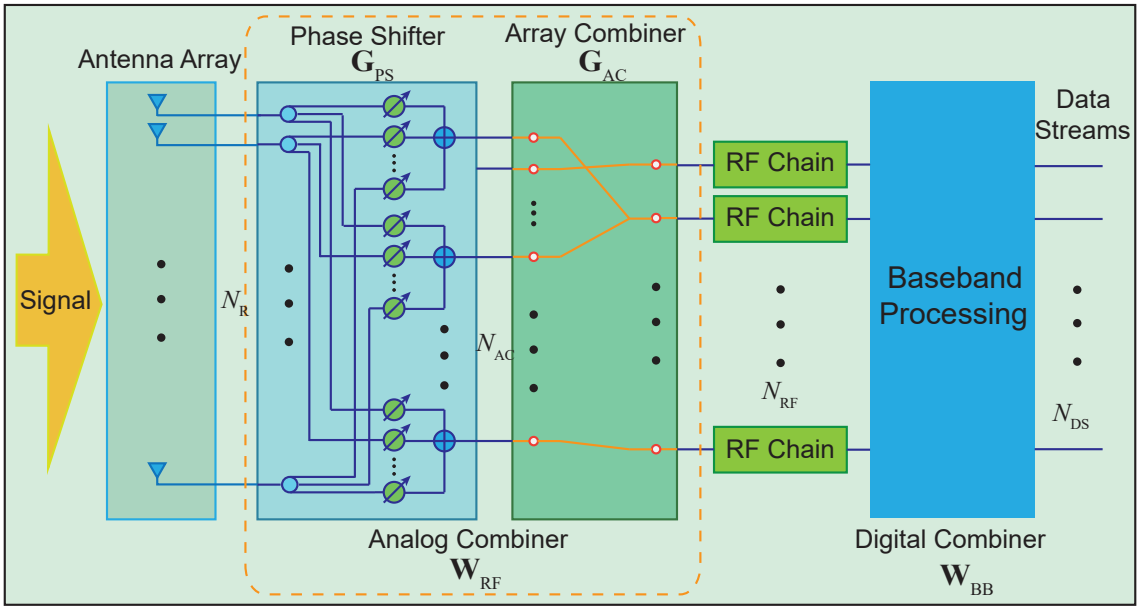


Figure 4.2 : The block diagram of hybrid beamforming architecture.

We consider a hybrid front-end architecture [107], as shown in Fig. 4.2(b). By applying a hybrid beamformer, $\mathbf{W} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{DS}}}$, to the received signal, $\mathbf{r}_m$, the output signal after beamforming can be expressed as

$$\mathbf{y}_m = \mathbf{W}^H \mathbf{r}_m = \mathbf{W}^H \mathbf{H}_m x_m + \mathbf{W}^H \mathbf{n}_m, \tag{4.6}$$

where the hybrid beamformer, $\mathbf{W} = \mathbf{W}_{\mathrm{RF}} \mathbf{W}_{\mathrm{BB}}$, is composed of an analog combiner, $\mathbf{W}_{\mathrm{RF}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{RF}}}$, and a digital combiner, $\mathbf{W_{BB}} \in \mathbb{C}^{N_{\mathbf{RF}} \times N_{\mathbf{DS}}}$. N_{RF} and N_{DS} are the numbers of RF chains and data streams, respectively.

We further divide the analog combiner, $\mathbf{W}_{\mathrm{RF}}$, into an array combiner set, $\mathbf{G}_{\mathrm{AC}} \in \mathbb{C}^{N_{\mathrm{AC}} \times N_{\mathrm{RF}}}$, and a phase shifter set, $\mathbf{G}_{\mathrm{PS}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{AS}}}$, i.e., $\mathbf{W}_{\mathrm{RF}} = \mathbf{G}_{\mathrm{PS}} \mathbf{G}_{\mathrm{AC}}$. N_{AC} is the number of the combiners deployed in the array combiner set. $N_{\mathrm{R}} \geq N_{\mathrm{AC}} \geq N_{\mathrm{RF}} \geq N_{\mathrm{DS}}$. As illustrated in Fig. 4.2, $\mathbf{G}_{\mathrm{PS}}$ is a phase shifter matrix with elements given by $[\mathbf{G}_{\mathrm{PS}}]_{n_{\mathrm{R}}, n_{\mathrm{AS}}} = \exp(j\xi)$ ($\xi \in \mathbb{R}$, $n_{\mathrm{R}} = 1, 2, \ldots, N_{\mathrm{R}}$, and $n_{\mathrm{AC}} = 1, 2, \ldots, N_{\mathrm{AC}}$). $\mathbf{G}_{\mathrm{AC}}$ is a binary matrix, and its entry $[\mathbf{G}_{\mathrm{AC}}]_{n_{\mathrm{AC}}, n_{\mathrm{RF}}} \in \{0, 1\} (n_{\mathrm{RF}} = 1, 2, \ldots, N_{\mathrm{RF}})$. Here, the role of $\mathbf{W}_{\mathrm{BB}}$ is to guarantee the power constraint.

4.3 Two-Step Wideband Hybrid Beamforming Strategy

In this section, new hybrid beamformers are designed to select the meaningful beams needed vertically for angle and delay estimation, and transform the received high-dimensional signals of horizontal UCAs to be low-dimensional by taking Q-DFT and the convergence property of the Bessel function. We prove that the number of low dimensions does not grow with the number of antennas per UCA. The minimal number of required RF chains is the product of the number of vertical beams and the number of low dimensions.

It is worth mentioning that all the beamformers we design here are linear transforms. Therefore, the critical invariance structure for the validity of ESPRIT for the angle and delay estimation, can be recovered losslessly between respective submatrices of the space-time response matrix for the subsequent angle and delay estimation.

4.3.1 Step 1: Vertical Beam Selection

We first propose a new hybrid beamformer, denoted by $\mathbf{W}_{\mathrm{s1}}$, in the vertical beamspace. By exploiting the sparsity (or low rank) nature of mmWave multi-antenna channels, the vertical beams can be selected: i) to estimate the number of paths; and ii) to determine the number of vertical beams needed for the angle and delay estimation (to be developed in Section 4.4).

The output signal after the vertical beamforming is $\mathbf{y}_{\mathrm{s1},m} = \mathbf{W}_{\mathrm{s1}}^{H}\mathbf{r}_m \in \mathbb{C}^{N_{\mathrm{DS},1}\times 1}$. The hybrid beamformer conducts vertical beamspace transforming and can be constructed as $\mathbf{W}_{\mathrm{s1}} = \mathbf{G}_{\mathrm{PS,s1}}\mathbf{G}_{\mathrm{AC,s1}}\mathbf{W}_{\mathrm{BB,s1}}$, where $\mathbf{W}_{\mathrm{BB,s1}} = \frac{1}{\sqrt{N_{\mathrm{V}}}}\mathbf{I}_{N_{\mathrm{V}}} \in \mathbb{C}^{N_{\mathrm{V}}\times N_{\mathrm{V}}}$, $\mathbf{G}_{\mathrm{PS,s1}} = \mathbf{U}_{\mathrm{d}} \otimes \mathbf{I}_{N_{\mathrm{H}}} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{R}}}$, and $\mathbf{G}_{\mathrm{AC,s1}} = \left[\mathbf{I}_{N_{\mathrm{V}}} \otimes \mathbf{1}_{N_{\mathrm{H}}}^{T}\right]^{T} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{V}}}$. Here, $\mathbf{U}_{\mathrm{d}}$ contains N_{V} orthogonal array response vectors corresponding to N_{V} vertically, an-

gularly evenly spaced beams. $\mathbf{U}_{\mathrm{d}} = [\mathbf{U}_{\mathrm{d},1}, \mathbf{U}_{\mathrm{d},2}, \ldots, \mathbf{U}_{\mathrm{d},N_{\mathrm{V}}}] \in \mathbb{C}^{N_{\mathrm{V}} \times N_{\mathrm{V}}}$, where

$$\begin{aligned}\mathbf{U}_{\mathrm{d},i} = [&\exp(-j\frac{2\pi}{N_{\mathrm{V}}}(-\frac{N_{\mathrm{V}}-1}{2})i), \exp(-j\frac{2\pi}{N_{\mathrm{V}}}(-\frac{N_{\mathrm{V}}-3}{2})i),\\ &\ldots, \exp(-j\frac{2\pi}{N_{\mathrm{V}}}(\frac{N_{\mathrm{V}}-1}{2})i)]^T,\ i = 1, 2, \ldots, N_{\mathrm{V}}.\end{aligned} \tag{4.7}$$

Thus at this step, the numbers of both data streams and RF chains are equal to that of beams, i.e., $N_{\mathrm{DS},1} = N_{\mathrm{RF},1} = N_{\mathrm{V}}$. The number of array combiners is equal to that of receive antennas, i.e., $N_{\mathrm{AC},1} = N_{\mathrm{R}}$. The i-th element of $\mathbf{y}_{\mathrm{s1},m}$ can be written as

$$[\mathbf{y}_{\mathrm{s1},m}]_{i,1} = \left[\left(\mathbf{G}_{\mathrm{PS,s1}}\mathbf{G}_{\mathrm{AC,s1}}\mathbf{W}_{\mathrm{BB,s1}}\right)^H \mathbf{r}_m\right]_i = \frac{1}{\sqrt{N_{\mathrm{V}}}}\mathbf{U}_{\mathrm{d},i}^H(\mathbf{I}_{N_{\mathrm{V}}} \otimes \mathbf{1}_{N_{\mathrm{H}}}^T)\mathbf{r}_m. \tag{4.8}$$

The total beam power at the m-th subcarrier is given by

$$\sigma_m^2 = \mathbf{y}_{\mathrm{s1},m}^H \mathbf{y}_{\mathrm{s1},m} = \sum_{i=1}^{N_{\mathrm{V}}} \sigma_{m,i}^2, \tag{4.9}$$

where $\sigma_{m,i}^2 = \left|[\mathbf{y}_{\mathrm{s1},m}]_{i,1}\right|^2$ is the power of the i-th beam which depends on the AoA of the impinging signal inside the beam. Given the sparsity of mmWave multi-antenna channels, the signal power is concentrated in a small number of beams. We select the dominant beams at the m-th subcarrier by defining an index selection set $\mathcal{U}_m$, as given by

$$\mathcal{U}_m \triangleq \{\eta(1), \eta(2), \ldots, \eta(N_{\mathrm{B},m})\}, \tag{4.10}$$

where $N_{\mathrm{B},m}$ is the number of selected beams, and $\eta(u_m)$ is the index for $\sigma_{m,\eta(u_m)}^2$ with $u_m = 1, 2, \ldots, N_{\mathrm{B},m}$. $\eta(u_m)$ can be obtained as

$$\begin{cases} \eta(1) = \arg\max_{i\in\{1,\ldots,N_{\mathrm{V}}\}} \sigma_{m,i}^2, \\ \eta(2) = \arg\max_{i\in\{1,\ldots,N_{\mathrm{V}}\}\setminus\{\eta(1)\}} \sigma_{m,i}^2, \\ \quad \vdots \\ \eta(N_{\mathrm{B},m}) = \arg\max_{i\in\{1,\ldots,N_{\mathrm{V}}\}\setminus\{\eta(1),\ldots,\eta(N_{\mathrm{B},m}-1)\}} \sigma_{m,i}^2. \end{cases} \tag{4.11}$$

The following criterion can be used to decide $N_{\mathrm{B},m}$ and select the $N_{\mathrm{B},m}$ strongest beams:

$$\sum_{u_m=1}^{N_{\mathrm{B},m}} \sigma^2_{m,\eta(u_m)} \geq \eta \sigma^2_m, \tag{4.12}$$

where η is a power threshold which can be empirically specified. η can be selected close to 1, e.g., $\eta = 0.9$, as paths reflected more than once, and diffuse scattering, account for less than 10% of the total energy, as found in [45][2]. Moreover, mmWave signals fade rapidly when reflecting off a surface [108], and become barely distinguishable from noises after two reflections [35, 45, 97].

There is dispersion in the angular domain across the bandwidth in multi-antenna wireless systems [39]. We first assume that the transmission channel at each subcarrier is narrowband. Because of small dispersion in narrowband systems, the number of orthogonal beams in the vertical beamspace is equal to the number of received paths, i.e., $N_{\mathrm{B},m} = N_{\mathrm{p}}$ [39]. However, the dispersion can have a non-negligible effect in broadband systems such as the one considered in this chapter, where a point source spreads across spatial angle and time. A strong dispersion would result in severe power loss and pulse distortion, if not addressed properly, and affect the follow-on angle and delay estimation. The dispersion effect can be characterized by the *channel dispersion factor*, γ, as specified by [39]

$$\gamma = \frac{1}{N_{\mathrm{p}}} \sum_{l=1}^{N_{\mathrm{p}}} \gamma_l = \frac{1}{N_{\mathrm{p}}} \sum_{l=1}^{N_{\mathrm{p}}} N_{\mathrm{V}} \alpha \left| \chi_{\mathrm{c},l} \right|, \tag{4.13}$$

where $\alpha = W/f_{\mathrm{c}}$ is the fractional bandwidth, $\chi_{\mathrm{c},l} = f_{\mathrm{c}} h \cos(\theta_{\mathrm{R},l})/c$ is the normalized beam angle, W is the signal bandwidth, and f_{c} is the center frequency.

To illustrate the impact of the dispersion, we assume that the system operates at f_{c} =30 GHz and the transmitted signal has unit amplitude. For simplicity, it

[2]It is shown in [45] that for mmWave systems, the contributions of paths reflected more than once, and the diffuse scattering components are weak, only accounting for less than 10% of the total energy.

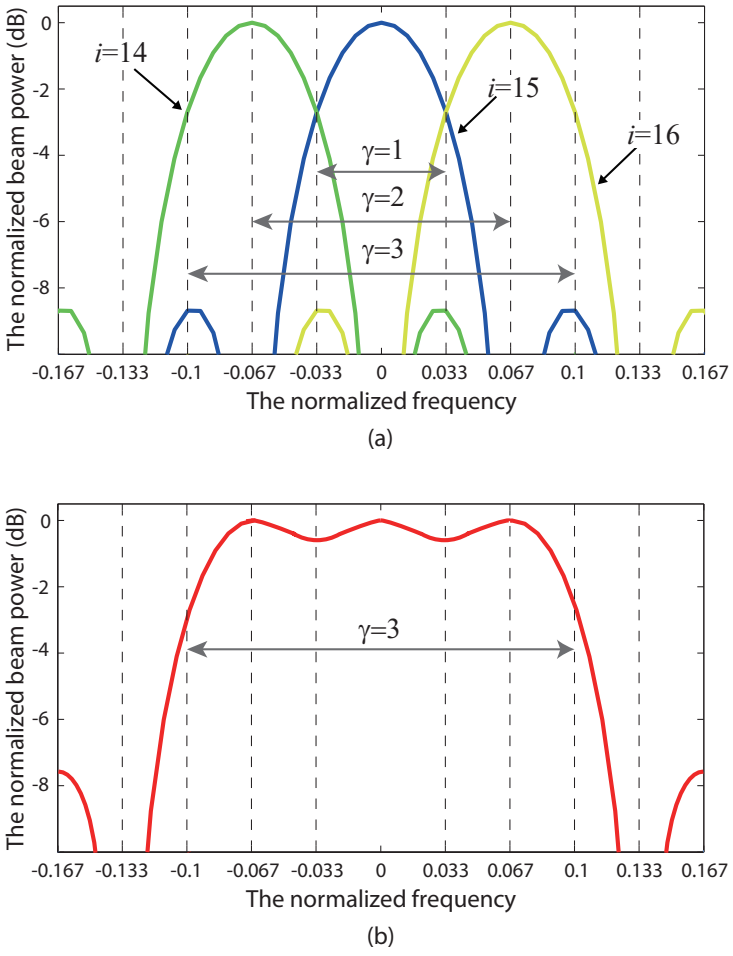


Figure 4.3 : The normalized beam power as a function of the normalized frequency. (a) The power of the14-th, 15-th, and 16-th beams; (b) The power of the combined three beams.

is assumed that only one path with $\beta_l = 1$ and $\theta_{\mathrm{R},l} = 60°$, and the number of beams is $N_{\mathrm{V}} = 60$. Thus we have $\chi_{\mathrm{c},l} \approx 0.25$, which corresponds to the $i_0 = 15$-th beam. Fig. 4.3(a) shows the normalized power, $|\mathbf{y}_{\mathrm{s1},i}(f)|^2/N_{\mathrm{V}}$, of the 14-th, 15-th, and 16-th beams versus the normalized frequency, $f_{\mathrm{normal}} = f/f_{\mathrm{c}}$, for $f \in [-W/2, W/2]$[3]. We can see that, if the normalized frequency $f_{\mathrm{normal}} < 0.033$, i.e., the channel dispersion factor $\gamma < 1$, the beams in Fig. 4.3(a) do not affect one another within the bandwidth, W. If γ is larger than 1, power loss and interference may occur. To prevent this from happening, the γ adjacent beams centered at i_0 need to be taken into consideration. In the case of $\gamma = 3$, Fig. 4.3(b) plots the normalized power of the combined one from three beams, i.e., the 14-th, 15-th, and 16-th beams. It can be seen that, by combining the three beams, the normalized power becomes approximately flat across the operating band.

Because of the dispersion, we have to jointly consider $N_{\mathrm{B}} = \gamma N_{\mathrm{p}}$ vertically spaced beams to include all possible beams, as the normalized beam angle, $\chi_{\mathrm{c},l}$, is unknown. The overall index selection set $\mathcal{U}$ is given by

$$\mathcal{U} = \mathcal{U}_0 \cup \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_{M-1}, \tag{4.14}$$

where the element of $\mathcal{U}$ is $\eta(u)$ with $u = 1, 2, \ldots, N_{\mathrm{B}}$. It is possible that the same indices are picked up at different subcarriers because of the dispersion, e.g., $\eta(u_m) = \eta(u_{m'})$ for $m' \in \{0, \ldots, M-1\} \setminus \{m\}$. We have $N_{\mathrm{B}} = \gamma N_{\mathrm{p}} \leq \sum_{m=0}^{M-1} N_{\mathrm{B},m}$ to avoid missing significant paths in the subsequent channel parameter estimation process. Algorithm 1[4] summarizes the procedure of the beam selection at this step.

[3]For convenience, here we only plot the beam power as a function of continuous frequency for illustration.

[4]Card($\mathcal{U}$) in Algorithm 1 denotes the cardinality of the set $\mathcal{U}$.

Algorithm 1 Beam selection algorithm

- **Input**: The processed signals, $\mathbf{y}_{\mathrm{s1},m}$,$(m = 0, 1, \ldots, M-1)$, the beam number, N_{V}, and the threshold, η.
- **Output**: The overall selection set, $\mathcal{U}$, and the estimated number of significant beams, N_{B}.
- **Initialization**: $\mathcal{U} = \mathcal{U}_0 = \mathcal{U}_1 = \ldots = \mathcal{U}_{M-1} = \emptyset$,
- **For** $m = 0$ **to** $M-1$ **do**
  - Set $\mathcal{V} = \emptyset$.
  - **For** $i = 1$ **to** N_{V} **do**
    - $\sigma^2_{m,i} = |\mathbf{y}_{\mathrm{s1},m,i}|^2$, and update$\mathcal{V} = \mathcal{V} \cup \{\sigma^2_{m,i}\}$.
  - **End for**
  - $\sigma^2_m = \sum_{i=1}^{N_{\mathrm{V}}} \sigma^2_{m,i}$.
  - **While** $\sum_{\eta(u_m) \in \mathcal{U}_m} \sigma^2_{m,\eta(u_m)} < \eta\sigma^2_m$ **do**
    - Find the largest $\sigma^2_{m,i}$ from $\mathcal{V}$, and update $\mathcal{U}_m = \mathcal{U}_m \cup \{i\}$ and $\mathcal{V} = \mathcal{V} \setminus \{\sigma^2_{m,i}\}$.
  - **End while**
  - Update $\mathcal{U} = \mathcal{U} \cup \mathcal{U}_m$.
- **End for**
- N_{B} =Card($\mathcal{U}$).

4.3.2 Step 2: Horizontal Q-DFT Beamforming

We proceed to design a new hybrid beamformer to transform the high-dimensional signals of each UCA to a low dimension requiring much fewer RF chains than array elements. This is achieved by first applying Q-DFT to the signals and then exploiting the convergence property of the Bessel function to remove insignificant dimensions.

We first derive an approximate expression for the array response vector to explain the design rationale of this step. According to the *Jacobi-Anger expansion* [109], the n_{H}-th array response vector on the horizontal plane can be written as

$$\begin{aligned}[\mathbf{a}_{\mathrm{H},m}(\phi_{\mathrm{R},l},\theta_{\mathrm{R},l})]_{n_{\mathrm{H}},1} &= \frac{1}{\sqrt{N_{\mathrm{H}}}} e^{j\varpi_{m,l}\cos(\phi_{\mathrm{R},l}-\varphi_{n_{\mathrm{H}}})} \\ &= \frac{1}{\sqrt{N_{\mathrm{H}}}} \sum_{q=-\infty}^{\infty} j^q J_q(\varpi_{m,l}) e^{jq(\phi_{\mathrm{R},l}-\varphi_{n_{\mathrm{H}}})}, \end{aligned} \tag{4.15}$$

where $\varpi_{m,l} = \frac{2\pi}{c} f_m r \sin(\theta_{\mathrm{R},l})$ and $J_q(\varpi_{m,l})$ is the Bessel function of the first kind of order q.

We notice that the last multiplier in (4.15), i.e., $e^{jq(\phi_{\mathrm{R},l}-\varphi_{n_{\mathrm{H}}})} = e^{jq\phi_{\mathrm{R},l}-j2\pi q(n_{\mathrm{H}}-1)/N_{\mathrm{H}}}$, which is of strong resemblance to the weight vectors in the DFT. We take the Q-DFT [110] to transform the horizontal array response vectors (4.15) to offset $\varphi_{n_{\mathrm{H}}}$. The p-th order Q-DFT of (4.15) can be expressed as

$$\begin{aligned} A_{\mathrm{PM},p} &= \sum_{n_{\mathrm{H}}=1}^{N_{\mathrm{H}}} \left([\mathbf{a}_{\mathrm{H},m}(\phi_{\mathrm{R},l},\theta_{\mathrm{R},l})]_{n_{\mathrm{H}},1}\right) e^{-j\frac{2\pi(n_{\mathrm{H}}-1)}{N_{\mathrm{H}}}p} \\ &= \frac{1}{\sqrt{N_{\mathrm{H}}}} \sum_{n_{\mathrm{H}}=1}^{N_{\mathrm{H}}} \sum_{q=-\infty}^{\infty} j^q J_q(\varpi_{m,l}) e^{jq(\phi_{\mathrm{R},l}-\varphi_{n_{\mathrm{H}}})-jp\varphi_{n_{\mathrm{H}}}} \\ &= \frac{1}{\sqrt{N_{\mathrm{H}}}} \sum_{n_{\mathrm{H}}=1}^{N_{\mathrm{H}}} \left(\sum_{q=-\infty}^{\infty} j^q J_q(\varpi_{m,l}) e^{-j\varphi_{n_{\mathrm{H}}}(p+q)} e^{jq\phi_{\mathrm{R},l}} \right). \end{aligned} \tag{4.16}$$

Let $p+q=QN_{\mathrm{H}}$, i.e., $q=QN_{\mathrm{H}}-p$. Then, (4.16) can be rewritten as

$$\begin{aligned} A_{\mathrm{PM},p} &= \frac{1}{\sqrt{N_{\mathrm{H}}}} \sum_{Q=-\infty}^{\infty} N_{\mathrm{H}} j^{(QN_{\mathrm{H}}-p)} J_{(QN_{\mathrm{H}}-p)}(\varpi_{m,l}) e^{j(QN_{\mathrm{H}}-p)\phi_{\mathrm{R},l}} \\ &\overset{(\star)}{=} \sqrt{N_{\mathrm{H}}} \left[j^p J_p(\varpi_{m,l}) e^{-jp\phi_{\mathrm{R},l}} + \sum_{Q=-\infty, Q\neq 0}^{\infty} \varepsilon_{p,Q}(\varpi_{m,l}, \phi_{\mathrm{R},l}) \right], \end{aligned} \tag{4.17}$$

where

$$\varepsilon_{p,Q}(\varpi_{m,l}, \phi_{\mathrm{R},l}) = j^{(QN_{\mathrm{H}}-p)} J_{(QN_{\mathrm{H}}-p)}(\varpi_{m,l}) e^{j(QN_{\mathrm{H}}-p)\phi_{\mathrm{R},l}}.$$

$(\star)$ is obtained by the property of the Bessel function $J_{-v}(x)=(-1)^v J_v(x)$ [109].

Lemma 1. *For the Bessel function $J_v(x)$, when its order v is larger than its argument x, i.e., $|v|>|x|$, the amplitude of $J_v(x)$ is so small and negligible, i.e., $|J_v(x)|\approx 0$.*

Proof. See Appendix C.1. □

From Lemma 1, we can derive the following theorem on the approximation of the horizontal array response vector.

Theorem 1. *If $N_{\mathrm{H}} \geq 2P$, the N_{H}-dimensional array response vectors on the horizontal plane can be transformed to a much smaller $(2P+1)$-dimensional space with negligible loss, i.e., $p \in [-P,P] \cap \mathbb{Z}$, where $P=\lfloor 2\pi f_0 r/c \rfloor$ is the highest order. The p-th order of the $(2P+1)$-dimensional vector, $A_{PM,p}$, can be approximated as $A_{PM,p} \approx \sqrt{N_H} j^p J_p(\varpi_{m,l}) \exp\left(-jp\phi_{R,l}\right)$.*

Proof. See Appendix C.2. □

According to Theorem 1, we see that, by using the Q-DFT, the N_{H}-dimensional array response vectors of the horizontal UCA can be transformed to only $(2P+1)$ dimensions, and each element of the vector can be approximately expressed as an

exponential function weighted by a Bessel function of the same order, as long as the conditions in Theorem 1, i.e., $N_{\rm H} \geq 2P$, is met[5].

We note that Q-DFT is a linear transform and hence can preserve the multiple-invariance structure of the array response vectors for the subsequent angle and delay estimation, as will be elaborated on in Section 4.4; see (4.33) and (4.40). Thus, combining with the beams selected at Step 1, we can design the values of the phase shifters based on the Q-DFT and the beamspace transform to convert the array response vectors to a low dimension. Only a small number of RF chains are needed for channel parameter estimation.

At this step, the hybrid beamformer is $\mathbf{W}_{\rm s2} = \mathbf{G}_{\rm PS,s2}\mathbf{G}_{\rm AC,s2}\mathbf{W}_{\rm BB,s2}$. Then we have

$$\mathbf{y}_{{\rm s2},m} = \mathbf{W}_{\rm s2}^H \mathbf{r}_m = (\mathbf{G}_{\rm PS,s2}\mathbf{G}_{\rm AC,s2}\mathbf{W}_{\rm BB,s2})^H \mathbf{r}_m \in \mathbb{C}^{N_{\rm DS,2}\times 1}, \tag{4.18}$$

where $\mathbf{W}_{\rm BB,s2} = \sqrt{N_{\rm V}/N_{\rm H}}\mathbf{I}_{(2P+1)N_{\rm B}} \in \mathbb{C}^{(2P+1)N_{\rm B}\times(2P+1)N_{\rm B}}$, $\mathbf{J}_{\rm B} = [\mathbf{J}_{{\rm B},1}, \mathbf{J}_{{\rm B},2}, \ldots, \mathbf{J}_{{\rm B},N_{\rm B}}] \in \mathbb{R}^{N_{\rm V}\times N_{\rm B}}$, and $\mathbf{G}_{\rm AC,s2} = \mathbf{J}_{\rm B} \otimes \mathbf{I}_{(2P+1)} \in \mathbb{R}^{(2P+1)N_{\rm V}\times(2P+1)N_{\rm B}}$. The element of $\mathbf{J}_{{\rm B},u} \in \mathbb{R}^{N_{\rm V}\times 1}$ is

$$[\mathbf{J}_{{\rm B},u}]_{n_{\rm V},1} = \begin{cases} 1, & \text{if } n_{\rm V} = \eta(u); \\ 0, & \text{otherwise.} \end{cases} \tag{4.19}$$

We design the phase shifter set of the analog part of the hybrid array, as $\mathbf{G}_{\rm PS,s2} = \mathbf{U}_{\rm d} \otimes \mathbf{U}_{\rm sH} \in \mathbb{C}^{N_{\rm R}\times(2P+1)N_{\rm V}}$, where $\mathbf{U}_{\rm d}$ is given in (4.7) and the element of $\mathbf{U}_{\rm sH} \in \mathbb{C}^{N_{\rm H}\times(2P+1)}$ can be expressed as $[\mathbf{U}_{\rm sH}]_{n_{\rm H},p+P+1} = e^{j2\pi(n_{\rm H}-1)p/N_{\rm H}}$. Hence, the analog combiner of the hybrid array can be constructed as

$$\begin{aligned} \mathbf{W}_{\rm RF,s2} = \mathbf{G}_{\rm PS,s2}\mathbf{G}_{\rm AC,s2} &= (\mathbf{U}_{\rm d} \otimes \mathbf{U}_{\rm sH})(\mathbf{J}_{\rm B} \otimes \mathbf{I}_{(2P+1)}) \\ &\overset{(\star)}{=} (\mathbf{U}_{\rm d}\mathbf{J}_{\rm B}) \otimes (\mathbf{U}_{\rm sH1}\mathbf{I}_{(2P+1)}) = \mathbf{U}_{\rm sV} \otimes \mathbf{U}_{\rm sH}, \end{aligned} \tag{4.20}$$

[5]In general, this condition can be met in large-scale antenna array systems, where a large number of antennas are deployed.

where $(\star)$ follows a property of the Khatri-Rao product, i.e., $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$. The element of $\mathbf{U}_{\mathrm{sV}} \in \mathbb{C}^{N_{\mathrm{V}} \times N_{\mathrm{B}}}$ can be calculated as $[\mathbf{U}_{\mathrm{sV}}]_{n_{\mathrm{V}},u} = \exp(-j\frac{2\pi}{N_{\mathrm{V}}}(\frac{N_{\mathrm{V}}+1}{2} - n_{\mathrm{V}})\eta(u))$. As a result, at this step we have $N_{\mathrm{DS},2} = N_{\mathrm{RF},2} = (2P+1)N_{\mathrm{B}}$ data streams and RF chains, and $N_{\mathrm{AC},2} = (2P+1)N_{\mathrm{V}}$ array combiners. The processed received signal in (4.18) can be written as

$$\begin{aligned}
&\mathbf{y}_{\mathrm{s2},m} = \mathbf{W}_{\mathrm{s2}}^H \mathbf{r}_m = (\mathbf{G}_{\mathrm{s2}}\mathbf{W}_{\mathrm{RF,s2}}\mathbf{W}_{\mathrm{BB,s2}})^H \mathbf{H}_m x_m + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m \\
&= \sqrt{\frac{N_{\mathrm{V}}}{N_{\mathrm{H}}}}(\mathbf{U}_{\mathrm{sV}} \otimes \mathbf{U}_{\mathrm{sH}})^H \sum_{l=1}^{N_{\mathrm{p}}} \beta_l x_m e^{-j2\pi f_m \tau_l} \mathbf{a}_m(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l}) + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m \\
&\overset{(\star)}{=} \sqrt{\frac{N_{\mathrm{V}}}{N_{\mathrm{H}}}} \sum_{l=1}^{N_{\mathrm{p}}} \beta_l x_m e^{-j2\pi f_m \tau_l} (\mathbf{U}_{\mathrm{sV}}^H \otimes \mathbf{U}_{\mathrm{sH}}^H)(\mathbf{a}_{\mathrm{V},m}(\theta_{\mathrm{R},l}) \otimes \mathbf{a}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})) + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m \\
&= \sum_{l=1}^{N_{\mathrm{p}}} \beta_l x_m e^{-j2\pi f_m \tau_l} (\sqrt{N_{\mathrm{V}}}\mathbf{U}_{\mathrm{sV}}^H \mathbf{a}_{\mathrm{V},m}(\theta_{\mathrm{R},l})) \otimes (\frac{1}{\sqrt{N_{\mathrm{H}}}}\mathbf{U}_{\mathrm{sH}}^H \mathbf{a}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})) + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m \\
&= \sum_{l=1}^{N_{\mathrm{p}}} \beta_l x_m e^{-j2\pi f_m \tau_l} (\tilde{\mathbf{a}}_{\mathrm{V},m}(\theta_{\mathrm{R},l}) \otimes \tilde{\mathbf{a}}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})) + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m, \qquad (4.21)
\end{aligned}$$

where $(\star)$ stems from another property of the Khatri-Rao product, i.e., $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$. According to Theorem 1, the elements of the resulting vertical and horizontal array response vectors $\tilde{\mathbf{a}}_{\mathrm{V},m}(\theta_{\mathrm{R},l})$ and $\tilde{\mathbf{a}}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})$ are given by

$$\begin{aligned}
&[\tilde{\mathbf{a}}_{\mathrm{V},m}(\theta_{\mathrm{R},l})]_{u,1} = \sqrt{N_{\mathrm{V}}}\mathbf{U}_{\mathrm{sV}}^H \mathbf{a}_{\mathrm{V},m}(\theta_{\mathrm{R},l}) \\
&= \sum_{n_{\mathrm{V}}=1}^{N_{\mathrm{V}}} \exp\left(-j\frac{2\pi}{c} f_m h (n_{\mathrm{V}} - \frac{N_{\mathrm{V}}+1}{2}) \cos(\theta_{\mathrm{R},l})\right) \exp\left(j\frac{2\pi}{N_{\mathrm{V}}}(\frac{N_{\mathrm{V}}+1}{2} - n_{\mathrm{V}})\eta(u)\right) \\
&= \frac{\sin\left(N_{\mathrm{V}}(2\pi f_m h \cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u)/N_{\mathrm{V}})/2\right)}{\sin\left((2\pi f_m h \cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u)/N_{\mathrm{V}})/2\right)} \qquad (4.22)
\end{aligned}$$

and

$$\begin{aligned}
&[\tilde{\mathbf{a}}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})]_{p+P+1,1} = \frac{1}{\sqrt{N_{\mathrm{H}}}}\mathbf{U}_{\mathrm{sH}}^H \mathbf{a}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l}) \\
&= \frac{1}{\sqrt{N_{\mathrm{H}}}} A_{\mathrm{PM},p} \approx j^p J_p(\varpi_{m,l}) e^{-jp\phi_{\mathrm{R},l}}. \qquad (4.23)
\end{aligned}$$

Steps 1 and 2 are indispensable, reducing the number of required RF chains substantially from N_{R} to $(2P+1)N_{\mathrm{B}}$.

4.3.3 Multidimensional Spatial Interpolation

When the fractional bandwidth or the scale of the antenna array is large, the aforementioned beam squint effect arises [39]. This is because the array response vectors (4.22) and (4.23) depend on the frequency of the specific subcarrier f_m. The beam squint effect would compromise the capability of jointly utilizing the received signals at all frequency bands to estimate the path parameters. As a result, the high temporal resolution of wideband mmWave systems could not be effectively exploited.

One could keep the array response matrices consistent across all frequencies, by transforming the array response vectors (4.22) and (4.23) associated with the frequency f_m, $\forall m = 0, 1, \ldots, M-1$, into the corresponding array response vectors at the reference frequency f_0 [71]. For continuous signals, this could be ideally achieved by the Shannon-Whittaker interpolation [111], which sets different vertical distances and radii at different frequencies, i.e., $h_{\mathrm{vi},m} = f_0 h / f_m$ and $r_{\mathrm{vi},m} = f_0 r / f_m$. Then, from (4.22) and (4.23), the virtual vertical and horizontal response vectors, $\dot{\mathbf{a}}_{\mathrm{V},m}(\theta_{\mathrm{R},l})$ and $\dot{\mathbf{a}}_{\mathrm{H},m}(\theta_{\mathrm{R},l})$, can be constructed as

$$\begin{aligned}
[\dot{\mathbf{a}}_{\mathrm{V},m}(\theta_{\mathrm{R},l})]_{u,1} &= \frac{\sin\left(N_{\mathrm{V}}(2\pi f_m h_{\mathrm{vi},m}\cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u)/N_{\mathrm{V}})/2\right)}{\sin\left((2\pi f_m h_{\mathrm{vi},m}\cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u)/N_{\mathrm{V}})/2\right)} \\
&= \frac{\sin\left(N_{\mathrm{V}}(2\pi f_0 h\cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u)/N_{\mathrm{V}})/2\right)}{\sin\left((2\pi f_0 h\cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u)/N_{\mathrm{V}})/2\right)} \\
&= [\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u,1}
\end{aligned} \tag{4.24}$$

and

$$\begin{aligned}
&[\dot{\mathbf{a}}_{\mathrm{H},m}(\phi_{\mathrm{R},l},\theta_{\mathrm{R},l})]_{p+P+1,1} = j^p J_p\left(\frac{2\pi}{c} f_m r_{\mathrm{vi},m}\sin(\theta_{\mathrm{R},l})\right)e^{-jp\phi_{\mathrm{R},l}} \\
&= j^p J_p\left(\frac{2\pi}{c} f_0 r\sin(\theta_{\mathrm{R},l})\right)e^{-jp\phi_{\mathrm{R},l}} = [\tilde{\mathbf{a}}_{\mathrm{H},0}(\phi_{\mathrm{R},l},\theta_{\mathrm{R},l})]_{p+P+1,1}.
\end{aligned} \tag{4.25}$$

The signal reconstructed by using (4.24) and (4.25) can be expressed as

$$\begin{aligned}\dot{\mathbf{y}}_{\mathrm{s2},m} &= \sum_{l=1}^{N_\mathrm{p}} \beta_l x_m e^{-j2\pi f_m \tau_l} (\dot{\mathbf{a}}_{\mathrm{V},m}(\theta_{\mathrm{R},l}) \otimes \dot{\mathbf{a}}_{\mathrm{H},m}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})) + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m \\ &= \sum_{l=1}^{N_\mathrm{p}} \beta_l x_m e^{-j2\pi f_m \tau_l} (\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l}) \otimes \tilde{\mathbf{a}}_{\mathrm{H},0}(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l})) + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m \\ &= \sum_{l=1}^{N_\mathrm{p}} \beta_l x_m e^{-j2\pi f_m \tau_l} \tilde{\mathbf{a}}_0(\phi_{\mathrm{R},l}, \theta_{\mathrm{R},l}) + \mathbf{W}_{\mathrm{s2},m}^H \mathbf{n}_m. \end{aligned} \tag{4.26}$$

In practice, unfortunately, the Shannon-Whittaker interpolation could hardly achieve perfect signal reconstruction for time-limited signals, and it also has a high computational complexity [111].

In this chapter, we extend linear interpolation [112] (which is a low-complexity and effective method for data point construction) to the multidimensional spatial interpolation. The multidimensional array response matrices consistent across all frequencies can be constructed by using the received time-limited signals. By applying the linear interpolation in both the vertical and horizontal spatial domains, we can reconstruct the signal in (4.21) and obtain an approximation of (4.26). The reconstructed signal is calculated as

$$[\tilde{\mathbf{y}}_{\mathrm{s2},m}]_{n_{\mathrm{DS},2},1} = [\mathbf{y}_{\mathrm{s2},m}]_{n_{\mathrm{DS},2},1} + \frac{r_{\mathrm{vi},m}}{r} \Delta_{\mathbf{y}_{\mathrm{s2,H},m}} + \frac{h_{\mathrm{vi},m}}{h} \Delta_{\mathbf{y}_{\mathrm{s2,V},m}}, \tag{4.27}$$

where $n_{\mathrm{DS},2} = (2P+1)(u-1)+p$. If $n_{\mathrm{DS},2} \leq (2P+1)(N_\mathrm{B}-1)$, $\Delta_{\mathbf{y}_{\mathrm{s2,H},m}}$ and $\Delta_{\mathbf{y}_{\mathrm{s2,V},m}}$ are constructed as $\Delta_{\mathbf{y}_{\mathrm{s2,H},m}} = [\mathbf{y}_{\mathrm{s2},m}]_{(n_{\mathrm{DS},2}+1),1} - [\mathbf{y}_{\mathrm{s2},m}]_{n_\mathrm{S},1}$ and $\Delta_{\mathbf{y}_{\mathrm{s2,V},m}} = [\mathbf{y}_{\mathrm{s2},m}]_{(n_{\mathrm{DS},2}+2P+1),1} - [\mathbf{y}_{\mathrm{s2},m}]_{n_{\mathrm{DS},2},1}$, respectively. Otherwise, $\Delta_{\mathbf{y}_{\mathrm{s2,H},m}} = [\mathbf{y}_{\mathrm{s2},m}]_{n_{\mathrm{DS},2},1} - [\mathbf{y}_{\mathrm{s2},m}]_{(n_{\mathrm{DS},2}-1),1}$ and $\Delta_{\mathbf{y}_{\mathrm{s2,V},m}} = [\mathbf{y}_{\mathrm{s2},m}]_{n_{\mathrm{DS},2},1} - [\mathbf{y}_{\mathrm{s2},m}]_{(n_{\mathrm{DS},2}-2P-1),1}$.

By using (4.27), all the subcarrier signals can be coherently combine, and thus the high temporal resolution offered by wideband mmWave systems can be utilized to improve the delay estimation accuracy. However, if we directly use the OFDM channel model (4.1), we process the signals at each subcarrier separately, which cannot exploit the high temporal resolution of wideband mmWave systems.

4.4 Wideband JDAE Algorithm

In this section, we estimate the path parameters of the MS based on the processed signals in Sections 4.3 and 4.3.3. Since the beamformers developed in Section 4.3 are linear transforms, the multiple-invariance structure required for ESPRIT can be recovered losslessly between respective submatrices of the space-time response matrix. By exploiting the recurrence relations in the multiple-invariance structure, the delay and elevation angle of each path can be estimated using ESPRIT. For the azimuth angles, the expression for the horizontal array response vectors (4.23) does not exhibit any recurrence. Hence the azimuth angles are estimated by using MUSIC after obtaining the corresponding elevation angles. The hardware and software complexities of the proposed estimation method are analyzed in the end.

Collecting the received signals at all frequencies, we have $\tilde{\mathbf{y}} = [\tilde{\mathbf{y}}_{\mathrm{s2},1}, \tilde{\mathbf{y}}_{\mathrm{s2},2}, \ldots, \tilde{\mathbf{y}}_{\mathrm{s2},M}]$. Assume that the same signals are transmitted at all subcarriers. We can vectorize $\tilde{\mathbf{y}}$ as

$$\tilde{\mathbf{y}}_{\mathrm{vec}} = \mathrm{vec}(\tilde{\mathbf{y}}) = \left[\mathbf{\Gamma} \diamond \tilde{\mathbf{A}}\right] \mathbf{d} + \mathrm{vec}(\tilde{\mathbf{n}}) = \mathbf{U}\mathbf{d} + \tilde{\mathbf{n}}_{\mathrm{v}}, \tag{4.28}$$

where $\tilde{\mathbf{A}} = \left[\tilde{\mathbf{a}}_0(\phi_{\mathrm{R},1}, \theta_{\mathrm{R},1}), \ldots, \tilde{\mathbf{a}}_0(\phi_{\mathrm{R},N_\mathrm{p}}, \theta_{\mathrm{R},N_\mathrm{p}})\right]$, $\tilde{\mathbf{n}} = \mathbf{W}_{\mathrm{s2}}^H [\mathbf{n}_1, \ldots, \mathbf{n}_M]$, $[\mathbf{\Gamma}]_{m,l} = e^{-j2\pi f_m \tau_l}$ and $\mathbf{d} = x\left[\beta_1, \beta_2, \ldots, \beta_{N_\mathrm{p}}\right]^T$. Here, $\mathbf{U} \in \mathbb{C}^{N_{\mathrm{DS},2}M \times N_\mathrm{p}}$, also known as the space-time response matrix in [113], collects the set of AoAs and path delays. The covariance matrix of $\tilde{\mathbf{y}}_{\mathrm{vec}}$ can be calculated as

$$\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec}}} = \mathrm{E}\left\{\tilde{\mathbf{y}}_{\mathrm{vec}} \tilde{\mathbf{y}}_{\mathrm{vec}}^H\right\} = \mathbf{U}\mathbf{\Lambda}_{\mathbf{d}}\mathbf{U}^H + \sigma_{\mathrm{n}}^2 \mathbf{I}_{(N_{\mathrm{DS},2}M)}, \tag{4.29}$$

where $\mathbf{\Lambda}_{\mathbf{d}} = \mathrm{E}\left\{\mathbf{d}\mathbf{d}^H\right\}$ is a diagonal matrix. The EVD of $\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec}}}$ can be obtained by

$$\begin{aligned}\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec}}} &= [\mathbf{E}_{\mathrm{s}}, \mathbf{E}_{\mathrm{n}}] \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{s}} & \mathbf{0}_{N_\mathrm{p} \times (N_{\mathrm{DS},2}M - N_\mathrm{p})} \\ \mathbf{0}_{(N_{\mathrm{DS},2}M - N_\mathrm{p}) \times N_\mathrm{p}} & \sigma_{\mathrm{n}}^2 \mathbf{I}_{N_{\mathrm{DS},2}M - N_\mathrm{p}} \end{bmatrix} \times [\mathbf{E}_{\mathrm{s}}, \mathbf{E}_{\mathrm{n}}]^H \\ &= \mathbf{E}_{\mathrm{s}} \mathbf{\Sigma}_{\mathrm{s}} \mathbf{E}_{\mathrm{s}}^H + \sigma_{\mathrm{n}}^2 \mathbf{E}_{\mathrm{n}} \mathbf{E}_{\mathrm{n}}^H, \end{aligned} \tag{4.30}$$

where $\mathbf{E}_{\mathrm{s}} \in \mathbb{C}^{N_{\mathrm{DS},2}M \times N_{\mathrm{p}}}$ and $\mathbf{E}_{\mathrm{n}} \in \mathbb{C}^{N_{\mathrm{DS},2}M \times (N_{\mathrm{DS},2}M - N_{\mathrm{p}})}$ correspond to the signal subspace and noise subspace, respectively. $\mathbf{\Sigma}_{\mathrm{s}} \in \mathbb{R}^{N_{\mathrm{p}} \times N_{\mathrm{p}}}$ is a diagonal matrix whose elements are the N_{p} largest eigenvalues of $\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec}}}$. Based on $\mathbf{E}_{\mathrm{n}}\mathbf{E}_{\mathrm{n}}^{H} + \mathbf{E}_{\mathrm{s}}\mathbf{E}_{\mathrm{s}}^{H} = \mathbf{I}_{N_{\mathrm{DS},2}M}$, (4.30) can be rewritten as

$$\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec}}} = \mathbf{E}_{\mathrm{s}}(\mathbf{\Sigma}_{\mathrm{s}} - \sigma_{\mathrm{n}}^{2}\mathbf{I}_{N_{\mathrm{p}}})\mathbf{E}_{\mathrm{s}}^{H} + \sigma_{\mathrm{n}}^{2}\mathbf{I}_{N_{\mathrm{DS},2}M}. \tag{4.31}$$

By letting (4.29) equal to (4.31), we obtain

$$\mathbf{E}_{\mathrm{s}} = \mathbf{U}\mathbf{T}, \tag{4.32}$$

where $\mathbf{T} \in \mathbb{C}^{N_{\mathrm{p}} \times N_{\mathrm{p}}}$ is a full rank matrix.

As noticed below, $\mathbf{U}$ in (4.32) has a multiple-invariance structure with a linear recurrence relationship. The relationship can facilitate utilizing the ESPRIT method to estimate the delay and elevation angle of each path.

4.4.1 Delay Estimation

Define the delay-selection matrix as $\mathbf{J}_{\mathrm{D}} = \mathrm{diag}\,(\mathbf{J}_{\mathrm{D},1}, \ldots, \mathbf{J}_{\mathrm{D},M}) \in \mathbb{R}^{M \times N_{\mathrm{DS},2}M}$, where $\mathbf{J}_{\mathrm{D},m} = \mathbf{1}_{N_{\mathrm{DS},2}}^{T}$. We can obtain the delay-related submatrix $\mathbf{U}_{\mathrm{D}} = \mathbf{J}_{\mathrm{D}}\mathbf{U} \in \mathbb{C}^{M \times N_{\mathrm{p}}}$. By defining $\tilde{\mathbf{J}}_{\mathrm{D},m} = [\mathbf{0}_{1\times(m-1)}, 1, \mathbf{0}_{1\times(M-m)}] \in \mathbb{R}^{1\times M}$, the delay-related submatrix associated with the frequency f_m can be calculated as $\mathbf{U}_{\mathrm{D},m} = \tilde{\mathbf{J}}_{\mathrm{D},m}\mathbf{U}_{\mathrm{D}} \in \mathbb{C}^{1\times N_{\mathrm{p}}}$. Thus, we can obtain a linear recurrence relation between the delay-related submatrices of each frequency as

$$\mathbf{U}_{\mathrm{D},\tilde{m}+1} = \mathbf{U}_{\mathrm{D},\tilde{m}}\mathbf{\Theta}_{\mathrm{D}}, \tag{4.33}$$

where $\mathbf{\Theta}_{\mathrm{D}} = \mathrm{diag}\left(e^{-j2\pi\Delta_{\mathrm{F}}\tau_1}, \ldots, e^{-j2\pi\Delta_{\mathrm{F}}\tau_{N_{\mathrm{p}}}}\right) \in \mathbb{C}^{N_{\mathrm{p}}\times N_{\mathrm{p}}}$ and $\tilde{m} = 1, 2, \ldots M-1$.

According to (4.32), the delay-related submatrix of the signal subspace matrix at the frequency f_m can be given by

$$\mathbf{E}_{\mathrm{D},m} = \tilde{\mathbf{J}}_{\mathrm{D},m}\mathbf{J}_{\mathrm{D}}\mathbf{E}_{\mathrm{s}} = \mathbf{U}_{\mathrm{D},m}\mathbf{T}. \tag{4.34}$$

Substituting (4.33) into (4.34), we obtain

$$\mathbf{E}_{\mathrm{D},\tilde{m}+1} = \mathbf{E}_{\mathrm{D},\tilde{m}}\mathbf{T}^{-1}\mathbf{\Theta}_{\mathrm{D}}\mathbf{T} = \mathbf{E}_{\mathrm{D},\tilde{m}}\mathbf{\Psi}_{\mathrm{D}}. \tag{4.35}$$

By using the total least-squares (TLS) criterion [37], we estimate $\Psi_{\mathrm{D}} = \mathbf{T}^{-1}\mathbf{\Theta}_{\mathrm{D}}\mathbf{T}$ as $\hat{\mathbf{\Psi}}_{\mathrm{D},\tilde{m}}$, each of which has in total N_{p} sorted eigenvalues, i.e., $\lambda_{\mathrm{D},\tilde{m},N_{\mathrm{p}}}$. Due to the fact that the eigenvalues of an upper triangular matrix are also diagonal elements of the matrix, we can obtain $(M-1)$ different estimates for each $\mathbf{\Theta}_{\mathrm{D}}$. As a result, the delay of the l-th path, τ_l, can be estimated as

$$\hat{\tau}_l = \frac{1}{M-1}\sum_{\tilde{m}}^{M-1}\left[j\ln(\lambda_{\mathrm{D},\tilde{m},l})/2\pi\Delta_{\mathrm{F}}\right]. \tag{4.36}$$

4.4.2 Angle Estimation

We first use the processed vertical array response vector (4.24) to estimate the elevation angle. According to (4.22), the $(u+1)$-th element of $\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})$ can be given by

$$[\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u+1,1} = \frac{\sin\left(N_{\mathrm{V}}(2\pi f_0 h\cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u+1)/N_{\mathrm{V}})/2\right)}{\sin\left((2\pi f_0 h\cos(\theta_{\mathrm{R},l})/c - 2\pi\eta(u+1)/N_{\mathrm{V}})/2\right)}. \tag{4.37}$$

Comparing $[\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u+1,1}$ with the u-th element in (4.24), we see that two successive components of the processed vertical array response vector, $[\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u,1}$ and $[\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u+1,1}$, are related as follows.

$$\begin{aligned}&(-1)^{\eta(u)}\sin\left((g(\theta_{\mathrm{R},l}) - 2\pi\eta(u)/N_{\mathrm{V}})/2\right)[\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u,1}\\ &= (-1)^{\eta(u+1)}\sin\left((g(\theta_{\mathrm{R},l}) - 2\pi\eta(u+1)/N_{\mathrm{V}})/2\right)\times[\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u+1,1},\end{aligned} \tag{4.38}$$

where $g(\theta_{\mathrm{R},l}) = \frac{2\pi}{c} f_0 h \cos(\theta_{\mathrm{R},l})$. By trigonometric manipulations, we rewrite (4.38) as

$$\begin{aligned}
&(-1)^{\eta(u)} \sin\left(\eta(u)\frac{\pi}{N_{\mathrm{V}}}\right) [\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u,1} \\
&\quad + (-1)^{\eta(u+1)+1} \sin\left(\eta(u+1)\frac{\pi}{N_{\mathrm{V}}}\right) [\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u+1,1} \\
&= \tan\left(\frac{g(\theta_{\mathrm{R},l})}{2}\right) \left[(-1)^{\eta(u)} \cos\left(\eta(u)\frac{\pi}{N_{\mathrm{V}}}\right) [\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u,1}\right. \\
&\quad \left. +(-1)^{\eta(u+1)+1} \cos\left(\eta(u+1)\frac{\pi}{N_{\mathrm{V}}}\right) [\tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l})]_{u+1,1}\right]. \qquad (4.39)
\end{aligned}$$

Stacking all $(N_{\mathrm{B}} - 1)$ equations together yields

$$\tan\left(\frac{g(\theta_{\mathrm{R},l})}{2}\right) \mathbf{F}_0 \tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l}) = \mathbf{F}_1 \tilde{\mathbf{a}}_{\mathrm{V},0}(\theta_{\mathrm{R},l}), \qquad (4.40)$$

where

$$[\mathbf{F}_0]_{\tilde{u},u} = \begin{cases} (-1)^{\eta(\tilde{u})} \cos(2\pi\eta(\tilde{u})/N_{\mathrm{V}}), & \text{if } u = \eta(\tilde{u}); \\ (-1)^{\eta(\tilde{u}+1)+1} \cos(2\pi\eta(\tilde{u}+1)/N_{\mathrm{V}}), & \text{if } u = \eta(\tilde{u}+1); \\ 0, & \text{otherwise.} \end{cases} \qquad (4.41)$$

$$[\mathbf{F}_1]_{\tilde{u},u} = \begin{cases} (-1)^{\eta(\tilde{u})} \sin(2\pi\eta(\tilde{u})/N_{\mathrm{V}}), & \text{if } u = \eta(\tilde{u}); \\ (-1)^{\eta(\tilde{u}+1)+1} \sin(2\pi\eta(\tilde{u}+1)/N_{\mathrm{V}}), & \text{if } u = \eta(\tilde{u}+1); \\ 0, & \text{otherwise.} \end{cases} \qquad (4.42)$$

with $\tilde{u} = 1, 2, \ldots, N_{\mathrm{B}} - 1$.

The processes of selecting the angle-related submatrices are similar to that of selecting the delay-related submatrices. Define the angle selection matrix as $\mathbf{J}_{\mathrm{A}} = \mathbf{1}_M^T \otimes \mathbf{I}_{N_{\mathrm{DS},2}} \in \mathbb{R}^{N_{\mathrm{DS},2} \times N_{\mathrm{DS},2} M}$. Then the angle-related submatrix can be formulated as $\mathbf{U}_{\mathrm{A}} = \mathbf{J}_{\mathrm{A}} \mathbf{U} \in \mathbb{C}^{N_{\mathrm{DS},2} \times N_{\mathrm{P}}}$. Based on the recurrence relation in (4.40), we can construct

$$\mathbf{F}_0 \mathbf{U}_{\mathrm{V}} \mathbf{\Theta}_{\mathrm{V}} = \mathbf{F}_1 \mathbf{U}_{\mathrm{V}}, \qquad (4.43)$$

where $\mathbf{\Theta}_{\mathrm{V}} = \mathrm{diag}\left(\tan(g(\theta_{\mathrm{R},1})/2), \ldots, \tan(g(\theta_{\mathrm{R},N_{\mathrm{P}}})/2)\right)$ and $\mathbf{U}_{\mathrm{V}} = \mathbf{J}_{\mathrm{V}}\mathbf{U}_{\mathrm{A}} \in \mathbb{C}^{N_{\mathrm{B}} \times N_{\mathrm{P}}}$ is a submatrix of $\mathbf{U}_{\mathrm{A}}$, where $\mathbf{J}_{\mathrm{V}} = \mathbf{I}_{(2P+1)} \otimes \mathbf{1}_{N_{\mathrm{B}}}^{T} \in \mathbb{R}^{N_{\mathrm{B}} \times N_{\mathrm{DS},2}}$. Thus, the vertical array response-related submatrix can be calculated as

$$\mathbf{E}_{\mathrm{V}} = \mathbf{J}_{\mathrm{V}}\mathbf{J}_{\mathrm{A}}\mathbf{E}_{\mathrm{s}} = \mathbf{J}_{\mathrm{V}}\mathbf{J}_{\mathrm{A}}\mathbf{U}\mathbf{T} = \mathbf{U}_{\mathrm{V}}\mathbf{T}. \tag{4.44}$$

Substituting (4.44) into (4.43), we can obtain

$$\mathbf{F}_0\mathbf{E}_{\mathrm{V}}\mathbf{T}^{-1}\mathbf{\Theta}_{\mathrm{V}}\mathbf{T} = \mathbf{F}_0\mathbf{E}_{\mathrm{V}}\mathbf{\Psi}_{\mathrm{V}} = \mathbf{F}_1\mathbf{E}_{\mathrm{V}}. \tag{4.45}$$

With reference to the delay estimation in Section 4.4.1, the elevation angle of the l-th path, $\hat{\theta}_l$, can be estimated as

$$\hat{\theta}_{\mathrm{R},l} = \arccos\left(\arctan(\lambda_{\mathrm{V},l})/\pi f_0 h\right), \tag{4.46}$$

where $\lambda_{\mathrm{V},l}$ is the l-th eigenvalue of $\hat{\mathbf{\Psi}}_{\mathrm{V}}$, and $\hat{\mathbf{\Psi}}_{\mathrm{V}}$ is the estimated matrix of $\mathbf{\Psi}_{\mathrm{V}} = \mathbf{T}^{-1}\mathbf{\Theta}_{\mathrm{V}}\mathbf{T}$.

According to (4.23), the expression for each horizontal response vector, which does not have the invariance structure, is an exponential function weighted by the Bessel function. There is no recursive relationship for the azimuth angle estimation. After obtaining the elevation angles, we use MUSIC to estimate their corresponding azimuth angles.

Define $\mathbf{J}_{\mathrm{H}} = \mathbf{J}_{\mathrm{HA}}\mathbf{J}_{\mathrm{A}} \in \mathbb{R}^{(2P+1)\times N_{\mathrm{DS},2}M}$, where $\mathbf{J}_{\mathrm{HA},u} = \left[\mathbf{I}_{(2P+1)}, \mathbf{0}_{(2P+1)\times(2P+1)(N_{\mathrm{B}}-1)}\right]$ $\in \mathbb{R}^{(2P+1)\times N_{\mathrm{DS},2}}$. We can obtain the corresponding horizontal signal $\tilde{\mathbf{y}}_{\mathrm{vec,H}} = \mathbf{J}_{\mathrm{H}}\tilde{\mathbf{y}}_{\mathrm{vec}} \in \mathbb{C}^{(2P+1)\times 1}$. As done in (4.30), the covariance matrix of $\tilde{\mathbf{y}}_{\mathrm{vec,H}}$ can be calculated as $\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec,H}}} = \mathbf{E}_{\mathrm{s,H}}\mathbf{\Sigma}_{\mathrm{s,H}}\mathbf{E}_{\mathrm{s,H}}^{H} + \sigma_{\mathrm{n}}^2\mathbf{E}_{\mathrm{n,H}}\mathbf{E}_{\mathrm{n,H}}^{H}$, where $\mathbf{E}_{\mathrm{s,H}}$ and $\mathbf{E}_{\mathrm{n,H}}$ are the signal and noise subspaces of $\tilde{\mathbf{y}}_{\mathrm{vec,H}}$, respectively.

By substituting the estimate of the l-th path, $\hat{\theta}_{\mathrm{R},l}$, in the MUSIC estimator, the azimuth angle of the path can be estimated by

$$\hat{\phi}_{\mathrm{R},l} = \arg\max_{\Phi} \left\| \mathbf{E}_{\mathrm{n,H}}^{H}\tilde{\mathbf{a}}_{\mathrm{H},0}(\Phi, \hat{\theta}_{\mathrm{R},l}) \right\|_{\mathrm{F}}^{-2}, \tag{4.47}$$

where Φ is the azimuth of the AoA, and can be estimated by 1D search.

4.4.3 Multipath Parameter Matching

As described above, the estimated channel parameters of each path can be matched automatically in the case of negligible noise. This is because they have the common factor $\mathbf{T}$, as shown in (4.32). In the presence of non-negligible noise, there can be a mismatch between the estimated parameters. We take the delay and the elevation AoA for an example. According to (4.35) and (4.45), we have $\mathbf{\Psi}_{\mathrm{D}} = \mathbf{T}_{\mathrm{D}}^{-1}\mathbf{\Theta}_{\mathrm{D}}\mathbf{T}_{\mathrm{D}}$ and $\mathbf{\Psi}_{\mathrm{V}} = \mathbf{T}_{\mathrm{V}}^{-1}\mathbf{\Theta}_{\mathrm{V}}\mathbf{T}_{\mathrm{V}}$, but $\mathbf{T}_{\mathrm{V}} \neq \mathbf{T}_{\mathrm{D}} \neq \mathbf{T}$ because of the noise. Most existing pair matching methods would require the approximate values of the estimates first, and then use an exhaustive search to match all possible parameter pairs [63]. Such methods would lead to prohibitive computational complexity if the numbers of paths and parameters are large.

We note that in our approach, the estimated elevation angles, $\hat{\theta}_{\mathrm{R},l}$, are used for the estimation of the azimuth counterparts, $\hat{\phi}_{\mathrm{R},l}$, so that the azimuth and elevation angles of each path are always matched; see (4.47). Since a mismatch of the estimated delays and angles is primarily caused by additive noises, the mismatch can be mitigated by adding perturbation matrices [114]. Provided two perturbation matrices $\mathbf{P}_{\mathrm{D}}$ and $\mathbf{P}_{\mathrm{V}}$ with $\tilde{\mathbf{\Psi}}_{\mathrm{D}} = \mathbf{\Psi}_{\mathrm{D}} + \mathbf{P}_{\mathrm{D}} = \tilde{\mathbf{T}}_{\mathrm{D}}^{-1}\mathbf{\Theta}_{\mathrm{D}}\tilde{\mathbf{T}}_{\mathrm{D}}$ and $\tilde{\mathbf{\Psi}}_{\mathrm{V}} = \mathbf{\Psi}_{\mathrm{V}} + \mathbf{P}_{\mathrm{V}} = \tilde{\mathbf{T}}_{\mathrm{V}}^{-1}\mathbf{\Theta}_{\mathrm{V}}\tilde{\mathbf{T}}_{\mathrm{V}}$, we can obtain $\tilde{\mathbf{T}}_{\mathrm{D}} = \tilde{\mathbf{T}}_{\mathrm{V}} = \tilde{\mathbf{T}}$. The parameter pair matching in (4.35) and (4.45) can be achieved. The perturbation matrices $\mathbf{P}_{\mathrm{D}}$ and $\mathbf{P}_{\mathrm{V}}$ can be obtained by solving the following problem [114]:

$$\min_{P_{\mathrm{D}}, P_{\mathrm{V}}} \|\mathbf{P}_{\mathrm{D}}\|_{\mathrm{F}}^2 + \|\mathbf{P}_{\mathrm{V}}\|_{\mathrm{F}}^2 \tag{4.48}$$

$$s.t. \left(\mathbf{\Psi}_{\mathrm{D}} + \mathbf{P}_{\mathrm{D}}\right)\left(\mathbf{\Psi}_{\mathrm{V}} + \mathbf{P}_{\mathrm{V}}\right) = \left(\mathbf{\Psi}_{\mathrm{D}} + \mathbf{P}_{\mathrm{V}}\right)\left(\mathbf{\Psi}_{\mathrm{V}} + \mathbf{P}_{\mathrm{D}}\right), \tag{4.49}$$

where (4.48) is formulated due to the fact that $\mathbf{P}_{\mathrm{D}}$ and $\mathbf{P}_{\mathrm{V}}$ need to obey the minimum Frobenius norm constraint [114]. The exact solution to this non-linearly constrained

problem (4.48) is hard to find. To solve the problem, we rewrite (4.49) as

$$\mathbf{P}_{\mathrm{D}}\mathbf{P}_{\mathrm{V}} - \mathbf{P}_{\mathrm{V}}\mathbf{P}_{\mathrm{D}} = \mathbf{\Psi}_{\mathrm{D}}\mathbf{\Psi}_{\mathrm{V}} + \mathbf{P}_{\mathrm{V}}\mathbf{\Psi}_{\mathrm{V}} + \mathbf{\Psi}_{\mathrm{D}}\mathbf{P}_{\mathrm{D}} - \mathbf{\Psi}_{\mathrm{D}}\mathbf{\Psi}_{\mathrm{V}} - \mathbf{\Psi}_{\mathrm{D}}\mathbf{P}_{\mathrm{V}} - \mathbf{P}_{\mathrm{D}}\mathbf{\Psi}_{\mathrm{V}}. \quad (4.50)$$

We assume that the perturbations are much smaller than $\mathbf{\Psi}_{\mathrm{D}}$ and $\mathbf{\Psi}_{\mathrm{V}}$, then the term $(\mathbf{P}_{\mathrm{D}}\mathbf{P}_{\mathrm{V}} - \mathbf{P}_{\mathrm{V}}\mathbf{P}_{\mathrm{D}})$ in (4.50) can be suppressed [115]. We only add the perturbation to $\mathbf{\Psi}_{\mathrm{D}}$, i.e., letting $\mathbf{P}_{\mathrm{V}} = \mathbf{0}$. Then $\mathbf{P}_{\mathrm{D}}$ can be obtained as

$$\mathrm{vec}(\mathbf{P}_{\mathrm{D}}) = [\mathbf{\Psi}_{\mathrm{V}}^{T} \oplus (-\mathbf{\Psi}_{\mathrm{V}})]^{\dagger}\mathrm{vec}(\mathbf{\Psi}_{\mathrm{V}}\mathbf{\Psi}_{\mathrm{D}} - \mathbf{\Psi}_{\mathrm{D}}\mathbf{\Psi}_{\mathrm{V}}). \quad (4.51)$$

By adding the perturbation matrix $\mathbf{P}_{\mathrm{D}}$ to the elevation angle eigenvalue matrix $\mathbf{\Psi}_{\mathrm{D}}$, the delay and the elevation angles can be matched. The parameters of each path can be associated correctly.

4.4.4 Complexity Analysis

We proceed to analyze the hardware and software complexities of the proposed channel parameter approach. For a large-scale antenna array system using fully digital beamforming, its hardware complexity is $O(N_{\mathrm{R}})$. In our proposed approach, the use of the hybrid beamformer allows for a dramatic reduction of the hardware complexity from $O(N_{\mathrm{R}})$ to $O(N_{\mathrm{RF}})$, where $N_{\mathrm{RF}} = \max(N_{\mathrm{V}}, (2P+1)N_{\mathrm{B}})$.

In terms of signal processing complexity, we compare the proposed approach with existing techniques, including quadric rotational invariance property-based method (QRIPM) [29], generalized beamspace method (GBM) [14], and Quasi-Maximum-Likelihood estimator (Q-MLE) [116]. For the proposed approach, after hybrid beamforming, the dimension of the received signal is reduced to $N_{\mathrm{DS},2}$, so the computational complexity of MDSL processing is $O(N_{\mathrm{DS},2}M) = O((2P+1)N_{\mathrm{B}}M) = O(\gamma P N_{\mathrm{p}} M)$. The computational complexity of calculating the covariance matrix, $\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec}}}$, in (4.29) and performing the EVD on $\mathbf{R}_{\tilde{\mathbf{y}}_{\mathrm{vec}}}$ according to (4.30) is $O(\gamma^2 P^2 N_{\mathrm{p}}^2 M^2 T_s)$ and $O(\gamma^3 P^3 N_{\mathrm{p}}^3 M^3)$, respectively, where T_s is the number of snapshots. The complexities of computing the delay $\hat{\tau}_l$ and the elevation angle, $\hat{\theta}_{\mathrm{R},l}$, are $O(MN_{\mathrm{p}}^3)$ and

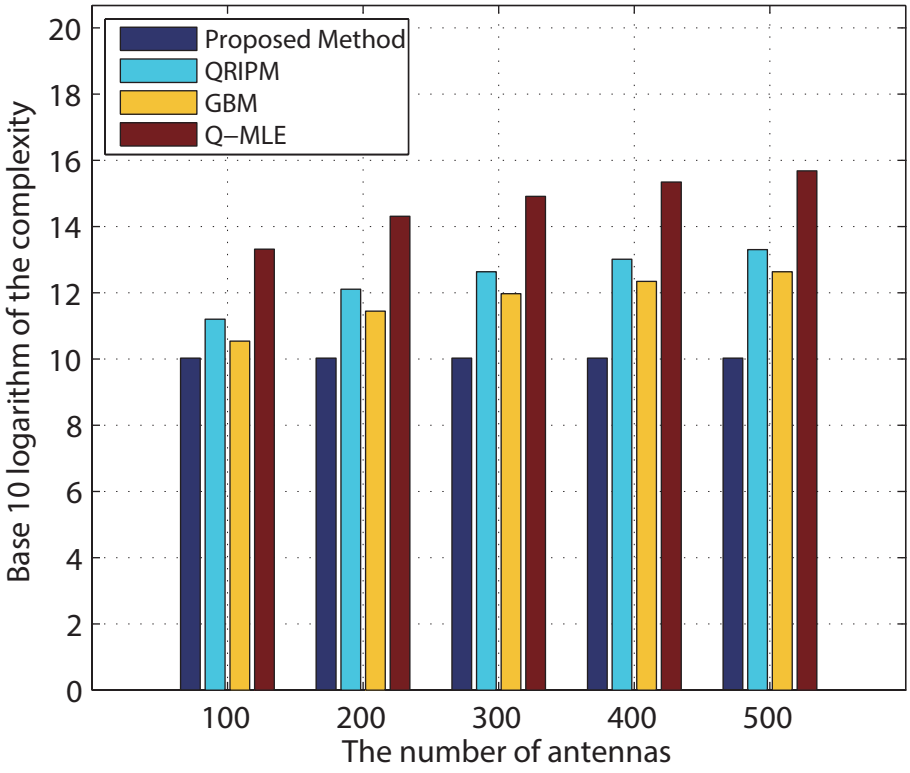


Figure 4.4 : Variation of the software complexity vs. the number of antennas.

$O(\gamma^2 N_{\mathrm{p}}^2 + N_{\mathrm{p}}^3)$, respectively. When estimating $\hat{\phi}_{\mathrm{R},l}$ with 1D search using (4.47), the computational complexity is $O(\gamma^2 N_{\mathrm{p}}^2 D)$, where D is the size of the search dimension. For the pair matching operation, the computational complexity is $O(N_{\mathrm{p}}^3)$. Thus, the overall computational complexity of our proposed estimation approach is $O(\gamma P N_{\mathrm{p}} M + \gamma^2 P^2 N_{\mathrm{p}}^2 M^2 T_s + \gamma^3 P^3 N_{\mathrm{p}}^3 M^3 + M N_{\mathrm{p}}^3 + \gamma^2 N_{\mathrm{p}}^2 + N_{\mathrm{p}}^3 + \gamma^2 N_{\mathrm{p}}^2 D + N_{\mathrm{p}}^3)$, which, in fact, has no dependence on the number of receive antenna N_{R}. The computational complexities of QRIPM and GBM increase rapidly, as the number of receive antennas increases. When the number of receive antennas N_{R} is large, the computational complexities of QRIPM and GBM are $O(N_{\mathrm{R}}^3 M^4)$ and $O(P^3 N_{\mathrm{V}}^3 M^4)$, respectively. The computational complexity of Q-MLE is $O(N_{\mathrm{R}}^2 M^2 N_{\mathrm{AZI}} N_{\mathrm{ELE}} N_{\mathrm{DEL}} + (N_{\mathrm{p}} N_{\mathrm{R}} M)^{3.5})$, where N_{AZI}, N_{ELE}, and N_{DEL} are the search grids of azimuth angle, elevation angle, and delay, respectively.

Fig. 4.4 compares the computational complexities of the four methods with the growing number of antennas $N_{\mathrm{R}} = N_{\mathrm{H}}N_{\mathrm{V}}$, where $\gamma = 2$, $N_{\mathrm{p}} = 3$, $M = 20$, and $P = 12$. We set $D = N_{\mathrm{AZI}} = N_{\mathrm{ELE}} = N_{\mathrm{DEL}} = 100$. The figure shows that, compared with the existing methods, the proposed approach has a substantially lower computational complexity. In addition, the reduction is increasingly significant with the growing number of receive antennas at the BS.

4.5 Simulation Results

This section presents simulation results to demonstrate the performance of the proposed approach under different parameters in the UMi scenario of future 5G/B5G systems. We set $f_0 = 30$ GHz and $B = 2$ GHz[6], and assume that there are a total of $N_{\mathrm{p}} = 3$ NLoS paths and $M = 20$ consecutive subcarriers. The distance, h, between adjacent receiving UCAs and the radius, r, of each UCA are $0.5\lambda_0$ and $2\lambda_0$, respectively.

Fig. 4.5 shows the RMSE of the angle and delay estimates, and compares the proposed approach with QRIPM [29], GBM [14], Q-MLE [116], and the CRLB[7]. Figs. 4.5(a) shows the RMSE of the estimated azimuth AoAs versus the number of receive antennas under different signal-to-noise ratio (SNR) conditions. In Fig. 4.5(a), we see that the RMSE of the estimated azimuth AoAs approaches the CRLB, as the average received SNR or the number of receive antenna increases. We observe that, when the number of antennas is not large, the RMSE of GBM and the proposed approach is worse than that of QRIPM. The reason is that the conditions in Theorem 1 may not be satisfied, and the approximation in (4.23) may not hold. However, when the number of antennas increases, the RMSE of GBM and the

[6]In this system, for a UCyA with 400 antennas (25×16), its dimension is about 40×80 mm^2, which is much smaller than the communication range. Therefore, the far-field condition is fulfilled.

[7]The CRLB is calculated according to [47, 117].

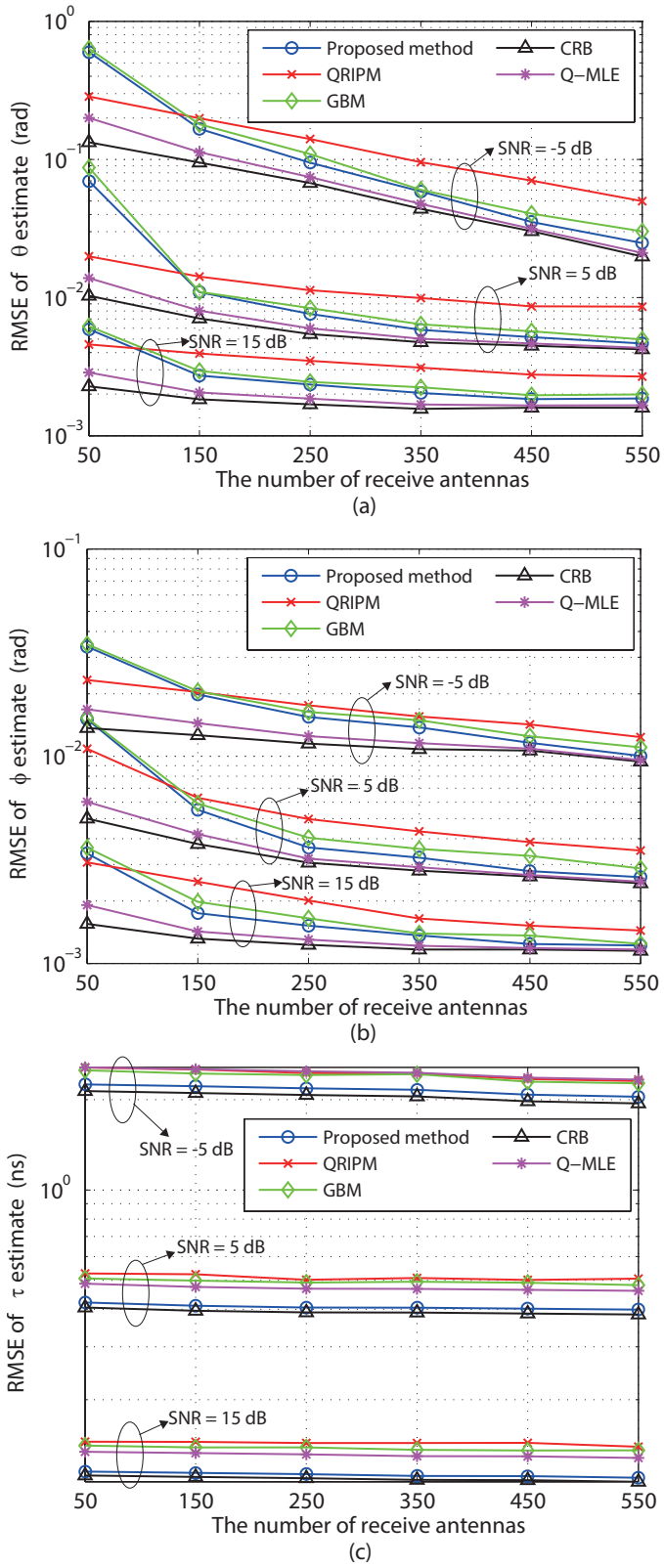


Figure 4.5 : RMSE vs. the number of BS antennas for the estimation of different parameters. (a) Azimuth AoA; (b) Elevation AoA; (c) Path delay.

proposed approach decreases faster than that of QRIPM, and our approach quickly outperforms QRIPM and GBM. The RMSE of the estimated elevation of the AoAs by using different estimation methods is shown in Fig. 4.5(b). We see that when the number of receive antennas is smaller than 100, GBM, and the proposed approach are still inferior to QRIPM. When the number of antennas is large, our proposed approach is better than QRIPM and GBM in terms of angle estimation accuracy. It is observed in Figs. 4.5(a) and 4.5(b) that Q-MLE outperforms the other three approaches, including the proposed approach, in terms of angle estimation. As analyzed in Section 4.4.4, Q-MLE has a much higher signal processing complexity than our approach. In addition, it is shown in Fig. 4.5(c) that our approach achieves the best delay estimation accuracy. The reason is that, by using the proposed multidimensional spatial interpolation, the high temporal resolution offered by wideband mmWave signals is exploited. As also shown in Fig. 4.5(c), we can see that the delay estimation accuracy does improve when the antenna number increases, but the improvement is not evident. This is because the delay estimation precision primarily depends on the signal bandwidth, instead of the number of receive antennas.

In order to validate Theorem 1, Fig. 4.6 plots the RMSE of the angle estimation versus the value of the highest order, P, under different numbers of horizontal array response vectors. We see that when the highest order $P \leq 11$, our proposed approach cannot perform satisfactorily, since the number of phase-mode vectors is not sufficient to represent the transformed array response vectors in Section 4.3.2. Fig. 4.6 also shows that, if $P \geq 12$, whatever the number of array response vectors is, increasing the phase-mode vectors has little influence on the angle estimation performance. This means that the number of phase-mode vectors needed in our approach does not depend on the number of array response vectors, which is important for complexity reduction, as discussed in Section 4.4.4. In addition, we also see that because the condition in Theorem 1, $N_{\mathrm{H}} \geq 2P$, is unlikely to be satisfied

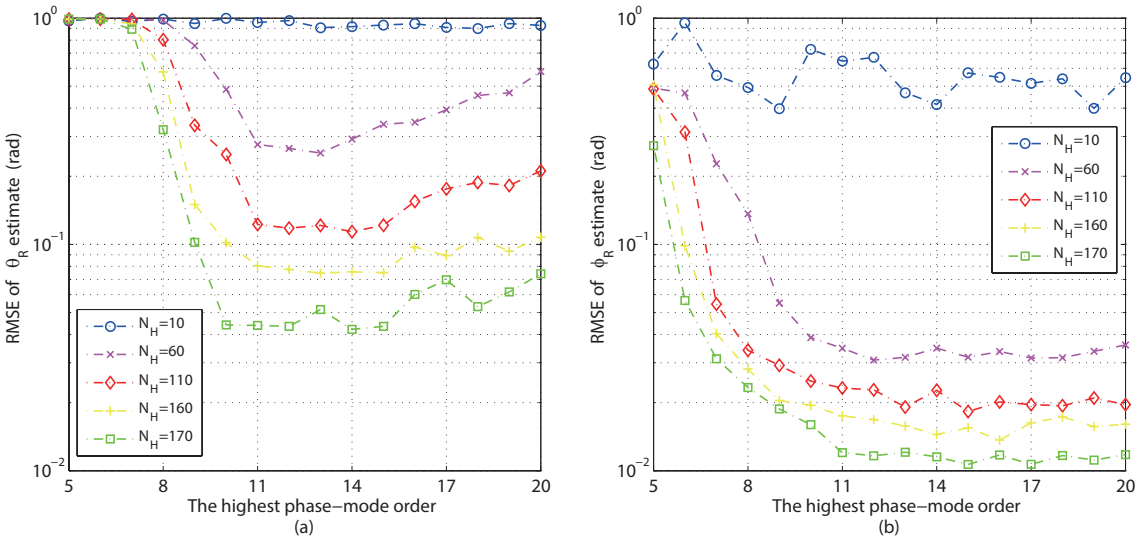


Figure 4.6 : The RMSE vs. the value of the highest order. (a) Azimuth AoA; (b) Elevation AoA.

when $N_{\mathrm{H}} = 10$, the RMSE is much poorer than those applying more array response vectors.

4.6 Summary

This chapter has presented a novel channel estimation method for mmWave UMi environments with wideband hybrid UCyAs. We have introduced a new 3D hybrid beamformer, which can reduce the number of required RF chains while maintaining the critical recursive property of the space-time response matrix for angle and delay estimation. We also have also generalized linear interpolation to reconstruct the output signals of the 3D hybrid beamformer. We have shown that this method can achieve consistent array response across the wideband and suppress the beam squint effect. As a result, the delay and the azimuth and elevation angles of every multi-path component can be estimated. Simulation results have shown that, in the UMi scenario of large-scale mmWave antenna array systems, when a large number of antennas is deployed, our proposed method is capable of precisely estimating the

channel parameters even under low SNR conditions. In this chapter, we assume that the bandwidth is continuous. In practical scenarios where the bandwidth is intermittent, we can first employ some filters to obtain the signals at equal-spacing subcarriers, and then use the proposed algorithm for parameter estimation.

Chapter 5

Tensor-based Parameter Estimation for Hybrid Directional UCyAs

This chapter is devoted to high-accuracy channel parameter estimation algorithm for hybrid UCyAs with directional beamforming. 5G mmWave communications, in general, is applied for short-distance communication applications, such as indoor environments and UMi, due to severe free-space pathloss. To support long-distance links in some specific mmWave communication scenarios, e.g., RMi, directional beamforming is deployed with large antenna arrays, which appears to be inevitable to provide sufficient signal transmission power. Moreover, the very short wavelengths of mm-Wave signals enables large antenna arrays to be packed into small form factors, which, fortunately, makes realization of the large arrays needed for high beamforming gains feasible.

In this chapter, we present a novel tensor-based method for multi-dimensional wideband channel estimation in large-scale mmWave hybrid UCyAs. We consider the RMi scenario in 5G/B5G mmWave systems, as illustrated in Fig. 5.1. We design a new resolution-preserving directional hybrid beamformer and a low-complexity

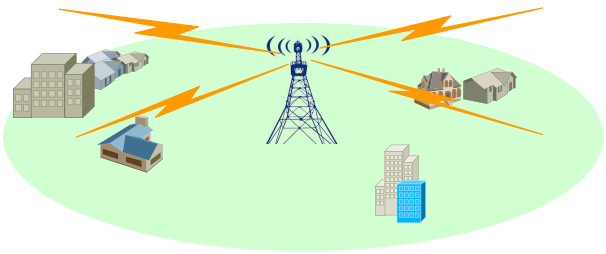


Figure 5.1 : The research scenario studied in Chapter 5.

beam squinting suppression method, and reveal the existence of shift-invariance relations in the tensor models of received array signals at the UCyA. Exploiting these relations, we propose a new tensor-based subspace estimation algorithm to suppress the receiver noises in all dimensions (time, frequency, and space). The algorithm can accurately estimate the channel parameters from both coherent and incoherent signals. Corroborated by the CRLB, simulation results show that, in the RMa scenario of future 5G/B5G systems, the proposed algorithm is able to achieve substantially higher estimation accuracy than existing matrix-based techniques, with a comparable computational complexity.

5.1 Motivation and Overview

In this chapter, we arrange and process the received signals in the tensor form to exploit the relations between each dimension/domain of the received signal. As stated in Section 5.1, tensor-based channel parameter estimations have been demonstrated to be more powerful than conventional matrix-based techniques [75–77]. However, traditional multiparameter estimation is typically matrix-based, which needs to stack and process the received multidimensional data in a two-dimensional space-time matrix. As a result, the relations between each dimension/domain (e.g., space, time, and frequency) of the received signal would be damaged [118]. On the other hand, since the parameters (e.g., AoA, AoD, and delay) are coupled in the space-time matrix, before using subspace-based algorithms, the space-time matrix has to be divided into multiple (≥ 6) subarrays to decouple them, increasing the estimation complexity.

Specifically, in Section 5.3, we first design a hybrid directional beamformer to synthesize the received signals. The beamformer sweeps on the vertical plane and operates omnidirectionally on the horizon plane. We use Q-DFT to reduce the dimension of the received signals with a negligible cost of the channel estimation

accuracy, and design the digital beamforming coefficients to make the beamformer sweep on the vertical plane to deal with the severe free-space pathloss of mmWave links. In this way, the signals can be processed with much fewer RF chains (than antennas) and an acceptable communication quality can be obtained.

We then develop a low-complexity UCAMI in Section 5.4, which can suppress the beam squinting effect and enable coherent combining of measurement signals across wideband. The conventional UCAMI [73] needs to solve a computationally expensive multi-dimensional optimization problem whose dimension is equal to the number of estimation parameters. After being processed by the RF network, we see that only the elevation angular-related array steering vectors depend on the subcarrier frequency. Therefore, the proposed new UCAMI only optimizes the focusing matrices in the elevation angular domain. There are only 1-D optimization problems in our method, whose computational complexity is much lower than the conventional UCAMI.

By stacking the received signal in the multidimensional tensor model, all the subsequent processing in this chapter are operated in the tensor form. In particular, we construct the truncated HOSVD model of the measure tensor in Section 5.5.1 and propose a new tensor-based joint delay-angle estimation algorithm to estimate the delay and azimuth and elevation angles in Sections 5.5.2 by revealing and exploiting the inherent linear recurrence relations in the first mode of the measure tensor. We also extend the spatial smoothing technique to the designed hybrid UCyA in Section 5.5.3 to decorrelate the coherent signals. By this means, the signal and noise subspaces of the received signal can be correctly decomposed in all dimensions.

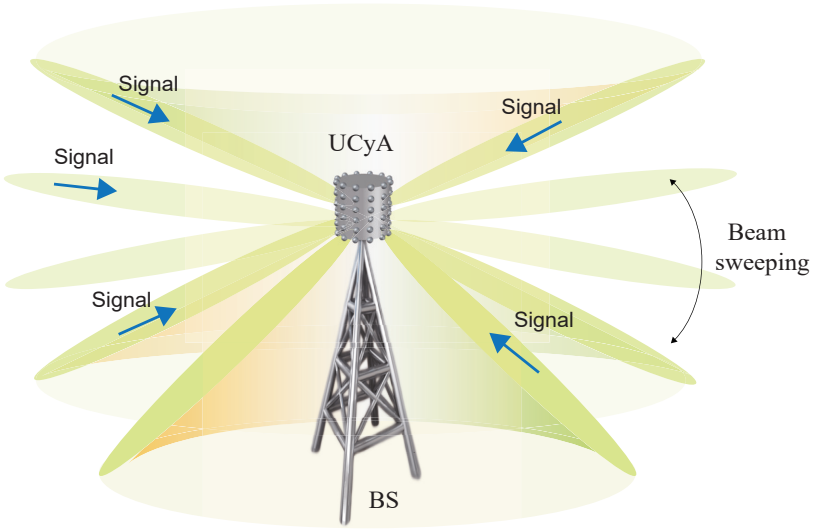
Simulation results are provided in Section 5.6 to demonstrate the performance of the proposed algorithm. Validated by the CRLB, simulation results show that, in the RMa scenario of large-scale mmWave array systems, the proposed algorithm is able

to achieve much higher accuracy than state-of-the-art matrix-based techniques for wideband mmWave hybrid UCyAs. This is because the noise is multi-dimensional with the same dimensions as the received signal. It is important to take all dimensions of the received signal into consideration, and suppress the noises in all the dimensions. However, when applying the matrix-based alternative, the noise is only suppressed in one of the dimensions (or modes) of the measurement tensor, hence degrading the estimation accuracy. We also plot the RMSE of the estimated parameters with an increasing number of received paths. We show that the performance gap between the matrix and tensor forms of the proposed algorithm decreases with the increasing number of received paths. This is because the noise components which can be suppressed by using the tensor-based algorithms in all modes of the measurement tensor, depend on the difference between the number of paths and the tensor dimension in each mode. As the number of received paths increases, the gain of the tensor-based algorithm diminishes.

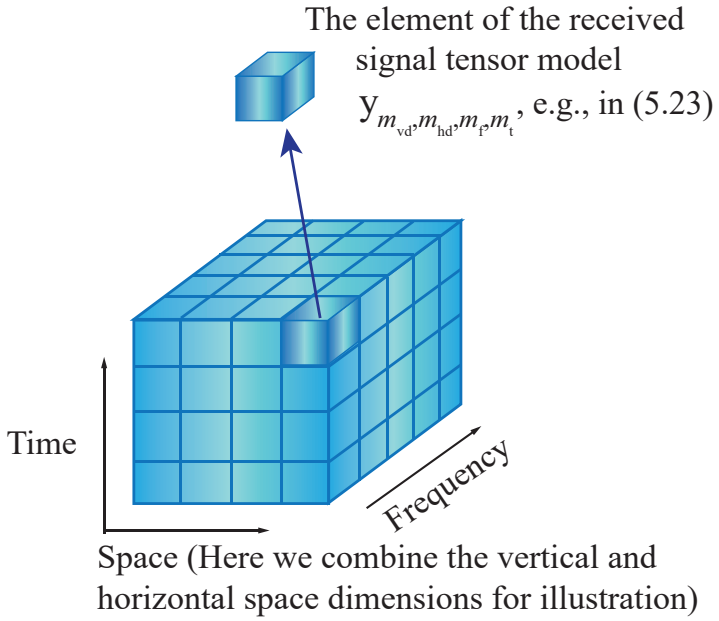
5.2 System Model

In our system, a BS is equipped with a large-scale hybrid mmWave UCyA with M_{bs} antennas, consisting of M_{v} vertically placed UCAs each with M_{h} antenna elements, and $M_{\mathrm{bs}} = M_{\mathrm{v}}M_{\mathrm{h}}$. Let r be the radius of the UCyA, and h be the vertical distance between any two adjacent vertical elements. A hybrid front-end is adopted (i.e., there are fewer RF chains than antennas) with consideration of hardware cost, energy consumption, weight and size. Consider a wideband OFDM system, with M_{f} subcarriers. There are a total of K 3-D sources, each of which is equipped with a single antenna with an isotropic beam pattern.

We apply vertical beam sweeping to obtain the signals from the sources, as shown in Fig.5.2(a). M_{b} evenly spaced elevation angles are swept successively. For each elevation angle, signal samples of M_{t} time frames are collected within a sweeping



(a)



(b)

Figure 5.2 : Illustration on the proposed system and signal models. (a) System configuration; (b) Signal tensor model.

time interval τ_{b}. In the m_{b}-th sweeping beam ($m_{\mathrm{b}} = 1, \ldots, M_{\mathrm{b}}$), the signals from $K_{m_{\mathrm{b}}}$ sources are captured at the BS (and $K \leq \sum_{m_{\mathrm{b}}=1}^{M_{\mathrm{b}}} K_{m_{\mathrm{b}}}$, due to the partially overlapping sweeping beams). The signal sample associated with the m_{f}-th subcarrier ($m_{\mathrm{f}} = 1, \ldots, M_{\mathrm{f}}$) at the m_{t}-th time frame ($m_{\mathrm{t}} = 1, \ldots, M_{\mathrm{t}}$) can be expressed as [64]:

$$\mathbf{x}_{m_{\mathrm{f}},m_{\mathrm{t}},m_{\mathrm{b}}} = \sum_{k_{m_{\mathrm{b}}}=1}^{K_{m_{\mathrm{b}}}} s_{m_{\mathrm{t}},k_{m_{\mathrm{b}}}} a_{\mathrm{fds},m_{\mathrm{f}},m_{\mathrm{b}}}(\tau_{k_{m_{\mathrm{b}}}}) \mathbf{B}_{m_{\mathrm{f}},m_{\mathrm{b}}}^{H} \mathbf{a}_{\mathrm{bs},m_{\mathrm{f}},m_{\mathrm{b}}}(\phi_{k_{m_{\mathrm{b}}}}, \theta_{k_{m_{\mathrm{b}}}}) + \mathbf{n}_{m_{\mathrm{f}},m_{\mathrm{t}},m_{\mathrm{b}}}, \tag{5.1}$$

where $\phi_{k_{m_{\mathrm{b}}}}$ and $\theta_{k_{m_{\mathrm{b}}}}$ are the azimuth and elevation angles-of-arrivals (AoAs) of the $k_{m_{\mathrm{b}}}$-th path, respectively; $\mathbf{a}_{\mathrm{bs},m_{\mathrm{f}},m_{\mathrm{b}}}(\phi_{k_{m_{\mathrm{b}}}}, \theta_{k_{m_{\mathrm{b}}}}) \in \mathbb{C}^{M_{\mathrm{bs}}}$ denotes the steering vector of the hybrid UCyA; $s_{m_{\mathrm{t}},k_{m_{\mathrm{b}}}} = \alpha_{k_{m_{\mathrm{b}}}} \tilde{s}_{m_{\mathrm{t}},k_{m_{\mathrm{b}}}} / \sqrt{\rho_{k_{m_{\mathrm{b}}}}}$, where $\tilde{s}_{m_{\mathrm{t}},k_{m_{\mathrm{b}}}}$ is the transmitted symbol, $\alpha_{k_{m_{\mathrm{b}}}}$ is the signal power, and $\rho_{k_{m_{\mathrm{b}}}}$ is the pathloss from the $k_{m_{\mathrm{b}}}$-th source to the BS; $\mathbf{n}_{m_{\mathrm{f}},m_{\mathrm{b}},m_{\mathrm{t}}} \in\in \mathbb{C}^{M_{\mathrm{bsd}}}$ denotes the additive white Gaussian noise (AWGN); $\mathbf{B}_{m_{\mathrm{f}},m_{\mathrm{b}}} = \mathbf{B}_{\mathrm{ab}} \mathbf{B}_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}} \in \mathbb{C}^{M_{\mathrm{bs}} \times M_{\mathrm{bsd}}}$ is the hybrid beamforming matrix, composed of an analog beamforming matrix $\mathbf{B}_{\mathrm{ab}} \in \mathbb{C}^{M_{\mathrm{bs}} \times M_{\mathrm{bsr}}}$ and a digital beamforming matrix $\mathbf{B}_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}} \in \mathbb{C}^{M_{\mathrm{bsr}} \times M_{\mathrm{bsd}}}$; M_{bsr} is the number of RF chains; M_{bsd} is the number of data streams after hybrid beamforming; and

$$a_{\mathrm{fbs},m_{\mathrm{f}},m_{\mathrm{b}}}(\tau_{k_{m_{\mathrm{b}}}}) = a_{\mathrm{f},m_{\mathrm{f}}}(\tau_{k_{m_{\mathrm{b}}}}) b_{\mathrm{f},m_{\mathrm{b}}}, \tag{5.2}$$

where $a_{\mathrm{f},m_{\mathrm{f}}}(\tau_{k_{m_{\mathrm{b}}}}) = e^{-j2\pi f_{m_{\mathrm{f}}} \tau_{k_{m_{\mathrm{b}}}}}$ and $b_{\mathrm{f},m_{\mathrm{b}}} = e^{-j2\pi f_{m_{\mathrm{f}}}(m_{\mathrm{b}}-1)\tau_{\mathrm{b}}}$ with $\tau_{k_{m_{\mathrm{b}}}}$ being the delay of the $k_{m_{\mathrm{b}}}$-th signal and $f_{m_{\mathrm{f}}}$ being the m_{f}-th subcarrier frequency. The delay $\tau_{k_{m_{\mathrm{b}}}}$ can be used to estimate the source distance.

Given the structure of UCyA, the array steering vector, i.e., $\mathbf{a}_{\mathrm{bs},m_{\mathrm{f}},m_{\mathrm{b}}}(\phi_{k_{m_{\mathrm{b}}}}, \theta_{k_{m_{\mathrm{b}}}})$, can be given as the Kronecker product of the vertical and horizontal array steering vectors:

$$\mathbf{a}_{\mathrm{bs},m_{\mathrm{f}},m_{\mathrm{b}}}(\phi_{k_{m_{\mathrm{b}}}}, \theta_{k_{m_{\mathrm{b}}}}) = \mathbf{a}_{\mathrm{v},m_{\mathrm{f}},m_{\mathrm{b}}}(\theta_{k_{m_{\mathrm{b}}}}) \otimes \mathbf{a}_{\mathrm{h},m_{\mathrm{f}},m_{\mathrm{b}}}(\theta_{k_{m_{\mathrm{b}}}}, \phi_{k_{m_{\mathrm{b}}}}). \tag{5.3}$$

The elements of $\mathbf{a}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})$ and $\mathbf{a}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})$ are:

$$\begin{aligned}\left[\mathbf{a}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})\right]_{m_\mathrm{v},1} &= a_{\mathrm{v},m_\mathrm{v},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}) \\ &= \frac{1}{\sqrt{M_\mathrm{v}}}\exp\left(-j\frac{2\pi}{c}f_{m_\mathrm{f}}h(m_\mathrm{v}-1)\cos(\theta_{k_{m_\mathrm{b}}})\right), \end{aligned} \tag{5.4}$$

$$\begin{aligned}\left[\mathbf{a}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})\right]_{m_\mathrm{h},1} &= a_{\mathrm{h},m_\mathrm{h},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) \\ &= \frac{1}{\sqrt{M_\mathrm{h}}}\exp\left(j\frac{2\pi}{c}f_{m_\mathrm{f}}r\sin(\theta_{k_{m_\mathrm{b}}})\cos(\phi_{k_{m_\mathrm{b}}}-\varphi_{m_\mathrm{h}})\right), \end{aligned} \tag{5.5}$$

where c is the speed of light, and $\varphi_{m_\mathrm{h}} = 2\pi(m_\mathrm{h}-1)/M_\mathrm{h}$ is the difference between the central angles of the m_h-th antenna and the first antenna of each UCA.

5.3 Hybrid Directional Beamforming Design

In this section, we design the analog and digital beamforming matrices, $\mathbf{B}_\mathrm{ab}$ and $\mathbf{B}_{\mathrm{db},m_\mathrm{f},m_\mathrm{b}}$, for the hybrid directional beamformer, as the first step shown in Fig. 5.2. The number of required RF chains is reduced while the angular resolution of the UCyA is not compromised as compared to its fully digital counterparts.

We decouple $\mathbf{B}_\mathrm{ab}$ between the vertical and horizontal planes, i.e., $\mathbf{B}_\mathrm{ab} = \mathbf{B}_\mathrm{vab} \otimes \mathbf{B}_\mathrm{hab}$ with $\mathbf{B}_\mathrm{vab} \in \mathbb{C}^{M_\mathrm{v}\times M_\mathrm{vr}}$ and $\mathbf{B}_\mathrm{hab} \in \mathbb{C}^{M_\mathrm{h}\times M_\mathrm{hr}}$. By decoupling the beamformers into the Kronecker products of horizontal and vertical matrices, we preserve the shift-invariance relations on the vertical and horizontal planes, as will be revealed later in Section 5.6. To maintain the angular resolution of the hybrid UCyA, we design $\mathbf{B}_\mathrm{hab}$ based on the following theorem.

Theorem 2. *Suppose that* $M_h \geq \lfloor 4\pi f_{m_f} r/c \rfloor$. *The array response vector* $\mathbf{a}_{h,m_f,m_b}(\theta_{k_{m_b}}, \phi_{k_{m_b}})$ *can be transformed into a beamspace by using Q-DFT. If the index for a beamspace dimension, p, is larger than* $\lfloor 2\pi f_{m_f} r/c \rfloor$, *the element in the dimension is negligible and can be suppressed. The expression for the elements in the other dimensions is given by:*

$$a_{QDFT,p,m_f,m_b}(\theta_{k_{m_b}},\phi_{k_{m_b}}) \approx \sqrt{M_h} j^p J_p\left(\gamma_{m_f}(\theta_{k_{m_b}})\right)\exp\left(-jp\phi_{k_{m_b}}\right), \tag{5.6}$$

where $\gamma_{m_f}(\theta_{k_{m_b}}) = 2\pi f_{m_f} r \sin(\theta_{k_{m_b}})/c$, $p = -P, -P+1, \ldots, P$, *and* $J_p\left(\gamma_{m_f}(\theta_{k_{m_b}})\right)$ *is the Bessel function of the first kind of order* p.

Proof. See Appendix D.1. □

Theorem 2 shows that, with the application of Q-DFT [110], the M_{h}-dimensional array response vector of each UCA, $\mathbf{a}_{\mathrm{h},m_{\mathrm{f}},m_{\mathrm{b}}}(\theta_{k_{m_{\mathrm{b}}}}, \phi_{k_{m_{\mathrm{b}}}})$, can be transformed to be $(2P+1)$-dimensional, where $P = \lfloor 2\pi f_{m_{\mathrm{f}}} r/c \rfloor$. As a result, only $M_{\mathrm{hr}} = (2P+1)$ RF chains are required on the horizontal plane. Specifically, according to Theorem 2, we design $\mathbf{B}_{\mathrm{hab}}$ as $[\mathbf{B}_{\mathrm{hab}}]_{m_{\mathrm{h}},m_{\mathrm{hr}}+P+1} = e^{-j2\pi(m_{\mathrm{h}}-1)m_{\mathrm{hr}}/M_{\mathrm{h}}}$, where $m_{\mathrm{hr}} = -P, -P+1, \ldots, P$. We set $\mathbf{B}_{\mathrm{vab}} = \mathbf{I}_{M_{\mathrm{v}}}$ to preserve the recurrence relation between the UCAs, i.e., the shift-invariance relation. The relation is crucial for the subspace-based estimation algorithms, and exploited to estimate the elevation AoAs in this chapter. With this design, the number of required RF chains is only $M_{\mathrm{bsr}} = M_{\mathrm{vr}} M_{\mathrm{hr}} = M_{\mathrm{v}}(2P+1)$.

Then, we design the digital beamformer $\mathbf{B}_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}}$ as

$$\mathbf{B}_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}} = \mathrm{diag}\left(b_{\mathrm{db},1,m_{\mathrm{f}},m_{\mathrm{b}}}, \ldots, b_{\mathrm{db},M_{\mathrm{bsr}},m_{\mathrm{f}},m_{\mathrm{b}}}\right), \tag{5.7}$$

where $b_{\mathrm{db},m_{\mathrm{bsr}},m_{\mathrm{f}},m_{\mathrm{b}}}$ $(m_{\mathrm{bsr}} = 1, 2, \ldots, M_{\mathrm{bsr}})$ is the beamforming weight coefficients. Since $\mathbf{B}_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}}$ is diagonal, we have $M_{\mathrm{bsr}} = M_{\mathrm{bsd}}$. Considering that sweeping beams on both the vertical and horizontal planes would take a longer time, we design the beamformers to sweep on the vertical plane only, and operate omnidirectionally on the horizon plane. The beamforming weight coefficients can be configured according to the beamforming response, $P_{m_{\mathrm{f}}}(\bar{\theta}_{m_{\mathrm{b}}})$, as given by

$$P_{m_{\mathrm{f}}}(\bar{\theta}_{m_{\mathrm{b}}}) = \mathbf{b}^H_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}} \mathbf{B}^H_{\mathrm{ab}} \mathbf{a}_{\mathrm{bs},m_{\mathrm{f}},m_{\mathrm{b}}}(\bar{\theta}_{m_{\mathrm{b}}}, \phi), \tag{5.8}$$

where

$$\mathbf{b}_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}} = \left[b_{\mathrm{db},1,m_{\mathrm{f}},m_{\mathrm{b}}}, \ldots, b_{\mathrm{db},M_{\mathrm{bsr}},m_{\mathrm{f}},m_{\mathrm{b}}}\right]^T \tag{5.9}$$

is the normalized digital beamforming vector, i.e., $\mathbf{b}_{\mathrm{db},m_\mathrm{f},m_\mathrm{b}}^H \mathbf{b}_{\mathrm{db},m_\mathrm{f},m_\mathrm{b}} = 1$, and $\bar{\theta}_{m_\mathrm{b}}$ is the m_b-th beamforming sweeping direction. Assume that the vertical angular sweeping interval is $\frac{\pi}{M_\mathrm{b}}$. The elevation angle of the m_b-th angular sample ranges from $\frac{\pi}{M_\mathrm{b}}(m_\mathrm{b}-1)$ to $\frac{\pi}{M_\mathrm{b}}m_\mathrm{b}$.

We also decouple the digital beamforming matrix $\mathbf{B}_{\mathrm{db},m_\mathrm{f},m_\mathrm{b}}$ in (5.7) between the vertical and horizontal planes, i.e., $\mathbf{B}_{\mathrm{db},m_\mathrm{f},m_\mathrm{b}} = \mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}} \otimes \mathbf{B}_{\mathrm{hdb},m_\mathrm{f},m_\mathrm{b}}$, where $\mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}} \in \mathbb{C}^{M_\mathrm{vd}\times M_\mathrm{vd}}$ and $\mathbf{B}_{\mathrm{hdb},m_\mathrm{f},m_\mathrm{b}} \in \mathbb{C}^{M_\mathrm{hd}\times M_\mathrm{hd}}$ are diagonal matrices with elements $b_{\mathrm{vdb},m_\mathrm{vd},m_\mathrm{f},m_\mathrm{b}}$ and $b_{\mathrm{hdb},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{b}}$, respectively. Thus, after hybrid beamforming, the array steering vectors $\mathbf{a}_{\mathrm{bs},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})$ can be written as:

$$
\begin{aligned}
&\mathbf{a}_{\mathrm{hd},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) = \mathbf{B}_{m_\mathrm{f},m_\mathrm{b}}^H \mathbf{a}_{\mathrm{bs},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) \\
&= \left((\mathbf{B}_{\mathrm{vab}} \otimes \mathbf{B}_{\mathrm{hab}})\left(\mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}} \otimes \mathbf{B}_{\mathrm{hdb},m_\mathrm{f},m_\mathrm{b}}\right)\right)^H \mathbf{a}_{\mathrm{bs},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) \\
&\overset{\text{(a)}}{=} \left((\mathbf{B}_{\mathrm{vab}}\mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}}) \otimes (\mathbf{B}_{\mathrm{hab}}\mathbf{B}_{\mathrm{hdb},m_\mathrm{f},m_\mathrm{b}})\right)^H \mathbf{a}_{\mathrm{bs},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) \\
&\overset{\text{(b)}}{=} \left((\mathbf{B}_{\mathrm{vab}}\mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}})^H \otimes (\mathbf{B}_{\mathrm{hab}}\mathbf{B}_{\mathrm{hdb},m_\mathrm{f},m_\mathrm{b}})^H\right)\left(\mathbf{a}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}) \otimes \mathbf{a}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})\right) \\
&= \left(\mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}}^H \mathbf{B}_{\mathrm{vab}}^H \mathbf{a}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})\right) \otimes \left(\mathbf{B}_{\mathrm{hdb},m_\mathrm{f},m_\mathrm{b}}^H \mathbf{B}_{\mathrm{hab}}^H \mathbf{a}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})\right) \\
&= \mathbf{a}_{\mathrm{vhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}) \otimes \mathbf{a}_{\mathrm{hhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}),
\end{aligned} \tag{5.10}
$$

where $\mathbf{a}_{\mathrm{vhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}) \in \mathbb{C}^{M_\mathrm{vd}}$, $\mathbf{a}_{\mathrm{hhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) \in \mathbb{C}^{M_\mathrm{hd}}$, $M_\mathrm{vd} = M_\mathrm{vr} = M_\mathrm{v}$, and $M_\mathrm{hd} = M_\mathrm{hr} = 2P+1$. In (5.10), (a) and (b) are based on two important properties of the Kronecker product, i.e., $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD}$ and $(\mathbf{A}\otimes\mathbf{B})^H = \mathbf{A}^H\otimes\mathbf{B}^H$ [119]. We have

$$
\mathbf{a}_{\mathrm{vhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}) = \mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}}^H \mathbf{a}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}). \tag{5.11}
$$

According to Theorem 2, the m_hd-th element of $\mathbf{a}_{\mathrm{hhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})$ is given by:

$$
\begin{aligned}
&a_{\mathrm{hhb},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) \\
&\approx \sqrt{M_\mathrm{h}} j^{m_\mathrm{hd}} b_{\mathrm{hdb},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{b}} J_{m_\mathrm{hd}}\left(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})\right) \exp\left(-jm_\mathrm{hd}\phi_{k_{m_\mathrm{b}}}\right).
\end{aligned} \tag{5.12}
$$

Given our hybrid beamforming design, we can present the beamspace signals of the mmWave UCyA in a tensor form. Considering the observations at all sweeping

intervals, subcarriers and time frames, the beamspace signals can be modeled as:

$$x_{m_{\mathrm{vd}},m_{\mathrm{hd}},m_{\mathrm{f}},m_{\mathrm{t}},m_{\mathrm{b}}} = \sum_{k_{m_{\mathrm{b}}}=1}^{K_{m_{\mathrm{b}}}} (s_{m_{\mathrm{t}},k_{m_{\mathrm{b}}}} a_{\mathrm{vhb},m_{\mathrm{vd}},m_{\mathrm{f}},m_{\mathrm{b}}}(\theta_{k_{m_{\mathrm{b}}}}) a_{\mathrm{hhb},m_{\mathrm{hd}},m_{\mathrm{f}},m_{\mathrm{b}}}(\theta_{k_{m_{\mathrm{b}}}}, \phi_{k_{m_{\mathrm{b}}}}) \times a_{\mathrm{fbs},m_{\mathrm{f}},m_{\mathrm{b}}}(\tau_{k_{m_{\mathrm{b}}}})) + n_{m_{\mathrm{vd}},m_{\mathrm{hd}},m_{\mathrm{f}},m_{\mathrm{t}},m_{\mathrm{b}}}, \tag{5.13}$$

where $n_{m_{\mathrm{vd}},m_{\mathrm{hd}},m_{\mathrm{f}},m_{\mathrm{t}},m_{\mathrm{b}}}$ is the additive noise.

We consider the samples from the m_{b}-th vertical sweeping beam, and (5.13) can be rewritten in the following tensor form [1]

$$\mathcal{X}_{:,:,:,:,m_{\mathrm{b}}} = \mathcal{A}_{m_{\mathrm{b}}} \times_4 \mathbf{S}_{m_{\mathrm{b}}} + \mathcal{N}_{m_{\mathrm{b}}} \in \mathbb{C}^{M_{\mathrm{vd}} \times M_{\mathrm{hd}} \times M_{\mathrm{f}} \times M_{\mathrm{t}}}, \tag{5.14}$$

where all the angle and delay parameters at the m_{b}-th sweeping beam are collected in the space-time response tensor $\mathcal{A}_{m_{\mathrm{b}}} \in \mathbb{C}^{M_{\mathrm{vd}} \times M_{\mathrm{hd}} \times M_{\mathrm{f}} \times K_{m_{\mathrm{b}}}}$; $\mathbf{S}_{m_{\mathrm{b}}} \in \mathbb{C}^{M_{\mathrm{t}} \times K_{m_{\mathrm{b}}}}$ collects the received symbols $s_{m_{\mathrm{t}},k_{m_{\mathrm{b}}}}$; and $\mathcal{N}_{m_{\mathrm{b}}} \in \mathbb{C}^{M_{\mathrm{vd}} \times M_{\mathrm{hd}} \times M_{\mathrm{f}} \times M_{\mathrm{t}}}$ collects the noise samples.

5.4 Low-Complexity Coherent Preprocessing for Wideband Signals

As the second step in Fig. 5.2, a new low-complexity UCAMI is developed in this section to suppress the beam squinting effect and enable coherent combining of measurement signals across wideband. The conventional UCAMI [73] needs to solve a computationally expensive multi-dimensional optimization problem whose dimension is equal to the number of estimation parameters. Different from the conventional UCAMI, there are only 1-D problems in our proposed approach.

As shown in (5.11) and (5.12), the array steering vectors depend on the frequency and so do the beamspace signals. As a consequence, the signals can suffer from the beam squinting effect, due to the wide bandwidth of mmWave signals. It is critical to preprocess the beamspace signals in order to suppress the frequency dependence of the array steering vectors. The suppression of frequency dependence is performed

by designing the so-called *focusing matrix*, which focuses the array steering vectors at each frequency to a reference frequency, denoted by f_0 [69, 72]. From (5.12), we see that after being processed by the RF network, the subcarrier frequency f_{m_f} in (5.5) is transformed into the Bessel function, $J_{m_\mathrm{hd}}\left(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})\right)$, which only depends on f_{m_f} and $\theta_{k_{m_\mathrm{b}}}$, and is independent of $\phi_{k_{m_\mathrm{b}}}$. We only need to optimize the focusing matrices in the elevation angular domain, since f_{m_f} is decoupled from the azimuth angle $\phi_{k_{m_\mathrm{b}}}$ in (5.12). Moreover, by taking the vertical array steering vector in (5.11) into consideration, we find that both $J_{m_\mathrm{hd}}\left(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})\right)$ and $a_{\mathrm{vhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})$ depend only on the elevation angle $\theta_{k_{m_\mathrm{b}}}$.

We first design the optimization problem for the horizontal array steering vectors in (5.12). Because the measurement samples in (5.14) are collected from the M_b vertical sweeping beams, the optimization can be conducted in each vertical angular sweeping interval separately. Define

$$\boldsymbol{g}_{m_\mathrm{f}}(\theta) = \left[J_{-P}\left(\gamma_{m_\mathrm{f}}(\theta)\right), J_{-P+1}\left(\gamma_{m_\mathrm{f}}(\theta)\right), \ldots, J_{P}\left(\gamma_{m_\mathrm{f}}(\theta)\right)\right]^T, \tag{5.15}$$

which collects all the Bessel functions in (5.12) at the m_f-th subcarrier. We discretize each sweeping interval into N_b elevation angular values. Then, the horizontal factor matrices associated with the subcarrier frequency, f_{m_f}, for the m_b-th sweeping interval can be written as:

$$\mathbf{G}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}} = \left[\boldsymbol{g}_{m_\mathrm{f}}(\theta_{m_\mathrm{b},1}), \boldsymbol{g}_{m_\mathrm{f}}(\theta_{m_\mathrm{b},2}), \ldots, \boldsymbol{g}_{m_\mathrm{f}}(\theta_{m_\mathrm{b},N_\mathrm{b}})\right], \tag{5.16}$$

where $\theta_{m_\mathrm{b},n_\mathrm{b}} = \frac{\pi}{M_\mathrm{b}}(m_\mathrm{b}-1) + \frac{\pi}{M_\mathrm{b}N_\mathrm{b}}(n_\mathrm{b}-1)$ is the discretized elevation angle.

We directly use $a_{\mathrm{vhb},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})$ to optimize the vertical array steering vectors by constructing

$$\mathbf{G}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}} = \left[\mathbf{a}_{\mathrm{vhb},m_\mathrm{f}}(\theta_{m_\mathrm{b},1}), \mathbf{a}_{\mathrm{vhb},m_\mathrm{f}}(\theta_{m_\mathrm{b},2}), \ldots, \mathbf{a}_{\mathrm{vhb},m_\mathrm{f}}(\theta_{m_\mathrm{b},N_\mathrm{b}})\right]. \tag{5.17}$$

We then obtain the focusing matrices on the vertical and horizontal planes, de-

noted by $\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}$ and $\mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}$, by formulating the following optimization problems:

$$\begin{aligned}\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}} = \arg\min_{\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}} &\left\|\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}\mathbf{G}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}} - \mathbf{G}_{\mathrm{v},m_{\mathrm{f}0},m_\mathrm{b}}\right\|_\mathrm{F}^2, \\ \text{s.t.} \quad &\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}^H\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}} = \mathbf{I}_{M_\mathrm{v}};\end{aligned} \tag{5.18}$$

$$\begin{aligned}\mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}} = \arg\min_{\mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}} &\left\|\mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}\mathbf{G}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}} - \mathbf{G}_{\mathrm{h},m_{\mathrm{f}0},m_\mathrm{b}}\right\|_\mathrm{F}^2, \\ \text{s.t.} \quad &\mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}^H\mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}} = \mathbf{I}_{2P+1},\end{aligned} \tag{5.19}$$

where $m_{\mathrm{f}0}$ is the index to the subcarriers at the reference frequency f_0, and the constraints prevent focusing losses [73].

The solutions to Problems (5.18) and (5.19) are given by [69]

$$\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}} = \mathbf{V}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}\mathbf{U}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}^H;\ \mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}} = \mathbf{V}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}\mathbf{U}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}^H, \tag{5.20}$$

where the columns of $\mathbf{U}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}$ (or $\mathbf{U}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}$) and $\mathbf{V}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}$ (or $\mathbf{V}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}$) are the left and right singular vectors of $\mathbf{G}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}\mathbf{G}_{\mathrm{v},m_{\mathrm{f}0},m_\mathrm{b}}^H$ (or $\mathbf{G}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}\mathbf{G}_{\mathrm{h},m_{\mathrm{f}0},m_\mathrm{b}}^H$), respectively.

We construct $\widetilde{b}_{\mathrm{f},m_\mathrm{b}} = b_{\mathrm{f},m_\mathrm{b}}^{-1}$, $\widetilde{\mathbf{B}}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}} = \mathbf{B}_{\mathrm{vdb},m_\mathrm{f},m_\mathrm{b}}^{-1}$, and $\widetilde{\mathbf{B}}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}} = \mathbf{B}_{\mathrm{hdb},m_\mathrm{f},m_\mathrm{b}}^{-1}$ to offset the impact of beam sweeping on the received signals. The focusing matrices (5.20) suppress the frequency dependence of the array steering vectors. After this coherent wideband processing, in the m_b-th sweeping beam, the received signal at the m_f-th subcarrier in (5.14) can be calculated as

$$\widetilde{\mathcal{X}}_{:,:,m_\mathrm{f},:,m_\mathrm{b}} = \mathcal{X}_{:,:,m_\mathrm{f},:,m_\mathrm{b}}\widetilde{b}_{\mathrm{f},m_\mathrm{b}} \times_1 \left(\mathbf{T}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}\widetilde{\mathbf{B}}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}\right) \times_2 \left(\mathbf{T}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}\widetilde{\mathbf{B}}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}\right). \tag{5.21}$$

The elements of $\widetilde{\mathcal{X}}_{:,:,m_\mathrm{f},:,m_\mathrm{b}}$ can be expressed as

$$\begin{aligned}\widetilde{x}_{m_\mathrm{vd},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{t},m_\mathrm{b}} = &\sum_{k_{m_\mathrm{b}}=1}^{K_{m_\mathrm{b}}} \widetilde{a}_{\mathrm{vhb},m_\mathrm{vd},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})\widetilde{a}_{\mathrm{hhb},m_\mathrm{hd},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})a_{\mathrm{f},m_\mathrm{f}}(\tau_{k_{m_\mathrm{b}}})s_{m_\mathrm{t},k_{m_\mathrm{b}}} \\ &+ \widetilde{n}_{m_\mathrm{vd},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{t},m_\mathrm{b}},\end{aligned} \tag{5.22}$$

where $\widetilde{a}_{\mathrm{vhb},m_\mathrm{vd},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})$ and $\widetilde{a}_{\mathrm{hhb},m_\mathrm{hd},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})$ are the resultant array manifolds in (5.13). $\widetilde{n}_{m_\mathrm{vd},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{t},m_\mathrm{b}}$ is the transformed noise sample, which still yields the

zero-mean Gaussian distribution due to the constraints on the beamforming weights and focusing matrices.

We note that there are two-dimensional variables, $\phi_{k_{m_\mathrm{b}}}$ and $\theta_{k_{m_\mathrm{b}}}$, in the frequency-dependent array steering vectors $\mathbf{a}_{\mathrm{v},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}})$ and $\mathbf{a}_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}})$. UCA-MI [73] would have to optimize the focusing matrices on the elevation and azimuth angular domains simultaneously, resulting in a two-dimensional problem with a high complexity. In contrast, our proposed method only needs a one-dimensional optimization problem, i.e., (5.18) and (5.19), reducing the complexity significantly.

5.5 Tensor-based Parameter Estimation

With the received signals preprocessed (in Sections 5.3 and 5.4), the resultant array steering vectors are frequency-independent in (5.22). Only the delay-dependent factor, $a_{\mathrm{f},m_\mathrm{f}}(\tau_{k_{m_\mathrm{b}}})$, depends on the carrier frequency. In this section, we propose a new tensor-based joint delay-angle estimation algorithm which is the last step in Fig. 1.5, and a new spatial smoothing method which is the second-to-last (optional) step in the figure. Despite the use of the existing truncated HOSVD, the proposed joint delay-angle estimation algorithm involves new estimation processes. Specifically, the matrix TLS problem formulation is generalized to the tensor case. The azimuth angles are estimated by substituting the estimated elevation angles, which avoids potential mismatches between the estimated results of the elevation and azimuth AoAs. By revealing and exploiting the recurrence relations between the UCAs at different layers of the UCyA, the proposed spatial smoothing method is developed to decorrelate the coherent signals to correctly decompose the signal and noise subspaces in all dimensions. The computational complexity of the proposed algorithm is analyzed at the end.

5.5.1 Truncated HOSVD Model of Measurement Samples

With no a-priori knowledge of the number of signals in each sweeping beam, K_{m_b}, we collect all the sweeping results in (5.22) to jointly process the signals from the K signal sources. The element of the received signal tensor model is given by

$$\begin{aligned} y_{m_\mathrm{vd},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{t}} &= \sum_{m_\mathrm{b}=1}^{M_\mathrm{b}} \widetilde{x}_{m_\mathrm{vd},m_\mathrm{hd},m_\mathrm{f},m_\mathrm{t},m_\mathrm{b}} \\ &= \sum_{m_\mathrm{b}=1}^{M_\mathrm{b}} \Bigg[\sum_{k_{m_\mathrm{b}}=1}^{K_{m_\mathrm{b}}} \widetilde{a}_{\mathrm{vhb},m_\mathrm{vd},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}) \widetilde{a}_{\mathrm{hhb},m_\mathrm{hd},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}},\phi_{k_{m_\mathrm{b}}}) \\ &\quad \times a_{\mathrm{f},m_\mathrm{f}}(\tau_{k_{m_\mathrm{b}}}) s_{m_\mathrm{t},k_{m_\mathrm{b}}} + \widetilde{n}_{m_\mathrm{v},m_\mathrm{p},m_\mathrm{f},m_\mathrm{b},m_\mathrm{t}} \Bigg] \\ &= \sum_{k=1}^{K} \widetilde{a}_{\mathrm{vhb},m_\mathrm{vd}}(\theta_k) \widetilde{a}_{\mathrm{hhb},m_\mathrm{hd}}(\theta_k,\phi_k) a_{\mathrm{f},m_\mathrm{f}}(\tau_k) s_{m_\mathrm{t},k} + \dot{n}_{m_\mathrm{v},m_\mathrm{p},m_\mathrm{f},m_\mathrm{t}}, \end{aligned} \tag{5.23}$$

which can be expressed concisely as:

$$\mathcal{Y} = \sum_{m_\mathrm{b}=1}^{M_\mathrm{b}} \widetilde{\mathcal{X}}_{:,:,:,:,m_\mathrm{b}} = \widetilde{\mathcal{A}} \times_4 \mathbf{S} + \dot{\mathcal{N}} \in \mathbb{C}^{M_\mathrm{vd} \times M_\mathrm{hd} \times M_\mathrm{f} \times M_\mathrm{t}}, \tag{5.24}$$

where $\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_{M_\mathrm{b}}] \in \mathbb{C}^{M_\mathrm{t} \times K}$ collects all the symbols and $\dot{\mathcal{N}} = \sum_{m_\mathrm{b}=1}^{M_\mathrm{b}} \mathcal{N}_{m_\mathrm{b}}$ collects all noise samples. An illustration of the received signal tensor model is shown in Fig. 5.2(b). In (5.24), $\widetilde{\mathcal{A}} \in \mathbb{C}^{M_\mathrm{vd} \times M_\mathrm{hd} \times M_\mathrm{f} \times K}$ is known as the space-time response tensor [113], and obtained by concatenating the K response tensors, $\widetilde{\mathcal{A}}_k \in \mathbb{C}^{M_\mathrm{vd} \times M_\mathrm{hd} \times M_\mathrm{f}}$, as given by:

$$\widetilde{\mathcal{A}} = \left[\widetilde{\mathcal{A}}_1 \sqcup_4 \widetilde{\mathcal{A}}_2 \sqcup_4 \ldots \sqcup_4 \widetilde{\mathcal{A}}_K \right]. \tag{5.25}$$

Because the array steering vectors are frequency-independent after the coherent wideband preprocessing (as described in Section 5.4), the space-time response tensor of the k-th signal source, $\widetilde{\mathcal{A}}_k$, is given by

$$\widetilde{\mathcal{A}}_k = \widetilde{\mathbf{a}}_\mathrm{vhb}(\theta_k) \circ \widetilde{\mathbf{a}}_\mathrm{hhb}(\theta_k,\phi_k) \circ \mathbf{a}_\mathrm{f}(\tau_k), \tag{5.26}$$

where $[\widetilde{\mathbf{a}}_\mathrm{vhb}(\theta_k)]_{m_\mathrm{vd},1} = \widetilde{a}_{\mathrm{vdb},m_\mathrm{vd}}(\theta_k)$, $[\widetilde{\mathbf{a}}_\mathrm{hhb}(\theta_k,\phi_k)]_{m_\mathrm{hd},1} = \widetilde{a}_{\mathrm{hhb},m_\mathrm{hd}}(\theta_k,\phi_k)$, and $[\mathbf{a}_\mathrm{f}(\tau_k)]_{m_\mathrm{f},1} = a_{\mathrm{f},m_\mathrm{f}}(\tau_k)$.

By substituting (5.25) and (5.26) into (5.24), we obtain

$$\mathcal{Y} = \sum_{k=1}^{K} \widetilde{\mathbf{a}}_{\mathrm{vhb}}(\theta_k) \circ \widetilde{\mathbf{a}}_{\mathrm{hhb}}(\theta_k, \phi_k) \circ \mathbf{a}_{\mathrm{f}}(\tau_k) \circ \mathbf{s}_k + \dot{\mathcal{N}}, \tag{5.27}$$

where $\mathbf{s}_k = [\mathbf{S}]_{:,k}$. (5.27) indicates that, in a noiseless case, $\mathcal{Y}$ can be regarded as the sum of K rank-one tensors. Therefore, (5.27) is the CP decomposition of $\mathcal{Y}$ (see Property 3 in Appendix A). $\mathrm{Rank}(\mathcal{Y}) = K$[1]. According to Property 3 in Appendix A, (5.27) can be written as

$$\mathcal{Y} = \left[\!\left[\mathcal{Z}_{\mathrm{s}}; \widetilde{\mathbf{A}}_{\mathrm{vhb}}, \widetilde{\mathbf{A}}_{\mathrm{hhb}}, \mathbf{A}_{\mathrm{f}}, \mathbf{S} \right]\!\right] + \dot{\mathcal{N}} \tag{5.28}$$

where $\left[\widetilde{\mathbf{A}}_{\mathrm{vhb}}\right]_{:,k} = \widetilde{\mathbf{a}}_{\mathrm{vhb}}(\theta_k)$, $\left[\widetilde{\mathbf{A}}_{\mathrm{hhb}}\right]_{:,k} = \widetilde{\mathbf{a}}_{\mathrm{hhb}}(\theta_k, \phi_k)$, $[\mathbf{A}_{\mathrm{f}}]_{:,k} = \mathbf{a}_{\mathrm{f}}(\tau_k)$, and $\mathcal{Z}_{\mathrm{s}} \in \mathbb{C}^{K \times K \times K \times K}$ is an identity superdiagonal tensor.

Given the typically sparse multipath propagation of mmWave, the number of received paths is much smaller than the numbers of antennas, subcarriers, and time frames, i.e., $K < \min(M_{\mathrm{vd}}, M_{\mathrm{hd}}, M_{\mathrm{f}}, M_{\mathrm{t}})$. Thus, the ranks of $\widetilde{\mathbf{A}}_{\mathrm{vhb}}$, $\widetilde{\mathbf{A}}_{\mathrm{hhb}}$, $\mathbf{A}_{\mathrm{f}}$ and $\mathbf{S}$ are all K. According to the CP model (5.28), in the presence of non-negligible noises, $\widetilde{\mathbf{A}}_{\mathrm{vhb}}$, $\widetilde{\mathbf{A}}_{\mathrm{hhb}}$, $\mathbf{A}_{\mathrm{f}}$ and $\mathbf{S}$ correspond to the factor matrix of the measurement tensor $\mathcal{Y}$ in each mode. The ranks of the mode-n unfoldings of tensor $\mathcal{Y}$, i.e., the n-ranks of $\mathcal{Y}$ ($n = 1, 2, 3, 4$), are all K.

As a high-dimensional generalization of matrix SVD, the HOSVD (see Property 2 in Appendix A) conducts the SVD of the unfolding of $\mathcal{Y}$ in each mode separately. This can suppress the received noise in each mode. The HOSVD of the measurement tensor $\mathcal{Y}$ is given by

$$\mathcal{Y} = \mathcal{L} \times_1 \mathbf{U}_{\mathrm{v}} \times_2 \mathbf{U}_{\mathrm{h}} \times_3 \mathbf{U}_{\mathrm{f}} \times_4 \mathbf{U}_{\mathrm{t}} = \left[\!\left[\mathcal{L}; \mathbf{U}_{\mathrm{v}}, \mathbf{U}_{\mathrm{h}}, \mathbf{U}_{\mathrm{f}}, \mathbf{U}_{\mathrm{t}} \right]\!\right], \tag{5.29}$$

where the unitary matrices, $\mathbf{U}_{\mathrm{v}} \in \mathbb{C}^{M_{\mathrm{vd}} \times M_{\mathrm{vd}}}$, $\mathbf{U}_{\mathrm{h}} \in \mathbb{C}^{M_{\mathrm{hd}} \times M_{\mathrm{hd}}}$, $\mathbf{U}_{\mathrm{f}} \in \mathbb{C}^{M_{\mathrm{f}} \times M_{\mathrm{f}}}$, and $\mathbf{U}_{\mathrm{t}} \in \mathbb{C}^{M_{\mathrm{t}} \times M_{\mathrm{t}}}$, are the left singular matrices of the mode-n unfoldings of tensor

[1]According to (5.27), we have $\mathrm{Rank}(\mathcal{Y}) \leq K$. $\mathrm{Rank}(\mathcal{Y}) < K$ only occurs when the locations of two coherent sources are the same, which rarely happens.

$\mathcal{Y}$, and the core tensor $\mathcal{L} \in \mathbb{C}^{M_{\rm vd}\times M_{\rm hd}\times M_{\rm f}\times M_{\rm t}}$ is obtained by moving the singular matrices to the left-hand side of (5.29):

$$\mathcal{L} = \mathcal{Y} \times_1 \mathbf{U}_{\rm v}^H \times_2 \mathbf{U}_{\rm h}^H \times_3 \mathbf{U}_{\rm f}^H \times_4 \mathbf{U}_{\rm t}^H. \tag{5.30}$$

Because the n-ranks of tensor $\mathcal{Y}$ are K, the SVD of the mode-1 unfolding $\mathbf{Y}_{(1)} \in \mathbb{C}^{M_{\rm vd}\times(M/M_{\rm vd})}$ can be written as

$$\mathbf{Y}_{(1)} = \mathbf{U}_{\rm v}\mathbf{\Sigma}_{\rm v}\mathbf{V}_{\rm v}^H = [\mathbf{U}_{\rm v,s}\ \mathbf{U}_{\rm v,n}] \begin{bmatrix} \mathbf{\Sigma}_{\rm v,s} & \mathbf{0}_{K\times(\frac{M}{M_{\rm vd}}-K_{\rm vd})} \\ \mathbf{0}_{(M_{\rm vd}-K_{\rm vd})\times K_{\rm vd}} & \mathbf{\Sigma}_{\rm v,n} \end{bmatrix} [\mathbf{V}_{\rm v,s}\ \mathbf{V}_{\rm v,n}]^H, \tag{5.31}$$

where $K_{\rm vd} = \min(K, M_{\rm vd})$ and $M = M_{\rm vd}M_{\rm hd}M_{\rm f}M_{\rm t}$. The signal subspace $\mathbf{U}_{\rm v,s} \in \mathbb{C}^{M_{\rm vd}\times K_{\rm vd}}$ and the noise subspace $\mathbf{U}_{\rm v,n} \in \mathbb{C}^{M_{\rm vd}\times(M_{\rm vd}-K_{\rm vd})}$ of the mode-1 unfolding $\mathbf{Y}_{(1)}$ correspond to the $K_{\rm vd}$ largest and the $(M_{\rm vd} - K_{\rm vd})$ smallest elements of the diagonal matrix $\mathbf{\Sigma}_{\rm v} = {\rm diag}(\sigma_{{\rm v},1}, \sigma_{{\rm v},2}, \ldots, \sigma_{{\rm v},M_{\rm vd}})$, respectively. $\sigma_{{\rm v},1}, \sigma_{{\rm v},2}, \ldots, \sigma_{{\rm v},M_{\rm vd}}$ are the non-zero singular values of the mode-1 unfolding $\mathbf{Y}_{(1)}$, and calculated by $\sigma_{{\rm v},m_{\rm vd}} = \|\mathcal{L}_{m_{\rm vd},:,:,:,:}\|$. The signal subspace matrices of the mode-2,3,4 unfoldings of $\mathcal{Y}$, i.e., $\mathbf{U}_{\rm h,s} \in \mathbb{C}^{M_{\rm hd}\times K_{\rm hd}}$, $\mathbf{U}_{\rm f,s} \in \mathbb{C}^{M_{\rm f}\times K_{\rm f}}$, and $\mathbf{U}_{\rm t,s} \in \mathbb{C}^{M_{\rm t}\times K}$ can be obtained in the same way, where $K_{\rm hd} = \min(K, M_{\rm hd})$ and $K_{\rm f} = \min(K, M_{\rm f})$.

By removing the noise subspace in each mode of $\mathcal{Y}$, we construct a low-rank truncated HOSVD model of the noise-free measurement tensor $\mathcal{Y}_{\rm s}$ [76], as given by

$$\mathcal{Y}_{\rm s} = \mathcal{L}_{\rm s} \times_1 \mathbf{U}_{\rm v,s} \times_2 \mathbf{U}_{\rm h,s} \times_3 \mathbf{U}_{\rm f,s} \times_4 \mathbf{U}_{\rm t,s} \in \mathbb{C}^{M_{\rm vd}\times M_{\rm hd}\times M_{\rm f}\times M_{\rm t}}, \tag{5.32}$$

where $\mathcal{L}_{\rm s} \in \mathbb{C}^{K_{\rm vd}\times K_{\rm hd}\times K_{\rm f}\times K}$ is obtained by discarding the insignificant singular values of the mode-n unfoldings of $\mathcal{Y}$.

5.5.2 Joint Angle-Delay Estimation

We propose a tensor-based joint delay-angle estimation algorithm by exploiting the shift-invariance relations between the elements in each mode of the measurement

tensor. By comparing (5.24) with (5.28), we first obtain

$$\widetilde{\mathcal{A}} = \mathcal{Z}_{\rm s} \times_1 \widetilde{\mathbf{A}}_{\rm vhb} \times_2 \widetilde{\mathbf{A}}_{\rm hhb} \times_3 \mathbf{A}_{\rm f}. \tag{5.33}$$

According to the truncated HOSVD model (5.32), we define the signal subspace tensor:

$$\mathcal{U}_{\rm s} = \mathcal{L}_{\rm s} \times_1 \mathbf{U}_{\rm v,s} \times_2 \mathbf{U}_{\rm h,s} \times_3 \mathbf{U}_{\rm f,s} \in \mathbb{C}^{M_{\rm vd} \times M_{\rm hd} \times M_{\rm f} \times K}. \tag{5.34}$$

By comparing (5.24), (5.32), (5.33) and (5.34), we have $\mathcal{U}_{\rm s} \times_4 \mathbf{U}_{\rm t,s} = \widetilde{\mathcal{A}} \times_4 \mathbf{S}$. Because $\mathbf{U}_{\rm t,s} \in \mathbb{C}^{M_{\rm t} \times K}$ and $\mathbf{S} \in \mathbb{C}^{M_{\rm t} \times K}$ are full column rank matrices, we obtain

$$\widetilde{\mathcal{A}} = \mathcal{U}_{\rm s} \times_4 \mathbf{D}, \tag{5.35}$$

where $\mathbf{D} \in \mathbb{C}^{K \times K}$ is a full rank matrix. Based on (5.35), we generalize the matrix-based subspace algorithm to the tensor, and estimate the delay and angles of each signal path.

Estimation of Elevation Angle

We first propose a tensor-based total-least-squares ESPRIT (TLS-ESPRIT) algorithm to estimate the elevation angle and delay. To estimate the elevation angle of each signal path, we first reveal and then exploit the shift-invariance relations underlying the vertical array steering matrix $\widetilde{\mathbf{A}}_{\rm v}$, according to (5.4) and (5.23).

To select the elevation angle-related subtensors, we define two selection matrices:

$$\mathbf{J}_{\rm v1} = [\mathbf{I}_{M_{\rm vd}-1}, \mathbf{0}_{(M_{\rm vd}-1)\times 1}] \in \mathbb{R}^{(M_{\rm vd}-1)\times M_{\rm vd}};\ \mathbf{J}_{\rm v2} = [\mathbf{0}_{(M_{\rm vd}-1)\times 1}, \mathbf{I}_{M_{\rm vd}-1}] \in \mathbb{R}^{(M_{\rm vd}-1)\times M_{\rm vd}}, \tag{5.36}$$

which are two auxiliary matrices. We reveal the following shift-invariance relation among the selected subtensors:

$$\widetilde{\mathcal{A}} \times_1 \mathbf{J}_{\rm v2} = \widetilde{\mathcal{A}} \times_1 \mathbf{J}_{\rm v1} \times_4 \mathbf{\Theta}_{\rm v}, \tag{5.37}$$

where $\mathbf{\Theta}_{\mathrm{v}} = \mathrm{diag}\left(e^{-j\frac{2\pi}{c}f_0 h\cos(\theta_1)}, \ldots, e^{-j\frac{2\pi}{c}f_0 h\cos(\theta_K)}\right) \in \mathbb{C}^{K\times K}$. The shift-invariance relation is the key to our design of the following tensor-based TLS-ESPRIT algorithm[2]. The algorithm estimates the elevation angle of each signal in the tensor form.

By substituting (5.35) into (5.37), we have

$$\mathcal{U}_{\mathrm{sv2}} \times_4 \mathbf{D} = \mathcal{U}_{\mathrm{sv1}} \times_4 (\mathbf{\Theta}_{\mathrm{v}}\mathbf{D}), \tag{5.38}$$

where $\mathcal{U}_{\mathrm{sv1}} = \mathcal{U}_{\mathrm{s}} \times_1 \mathbf{J}_{\mathrm{v1}} \in \mathbb{C}^{(M_{\mathrm{vd}}-1)\times M_{\mathrm{hd}}\times M_{\mathrm{f}}\times K}$ and $\mathcal{U}_{\mathrm{sv2}} = \mathcal{U}_{\mathrm{s}} \times_1 \mathbf{J}_{\mathrm{v2}} \in \mathbb{C}^{(M_{\mathrm{vd}}-1)\times M_{\mathrm{hd}}\times M_{\mathrm{f}}\times K}$ are the selected subtensors of the signal subspace tensor $\mathcal{U}_{\mathrm{s}}$. Since $\mathbf{D}$ is a full rank matrix, we can left-multiply its inverse to both sides of (5.38) and obtain

$$\mathcal{U}_{\mathrm{sv2}} = \mathcal{U}_{\mathrm{sv1}} \times_4 \mathbf{\Psi}_{\mathrm{v}}, \tag{5.39}$$

where $\mathbf{\Psi}_{\mathrm{v}} = \mathbf{D}^{-1}\mathbf{\Theta}_{\mathrm{v}}\mathbf{D} \in \mathbb{C}^{K\times K}$.

To obtain the estimate of $\mathbf{\Psi}_{\mathrm{v}}$ in (5.39), we define $\mathbf{\Upsilon}_{\mathrm{v}} = [\mathbf{\Upsilon}_{\mathrm{v1}} \quad \mathbf{\Upsilon}_{\mathrm{v2}}] \in \mathbb{C}^{K\times 2K}$. According to the standard TLS [119], the estimate of $\mathbf{\Psi}_{\mathrm{v}}$ is $\hat{\mathbf{\Psi}}_{\mathrm{v}} = -\hat{\mathbf{\Upsilon}}_{\mathrm{v1}}\hat{\mathbf{\Upsilon}}_{\mathrm{v2}}^{-1}$, where the K eigenvalues of $\hat{\mathbf{\Psi}}_{\mathrm{v}}$, i.e., $\lambda_{\mathrm{v},k}$, $k = 1, 2, \ldots, K$, are sorted in descending order. We now generalize the matrix TLS problem formulation [119] to the tensor case, as given by:

$$\hat{\mathbf{\Upsilon}}_{\mathrm{v}} = \arg\min_{\mathbf{\Upsilon}_{\mathrm{v}}} \left\|\mathcal{U}_{\mathrm{sv1}} \times_4 \mathbf{\Upsilon}_{\mathrm{v1}} + \mathcal{U}_{\mathrm{sv2}} \times_4 \mathbf{\Upsilon}_{\mathrm{v2}}\right\|, \quad \text{s.t.} \quad \mathbf{\Upsilon}_{\mathrm{v}}\mathbf{\Upsilon}_{\mathrm{v}}^H = \mathbf{I}_K, \tag{5.40}$$

which finds a unitary matrix $\mathbf{\Upsilon}_{\mathrm{v}}$ whose submatrices are orthogonal to $\mathcal{U}_{\mathrm{sv1}}$ and $\mathcal{U}_{\mathrm{sv2}}$ in mode-4.

According to (A.7), the mode-4 unfoldings of $\mathcal{U}_{\mathrm{sv1}}$ is given by

$$\mathbf{U}_{\mathrm{sv1}(4)} = \mathbf{U}_{\mathrm{s}(4)} \left(\mathbf{J}_{\mathrm{v1}} \otimes \mathbf{I}_{M_{\mathrm{hd}}} \otimes \mathbf{I}_{M_{\mathrm{f}}}\right)^T, \tag{5.41}$$

[2]The least-squares (LS) procedure can also be used for solving the invariance equation (5.39), but has slightly lower accuracy than TLS. Section 5.6 will provide the results of performance comparison between the proposed algorithm (T-CTLS), which applies TLS-ESPRIT for parameter estimation, with its variation (T-CLS), which uses LS-ESPRIT.

where $\mathbf{U}_{\mathrm{s}(4)} \in \mathbb{C}^{K \times M_{\mathrm{vd}} M_{\mathrm{hd}} M_{\mathrm{f}}}$ is the mode-4 unfolding of $\mathcal{U}_{\mathrm{s}}$. The mode-4 unfoldings of $\mathcal{U}_{\mathrm{sv1}}$ can be formulated in the same way. Since $\|\mathcal{A}\| = \left\|\mathbf{A}_{(n)}\right\|_{\mathrm{F}}$ $(n = 1, 2, \ldots, N)$ [76], we rewrite the tensor TLS problem (5.40) in a matrix format as:

$$\begin{aligned}\hat{\boldsymbol{\Upsilon}}_{\mathrm{v}} &= \arg\min_{\boldsymbol{\Upsilon}_{\mathrm{v}}} \left\| \boldsymbol{\Upsilon}_{\mathrm{v1}} \mathbf{U}_{\mathrm{s}(4)} \left(\mathbf{J}_{\mathrm{v1}} \otimes \mathbf{I}_{M_{\mathrm{hd}}} \otimes \mathbf{I}_{M_{\mathrm{f}}}\right)^T + \boldsymbol{\Upsilon}_{\mathrm{v2}} \mathbf{U}_{\mathrm{s}(4)} \left(\mathbf{J}_{\mathrm{v2}} \otimes \mathbf{I}_{M_{\mathrm{hd}}} \otimes \mathbf{I}_{M_{\mathrm{f}}}\right)^T \right\|_{\mathrm{F}} \\ &= \arg\min_{\boldsymbol{\Upsilon}_{\mathrm{v}}} \left\| \mathbf{W}_{\mathrm{v}} \boldsymbol{\Upsilon}_{\mathrm{v}}^T \right\|_{\mathrm{F}}, \end{aligned} \tag{5.42}$$

where

$$\mathbf{W}_{\mathrm{v}} = \left[\left(\mathbf{J}_{\mathrm{v1}} \otimes \mathbf{I}_{M_{\mathrm{hd}}} \otimes \mathbf{I}_{M_{\mathrm{f}}}\right) \mathbf{U}_{\mathrm{s}(4)}^T \quad \left(\mathbf{J}_{\mathrm{v2}} \otimes \mathbf{I}_{M_{\mathrm{hd}}} \otimes \mathbf{I}_{M_{\mathrm{f}}}\right) \mathbf{U}_{\mathrm{s}(4)}^T \right] \in \mathbb{C}^{(M_{\mathrm{vd}}-1) M_{\mathrm{hd}} M_{\mathrm{f}} \times 2K}. \tag{5.43}$$

The SVD of $\mathbf{W}_{\mathrm{v}}^H \mathbf{W}_{\mathrm{v}}$ is written as $\mathbf{W}_{\mathrm{v}}^H \mathbf{W}_{\mathrm{v}} = \dot{\mathbf{U}}_{\mathrm{v}} \dot{\boldsymbol{\Lambda}}_{\mathrm{v}} \dot{\mathbf{V}}_{\mathrm{v}}$, where $\dot{\mathbf{U}}_{\mathrm{v}} \in \mathbb{C}^{2K \times 2K}$ and $\dot{\mathbf{V}}_{\mathrm{v}} \in \mathbb{C}^{2K \times 2K}$ are the left and right singular matrices, respectively; and $\dot{\boldsymbol{\Lambda}}_{\mathrm{v}} \in \mathbb{C}^{2K \times 2K}$ contains singular values. We partition $\dot{\mathbf{U}}_{\mathrm{v}}$ into four blocks:

$$\dot{\mathbf{U}}_{\mathrm{v}} = \begin{bmatrix} \dot{\mathbf{U}}_{\mathrm{v11}} & \dot{\mathbf{U}}_{\mathrm{v12}} \\ \dot{\mathbf{U}}_{\mathrm{v21}} & \dot{\mathbf{U}}_{\mathrm{v22}} \end{bmatrix} \in \mathbb{C}^{2K \times 2K}. \tag{5.44}$$

Let $\hat{\boldsymbol{\Upsilon}}_{\mathrm{v1}} = \dot{\mathbf{U}}_{\mathrm{v12}}^T \in \mathbb{C}^{K \times K}$ and $\hat{\boldsymbol{\Upsilon}}_{\mathrm{v2}} = \dot{\mathbf{U}}_{\mathrm{v22}}^T \in \mathbb{C}^{K \times K}$.

According to the array steering expression in (5.4), the elevation angle of the k-th path can be finally estimated as

$$\hat{\theta}_k = \arccos\left(\frac{jc \ln(\lambda_{\mathrm{v},k})}{2\pi f_0 h} \right). \tag{5.45}$$

Estimation of Delay

We can estimate the delays by exploiting the shift-invariance relation between the delay-related subtensors. We express the delay-dependent shift-invariance relation, as follows.

$$\widetilde{\mathcal{A}} \times_3 \mathbf{J}_{\mathrm{f2}} = \widetilde{\mathcal{A}} \times_3 \mathbf{J}_{\mathrm{f1}} \times_4 \boldsymbol{\Theta}_{\mathrm{f}}, \tag{5.46}$$

where $\boldsymbol{\Theta}_{\mathrm{f}} = \mathrm{diag}\left(e^{-j2\pi\Delta_{\mathrm{F}}\tau_1}, \ldots, e^{-j2\pi\Delta_{\mathrm{F}}\tau_K}\right)$ with Δ_{F} being the subcarrier spacing. $\mathbf{J}_{\mathrm{f1}}$ and $\mathbf{J}_{\mathrm{f2}}$ are two selection matrices to select the delay-related subtensors. $\mathbf{J}_{\mathrm{f1}}$ and

$\mathbf{J}_{\mathrm{f}2}$ can be constructed in the same way as in (5.36). By using TLS-ESPRIT (5.40), the delay of the k-th path, τ_k, can be estimated as

$$\hat{\tau}_k = \frac{j \ln(\lambda_{\mathrm{f},k})}{2\pi\Delta_{\mathrm{F}}}, \tag{5.47}$$

where $\lambda_{\mathrm{f},k}$ is an eigenvalue of the delay-related matrix $\mathbf{\Psi}_{\mathrm{f}} = \mathbf{D}\mathbf{\Theta}_{\mathrm{f}}\mathbf{D}^{-1}$. In the presence of non-negligible noises, the estimates of the elevation angle and delay of each source may be paired incorrectly. After obtaining the estimates of $\hat{\mathbf{\Psi}}_{\mathrm{v}}$ and $\hat{\mathbf{\Psi}}_{\mathrm{f}}$ with (5.40), joint SVD methods [120] can be used to obtain the joint eigenvalues of $\hat{\mathbf{\Psi}}_{\mathrm{v}}$ and $\hat{\mathbf{\Psi}}_{\mathrm{f}}$, and then the correctly matched pairs of estimated parameters can be obtained.

Estimation of Azimuth Angle

We design the tensor-MUSIC algorithm [78] to estimate the azimuth angle of each path. From (5.12), there are nonlinear Bessel functions in the expression for the horizontal array steering matrix $\widetilde{\mathbf{A}}_{\mathrm{h}}$, and therefore there is no shift-invariance relation for the azimuth angle estimation, as opposed to (5.37).

According to (5.32), we discard the largest K singular values of the mode-n unfoldings of the measurement tensor $\mathcal{Y}$, i.e., setting the corresponding parts of $\mathcal{L}$ to zero. Then we obtain the noise subspace tensor as[3]:

$$\mathcal{Y}_{\mathrm{n}} = \mathcal{L}_{\mathrm{n}} \times_1 \mathbf{U}_{\mathrm{v,n}} \times_2 \mathbf{U}_{\mathrm{h,n}} \times_3 \mathbf{U}_{\mathrm{f,n}} \times_4 \mathbf{U}_{\mathrm{t,n}}, \tag{5.48}$$

where $\mathbf{U}_{\mathrm{v,n}} \in \mathbb{C}^{M_{\mathrm{vd}} \times (M_{\mathrm{vd}} - K)}$ is constructed by the last $(M_{\mathrm{vd}} - K)$ columns of $\mathbf{U}_{\mathrm{v}}$; $\mathbf{U}_{\mathrm{h,n}} \in \mathbb{C}^{M_{\mathrm{hd}} \times (M_{\mathrm{hd}} - K)}$ is the last $(M_{\mathrm{hd}} - K)$ columns of $\mathbf{U}_{\mathrm{h}}$; $\mathbf{U}_{\mathrm{f,n}} \in \mathbb{C}^{M_{\mathrm{f}} \times (M_{\mathrm{f}} - K)}$ is the last $(M_{\mathrm{f}} - K)$ columns of $\mathbf{U}_{\mathrm{f}}$; and $\mathbf{U}_{\mathrm{t,n}} \in \mathbb{C}^{M_{\mathrm{t}} \times (M_{\mathrm{t}} - K)}$ is the last $(M_{\mathrm{t}} - K)$ columns of $\mathbf{U}_{\mathrm{t}}$. The core $\mathcal{L}_{\mathrm{n}}$ can be evaluated by

$$\mathcal{L}_{\mathrm{n}} = \mathcal{Y}_{\mathrm{n}} \times_1 \mathbf{U}_{\mathrm{v,n}}^H \times_2 \mathbf{U}_{\mathrm{h,n}}^H \times_3 \mathbf{U}_{\mathrm{f,n}}^H \times_4 \mathbf{U}_{\mathrm{t,n}}^H. \tag{5.49}$$

[3]It is well known that this solution for estimating the noise subspace is not optimal in the least squares sense. However, it is a good approximation in most cases [2,75] and it is easy to implement.

Based on the subspace estimation of $\mathcal{Y}$ (5.34), we generalize the matrix-based MUSIC, and the tensor MUSIC spectrum of the azimuth angle is defined as

$$\mathrm{SP}_{\mathrm{MUSIC}}(\mathbf{\Phi}) = \left\| \widetilde{\mathcal{A}} \times_2 \mathbf{U}_{\mathrm{h,n}}^H \right\|^{-2}, \tag{5.50}$$

where $\mathbf{\Phi} = [\phi_1, \phi_2, \ldots, \phi_K]$.

According to (A.7), the mode-2 matricization of $\widetilde{\mathcal{A}}$ in (5.33) can be expressed as

$$\widetilde{\mathbf{A}}_{(2)} = \widetilde{\mathbf{A}}_{\mathrm{hhb}} Z_{\mathrm{s}(2)} \left(\mathbf{A}_{\mathrm{f}} \otimes \mathbf{I}_{M_{\mathrm{t}}} \otimes \widetilde{\mathbf{A}}_{\mathrm{vhb}} \right)^T. \tag{5.51}$$

We substitute (5.51) into (5.50) and obtain the mode-2 matricization of (5.50), as given by

$$\mathrm{SP}_{\mathrm{MUSIC}}(\mathbf{\Phi}) = \left\| \mathbf{U}_{\mathrm{h,n}}^H \widetilde{\mathbf{A}}_{\mathrm{hhb}} Z_{\mathrm{s}(2)} \left(\mathbf{A}_{\mathrm{f}} \otimes \mathbf{I}_{M_{\mathrm{t}}} \otimes \widetilde{\mathbf{A}}_{\mathrm{vhb}} \right)^T \right\|_{\mathrm{F}}^{-2}. \tag{5.52}$$

By substituting the estimated elevation angle of each path, i.e., (5.45), into (5.52), the corresponding azimuth angle ϕ_k can be estimated by searching the prominent peaks of the tensor MUSIC spectrum (5.52).

Remark 1. *When applying the tensor-based TLS-ESPRIT and MUSIC algorithms to estimate the parameters, we first apply the HOSVD evaluates the SVD of the unfoldings of $\mathcal{Y}$ in all modes, and then suppress the noise components by discarding the singular vectors and slices of the core tensor that correspond to insignificant singular values of the matricized tensor in each mode. The uniqueness and identifiability of the proposed algorithm inherits from that of the matrix-based counterpart of the algorithm, due to the fact that the proposed algorithm can be regarded as the high-dimensional generalization of the matrix-based counterpart [76]. In particular, to achieve the unique parameter estimates of the K sources would need to construct the signal subspace tensor $\mathcal{U}_s$ with a smaller number of sources K than time frames M_t. Our method is suitable for multi-dimensional parameter estimation problems in mmWave systems, where $K \ll \min(M_{vd}, M_{hd}, M_f, M_t)$ due to the sparsity of*

mmWave[4].

When applying the matrix-based alternative, the noise is only suppressed in one of the dimensions (or modes) of the measurement tensor, hence degrading the estimation accuracy. This is because the noise is multi-dimensional with the same dimensions as the received signal. It is important to take all dimensions of the received signal into consideration, and suppress the noises in all the dimensions. Thus, the use of tensors can better suppress the noises than matrices, hence improving the estimation accuracy of the elevation and azimuth angles and the delay, i.e., $\hat{\theta}_k$, $\hat{\phi}_k$, *and* $\hat{\tau}_k$.

5.5.3 Tensor-based Spatial Smoothing for UCyA

The parameter estimation presented in Sections 5.5.1 and 5.5.2 is actually the last step in Fig. 1.5. In this subsection, we propose the necessary optional second-to-last step. The decomposition of the signal and noise subspaces in (5.31) is under the assumption that all the received signals are incoherent, as typically required in the subpace-based parameter estimation algorithms, such as MUSIC [29] and ESPRIT [63]. The rank of the signal subspace is assumed to be the number of received signals K. In practice, coherent signals are often received. The rank of the signal subspace decreases, leading to incorrect decomposition of the subspaces. An effective method to restore the rank is a spatial smoothing technique [119] which divides an antenna array into several subarrays and exploits the inherent linear recurrence relations (i.e., shift invariances) among the subarrays to decorrelate the coherent signals. Unfortunately, the spatial smoothing technique is only applicable

[4]In rich multipath environments, i.e., $K \geq \max(M_{\mathrm{vd}}, M_{\mathrm{hd}}, M_{\mathrm{f}})$, no singular values and core slices of the mode-n unfoldings can be discarded, because all these belong to the signal subspace. Thus, in this case, the tensor-based subspace estimation is equivalent to the matrix-based counterpart [118].

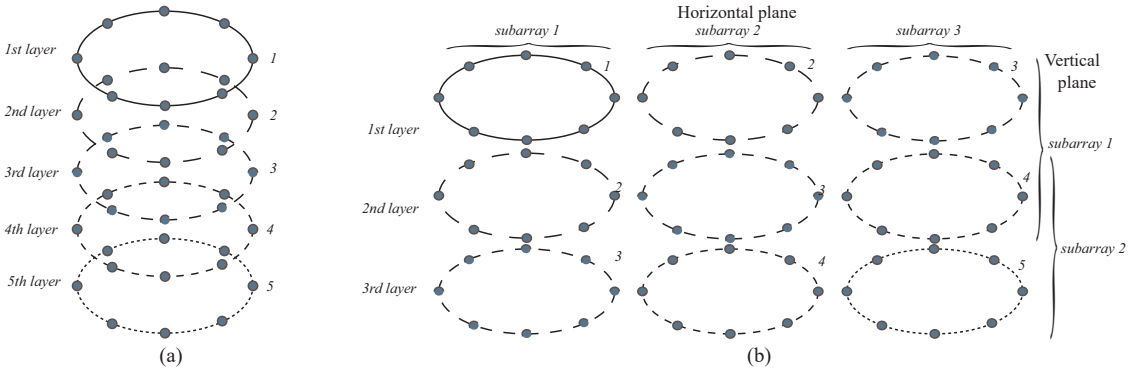


Figure 5.3 : An illustration of the proposed spatial smoothing for a five-layer UCyA, where we need to construct three "subarrays" on the horizontal plane, the second and third UCAs are seen as the translations of the first UCA at the same layer. After the spatial smoothing, the original first, second, and third UCAs are at the first layer of the "new" UCyA, the second layer accommodates the original second, third and fourth UCAs, and the third layer of the "new" UCyA accommodates the original third, fourth and fifth UCAs.

to systems with uniformly and linearly spaced antenna elements [119].

We extend the spatial smoothing technique to our hybrid UCyA to decorrelate coherent signals. This is not trivial, as the array manifolds of the UCyA in the horizontal space domain (i.e., the second mode of $\mathcal{Y}$) are UCAs, not linear arrays. It is difficult to split subarrays and obtain the required recurrence relations, as existing spatial smoothing techniques would require. We propose to utilize the recurrence relations between the UCAs at different layers of the UCyA to create the required recurrence-relation subarrays in the horizontal space domain. In other words, we regard each UCA as a subarray, and use these vertically arranged and coaxially aligned subarrays to construct the "virtual" subarrays in the horizontal space domain. The n_{h}-th subarray in the horizontal space domain can be constructed

as

$$\mathcal{Y}_{\rm ss}^{(n_{\rm h})} = \mathcal{Y} \times_1 \mathbf{J}_{{\rm ssh},n_{\rm h}}, \tag{5.53}$$

where $\mathbf{J}_{{\rm ssh},n_{\rm h}} = [\mathbf{0}_{M_{\rm hd}\times(n_{\rm h}-1)}, \mathbf{I}_{M_{\rm hd}}, \mathbf{0}_{M_{\rm hd}\times(N_{\rm h}-n_{\rm h})}]$.

Then, we can generate the required linear recurrence relation between two adjacent subarrays: $\mathcal{Y}_{\rm ss}^{(n_{\rm h}+1)} = \mathcal{Y}_{\rm ss}^{(n_{\rm h})} \times_4 \mathbf{\Theta}_{\rm h}$, where

$$\mathbf{\Theta}_{\rm h} = {\rm diag}\left(e^{-j\frac{2\pi}{c}f_0 h\cos(\theta_1)}, \ldots, e^{-j\frac{2\pi}{c}f_0 h\cos(\theta_K)}\right). \tag{5.54}$$

The numbers of subarrays and elements per subarray are determined in the following theorem:

Theorem 3. *If both the numbers of subarrays and elements per subarray are larger than the number of signals, i.e., $N_h \geq K$ and $M_{hd} \geq K$, the rank of the signal subspace in the mode-2 of the concatenated tensor $\mathcal{Y}_{ssh} = \left[\underset{n_h=1,\ldots,N_h}{\sqcup_4} \mathcal{Y}_{ss}^{(n_h)}\right]$ is K.*

Proof. The proof can be developed in the same way as in [1], and hence omitted. □

According to Theorem 3, we need to construct subarrays in all domains for the correct decomposition of the subspaces, and apply the HOSVD in all modes of $\mathcal{Y}$. Because some of the vertically arranged UCAs are used to construct the "virtual" subarrays in the horizontal space domain, the number of subarrays in the vertical space domain decreases. Take the five-layer UCyA in Fig. 5.3(a) for an example. The original five-layer UCyA shown in Fig. 5.3(a) becomes a three-layer virtual array, which constructs the subarrays in the vertical space domain, as shown in Fig. 5.3(b).

We propose to meticulously arrange the virtual subarrays. $N_{\rm v}$ subarrays are constructed in the vertical space with $\tilde{M}_{\rm v}$ elements per subarray, and $N_{\rm h}$ subarrays are constructed in the horizontal space with $M_{\rm hd}$ elements per subarray. Because there are linear recurrence relations among subcarrier frequencies, the standard spatial

smoothing technique can be used in the frequency domain (i.e., the mode-3 of $\mathcal{Y}$). We decouple the mode-3 of $\mathcal{Y}$ into $N_{\rm f}$ subarrays with $\tilde{M}_{\rm f}$ elements each. As a result, the spatially smoothed tensor is given by

$$\mathcal{Y}_{\rm ss} = \left[\bigsqcup_{n_{\rm v}=1,\ldots,N_{\rm v}}^{4} \bigsqcup_{n_{\rm h}=1,\ldots,N_{\rm h}}^{4} \bigsqcup_{n_{\rm f}=1,\ldots,N_{\rm f}}^{4} \mathcal{Y}_{\rm ss}^{(n_{\rm v},n_{\rm h},n_{\rm f})} \right] \in \mathbb{C}^{\tilde{M}_{\rm v}\times M_{\rm hd}\times \tilde{M}_{\rm f}\times (M_{\rm t}N_{\rm v}N_{\rm h}N_{\rm f})}, \tag{5.55}$$

which is obtained by concatenating the subarrays in mode-4:

$$\mathcal{Y}_{\rm ss}^{(n_{\rm v},n_{\rm h},n_{\rm f})} = \mathcal{Y} \times_1 \mathbf{J}_{{\rm ssvh},n_{\rm vh}} \times_3 \mathbf{J}_{{\rm ssf},n_{\rm f}}, \tag{5.56}$$

where $n_{\rm vh} = n_{\rm v} + n_{\rm h} - 1$. $\mathbf{J}_{{\rm ssvh},n_{\rm vh}}$ and $\mathbf{J}_{{\rm ssf},n_{\rm f}}$ are two subtensor selection matrices, as given respectively by

$$\mathbf{J}_{{\rm ssvh},n_{\rm vh}} = [\mathbf{0}_{\tilde{M}_{\rm v}\times(n_{\rm vh}-1)}, \mathbf{I}_{\tilde{M}_{\rm v}}, \mathbf{0}_{\tilde{M}_{\rm v}\times(N_{\rm vh}-n_{\rm vh})}];\ \mathbf{J}_{{\rm ssf},n_{\rm f}} = [\mathbf{0}_{\tilde{M}_{\rm f}\times(n_{\rm f}-1)}, \mathbf{I}_{\tilde{M}_{\rm f}}, \mathbf{0}_{\tilde{M}_{\rm f}\times(N_{\rm f}-n_{\rm f})}]. \tag{5.57}$$

The number of subarray elements in the mode-1 and mode-3 can be computed by $\tilde{M}_{\rm v} = M_{\rm vd} - N_{\rm v} - N_{\rm h} + 2$ and $\tilde{M}_{\rm f} = M_{\rm f} - N_{\rm f} + 1$, respectively. To decorrelate coherent signals in each domain, we use $\mathcal{Y}_{\rm ss}$ to replace $\mathcal{Y}$ in (5.24). The parameter estimation of coherent signals follows the rest of the steps recorded in the earlier part of Section 5.5, which is the last step in Fig. 5.2.

Note that the proposed smoothing method is needed to guarantee that the rank used for parameter estimation is the actual rank. If we conduct the HOSVD based on a smaller rank (due to coherent signals) than the actual rank, the estimation performance of the azimuth and elevation angles, and delays would degrade. This is because when the smaller rank is used, signal components can be incorrectly decomposed into the noise subspace, reducing the dimensions of the constructed truncated HOSVD model of $\mathcal{Y}_{\rm s}$ in all modes. As a result, we would not be able to correctly estimate the azimuth and elevation angles, and delays.

Also note that by using the proposed method, the antenna apertures in the first and third modes are reduced, as the elements in the two modes of the original measurement tensor $\mathcal{Y}$ are used to construct a sufficient number of subarrays according

to Theorem 3. The loss of the antenna aperture in the first mode is nearly one third. The antenna aperture in the second mode does not change, because the subarrays in the mode are constructed by the the spatial shift of UCAs at the other layers. Algorithm 2 summarizes the procedure of the proposed tensor-based subspace estimation algorithm.

5.5.4 Complexity Analysis

The hardware and software complexity of the proposed tensor-based parameter estimation algorithm is analyzed. The proposed hybrid beamformers reduces the hardware complexity to $O(M_{\mathrm{bsr}}) = O(PM_{\mathrm{v}})$, while fully digital beamformers using the same number of antennas have hardware complexity $O(M_{\mathrm{bs}})$.

As for signal processing complexity, we compare the computational complexity of the proposed tensor-based algorithm with its matrix-based counterpart and the state-of-the-art CP-based orthogonal matching pursuit (CP-OMP) algorithm. For matrix-based algorithms, the computational complexity of performing SVD on the measurement sample matrix and truncating its rank to K is $O(PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}K)$. The complexities of estimating the delay, elevation angle, and azimuth angle are $O(K^3 + PM_{\mathrm{v}}M_{\mathrm{f}})$, $O(K^3 + PM_{\mathrm{v}})$, and $O(PK^2 + P^2KD)$, respectively. D is the size of search dimension. Thus, the overall complexity of the matrix-based estimation is $O(PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}K + PM_{\mathrm{v}}M_{\mathrm{f}} + K^3 + PM_{\mathrm{v}} + PK^2 + P^2KD)$. For the proposed tensor-based algorithm, the truncated HOSVD of the measurement tensor evaluates the SVD of its matricized form in each mode and discards insignificant singular vectors. The complexity is $O(4PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}K) = O(PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}K)$. The complexity of computing the core $\mathcal{L}_{\mathrm{n}}$ and the tensor signal subspace $\mathcal{U}_{\mathrm{s}}$ in (5.34) is $O(PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}K + PM_{\mathrm{v}}M_{\mathrm{f}}K^2)$. The complexities of estimating delay, elevation and azimuth angles are $O(PM_{\mathrm{v}}M_{\mathrm{f}}{+}K^3)$, $O(PM_{\mathrm{v}}M_{\mathrm{f}}{+}K^3)$ and $O(PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}K{+}P^2KD)$, respectively. The tensor-based algorithm needs slightly more computations, but its

Algorithm 2 Tensor-based subspace estimation algorithm

- **Input**: The received signals, $\mathbf{x}_{m_{\mathrm{f}},m_{\mathrm{t}},m_{\mathrm{b}}}$ ($m_{\mathrm{b}} = 1,\ldots,M_{\mathrm{b}}$, $m_{\mathrm{t}} = 1,\ldots,M_{\mathrm{t}}, m_{\mathrm{f}} = 1,\ldots,M_{\mathrm{f}}$), the number of sources, K, and geometrical parameters of the UCyA.

- **Output**: The estimated delay, elevation and azimuth angles, $\hat{\tau}_k$, $\hat{\theta}_k$, and $\hat{\phi}_k$, $k = 1,2,\ldots,K$.

- Design the analog and digital beamforming matrices, $\mathbf{B}_{\mathrm{ab}}$ and $\mathbf{B}_{\mathrm{db},m_{\mathrm{f}},m_{\mathrm{b}}}$, and model the beamspace signals according to (5.14).

- Calculate the focusing matrices, $\mathbf{T}_{\mathrm{v},m_{\mathrm{f}},m_{\mathrm{b}}}$ and $\mathbf{T}_{\mathrm{h},m_{\mathrm{f}},m_{\mathrm{b}}}$, by solving (5.18) and (5.19), and formulate the signals according to (5.21).

- Collect all the sweeping results in (5.22) and formulate them as $\mathcal{Y}$.

- Construct the spatially smoothed tensor $\mathcal{Y}_{\mathrm{ss}}$ by using (5.55).

- Take HOSVD of $\mathcal{Y}_{\mathrm{ss}}$ and get $\mathcal{U}_{\mathrm{s}}$ according to (5.32) and (5.34).

- Use TLS-ESPRIT (5.38)-(5.44), and estimate $\hat{\theta}_k$ and τ_k by using (5.45) and (5.47), respectively.

- Calculate the noise subspace tensor $\mathcal{Y}_{\mathrm{n}}$ in (5.48) and estimate $\hat{\phi}_k$ by searching the prominent peaks of (5.52).

complexity is still in the same order with that of its matrix-based counterpart. The CP-OMP algorithm [78] applies CP decomposition to decompose the received signal tensor model, and then uses OMP to estimate the parameters. The complexities are $O(PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}K + PM_{\mathrm{v}}M_{\mathrm{f}}K^2 + K^4)$ and $O(PM_{\mathrm{v}}M_{\mathrm{f}}M_{\mathrm{t}}(N_1 + N_2 + N_3 + N_4))$, respectively, where $N_1 \gg K$, $N_2 \gg K$, $N_3 \gg K$ and $N_4 \gg K$ are the dimensions of the OMP grid. The CP-OMP algorithm has a much higher complexity than that of our algorithm.

5.6 Simulation Results

In this section, simulation results are provided to demonstrate the performance of the proposed algorithm in the RMa scenario of future 5G/B5G systems. We simulate a system with 2 GHz bandwidth and a total of 2,000 subcarriers. Out of the total 2,000 subcarriers, $M_{\mathrm{f}} = 20$ evenly spaced subcarriers are selected for the proposed channel parameter estimation. Each of the subcarriers undergoes flat fading. The reference frequency $f_0 = 28$ GHz, and the number of time frames is $M_{\mathrm{t}}= 20$. To evaluate the performance of the proposed algorithm in typical mmWave channels, all the channel parameters are set according to 3GPP TR 38.901 [121]. A RMa scenario is considered in our simulation, and thus, the RMa pathloss model presented in [121] is applied. We set both the azimuth and elevation angle spreads to be 31.6° and the delay spread to be 32.3 ns. The number of time frames is set to $M_{\mathrm{t}}= 20$. We assume that there are $K = 5$ signals, two of which are coherent. The actual azimuth angles, elevation angles, and delays of the signals are set up randomly each time. The distance between vertically adjacent UCAs is $h = 0.5\lambda_0$ and the radius of the UCyA is $r = 2\lambda_0$, where $\lambda_0 = c/f_0$.

We compare the proposed tensor-based coherent TLS (T-CTLS) algorithm with its variation (T-CLS) which applies the LS procedure for solving the invariance equation (5.39); its variation without using smoothing (T-CTLS w/o S); its reduced

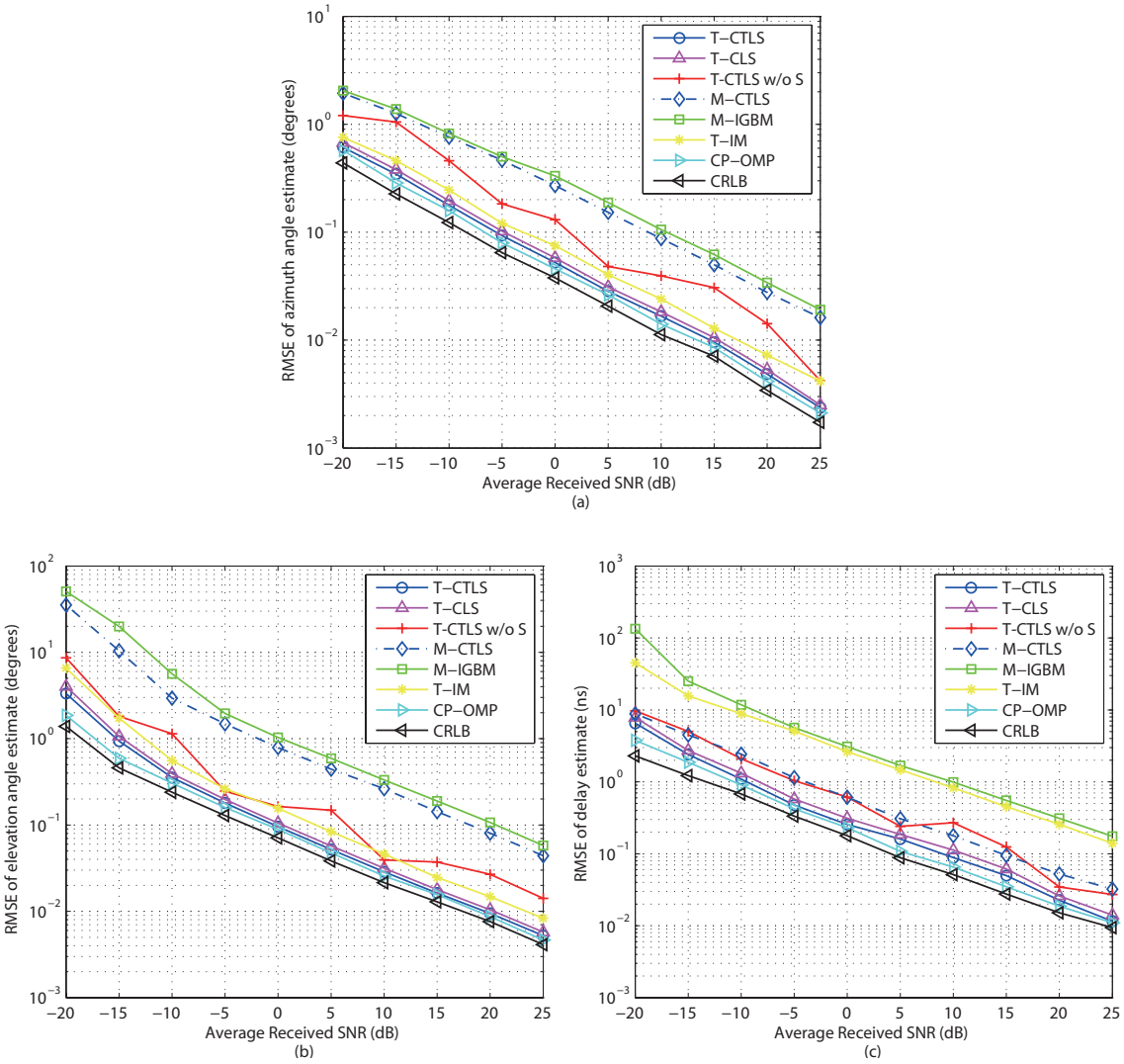


Figure 5.4 : The RMSE vs. the average received SNR for the estimation of different parameters. (a) Azimuth angle; (b) Elevation angle; (c) Delay.

version in the matrix form (M-CTLS); the state-of-the-art matrix-based incoherent generalized beamspace MUSIC (M-IGBM) [14]; the tensor-based incoherent MUSIC (T-IM) [2]; and the state-of-the-art CP-OMP [78]. The CRLB is derived according to [47, 117]. Note that both CP-OMP and our proposed parameter estimation algorithms are only applicable for additive Gaussian noises, where the noises are independent between different antennas and the noise power is identical at the antennas. This is because the algorithms which exploit the second-order statistics of the received signals cannot correctly decompose the signal and (non-Gaussian) noise subspaces.

Fig. 5.4 plots the RMSEs for the estimates of azimuth angles, elevation angles, and delays of the signals versus the average received SNR, where the BS has 400 receive antennas. Fig. 5.4 shows that our proposed T-CTLS algorithm outperforms the other algorithms, and its RMSE approaches the CRLB. In Figs. 5.4(a) and (b), we see that the tensor-based algorithms provide higher accuracy than their matrix-based counterparts, especially in low SNR regimes. The matrix-based algorithms are less robust to noises than the proposed tensor-based algorithms. We also see that CP-OMP has slightly better performance than our proposed algorithm, due to the fact that CP decomposition can be regarded as a maximum likelihood method under the additive Gaussian noise. However, its performance improvement is limited since OMP can only generate discrete estimates. In addition, CP-OMP also has a much higher complexity than our algorithm, as analyzed in Section 5.5.4. Fig. 5.4(c) shows that the methods applying coherent wideband signal preprocessing outperform those employing incoherent wideband preprocessing, in terms of delay estimation, because the former fully exploits the high temporal resolution offered by wideband mmWave systems.

Fig. 5.5 shows the RMSEs versus the number of receive antennas under -5 dB SNR. It is seen that the RMSE of the estimated parameters approaches the

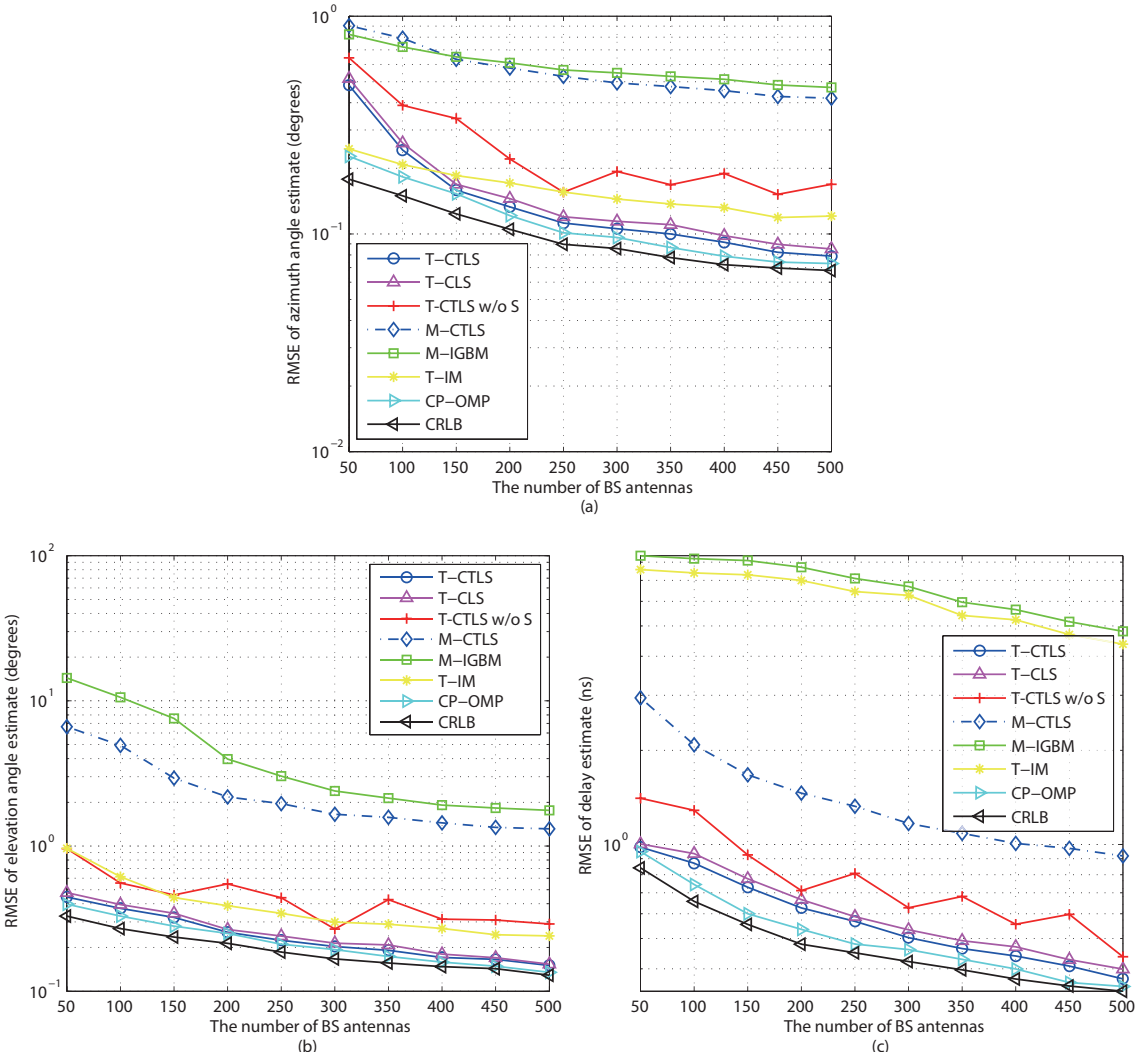


Figure 5.5 : The RMSE vs. the number of BS antennas for the estimation of different parameters. (a) Azimuth angle; (b) Elevation angle; (c) Delay.

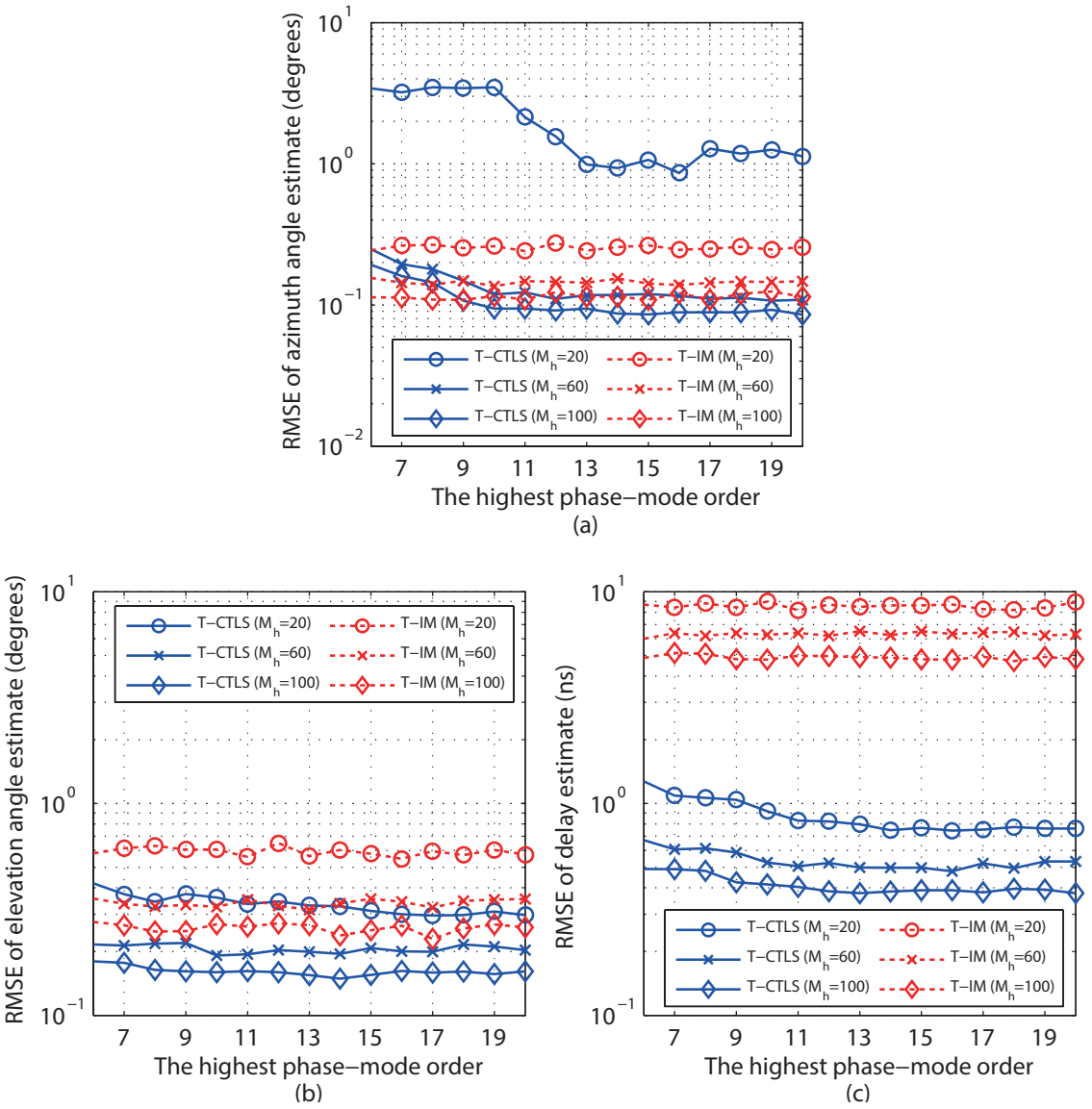


Figure 5.6 : The RMSE vs. the highest beamspace dimension. (a) Azimuth angle; (b) Elevation angle; (c) Delay.

CRLB, as the number of antennas increases. However, when the number of antennas is not very large, e.g. less than 100, the algorithms, including T-CTLS, T-CTL, and M-CTLS, cannot achieve accurate azimuth angle estimation, as shown in Fig. 5.5(a). The reason is that the conditions of Theorem 2 may not be met, and thus the approximation in (5.6) becomes inaccurate. Nevertheless, when the number of antennas is large, the RMSEs of these three algorithms decrease fast, and T-CTLS rapidly outperforms the others. By comparing Figs. 5.4 and 5.5, we also see that if the proposed spatial smoothing technique is not applied, the estimation accuracy of the proposed algorithm decreases noticeably. This is because two coherent signals are decorrelated, the signal and noise subspaces can be incorrectly decoupled without spatial smoothing, and the parameters of the coherent signals cannot be precisely estimated.

In order to validate Theorem 2, Fig. 5.6 plots the RMSE of the parameter estimation versus the highest order, P, with different numbers of horizontal array steering vectors. The SNR is -5 dB. We see that when P is less than 10 or the number of the horizontal array steering vectors in (5.5) is 20, the algorithms applying Theorem 2 to design the hybrid beamformers (i.e., T-CTLS and M-CTLS), cannot achieve satisfactory estimation, because the number of the transformed beamspace vectors (5.6) is not sufficient to represent the array response vectors. When $P \geq 12$, regardless of the number of array response vectors, increasing the beamspace vectors has little impact on the estimation. By exploiting this property, we can reduce the number of required RF chains and, in turn, the hardware cost.

Fig. 5.7 shows the RMSE of the estimated azimuth angles, elevation angles, and delays, with an increasing number of received paths. T-CTLS and M-CTLS are tested. We set SNR to -5 dB and $M_{\mathrm{f}} = 8$. We observe that the performance gap between the matrix and tensor forms of the proposed algorithm, i.e., M-CTLS and T-CTLS, decreases with the increasing number of received paths. This is because the

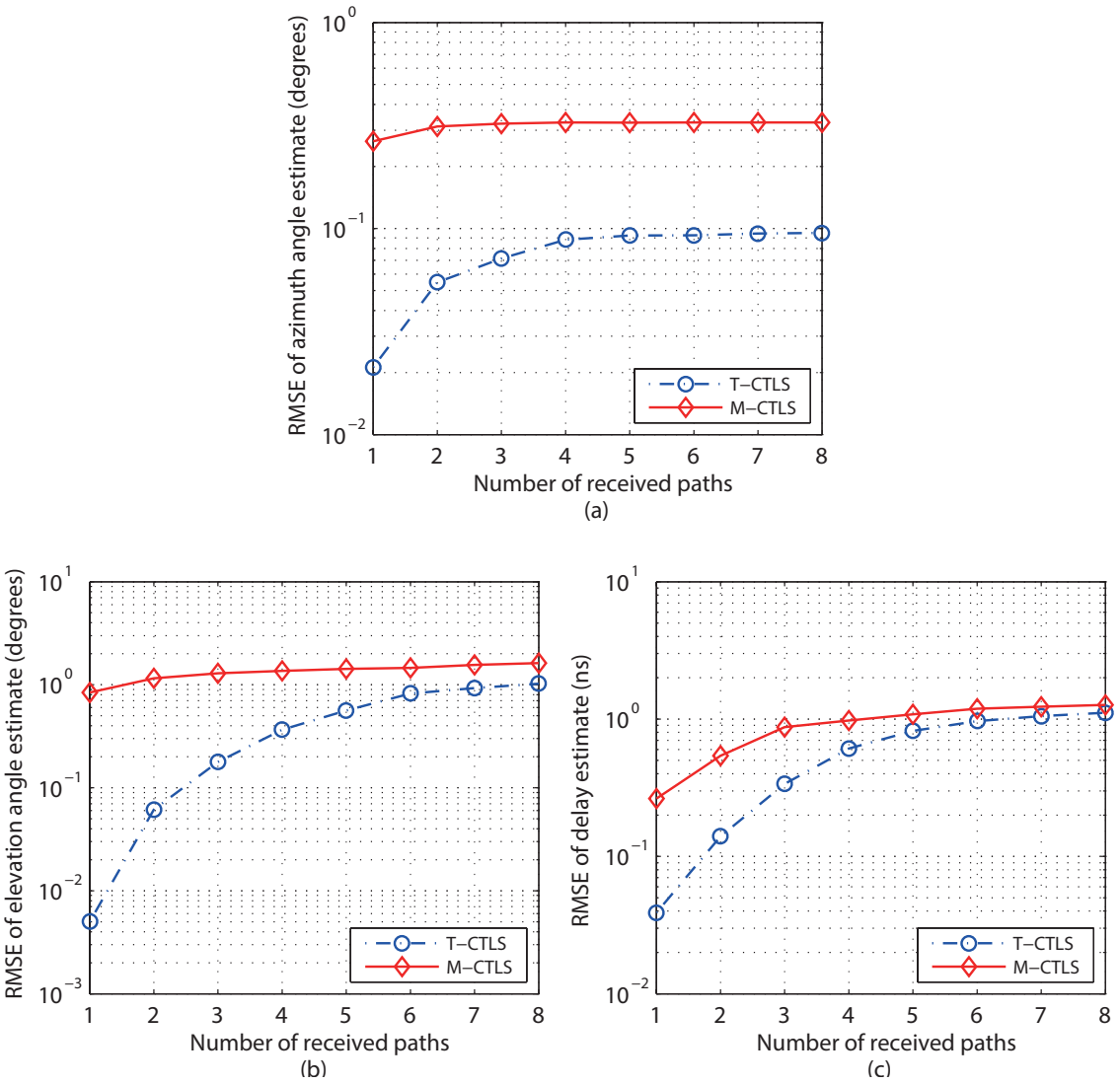


Figure 5.7 : The RMSE vs. the number of received paths. (a) Azimuth angle; (b) Elevation angle; (c) Delay.

noise components which can be suppressed by using the tensor-based algorithms in the first, second, and third modes of $\mathcal{Y}$, depend on the difference between the number of paths and the tensor dimension in each mode of $\mathcal{Y}$. As the number of received paths increases, the gain of the tensor-based algorithm, T-CTLS, diminishes. The performance gap remains consistent in Fig. 5.7(a) though. This is because, despite the number of paths increases, the dimension in the first mode of $\mathcal{Y}$, i.e., $M_{\text{vd}} = 2P + 1$, is still much larger than the number of paths. Moreover, we estimate the azimuth angles with tensor-MUSIC in (5.50). The method involves peak search, and is hardly affected by the number of paths. In conclusion, the new tensor-based algorithm, T-CTLS, can achieve much better performance than its matrix-based counterpart, especially under B5G settings where the number of received paths is small due to the sparsity of mmWave propagation.

5.7 Summary

In this chapter, we have presented a new tensor-based multi-dimensional channel parameter estimation algorithm for wideband directional hybrid UCyAs. By exploiting the multidimensional structure of the received signals, the algorithm suppresses the noises across all domains of the received signals, improving estimation accuracy. Specifically, we have designed the hybrid beamformers, which maintains the angular resolution and suppresses the beam squint effect. We have also formulated the received signals in the tensor form. We have shown that by applying tensor signal processing and developing the new HOSVD model, the noise in all the domains can be suppressed, and the shift-invariance relations can be revealed for parameter estimation. Given the relations, we have designed the new tensor-based TLS-ESPRIT algorithm, which can accurately estimate the channel parameters from both coherent and incoherent signals. By applying the channel parameters presented by 3GPP TR 38.901 [121], simulations have shown that, in the RMa scenario of future

5G/B5G systems, the proposed tensor-based algorithm can accurately estimate the multi-dimensional parameters in typically used mmWave channels, and can achieve much better performance than its matrix-based counterpart.

Chapter 6

Nested Hybrid UCyA Design and DoA Estimation

This chapter is devoted to sparse array-based hybrid beamforming and its corresponding angle estimation algorithm for mIoT networks, as illustrated in Fig. 6.1. Specifically, we propose a new nested massive hybrid UCyA design and the corresponding tensor-based angle estimation algorithm for of mIoT networks. By exploiting the sparse array techniques, the proposed hybrid antenna array enables the BS to estimate the DoAs of a large number of devices with much fewer RF chains than antennas. As a result, the massive access requirement of mIoT can be met, with significantly reduced hardware cost and network overhead. Tailored for the new hybrid UCyA array, we also propose a new tensor n-rank enhancement method and a new tensor-based 2-D DoA estimation algorithm. The algorithm suppresses the noise components in all tensor modes and operates on the signal data model directly, hence improving estimation accuracy with an affordable computational

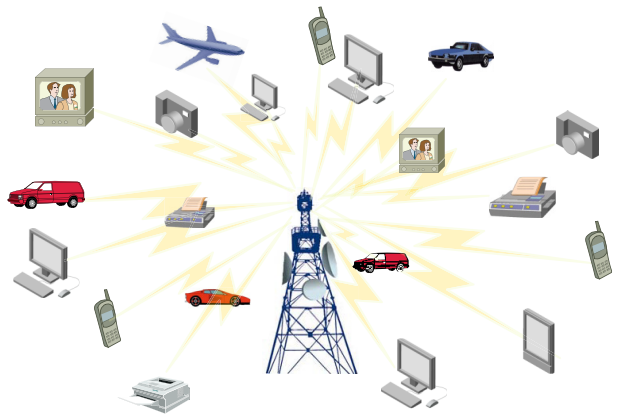


Figure 6.1 : The research scenario studied in Chapter 6.

complexity. Corroborated by a CRLB analysis, simulation results show that in the mIoT networks, the proposed hybrid UCyA array and the DoA estimation algorithm can accurately estimate the 2-D DoAs of a large number of IoT devices.

6.1 Motivation and Overview

MIoT networks, which enables to connect a large number of IoT devices to an internet-enabled system [84,85], require to reduce cost and power consumption while maintaining a high network access capability. Hybrid beamforming is a cost- and energy-efficient architecture to meet the requirement, but most of them apply RF networks to directly combine the received signals from multiple antennas, which would penalize system DoF and channel estimation accuracy [18, 38].

In this chapter, we apply the sparse array technique to design the hybrid beamforming, which enables the DoAs of a large number of devices to be estimated with a marginal accuracy loss while significantly reducing the number of required R-F chains. The details of the proposed hybrid beamforming design are presented in Section 6.3, where we design the hybrid beamforming based on the 2-D nested array. Compared to other sparse arrays, such as minimum redundancy array (MRA) [88], minimum hole array (MHA) [89] and coprime array [90], nested arrays can generate larger hole-free difference coarrays than coprime arrays under the same setting, and have simple closed-form expressions for a large number of elements, which cannot be achieved in MRA and MHA. In particular, we first design the phase shifter matrix to transform the nonlinear phase of the UCyA steering vectors to be linear to the element locations, so that the horizontal symmetric structure of UCyA can be preserved. We then flatten the 3-D RF-chain connection network of UCyA into a 2-D plane, and design the sparse RF-chain connection network developed from the "Configuration II" nested array [48]. By deploying the proposed sparse RF-chain connection network, we show that the proposed network can provide larger DoF

than the fully connected beamforming array.

To improve the accuracy of channel parameter estimation, we formulate the received signals in the tensor form and propose a spatial smoothing-based method to enhance the n-rank of the constructed the second-order statistics of the signal tensor model. In Section 6.4, we verify that we can build a signal tensor model that provides a large enough rank in each mode to perform the DoA estimation of all the devices, even when all the received signal powers, i.e., the second-order statistics of the signals, are equal. A new tensor-based subspace 2-D DoA estimation algorithm for the designed nested hybrid UCyA is developed in Section 6.5, where we combine the tensor tool with ESPRIT to estimate the elevation angles, and substitute the estimated elevation angles to derive the azimuth angles by using tensor MUSIC. The hardware and software complexity of the proposed tensor-based parameter estimation algorithm is also analyzed in 6.5.3.

In Section 6.6, we provide simulation results to evaluate the performance of our proposed algorithm in the mIoT networks. We show that although our proposed tensor-based algorithm provides a better parameter estimation accuracy than its matrix-based counterparts, because it can suppress much more noise components, it cannot achieve the same accuracy as the CP-based algorithms. However, the performance gap is small and the CP-based algorithms have much higher computational complexities than our algorithm, as discussed in Section 6.5.3. We also compare the proposed algorithm with the tensor-based 2-D MUSIC algorithm, and show that the performance degeneration of the tensor-based 2-D MUSIC is larger than the proposed algorithm when SNR decreases. This is because the 2-D MUSIC uses signal covariance tensors for the 2-D DoA estimation, and its MUSIC spectrum is a product of multiple separable second-order mode-n spectra, which results in undesirable cross-terms [94] and compromises the estimation accuracy.

6.2 System Model

In this chapter, we consider a BS equipped with an M_{bs}-antenna large-scale hybrid mmWave UCyA consisting of M_{v} vertically placed UCAs. Each of the UCAs is on a horizontal plane with M_{h} elements, and $M_{\text{bs}} = M_{\text{v}}M_{\text{h}}$. Let r be the radius of the UCyA, and h be the vertical spacing between any two adjacent vertical elements. We assume that there are K IoT devices, each equipped with a single antenna[1]. Each device has a dominating path and different devices have separable and resolvable paths. Hence, K signal paths are received by the BS. The received signal sample at the m_{t}-th time frame ($m_{\text{t}} = 1, \ldots, M_{\text{t}}$) can be expressed as [122]

$$\mathbf{x}_{m_{\text{t}}} = \sum_{k=1}^{K} s_{m_{\text{t}},k}\mathbf{B}^H\mathbf{a}_{\text{bs}}(\phi_k, \theta_k) + \mathbf{n}_{m_{\text{t}}}, \tag{6.1}$$

where ϕ_k and θ_k are the azimuth and elevation DoAs of the k-th device, respectively; $\mathbf{a}_{\text{bs}}(\phi_k, \theta_k) \in \mathbb{C}^{M_{\text{bs}}\times 1}$ denotes the steering vector of the hybrid UCyA; $s_{m_{\text{t}},k}$ is the received symbol of the k-th device at the m_{t}-th time frame; $\mathbf{n}_{m_{\text{t}}} \in \mathbb{C}^{M_{\text{bs}}\times 1}$ denotes the AWGN; $\mathbf{B} \in \mathbb{C}^{M_{\text{bs}}\times M_{\text{bsd}}}$ is the hybrid beamforming matrix; and M_{bsd} is the number of data streams.

Given the structure of the UCyA, the array steering vector $\mathbf{a}_{\text{bs}}(\phi_k, \theta_k)$ can be written as $\mathbf{a}_{\text{bs}}(\phi_k, \theta_k) = \mathbf{a}_{\text{v}}(\theta_k) \otimes \mathbf{a}_{\text{h}}(\theta_k, \phi_k)$, where $\mathbf{a}_{\text{v}}(\theta_k)$ and $\mathbf{a}_{\text{h}}(\theta_k, \phi_k)$ are the vertical and horizontal array steering vectors with their elements given by

$$[\mathbf{a}_{\text{v}}(\theta_k)]_{m_{\text{v}}} = a_{\text{v},m_{\text{v}}}(\theta_k) = \frac{1}{\sqrt{M_{\text{v}}}} \exp\left(-j\frac{2\pi}{\lambda}h(m_{\text{v}} - 1)\cos(\theta_k)\right), \tag{6.2}$$

$$\begin{aligned}[\mathbf{a}_{\text{h}}(\theta_k, \phi_k)]_{m_{\text{h}}} &= a_{\text{h},m_{\text{h}}}(\theta_k, \phi_k) \\ &= \frac{1}{\sqrt{M_{\text{h}}}} \exp\left(j\frac{2\pi}{\lambda}r\sin(\theta_k)\cos(\phi_k - \varphi_{m_{\text{h}}})\right),\end{aligned} \tag{6.3}$$

[1]The proposed technique can be readily applied when multiple antennas are deployed at a device. In that case, the paths originating from different antennas can be distinguished by transmitting different pilot signals.

where λ is the wavelength, $m_{\rm v} = 1, \ldots, M_{\rm v}$ and $m_{\rm h} = 1, \ldots, M_{\rm h}$. $\varphi_{m_{\rm h}} = 2\pi(m_{\rm h} - 1)/M_{\rm h}$ is the difference of the central angles between the $m_{\rm h}$-th antenna and the first antenna of each UCA. In this chapter, the antenna array can be reasonably treated as a phased array because the signal bandwidth B is much smaller than the carrier frequency f, i.e. $B \ll f$[2], and the signals are narrowband.

6.3 Proposed Nested 3-D Hybrid UCyA

In this section, we design the hybrid beamformer $\mathbf{B}$ for performing channel estimation. $\mathbf{B} = \mathbf{B}_{\rm rf}\mathbf{B}_{\rm bb} \in \mathbb{C}^{M_{\rm bs} \times M_{\rm bsd}}$ can be decoupled between an analog beamforming matrix $\mathbf{B}_{\rm rf} \in \mathbb{C}^{M_{\rm bs} \times M_{\rm rf}}$ and a digital beamforming matrix $\mathbf{B}_{\rm bb} \in \mathbb{C}^{M_{\rm rf} \times M_{\rm bsd}}$. Here, $M_{\rm rf}$ is the number of RF chains. We first briefly review the concept of difference coarray and sparse array, which are heavily used in this chapter. Then, we introduce the $\mathbf{B}$ design process in detail.

6.3.1 Review of Sparse Arrays

Definition 1 (Difference Coarray): For an antenna array with N elements, $\mathbf{w}_n$ is the position of its n-th element, $n = 1, 2, \ldots, N$. Let $\mathbf{w}_n \in \mathbb{C}^{3\times 1}$ denote the 3-D coordinate of the n-th antenna array element. The locations of all array elements are collected in the set $\mathbb{D}_{\rm a}$, i.e., $\mathbb{D}_{\rm a} = \{\mathbf{w}_n\}$. The difference coarray of the antenna array is an (virtual) array with element positions given by the set $\mathbb{D}_{\rm dc}$:

$$\mathbb{D}_{\rm dc} = \{\mathbf{w}_{n_1} - \mathbf{w}_{n_2}\}, \ \forall n_1, n_2 = 1, 2, \ldots, N. \tag{6.4}$$

According to (6.4), the element positions of the difference coarray are the *(self) differences* between the locations of original physical antenna elements.

[2]Much smaller is defined by $|(f \pm B)/f| \approx 1$. When $f = 60$ GHz and $B \leq 2$ GHz, it has $|(f \pm B)/f| \in [0.97, 1.03]$.

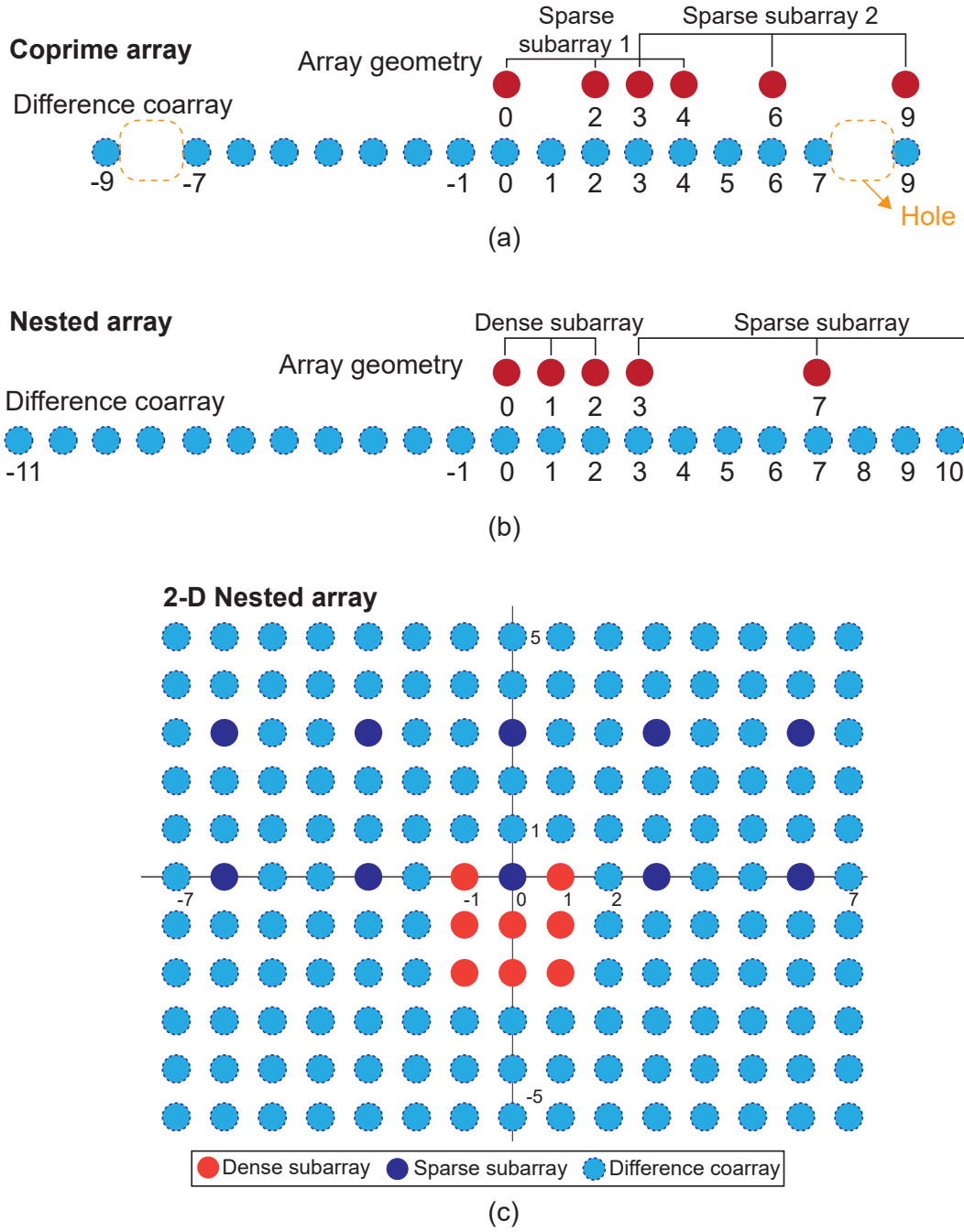


Figure 6.2 : An example of (a) coprime array, which is composed of two sparse subarrays: one with $N = 3$ elements and separation $M = 2$, and another one with $2M - 1$ elements and separation N; (b) nested array, which is composed of a dense subarray with $N_1 = 3$ elements and separation 1, and a sparse subarray with $N_2 = 3$ elements and separation $N_1 + 1$; and (c) 2-D nested array, which is composed of a 3×3 dense subarray and a 5×2 sparse subarray.

Based on the definition of difference coarray, we can define a cross difference coarray, which corresponds to the cross differences between the element locations of two arrays with N and M elements:

$$\mathbb{D}_{\text{cdc}} = \pm \{\mathbf{w}_n - \mathbf{w}_m\}, \ \forall n = 1, 2, \ldots, N, \ m = 1, 2, \ldots, M. \tag{6.5}$$

According to (6.4) and (6.5), we can see that the concept of difference coarray arises naturally in the second-order statistics of the impinging signals. For example, we consider that a signal $\mathbf{x}_{\mathbf{w}_n} \in \mathbb{C}^{N\times 1}$ is received at the n-th element of an antenna array. The cross-correlation between the signals received at the n_1-th and n_2-th elements of the array is given by

$$\mathrm{E}\left\{\mathbf{x}_{\mathbf{w}_{n_1}} \mathbf{x}^H_{\mathbf{w}_{n_2}}\right\} = \mathbf{R}_{(\mathbf{w}_{n_1} - \mathbf{w}_{n_2})} \in \mathbb{C}^{N\times N}, \ \mathbf{w}_{n_1}, \mathbf{w}_{n_2} \in \mathbb{D}_{\text{a}} \tag{6.6}$$

where $\mathbf{R}_{(\mathbf{w}_{n_1} - \mathbf{w}_{n_2})}$ can be viewed as a signal sample received by a (larger) difference coarray with virtual array elements located at $(\mathbf{w}_{n_1} - \mathbf{w}_{n_2}) \in \mathbb{D}_{\text{dc}}$ [87].

By adequately designing the element locations, i.e., $\mathbb{D}_{\text{a}}$, we can increase the number of virtual elements in the difference coarray after computing the autocorrelation. If we use the samples from the difference coarray to perform spectral estimation, the parameters of much more targets can be estimated.

We proceed to introduce the concept of sparse array. An array is said to be sparse if the spacing between a majority or all of adjacent elements is more than one (half-wavelength) [87, 90]. By applying the concept of sparse array to antenna design, we can significantly improve the number of distinguishable targets using a small number of physical antenna elements [123]. Some well-known 1-D sparse arrays include MRA [88], MHA [89], nested arrays [87], and coprime arrays [90]. With $O(N)$ physical array elements, both MRA and MHA can construct difference coarrays with the size of $O(N^2)$. However, their geometries need to be constructed by using searching algorithms, e.g., integer programming [89,124]. Nested and coprime

arrays were proposed in [87, 90] with closed-form expressions for element locations, and both of them can construct difference coarrays with the same DoF as MRA and MHA. An example of nested and coprime arrays, and their difference coarrays are shown in Figs. 6.2(a) and 6.2(b). Nested arrays can offer larger difference coarray than coprime arrays, as shown in Fig. 6.2(b), where both of them have six physical elements. In addition, the difference coarrays of nested arrays consist of evenly spaced virtual elements with no holes, so that the subspace-based estimation algorithms, such as MUSIC and ESPRIT, can be utilized on the coarray domain without creating ambiguities [124]. For the details of these arrays, interested readers can refer to [87–90].

In the next subsection, we design the phase shifter matrix $\mathbf{B}_{\text{ps}}$. $\mathbf{B}_{\text{ps}}$ can transform the UCyA steering vectors from the element space into a phase space, where the phases of the array steering vectors are linear to the element locations. From (6.6), we see that if we want to construct a difference coarray with a similar geometry to that of the original array, e.g., the UCyA in our system, the phase of the array steering vectors should vary linearly with the element locations. However, due to the special geometry of the UCyA, if we directly calculate the cross correlation of the array steering vectors, it would generate a virtual non-UCyA composed of multiple non-UCAs [92], leading to an increased computational complexity and degraded the estimation accuracy.

6.3.2 Phase-Space Transformation

The analog beamforming matrix $\mathbf{B}_{\text{rf}} = \mathbf{B}_{\text{ps}}\mathbf{B}_{\text{rfc}} \in \mathbb{C}^{M_{\text{bs}} \times M_{\text{rf}}}$ is composed of a phase shifter matrix $\mathbf{B}_{\text{ps}} \in \mathbb{C}^{M_{\text{bs}} \times M_{\text{bsr}}}$ and an RF-chain connection matrix $\mathbf{B}_{\text{rfc}} \in \mathbb{C}^{M_{\text{bsr}} \times M_{\text{rf}}}$, where M_{bsr} is the number of output ports of the phase-shifter matrix. An illustration of the RF front-end structure is shown in Fig. 6.3. Here, we design $\mathbf{B}_{\text{ps}}$ based on circular phase-space transformation [110], to transform the nonlinear

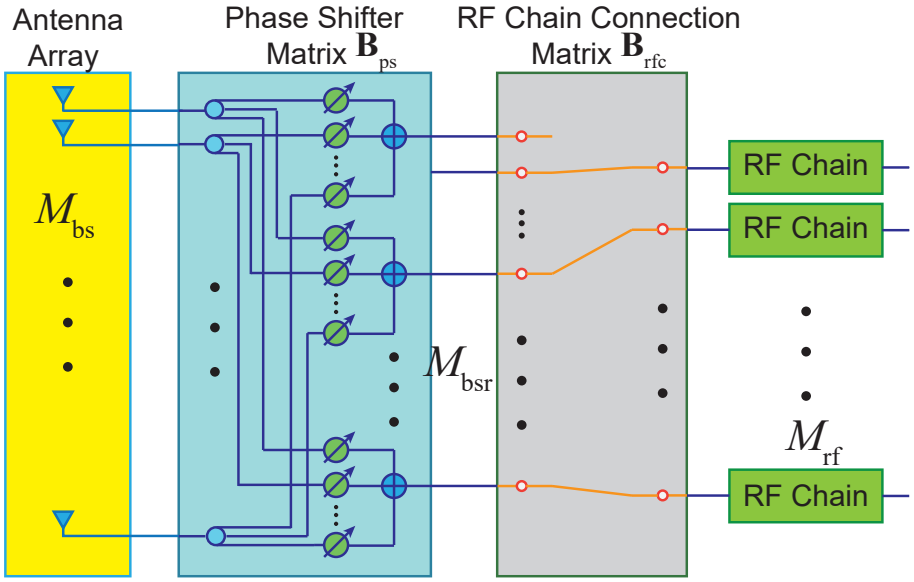


Figure 6.3 : The block diagram of RF front-end structure.

phase of UCyA steering vectors to be linear to the element locations.

We decouple $\mathbf{B}_{\text{ps}}$ between the vertical and horizontal planes, i.e., $\mathbf{B}_{\text{ps}} = \mathbf{B}_{\text{vps}} \otimes \mathbf{B}_{\text{hps}}$ with $\mathbf{B}_{\text{vps}} \in \mathbb{C}^{M_{\text{v}} \times M_{\text{vr}}}$ and $\mathbf{B}_{\text{hps}} \in \mathbb{C}^{M_{\text{h}} \times M_{\text{hr}}}$, and thus $M_{\text{bsr}} = M_{\text{vr}} M_{\text{hr}}$, where M_{vr} and M_{hr} are the number of the phase-shifter output ports along the vertical and horizontal directions, respectively. According to the phase-space transformation of UCAs [110], we design $\mathbf{B}_{\text{hps}}$ as $[\mathbf{B}_{\text{hps}}]_{m_{\text{h}}, m_{\text{hr}}+P+1} = e^{-j\frac{2\pi(m_{\text{h}}-1)}{M_{\text{h}}} m_{\text{hr}}}$, where $M_{\text{hr}} = 2P+1$, $m_{\text{hr}} = -P, -P+1, \ldots, P$, and P is the highest phase-space dimension. Thus, the M_{h}-dimensional array steering vector $\mathbf{a}_{\text{h}}(\theta_k, \phi_k)$ can be transformed into a $(2P+1)$-dimensional phase space, i.e., $\mathbf{a}_{\text{hps}}(\theta_k, \phi_k) = \mathbf{B}_{\text{hps}}^H \mathbf{a}_{\text{h}}(\theta_k, \phi_k) \in \mathbb{C}^{(2P+1)\times 1}$. The value of the highest phase-space dimension, P, can be configured based on the following theorem.

Theorem 4. *Suppose that $M_h \geq \lfloor 4\pi r/\lambda \rfloor$. If the highest phase-space dimension, P, is larger than $\lfloor 2\pi r/\lambda \rfloor$ and smaller than $M_{hr}/2$, then the elements in the phase-space response can be approximated by*

$$a_{hps,p}(\theta_k, \phi_k) \approx \sqrt{M_h} j^p J_p\left(\gamma(\theta_k)\right) \exp\left(-jp\phi_k\right), \tag{6.7}$$

where $\gamma(\theta_k) = 2\pi r \sin(\theta_k)/\lambda$, $p = -P, -P+1, \ldots, P$, *and* $J_p(\gamma(\theta_k))$ *is the Bessel function of the first kind of order* p.

Proof. See Appendix E.1. □

We set $\mathbf{B}_{\text{vps}} = \mathbf{I}_{M_\text{v}}$ to preserve the recurrence relations among UCAs. According to Theorem 4, the array steering vectors $\mathbf{a}_{\text{bs}}(\theta_k, \phi_k)$ after the hybrid beamformer is given by

$$\begin{aligned}
\mathbf{a}_{\text{bd}}(\theta_k, \phi_k) &= \mathbf{B}^H \mathbf{a}_{\text{bs}}(\theta_k, \phi_k) \\
&= (\mathbf{B}_{\text{rf}}\mathbf{B}_{\text{bb}})^H \mathbf{a}_{\text{bs}}(\theta_k, \phi_k) \\
&= ((\mathbf{B}_{\text{vps}} \otimes \mathbf{B}_{\text{hps}})\mathbf{B}_{\text{rfc}}\mathbf{I}_{M_{\text{bsr}}})^H \mathbf{a}_{\text{bs}}(\theta_k, \phi_k) \\
&= \mathbf{B}_{\text{rfc}}^H (\mathbf{B}_{\text{vps}} \otimes \mathbf{B}_{\text{hps}})^H \mathbf{a}_{\text{bs}}(\theta_k, \phi_k),
\end{aligned} \tag{6.8}$$

where $\mathbf{B}_{\text{bb}}$ is a diagonal matrix used to guarantee the power constraint [125]. Without loss of generality, we set $\mathbf{B}_{\text{bb}} = \mathbf{I}_{M_{\text{bsr}}}$ in this chapter. According to two properties of the Khatri-Rao product: $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$ and $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ [119], (6.8) can be rewritten as

$$\begin{aligned}
\mathbf{a}_{\text{bd}}(\theta_k, \phi_k) &= \mathbf{B}_{\text{rfc}}^H \left(\mathbf{B}_{\text{vps}}^H \otimes \mathbf{B}_{\text{hps}}^H\right) (\mathbf{a}_\text{v}(\theta_k) \otimes \mathbf{a}_\text{h}(\theta_k, \phi_k)) \\
&= \mathbf{B}_{\text{rfc}}^H \left[\left(\mathbf{B}_{\text{vps}}^H \mathbf{a}_\text{v}(\theta_k)\right) \otimes \left(\mathbf{B}_{\text{hps}}^H \mathbf{a}_\text{h}(\theta_k, \phi_k)\right)\right] \\
&= \mathbf{B}_{\text{rfc}}^H \left[\mathbf{a}_{\text{vps}}(\theta_k) \otimes \mathbf{a}_{\text{hps}}(\theta_k, \phi_k)\right],
\end{aligned} \tag{6.9}$$

where $\mathbf{a}_{\text{vps}}(\theta_k) \in \mathbb{C}^{M_{\text{vr}} \times 1}$, $\mathbf{a}_{\text{hps}}(\theta_k, \phi_k) \in \mathbb{C}^{M_{\text{hr}} \times 1}$, $M_{\text{vr}} = M_\text{v}$, and $M_{\text{hr}} = 2P + 1$.

According to Theorem 4, we have

$$\mathbf{a}_{\text{vps}}(\theta_k) = \mathbf{a}_\text{v}(\theta_k) = \mathbf{I}_{M_\text{v}} \mathbf{a}_\text{v}(\theta_k), \tag{6.10}$$

$$a_{\text{hps},m_{\text{hr}}}(\theta_k, \phi_k) = \left[\mathbf{B}_{\text{hps}}^H\right]_{m_{\text{hr}}+P+1,:} \mathbf{a}_\text{h}(\theta_k, \phi_k \approx \sqrt{M_\text{h}} j^{m_{\text{hr}}} J_{m_{\text{hr}}}(\gamma(\theta_k)) \exp(-j m_{\text{hr}} \phi_k). \tag{6.11}$$

From (6.10) and (6.11), we see that, through the proposed $\mathbf{B}_{\text{ps}}$, the phases of the array steering vectors become linear to the element locations. This is important to exploit the property of the sparse array theory to design the RF-chain connection matrix $\mathbf{B}_{\text{rfc}}$.

6.3.3 RF-Chain Connection Network Design

In this subsection, we apply the sparse array technique to design $\mathbf{B}_{\text{rfc}}$, which enables the DoAs of a large number of devices to be estimated with a marginal accuracy loss while significantly reducing the number of required RF chains. We aim to use as few RF chains as possible to achieve the same, or even larger, DoF than the fully connected beamforming array[3]. This objective is different from the previous sparse array researches, which have typically focused on maximizing the size of difference coarrays under the constraint of a fixed number of physical antenna elements.

We first flatten the 3-D RF-chain connection network of UCyA into a 2-D plane, as shown in Fig. 6.4, by disjoining the RF-chain connection network at the first column phase shifters of every UCA. Different from typical 2-D arrays, due to the periodicity of UCAs, the first and the last phase-shifter output ports of every row in the flattened 2-D RF-chain connection network are identical, as shown in Fig. 6.4, where the dotted circles denote the last-column phase shifters.

After the flattening processing, the 3-D RF-chain connection network becomes a quasi-2-D rectangular array with size of $(M_{\text{hr}}+1)\times M_{\text{vr}}$, where the increased dimension is due to the repeated phase-shifter output ports, as shown in Fig. 6.4. The idea of 2-D sparse arrays can be applied to design a 3-D RF-chain connection network of

[3]Due to the use of phase shifter network, the antenna DoF of UCyA depends on the scale of the phase shifter network. Thus, if we use a fully connected beamforming array, $(2P+1)M_{\text{v}}$ RF chains are needed, which can provide $\mathcal{O}(PM_{\text{v}})$ DoFs.

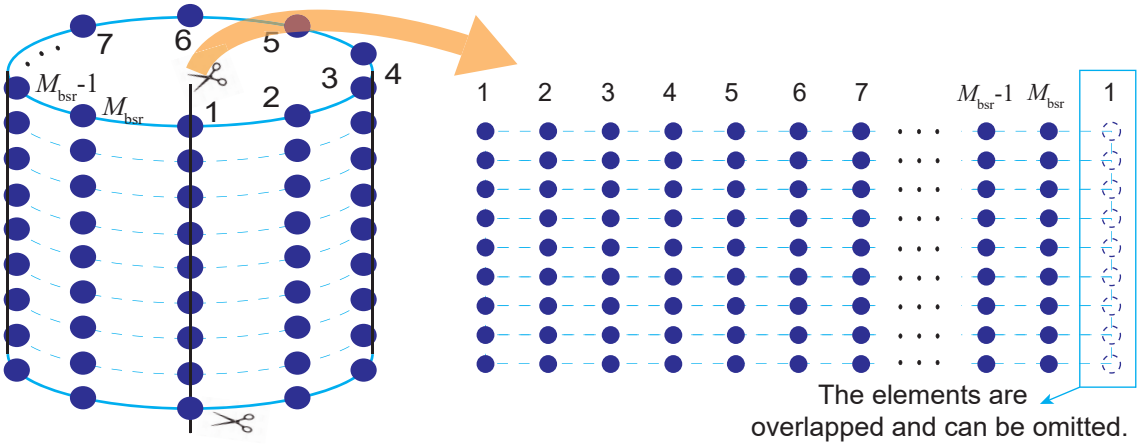


Figure 6.4 : An illustration of unfolding phase-shifter output ports of a 3-D UCyA to be a 2-D array.

UCyA[4]. In this chapter, we design the RF-chain connection network based on a 2-D nested array. This is because (1) a nested array can generate larger hole-free difference coarrays than a coprime array under the same setting, as discussed in Section 6.3.1; and (2) it has simple closed-form expressions for a large number of elements, which cannot be achieved in MRA and MHA. There are also some other frequently-used 2-D sparse arrays, e.g., hourglass arrays and open box arrays (OBAs) [124,126]. We will compare the RF-chain connection networks designed based on those array geometries with our design in Section 6.6.

Our proposed sparse RF-chain connection network is developed from the "Configuration II" nested array [48]. In the general "Configuration II" nested array, when there are $N_{\text{dense}} = N_{\text{vd}}N_{\text{hd}} - 1$ and $N_{\text{sparse}} = N_{\text{vs}}N_{\text{hs}}$ elements in the dense and sparse subarrays, respectively, the constructed hole-free difference coarray has $N_{\text{dc}} = N_{\text{vdc}}N_{\text{hdc}} = (2N_{\text{vd}}N_{\text{vs}}-1)N_{\text{hd}}N_{\text{hs}}$ elements [48], as shown in Fig. 6.2(c). Here, N_{vd} and N_{hd} are the numbers of elements in the dense subarray along the vertical

[4]Although the RF-chain connection network actually does not have the exact shape, according to the array steering vectors in (6.10) and (6.11), we can also regard the network as an UCyA.

and horizontal directions, respectively; N_{vs} and N_{hs} are the numbers of elements in the sparse subarray along the vertical and horizontal directions, respectively; and N_{vdc} and N_{hdc} are the numbers of elements in the difference coarray along the vertical and horizontal directions, respectively. We wish to find the distribution of the RF chains between the sparse and the dense arrays that use as few RF chains as possible to achieve the same DoF as the fully connected beamforming array.

Due to the above-mentioned periodicity of UCAs, when we apply the sparse array technique into our hybrid front-end design, two cases need to be considered for UCyAs. Fig. 6.5(a) shows the first case, where the first and the last columns of the difference coarray overlap. From Fig. 6.5(a), we can see that because one column (first or last) of the difference coarray is redundant, two elements of the sparse subarray can be omitted to reduce the element number. A drawback is that this case requires $M_{\text{hr}} = (N_{\text{hs}} - 1)N_{\text{hd}}$, which would impose a strict requirement on the number of phase shifters on the horizontal plane. The second case is shown in Fig. 6.5(b), which requires the constructed difference coarray to be larger than the original UCyA, to achieve the same DoF as a fully connected beamforming array on the horizontal space. In this case, $N_{\text{hd}}N_{\text{hs}} \geq M_{\text{hr}} > (N_{\text{hs}} - 1)N_{\text{hd}}/2$.

In our system, due to the new phase shifter network designed in Section 6.3.2, we have $M_{\text{hr}} = 2P + 1$ (which is an odd number). Since in the Configuration II nested array, the dense and sparse subarrays are symmetric, i.e., both N_{hs} and N_{hd} are odd, we have $(N_{\text{hs}} - 1)N_{\text{hd}}$ is even, and only the above-mentioned second case

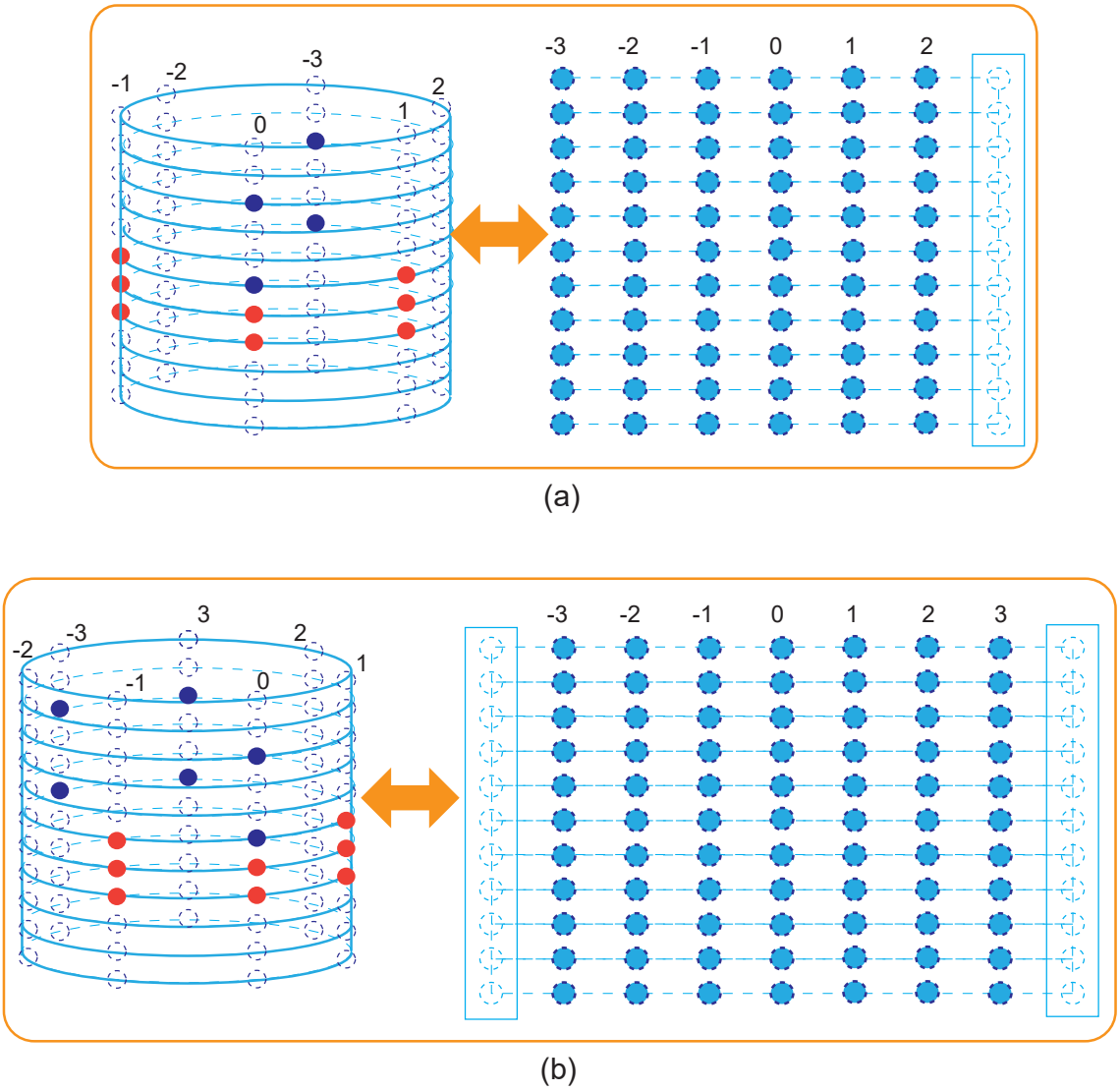


Figure 6.5 : Two cases of 3-D UCyA unfolding. The locations of the cylindrical post-phase-shifting ports in sparse and dense subarrays are highlighted with blue and red dots.

needs to be considered in our system. We formulate the optimization problem as

$$\min_{N_{\rm vd},N_{\rm hd},N_{\rm vs},N_{\rm hs}\in\mathbb{Z}^+} M_{\rm rf} = N_{\rm vd}N_{\rm hd} + N_{\rm vs}N_{\rm hs} - 1 \tag{6.12}$$

$$\text{s.t. (C1): } 2N_{\rm vd}N_{\rm vs} - 1 \geq M_{\rm vr},$$

$$\text{(C2): } N_{\rm hd}N_{\rm hs} \geq M_{\rm hr},$$

$$\text{(C3): } N_{\rm vd}, N_{\rm hd} > 1,$$

$$\text{(C4): } N_{\rm hd} \text{ is odd and } N_{\rm vd}/N_{\rm hd} \in \mathbb{Z}^+.$$

(C1) and (C2) guarantee the DoF requirements for the constructed difference coarray of the RF-chain connection network. (C3) avoids solutions that degenerate to 1-D arrays. (C4) is due to the fact that the dense array in Configuration II is symmetrical, and $N_{\rm hd}$ and $N_{\rm vd}$ are invariant factors [93] of the array distribution matrix.

The solution for (6.12) can be obtained by adopting the following strategy. According to (C1) and (C2), since $N_{\rm vs}, N_{\rm hs} \in \mathbb{Z}^+$, we can obtain $N_{\rm vs} = \lceil (M_{\rm vr}-1)/2N_{\rm vd} \rceil$ and $N_{\rm hs} = \lceil M_{\rm hr}/N_{\rm hd} \rceil$. The optimization problem (6.12) becomes

$$\min_{N_{\rm vd},N_{\rm hd}\in\mathbb{Z}^+} M_{\rm rf} = N_{\rm vd}N_{\rm hd} + \left\lceil \frac{(M_{\rm vr}-1)/2}{N_{\rm vd}} \right\rceil \left\lceil \frac{M_{\rm hr}}{N_{\rm hd}} \right\rceil \tag{6.13}$$

$$\text{s.t. (C3) and (C4)}$$

Given $M_{\rm vr}$ and $M_{\rm hr}$, we see that (6.13) exhibits the form of $y = x + \frac{a}{x}$, where $a > 0$ is a constant and $y = x + \frac{a}{x} \geq 2\sqrt{a}$. Because $y = 2\sqrt{a}$ iff $x = \frac{a}{x}$, the minimum $M_{\rm rf}$ can be obtained when the difference between $N_{\rm vd}N_{\rm hd}$ and $\lceil (M_{\rm vr}-1)/2N_{\rm vd} \rceil \lceil M_{\rm hr}/N_{\rm hd} \rceil$ is the smallest. Since $N_{\rm vd}, N_{\rm hd}, N_{\rm vs}, N_{\rm hs} \in \mathbb{Z}^+$, we can determine the approximate value ranges of $N_{\rm vd}N_{\rm hd}$ and $\lceil (M_{\rm vr}-1)/2N_{\rm vd} \rceil \lceil M_{\rm hr}/N_{\rm hd} \rceil$, and (6.13) is an integer programming problem. According to (C3) and (C4), the optimal solutions of $N_{\rm vd}, N_{\rm hd}, N_{\rm vs}$, and $N_{\rm hs}$ to (6.13) can be obtained by using brute-force search with the value range between $N_{\rm vd}N_{\rm hd}$ and $\lceil (M_{\rm vr}-1)/2N_{\rm vd} \rceil \lceil M_{\rm hr}/N_{\rm hd} \rceil$.

In the proposed sparse RF-chain connection network, the RF chains only need to connect the phase shifters located in the dense and sparse subarrays. Based on the calculated values of $N_{\rm vd}, N_{\rm hd}, N_{\rm vs}$, and $N_{\rm hs}$, now we provide the element locations in the dense and sparse arrays. For illustration convenience, we define the overlapping point of the sparse and dense arrays as the origin of the nested array[5]. Let $\mathbf{m}_{\rm sp} = (m_{\rm v\_sp}, m_{\rm h\_sp})$ and $\mathbf{m}_{\rm de} = (m_{\rm v\_de}, m_{\rm h\_de})$ as the locations of elements in the dense and sparse arrays (to which the RF chains connect), respectively. We have

$$\begin{cases} m_{\rm v\_sp} = N_{\rm vd}(n_{\rm vs} - 1), \\ m_{\rm h\_sp} = N_{\rm hd}(-N_{\rm hs}/2 + n_{\rm hs} - 1/2), \end{cases} \tag{6.14}$$

$$\begin{cases} m_{\rm v\_de} = -N_{\rm vd} + n_{\rm vd}, \\ m_{\rm h\_de} = -(N_{\rm hd} - 1)/2 + n_{\rm hd} - 1, \end{cases} \tag{6.15}$$

where $n_{\rm vs} = 1, 2, \ldots, N_{\rm vs}$; $n_{\rm hs} = 1, 2, \ldots, N_{\rm hs}$; $n_{\rm vd} = 1, 2, \ldots, N_{\rm vd}$; and $n_{\rm hd} = 1, 2, \ldots, N_{\rm hd}$. Let $\mathbb{M}_{\rm rf\_d} = \{\mathbf{m}_{\rm v\_de} \otimes \mathbf{m}_{\rm h\_de}\}$ and $\mathbb{M}_{\rm rf\_s} = \{\mathbf{m}_{\rm v\_sp} \otimes \mathbf{m}_{\rm h\_sp}\}$ denote the sets of the RF-chain connection points in the dense and sparse arrays, respectively. $\mathbf{m}_{\rm v\_de} \in \mathbb{C}^{N_{\rm vd} \times 1}$ and $\mathbf{m}_{\rm h\_de} \in \mathbb{C}^{N_{\rm hd} \times 1}$ are the element locations of the dense array along the vertical and horizontal directions, respectively. $\mathbf{m}_{\rm v\_sp} \in \mathbb{C}^{N_{\rm vs} \times 1}$ and $\mathbf{m}_{\rm h\_sp} \in \mathbb{C}^{N_{\rm hs} \times 1}$ are the element locations of the sparse array along the vertical and horizontal directions, respectively. The set of RF-chain connection points is $\mathbb{M}_{\rm rfc} = \{\mathbb{M}_{\rm rf\_d} \cup \mathbb{M}_{\rm rf\_s}\}$. The constructed RF-chain connection matrix $\mathbf{B}_{\rm rfc}$ is given by

$$[\mathbf{B}_{\rm rfc}]_{m_{\rm bsr}, m_{\rm rf}} = \begin{cases} 1, & \text{if } m_{\rm bsr} \in \mathbb{M}_{\rm rfc} \text{ and} \\ & [\mathbf{B}_{\rm rfc}]_{m'_{\rm bsr} \neq m_{\rm bsr}, m_{\rm rf}} = [\mathbf{B}_{\rm rfc}]_{m_{\rm bsr}, m'_{\rm rf} \neq m_{\rm rf}} = 0; \\ 0, & \text{otherwise.} \end{cases} \tag{6.16}$$

[5]Because the parameter estimation depends on the difference between array elements, the changed absolute positions of array elements does not effect the estimation performance.

By deploying the proposed sparse RF-chain connection network and using the second-order statistics of the received signal for channel estimation, the DoF of the proposed network is $O((2N_{\rm vd}N_{\rm vs}-1)\times M_{\rm hr})$. In other words, according to [93], we can estimate the channel parameters of $(N_{\rm vdc}-1)(N_{\rm hdc}-1)$ devices by using the proposed hybrid front-end. Because there is an overlapping point at the origin, the total number of RF chains required in our network is $M_{\rm rf}=N_{\rm vd}N_{\rm hd}+N_{\rm vs}N_{\rm hs}-1$, as shown in (6.12)[6]. Due to the periodicity of UCAs, there are only up to $N_{\rm hdc}=M_{\rm hr}$ virtual elements on the horizontal plane of the constructed 3-D difference coarray of the RF-chain connection network. Along the vertical direction, the number of virtual elements is $N_{\rm vdc}=2N_{\rm vd}N_{\rm vs}-1$.

According to the element locations of the dense and sparse arrays in (6.14) and (6.15), we also provide the element locations of the constructed difference coarray. Let $\mathbf{m}_{\rm dc}=(m_{\rm v\_dc},m_{\rm h\_dc})$, where $m_{\rm v\_dc}$ and $m_{\rm h\_dc}$ correspond to the locations along the vertical and horizontal directions, respectively. We have

$$m_{\rm v\_dc}=-(N_{\rm vdc}-1)/2+n_{\rm vdc}-1=-N_{\rm vd}N_{\rm vs}+n_{\rm vdc}, \tag{6.17}$$

$$m_{\rm h\_dc}=-(N_{\rm hdc}-1)/2+n_{\rm hdc}-1=-P+n_{\rm hdc}-1, \tag{6.18}$$

where $n_{\rm vdc}=1,2,\ldots,N_{\rm vdc}$ and $n_{\rm hdc}=1,2,\ldots,N_{\rm hdc}$. The shape of the constructed 3-D difference coarray RF-chain connection network is the same as the UCyA, but the former has a larger DoF.

The signals through the proposed RF-chain connection network are given by

$$\mathbf{x}_{{\rm sn},m_{\rm t}}=\sum_{k=1}^{K}s_{m_{\rm t},k}\mathbf{a}_{\rm sn}(\phi_k,\theta_k)+\mathbf{n}_{{\rm sn},m_{\rm t}}, \tag{6.19}$$

[6]It can also be proved that the phase shifter at this location is useless and does not need to be connected [48]. However, for ease of description, here we assume that this phase shifter is connected in our network, which does not affect the results in this chapter.

where

$$\mathbf{a}_{\mathrm{sn}}(\phi_k,\theta_k)=\begin{bmatrix}\mathbf{a}_{\mathrm{sn,s}}(\phi_k,\theta_k)\\ \mathbf{a}_{\mathrm{sn,d}}(\phi_k,\theta_k)\end{bmatrix}=\begin{bmatrix}\mathbf{a}_{\mathrm{sn,sv}}(\theta_k)\otimes\mathbf{a}_{\mathrm{sn,sh}}(\theta_k,\phi_k)\\ \mathbf{a}_{\mathrm{sn,dv}}(\theta_k)\otimes\mathbf{a}_{\mathrm{sn,dh}}(\theta_k,\phi_k)\end{bmatrix}. \tag{6.20}$$

The elements of $\mathbf{a}_{\mathrm{sn,sv}}(\theta_k)\in\mathbb{C}^{N_{\mathrm{vs}}\times 1}$ and $\mathbf{a}_{\mathrm{sn,sh}}(\theta_k,\phi_k)\in\mathbb{C}^{N_{\mathrm{hs}}\times 1}$ are $a_{\mathrm{sn,sv},n_{\mathrm{vs}}}(\theta_k)=a_{\mathrm{vs},m_{\mathrm{v\_sp}}}(\theta_k)$ and $a_{\mathrm{sn,sh},n_{\mathrm{hs}}}(\theta_k,\phi_k)=a_{\mathrm{hs},m_{\mathrm{h\_sp}}}(\theta_k,\phi_k)$, respectively, where $a_{\mathrm{vs},m_{\mathrm{v\_sp}}}(\theta_k)$ and $a_{\mathrm{hs},m_{\mathrm{h\_sp}}}(\theta_k,\phi_k)$ are the array steering vectors of the sparse subarray along the vertical and horizontal directions, respectively. The elements of the array steering vectors of the dense subarray, i.e., $\mathbf{a}_{\mathrm{sn,dv}}(\theta_k)\in\mathbb{C}^{N_{\mathrm{vd}}\times 1}$ and $\mathbf{a}_{\mathrm{sn,dh}}(\theta_k,\phi_k)\in\mathbb{C}^{N_{\mathrm{hd}}\times 1}$, can be written in the same way. Here, $\mathbf{n}_{\mathrm{sn},m_{\mathrm{t}}}\in\mathbb{C}^{M_{\mathrm{rf}}\times 1}$ is the noise component through the RF-chain connection network.

The signal model (6.19) can also be rewritten as

$$\mathbf{x}_{\mathrm{sn},m_{\mathrm{t}}}=\mathbf{A}_{\mathrm{sn}}\mathbf{s}_{m_{\mathrm{t}}}+\mathbf{n}_{\mathrm{sn},m_{\mathrm{t}}}, \tag{6.21}$$

where $\mathbf{A}_{\mathrm{sn}}=[\mathbf{a}_{\mathrm{sn}}(\phi_1,\theta_1),\mathbf{a}_{\mathrm{sn}}(\phi_2,\theta_2),\ldots,\mathbf{a}_{\mathrm{sn}}(\phi_K,\theta_K)]\in\mathbb{C}^{M_{\mathrm{rf}}\times K}$ and $\mathbf{s}_{m_{\mathrm{t}}}=[s_{m_{\mathrm{t}},1},s_{m_{\mathrm{t}},2},\ldots,s_{m_{\mathrm{t}},K}]^T\in\mathbb{C}^{K\times 1}$.

By calculating the autocorrelation of $\mathbf{x}_{\mathrm{sn},m_{\mathrm{t}}}$, we have

$$\mathbf{R}_{\mathrm{sn},m_{\mathrm{t}}}=\mathrm{E}\left\{\mathbf{x}_{\mathrm{sn},m_{\mathrm{t}}}\mathbf{x}_{\mathrm{sn},m_{\mathrm{t}}}^H\right\}=\mathbf{A}_{\mathrm{sn}}\mathbf{R}_{\mathrm{ss},m_{\mathrm{t}}}\mathbf{A}_{\mathrm{sn}}^H+\mathbf{R}_{\mathrm{nn},m_{\mathrm{t}}}, \tag{6.22}$$

where $\mathbf{R}_{\mathrm{ss},m_{\mathrm{t}}}=\mathrm{diag}\left(\sigma^2_{\mathrm{s},m_{\mathrm{t}},1},\ldots,\sigma^2_{\mathrm{s},m_{\mathrm{t}},K}\right)$ and $\mathbf{R}_{\mathrm{nn},m_{\mathrm{t}}}=\mathrm{diag}\left(\sigma^2_{\mathrm{n},m_{\mathrm{t}},1},\ldots,\sigma^2_{\mathrm{n},m_{\mathrm{t}},K}\right)$ are the autocorrelation matrices of $\mathbf{s}_{m_{\mathrm{t}}}$ and $\mathbf{n}_{\mathrm{sn},m_{\mathrm{t}}}$, respectively.

We vectorize $\mathbf{R}_{\mathrm{sn},m_{\mathrm{t}}}$ as

$$\mathbf{y}_{\mathrm{vR},m_{\mathrm{t}}}=\mathrm{vec}(\mathbf{R}_{\mathrm{sn},m_{\mathrm{t}}})=[\mathbf{A}_{\mathrm{sn}}^*\diamond\mathbf{A}_{\mathrm{sn}}]\,\mathbf{d}_{m_{\mathrm{t}}}+\mathrm{vec}(\mathbf{R}_{\mathrm{nn},m_{\mathrm{t}}}), \tag{6.23}$$

where $[\mathbf{d}_{m_{\mathrm{t}}}]_{k,1}=\sigma^2_{\mathrm{s},m_{\mathrm{t}},k}$, and $\sigma^2_{\mathrm{s},m_{\mathrm{t}},k}$ is the power of the k-th signal. The k-th column of the matrix $[\mathbf{A}_{\mathrm{sn}}^*\diamond\mathbf{A}_{\mathrm{sn}}]$ contains elements representing the cross-differences between sparse and dense subarrays, i.e., $a^*_{\mathrm{sn,s},\mathbf{m}_{\mathrm{sp}}}(\phi_k,\theta_k)a_{\mathrm{sn,d},\mathbf{m}_{\mathrm{de}}}(\phi_k,\theta_k)$ and $a^*_{\mathrm{sn,s},\mathbf{m}_{\mathrm{de}}}(\phi_k,\theta_k)a_{\mathrm{sn,d},\mathbf{m}_{\mathrm{sp}}}(\phi_k,\theta_k)$, and the self-differences of sparse and dense subarrays, i.e., $a^*_{\mathrm{sn,s},\mathbf{m}_{\mathrm{sp},1}}$

$(\phi_k, \theta_k) a_{\text{sn,s},\mathbf{m}_{\text{sp},2}}(\phi_k, \theta_k)$ and $a^*_{\text{sn,d},\mathbf{m}_{\text{de},1}}(\phi_k, \theta_k) a_{\text{sn,d},\mathbf{m}_{\text{de},2}}(\phi_k, \theta_k)$. Here, $\mathbf{m}_{\text{sp},1}$ and $\mathbf{m}_{\text{sp},2}$ denote that $a_{\text{sn,s},\mathbf{m}_{\text{sp},1}}(\phi_k, \theta_k)$ and $a_{\text{sn,s},\mathbf{m}_{\text{sp},2}}(\phi_k, \theta_k)$ are different elements in the sparse subarray, and $\mathbf{m}_{\text{de},1}$ and $\mathbf{m}_{\text{de},2}$ denote that $a_{\text{sn,d},\mathbf{m}_{\text{de},1}}(\phi_k, \theta_k)$ and $a_{\text{sn,d},\mathbf{m}_{\text{de},2}}(\phi_k, \theta_k)$ are different elements in the dense subarray.

We sort the rows of $\mathbf{y}_{\text{vR},m_\text{t}}$ in the ascending order of their phases, and then remove the redundant rows with the same phases. Then, we can obtain the array steering vector of the difference coarray $\mathbf{A}_{\text{df}} \in \mathbb{C}^{N_{\text{vdc}} N_{\text{hdc}} \times 1}$ from $[\mathbf{A}^*_{\text{sn}} \diamond \mathbf{A}_{\text{sn}}]$. We also calculate and store the mean of the "nonzero" rows of $\text{vec}(\mathbf{R}_{\text{nn},m_\text{t}})$[7], and obtain

$$\mathbf{y}_{\text{df},m_\text{t}} = \mathbf{A}_{\text{df}} \mathbf{d}_{m_\text{t}} + \sigma^2_{\text{n}} \mathbf{e}_{\text{df}}, \tag{6.24}$$

where $[\mathbf{A}_{\text{df}}]_{:,k} = \mathbf{a}_{\text{df}}(\phi_k, \theta_k) = \mathbf{a}_{\text{dfv}}(\theta_k) \otimes \mathbf{a}_{\text{dfh}}(\theta_k, \phi_k)$ and $\mathbf{e}_{\text{df}} \in \mathbb{C}^{N_{\text{vdc}} N_{\text{hdc}} \times 1}$ is a vector of all zeros except a "1" at the $(P+1)N_{\text{vd}} N_{\text{vs}}$-th entry. The elements of $\mathbf{a}_{\text{dfv}}(\theta_k) \in \mathbb{C}^{N_{\text{vdc}} \times 1}$ and $\mathbf{a}_{\text{dfh}}(\theta_k, \phi_k) \in \mathbb{C}^{N_{\text{hdc}} \times 1}$ are given by

$$a_{\text{dfv},n_{\text{vdc}}}(\theta_k) = \frac{1}{M_\text{v}} \exp\left(-j \frac{2\pi}{\lambda} h m_{\text{v\_dc}} \cos(\theta_k)\right), \tag{6.25}$$

$$a_{\text{dfh},n_{\text{hdc}}}(\theta_k, \phi_k) = \xi_{m_{\text{h\_dc}}}(\theta_k) \exp\left(-j m_{\text{h\_dc}} \phi_k\right), \tag{6.26}$$

where $\xi_{m_{\text{h\_dc}}}(\theta_k) = M_\text{h} j^{m_{\text{h\_dc}}} J_{m_{\text{h\_ds},1}}(\gamma(\theta_k)) J_{m_{\text{h\_ds},2}}(\gamma(\theta_k))$, and $m_{\text{h\_dc}} = m_{\text{h\_ds},1} - m_{\text{h\_ds},2}$ $(m_{\text{h\_ds},1}, m_{\text{h\_ds},2} \in \mathbb{M}_{\text{h\_ds}})$. $\mathbb{M}_{\text{h\_ds}} = \{m_{\text{h\_de}}, m_{\text{h\_sp}}\}$ collects the horizontal locations of the elements in the dense and sparse arrays. (6.24) can be viewed as the signal $\mathbf{d}_{m_\text{t}}$ received at an array with steering matrix $\mathbf{A}_{\text{df}}$.

Now we formulate the received samples in the tensor form. We first decompose $\mathbf{y}_{\text{df},m_\text{t}}$ into the vertical and horizontal domains (corresponding to the first and second modes of the tensor model), as given by $\mathbf{Y}_{\text{df},m_\text{t}} = \text{invec}(\mathbf{y}_{\text{df},m_\text{t}}) \in \mathbb{C}^{(2N_{\text{vd}} N_{\text{vs}} - 1) \times M_{\text{hr}}}$.

[7] All the rows of $\text{vec}(\mathbf{R}_{\text{nn},m_\text{t}})$ with nonzero value correspond to the phase difference of 0 in the difference coarray, which are produced by the self difference of sparse and dense subarrays. Because the noise is temporally and spatially white with power σ^2_{n}, by averaging the value of these rows, we have $\sigma^2_{\text{n}} = \sum_{k=1}^{K} \sigma^2_{\text{n},m_\text{t},k}$.

Then, we collect $\mathbf{Y}_{\mathrm{df},m_\mathrm{t}}$ at all time frames, and store them in the time domain (corresponding to the third mode of the tensor model). Thus, the received samples can be expressed as

$$\mathcal{Y}_\mathrm{df} = [\mathbf{Y}_{\mathrm{df},1} \sqcup_3 \mathbf{Y}_{\mathrm{df},2} \sqcup_3 \ldots \sqcup_3 \mathbf{Y}_{\mathrm{df},M_\mathrm{t}}] = \mathcal{A}_\mathrm{df} \times_3 \mathbf{D} + \mathcal{N}_\mathrm{df} \in \mathbb{C}^{N_\mathrm{vdc} \times N_\mathrm{hdc} \times M_\mathrm{t}}, \quad (6.27)$$

where $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{M_\mathrm{t}}]^T \in \mathbb{C}^{M_\mathrm{t} \times K}$, $\mathcal{A}_\mathrm{df} \in \mathbb{C}^{N_\mathrm{vdc} \times N_\mathrm{hdc} \times K}$ is known as the space-time response tensor [113], and $\mathcal{N}_\mathrm{df}$ is the noise tensor model. Due to the above-mentioned process (6.24)-(6.27), the elements of $\mathcal{N}_\mathrm{df}$ are all zeros except σ_n^2 at $(0, 0, m_\mathrm{t})$, $m_\mathrm{t} = 1, 2, \ldots, M_\mathrm{t}$. In (6.27), $\mathcal{A}_\mathrm{df}$ is obtained as

$$\begin{aligned}\mathcal{A}_\mathrm{df} = [&\mathbf{a}_\mathrm{dfv}(\theta_1) \circ \mathbf{a}_\mathrm{dfh}(\theta_1, \phi_1) \sqcup_3 \mathbf{a}_\mathrm{dfv}(\theta_2) \circ \mathbf{a}_\mathrm{dfh}(\theta_2, \phi_2) \\ &\sqcup_3 \ldots \sqcup_3 \mathbf{a}_\mathrm{dfv}(\theta_K) \circ \mathbf{a}_\mathrm{dfh}(\theta_K, \phi_K)]. \end{aligned} \quad (6.28)$$

By substituting (6.28) into (6.27), we obtain

$$\mathcal{Y}_\mathrm{df} = \sum_{k=1}^{K} \mathbf{a}_\mathrm{dfv}(\theta_k) \circ \mathbf{a}_\mathrm{dfh}(\theta_k, \phi_k) \circ [\mathbf{D}]_{:,k} + \mathcal{N}_\mathrm{df} = [\![\mathcal{Z}_\mathrm{df}; \mathbf{A}_\mathrm{dfv}, \mathbf{A}_\mathrm{dfh}, \mathbf{D}]\!] + \mathcal{N}_\mathrm{df}, \quad (6.29)$$

where $[\mathbf{A}_\mathrm{dfv}]_{:,k} = \mathbf{a}_\mathrm{dfv}(\theta_k)$, $[\mathbf{A}_\mathrm{dfh}]_{:,k} = \mathbf{a}_\mathrm{dfh}(\theta_k, \phi_k)$, and $\mathcal{Z}_\mathrm{df} \in \mathbb{C}^{K \times K \times K}$ is an order-3 identity superdiagonal tensor[8].

Eq. (6.27) shows that the elements of the equivalent signal matrix $\mathbf{D} \in \mathbb{C}^{M_\mathrm{t} \times K}$ are actually the received signal powers due to the autocorrelation calculation (6.22). To build a full-rank matrix $\mathbf{D}$ for DoA estimation, one would need to assume that the received signal powers change over time, and the power of every signal is different from each other, as assumed in [127]. However, such assumption is unrealistic in practice. It is possible that the rank of the equivalent device signal matrix $\mathbf{D}$ is smaller than the number of devices K, i.e., $\mathrm{Rank}(\mathbf{D}) < K$, which behaves as if some of the received signals are coherent, leading to incorrect channel estimation.

[8] A tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is diagonal if $a_{i_1 i_2 \cdots i_N} \neq 0$ only if $i_1 = i_2 = \cdots = i_N$. When $I_1 = I_2 = \cdots = I_N$, $\mathcal{A}$ is called as superdiagonal.

Prevent possible coherent signals, we propose a novel approach in the next section to construct a signal tensor model with suitable n-ranks in all modes. This allows us to estimate the 2-D DoAs of K devices.

6.4 Spatial smoothing-based tensor n-rank enhancement

In this section, we analyze the relationship between the rank of $\mathbf{D}$ and the n-rank of $\mathcal{Y}_{\text{df}}$. We propose a spatial smoothing-based method to enhance the n-rank of $\mathcal{Y}_{\text{df}}$. By using the proposed method, we verify that one can build a signal tensor model that provides a large enough rank in each mode to perform the DoA estimation of K devices, even when the received signal powers of all the devices are equal. These powers are steady temporally across all time frames.

As discussed in Section 6.3.3, the rank of $\mathbf{D}$ in (6.29) is typically smaller than the number of devices, K, in practice. Based on the uniqueness condition of tensor CP decomposition [75], we first provide the following theorem to evaluate the impact of Rank($\mathbf{D}$) on the n-ranks of the tensor $\mathcal{Y}_{\text{df}}$.

Theorem 5. *For $\mathcal{Y}_{df} = [\![\mathcal{Z}_{df}; \mathbf{A}_{dfv}, \mathbf{A}_{dfh}, \mathbf{D}]\!] + \mathcal{N}_{df}$, if Rank($\mathbf{D}$) $< K$, the ranks of the signal spaces of $\mathcal{Y}_{df}$ in all modes are smaller than the number of devices K, i.e., Rank($\mathbf{U}_{v,n}$) $< K$, $n = 1, 2, 3$, where $\mathbf{U}_{v,n}$ is the mode-n signal subspace of $\mathcal{Y}_{df}$ with $\mathbf{U}_{v,1} \in \mathbb{C}^{N_{vdc} \times K}$, $\mathbf{U}_{v,2} \in \mathbb{C}^{N_{hdc} \times K}$, and $\mathbf{U}_{v,3} \in \mathbb{C}^{M_t \times K}$.*

Proof. See Appendix E.2. □

According to Theorem 5, when the rank of $\mathbf{D}$ in (6.29) is smaller than the number of devices K, we cannot decompose the tensor model (6.27) into the signal and noise spaces in all modes. As a result, the subspace-based algorithms cannot be used to estimate the angles of the devices. To enhance the n-rank of the signal tensor model, we apply spatial smoothing techniques [72] to build up a sample tensor model whose signal subspace is full rank in each mode.

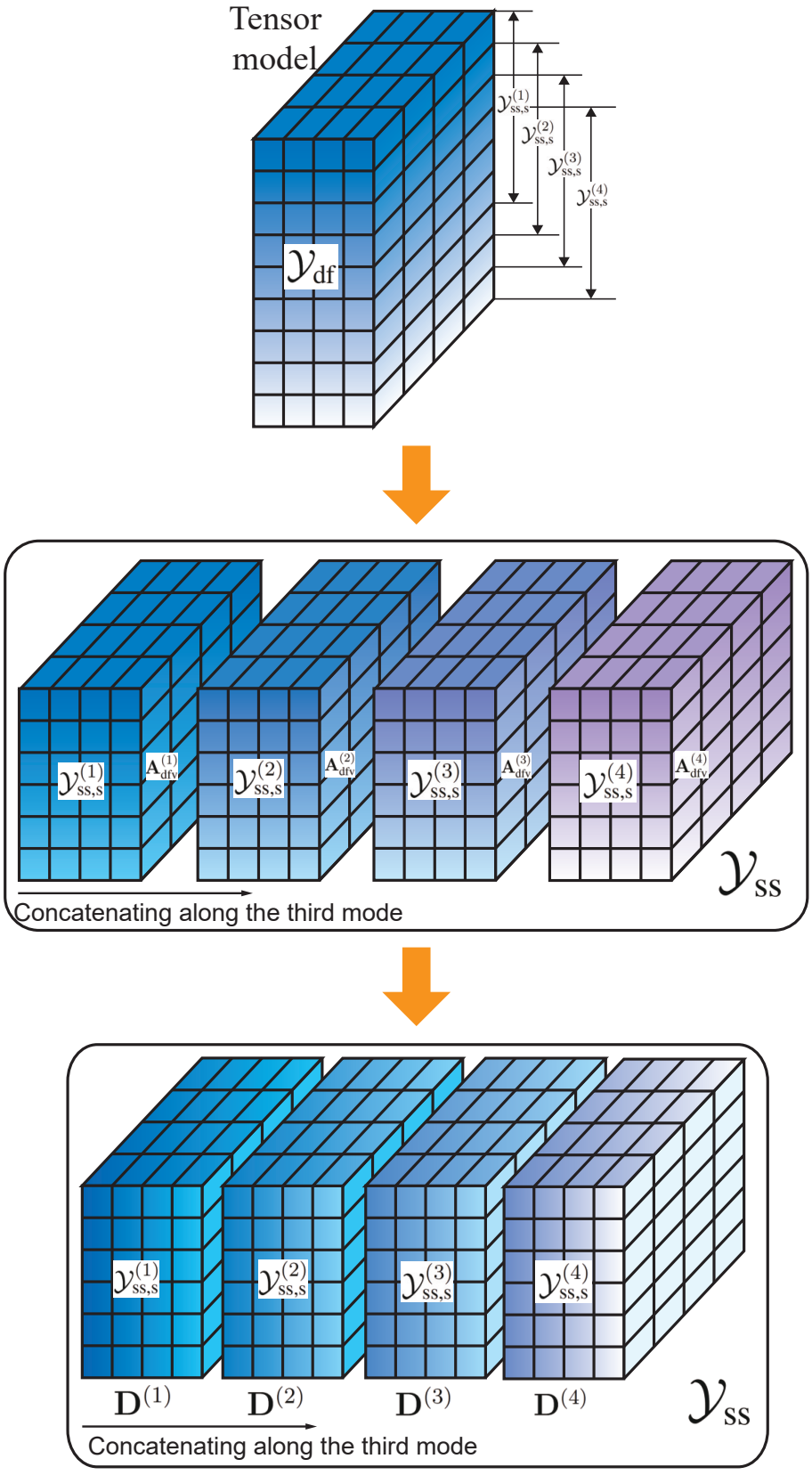


Figure 6.6 : An illustration of the proposed tensor n-rank enhancement method.

We divide $\mathcal{Y}_{\rm df}$ in (6.27) into $N_{\rm is}$ identical subtensors in its first mode, as shown in the left-hand side of Fig. 6.6. The $n_{\rm is}$-th subtensor ($n_{\rm is} = 1, 2, \ldots, N_{\rm is}$) can be constructed as

$$\mathcal{Y}_{\rm ss}^{(n_{\rm is})} = \mathcal{Y}_{\rm df} \times_1 \mathbf{J}_{{\rm ss},n_{\rm is}} \in \mathbb{C}^{N_{\rm ss} \times N_{\rm hdc} \times M_{\rm t}}, \tag{6.30}$$

where $\mathbf{J}_{{\rm ss},n_{\rm is}} = [\mathbf{0}_{N_{\rm ss} \times (n_{\rm is}-1)}, \mathbf{I}_{N_{\rm ss}}, \mathbf{0}_{N_{\rm ss} \times (N_{\rm is}-n_{\rm is})}]$ and $N_{\rm ss} = 2N_{\rm vd}N_{\rm vs} - N_{\rm is}$.

We can see that

$$\mathcal{Y}_{\rm ss}^{(n_{\rm is})} = \left[\!\left[\mathcal{Z}_{\rm df}; \mathbf{A}_{\rm dfv}^{(n_{\rm is})}, \mathbf{A}_{\rm dfh}, \mathbf{D} \right]\!\right] + \mathcal{N}^{(n_{\rm is})}, \tag{6.31}$$

where $\mathbf{A}_{\rm dfv}^{(n_{\rm is})} = \mathbf{J}_{{\rm ss},n_{\rm is}} \mathbf{A}_{\rm dfv} = \mathbf{A}_{\rm dfv}^{(1)} \mathbf{Q}_{\rm ss}^{n_{\rm is}-1} \in \mathbb{C}^{N_{\rm ss} \times K}$, $\mathbf{Q}_{\rm ss} = {\rm diag}\left(q_{{\rm ss},1}, q_{{\rm ss},2}, \ldots, q_{{\rm ss},K}\right) \in \mathbb{C}^{K \times K}$, $q_{{\rm ss},k} = e^{j\frac{2\pi}{\lambda} h \cos(\theta_k)}$, and $\mathcal{N}^{(n_{\rm is})} = \mathcal{N}_{\rm df} \times_1 \mathbf{J}_{{\rm ss},n_{\rm is}} \in \mathbb{C}^{N_{\rm ss} \times N_{\rm hdc} \times M_{\rm t}}$ is the selected subtensor of the noise model. We can verify that only when $N_{\rm vd}N_{\rm vs} - N_{\rm ss} + 1 \leq n_{\rm is} \leq N_{\rm vd}N_{\rm vs}$, $\mathcal{N}^{(n_{\rm is})}$ has $\sigma_{\rm n}^2$ at the $n_{\rm ss} = (N_{\rm vd}N_{\rm vs} - n_{\rm is} + 1)$-th entry of the first mode, while $n_{\rm hdc} = P + 1$ and $m_{\rm t} = 1, 2, \ldots, M_{\rm t}$. In all other cases, $\mathcal{N}^{(n_{\rm is})} = 0$.

By concatenating the $N_{\rm is}$ identical subtensors $\mathcal{Y}_{\rm ss}^{(n_{\rm is})}$, $n_{\rm is} = 1, \ldots, N_{\rm is}$, as shown in the middle block of Fig. 6.6, the spatially smoothed signal tensor model can be constructed as

$$\mathcal{Y}_{\rm ss} = \left[\bigsqcup_{n_{\rm is}=1,\ldots,N_{\rm is}}^{3} \mathcal{Y}_{\rm ss}^{(n_{\rm is})} \right] \in \mathbb{C}^{N_{\rm ss} \times N_{\rm hdc} \times (M_{\rm t}N_{\rm is})}, \tag{6.32}$$

which has a rank large enough in each mode to perform DoA estimation of the K devices.

Now, we proceed to verify the n-ranks of $\mathcal{Y}_{\rm ss}$. Define $\mathcal{Y}_{\rm ss,s}^{(n_{\rm is})} = \left[\!\left[\mathcal{Z}_{\rm df}; \mathbf{A}_{\rm dfv}^{(n_{\rm is})}, \mathbf{A}_{\rm dfh}, \mathbf{D} \right]\!\right]$. We have

$$\begin{aligned}
\left[\mathcal{Y}_{\rm ss,s}^{(n_{\rm is})}\right]_{n_{\rm ss},n_{\rm hdc},m_t} &= \sum_{k=1}^{K} \left[\mathbf{A}_{\rm dfv}^{(n_{\rm is})}\right]_{n_{\rm ss},k} \left[\mathbf{A}_{\rm dfh}\right]_{n_{\rm hdc},k} \left[\mathbf{D}\right]_{m_{\rm t},k} \\
&= \sum_{k=1}^{K} \left(\left[\mathbf{A}_{\rm dfv}^{(1)}\right]_{n_{\rm ss},k} q_{{\rm ss},k}^{n_{\rm is}-1} \right) \left[\mathbf{A}_{\rm dfh}\right]_{n_{\rm hdc},k} \left[\mathbf{D}\right]_{m_{\rm t},k} \\
&= \sum_{k=1}^{K} \left[\mathbf{A}_{\rm dfv}^{(1)}\right]_{n_{\rm ss},k} \left[\mathbf{A}_{\rm dfh}\right]_{n_{\rm hdc},k} \left(\left[\mathbf{D}\right]_{m_{\rm t},k} q_{{\rm ss},k}^{n_{\rm is}-1} \right).
\end{aligned} \tag{6.33}$$

Hence,

$$\mathcal{Y}_{\text{ss,s}}^{(n_{\text{is}})} = \left[\!\left[\mathcal{Z}_{\text{df}}; \mathbf{A}_{\text{dfv}}^{(n_{\text{is}})}, \mathbf{A}_{\text{dfh}}, \mathbf{D} \right]\!\right] = \left[\!\left[\mathcal{Z}_{\text{df}}; \mathbf{A}_{\text{dfv0}}, \mathbf{A}_{\text{dfh}}, \mathbf{D}^{(n_{\text{is}})} \right]\!\right], \tag{6.34}$$

where $\mathbf{A}_{\text{dfv0}} = \mathbf{A}_{\text{dfv}}^{(1)}$ and $\mathbf{D}^{(n_{\text{is}})} = \mathbf{D}\mathbf{Q}_{\text{ss}}^{n_{\text{is}}-1}$. Therefore, (6.32) can be rewritten as

$$\mathcal{Y}_{\text{ss}} = \left[\underset{n_{\text{is}}=1,\ldots,N_{\text{is}}}{\sqcup_3} \mathcal{Y}_{\text{ss}}^{(n_{\text{is}})} \right] = [\![\mathcal{Z}_{\text{df}}; \mathbf{A}_{\text{dfv0}}, \mathbf{A}_{\text{dfh}}, \mathbf{D}_{\text{ss}}]\!] + \mathcal{N}_{\text{ss}}, \tag{6.35}$$

where

$$\mathbf{D}_{\text{ss}} = \left[\left(\mathbf{D}^{(1)}\right)^T, \left(\mathbf{D}^{(2)}\right)^T, \ldots, \left(\mathbf{D}^{(N_{\text{is}})}\right)^T \right]^T \in \mathbb{C}^{(M_{\text{t}}N_{\text{is}})\times K}, \tag{6.36}$$

and $\mathcal{N}_{\text{ss}} = \left[\underset{n_{\text{is}}=1,\ldots,N_{\text{is}}}{\sqcup_3} \mathcal{N}^{(n_{\text{is}})} \right] \in \mathbb{C}^{N_{\text{ss}}\times N_{\text{hdc}}\times(M_{\text{t}}N_{\text{is}})}$ is a tensor of all zeros except σ_{n}^2 at $(n_{\text{ss}}, 0, \tilde{m}_{\text{t}})$, where $n_{\text{ss}} = (N_{\text{vd}}N_{\text{vs}} - n_{\text{is}} + 1)$, $M_{\text{t}}(n_{\text{is}} - 1) \leq \tilde{m}_{\text{t}} \leq M_{\text{t}}n_{\text{is}}$, and $n_{\text{is}} = 1, \ldots, N_{\text{is}}$.

An illustration of (6.35) is shown at the right of Fig. 6.6, where the recurrence relations among the divided subtensors in the mode-1 is equivalent to those in the mode-3. $\mathbf{A}_{\text{dfv}}^{(n_{\text{is}})}$ and $\mathbf{D}^{(n_{\text{is}})}$ are the factor matrices [76] of mode-1 and mode-3, respectively. This property can be used to enhance the n-ranks of the signal tensor model.

We consider the extreme case where the received powers of all the devices are equal, and these powers are steady temporally across all time frames, i.e., $\sigma_{\text{s},m_{\text{t}},k}^2 = \sigma_{\text{s}}^2$, $m_{\text{t}} = 1, 2, \ldots, M_{\text{t}}$, and $k = 1, 2, \ldots, K$. Then, $\mathbf{D} = \sigma_{\text{s}}^2 \mathbf{1}_{M_{\text{t}}\times K}$. As a result, $\mathbf{D}_{\text{ss}}$ can be rewritten as $\mathbf{D}_{\text{ss}} = \sigma_{\text{s}}^2 \tilde{\mathbf{Q}}_{\text{ss}} \otimes \mathbf{1}_{M_{\text{t}}}$, where $\tilde{\mathbf{Q}}_{\text{ss}} = [\mathbf{1}_K, \mathbf{q}_{\text{ss},1}, \mathbf{q}_{\text{ss},2}, \ldots, \mathbf{q}_{\text{ss},N_{\text{is}}-1}]^T \in \mathbb{C}^{N_{\text{is}}\times K}$, $\mathbf{q}_{\text{ss},n'_{\text{is}}} = \left[q_{\text{ss},1}^{n'_{\text{is}}}, q_{\text{ss},2}^{n'_{\text{is}}}, \ldots, q_{\text{ss},K}^{n'_{\text{is}}} \right]^T$, and $n'_{\text{is}} = 1, 2, \ldots, N_{\text{is}} - 1$.

Because the paths are from different directions, $\tilde{\mathbf{Q}}_{\text{ss}}$ is an $N_{\text{is}} \times K$ Vandermonde matrix and $\text{Rank}(\tilde{\mathbf{Q}}_{\text{ss}}) = \min(N_{\text{is}}, K)$. $\text{Rank}(\tilde{\mathbf{Q}}_{\text{ss}}) = K$ iff $N_{\text{is}} \geq K$. According to Theorem 5, $\text{Rank}_n(\tilde{\mathcal{Y}}_{\text{ss}}) = K$, when $\text{Rank}(\mathbf{D}_{\text{ss}}) = K$. Thus, the signal and noise spaces of $\mathcal{Y}_{\text{ss}}$ in (6.32) can be decomposed in each mode.

We note that the number of $\mathcal{Y}_{\text{ss}}^{(n_{\text{is}})}$ needs to be larger than the number of devices, i.e., $N_{\text{is}} \geq K$, to guarantee that $\mathcal{Y}_{\text{ss}}$ is full rank. Also, the system DoF available

after spatial smoothing is proportional to the size of $\mathcal{Y}_{\rm ss}^{(n_{\rm is})}$. Since the total number of elements in $\mathcal{Y}_{\rm df}$ is constant, increasing the number of $\mathcal{Y}_{\rm ss}^{(n_{\rm is})}$ implies that the size of each $\mathcal{Y}_{\rm ss}^{(n_{\rm is})}$ is smaller, while a larger size of each $\mathcal{Y}_{\rm ss}^{(n_{\rm is})}$ means there is a smaller number of recurrence shifts available. In this sense, the best strategy is to minimize the difference between $N_{\rm is}$ and $N_{\rm ss}$. Since in our system, we have $N_{\rm is} + N_{\rm ss} - 1 = N_{\rm vdc} = 2N_{\rm vd}N_{\rm vs} - 1$, we set $N_{\rm ss} = N_{\rm is} = N_{\rm vd}N_{\rm vs}$.

Remark 2. *After spatial smoothing, the system DoF becomes half of that in* (6.12)*, because we divide $\mathcal{Y}_{df}$ into multiple $\mathcal{Y}_{ss}^{(n_{is})}$. Therefore, to prevent the system DoF from decreasing and achieve the target set in Section 6.3.3, we modify (C1) in the optimization problem* (6.12) *to $N_{vd}N_{vs} \geq M_{vr}$. Applying the analytical strategy developed in Section 6.3.3, we formulate the modified optimization problem* (6.12) *as*

$$\min_{N_{vd},N_{hd}\in\mathbb{Z}^+} M_{rf} = N_{vd}N_{hd} + \left\lceil \frac{M_{vr}}{N_{vd}} \right\rceil \left\lceil \frac{M_{hr}}{N_{hd}} \right\rceil \tag{6.37}$$

$$\textit{s.t. (C3) and (C4).}$$

We see that the minimum M_{rf} can be obtained when $N_{vd}N_{hd}$ and $\lceil M_{vr}/N_{vd} \rceil \lceil M_{hr}/N_{hd} \rceil$ are close or equal. Because $N_{vd}, N_{hd}, N_{vs}, N_{hs} \in \mathbb{Z}^+$, the optimal value of N_{vd}, N_{hd}, N_{vs}, and N_{hs} can be obtained.

6.5 2-D DoA estimation

In this section, the 2-D DoAs are estimated by developing a new tensor-based subspace estimation algorithm. By exploiting the recurrence relations among the UCAs, the elevation DoAs are estimated first, and then the corresponding azimuth angles are estimated by using the tensor MUSIC. The hardware and software complexities are analyzed in the end.

6.5.1 Estimation of Elevation Angle

We first propose a tensor-based total least-squares (TLS)-ESPRIT algorithm to estimate the elevation angle of each device. The HOSVD of the measurement tensor $\mathcal{Y}_{\rm ss}$ is given by

$$\mathcal{Y}_{\rm ss} = \mathcal{L} \times_1 \mathbf{U}_{\rm dfv0} \times_2 \mathbf{U}_{\rm dfh} \times_3 \mathbf{U}_{\rm ss} = [\![\mathcal{L}; \mathbf{U}_{\rm dfv0}, \mathbf{U}_{\rm dfh}, \mathbf{U}_{\rm ss}]\!] \in \mathbb{C}^{N_{\rm ss} \times N_{\rm hdc} \times (M_{\rm t} N_{\rm is})}, \quad (6.38)$$

where $\mathbf{U}_{\rm dfv0} \in \mathbb{C}^{N_{\rm ss} \times N_{\rm ss}}$, $\mathbf{U}_{\rm dfh} \in \mathbb{C}^{N_{\rm hdc} \times N_{\rm hdc}}$, and $\mathbf{U}_{\rm ss} \in \mathbb{C}^{(M_{\rm t} N_{\rm is}) \times (M_{\rm t} N_{\rm is})}$, are the left singular matrices of the mode-n unfoldings of tensor $\mathcal{Y}_{\rm ss}$, and the core tensor $\mathcal{L} \in \mathbb{C}^{N_{\rm ss} \times N_{\rm hdc} \times (M_{\rm t} N_{\rm is})}$ is obtained by moving the singular matrices to the left-hand side of (6.38):

$$\mathcal{L} = \mathcal{Y}_{\rm ss} \times_1 \mathbf{U}_{\rm dfv0}^H \times_2 \mathbf{U}_{\rm dfh}^H \times_3 \mathbf{U}_{\rm ss}^H. \quad (6.39)$$

Define $\tilde{\mathcal{Y}}_{\rm ss} = [\![\mathcal{Z}_{\rm df}; \mathbf{A}_{\rm dfv0}, \mathbf{A}_{\rm dfh}, \mathbf{D}_{\rm ss}]\!]$, which contains the noise-free components of $\mathcal{Y}_{\rm ss}$. By removing the noise subspace component in each mode, we obtain the HOSVD model of $\tilde{\mathcal{Y}}_{\rm ss}$, as given by

$$\tilde{\mathcal{Y}}_{\rm ss} = \mathcal{L}_{\rm ss} \times_1 \mathbf{U}_{\rm dfv0,s} \times_2 \mathbf{U}_{\rm dfh,s} \times_3 \mathbf{U}_{\rm ss,s} \in \mathbb{C}^{N_{\rm ss} \times N_{\rm hdc} \times (M_{\rm t} N_{\rm is})}, \quad (6.40)$$

where $\mathbf{U}_{\rm dfv0,s} \in \mathbb{C}^{N_{\rm ss} \times K}$, $\mathbf{U}_{\rm dfh,s} \in \mathbb{C}^{N_{\rm hdc} \times K}$, and $\mathbf{U}_{\rm ss,s} \in \mathbb{C}^{(M_{\rm t} N_{\rm is}) \times K}$ are the signal subspaces in the first, second, and third modes, respectively; and $\mathcal{L}_{\rm ss} \in \mathbb{C}^{K \times K \times K}$ is obtained by discarding insignificant singular values of $\mathcal{Y}_{\rm ss}$ in all the modes.

Define the signal subspace as

$$\mathcal{U}_{\rm s} = \mathcal{L}_{\rm ss} \times_1 \mathbf{U}_{\rm dfv0,s} \times_2 \mathbf{U}_{\rm dfh,s} \in \mathbb{C}^{N_{\rm ss} \times N_{\rm hdc} \times K}. \quad (6.41)$$

Because $\tilde{\mathcal{Y}}_{\rm ss}$ can be rewritten as $\tilde{\mathcal{Y}}_{\rm ss} = \mathcal{A}_{\rm ss} \times_3 \mathbf{D}_{\rm ss}$ with $\mathcal{A}_{\rm ss} = \mathcal{Z}_{\rm df} \times_1 \mathbf{A}_{\rm dfv0} \times_2 \mathbf{A}_{\rm dfh}$, we obtain

$$\mathcal{A}_{\rm ss} = \mathcal{U}_{\rm s} \times_3 \mathbf{D}_{\rm ss}. \quad (6.42)$$

where $\mathbf{D}_{\rm ss} \in \mathbb{C}^{(M_{\rm t} N_{\rm is}) \times K}$ is a full column rank matrix. According to the shift-invariance relation among the subtensors in mode-1, we have

$$\mathcal{A}_{\rm ss} \times_1 \mathbf{J}_{\rm v2} = \mathcal{A}_{\rm ss} \times_1 \mathbf{J}_{\rm v1} \times_3 \mathbf{\Theta}_{\rm v}, \quad (6.43)$$

where $\mathbf{\Theta}_{\rm v} = {\rm diag}\left(e^{-j\frac{2\pi}{\lambda}h\cos(\theta_1)}, \ldots, e^{-j\frac{2\pi}{\lambda}h\cos(\theta_K)}\right)$, $\mathbf{J}_{\rm v1} = [\mathbf{I}_{M_{\rm vr}-1}, \mathbf{0}_{(M_{\rm vr}-1)\times 1}]$, and $\mathbf{J}_{\rm v2} = [\mathbf{0}_{(M_{\rm vr}-1)\times 1}, \mathbf{I}_{M_{\rm vr}-1}]$. Let

$$\mathcal{U}_{\rm sv1} = \mathcal{U}_{\rm s} \times_1 \mathbf{J}_{\rm v1} \text{ and } \mathcal{U}_{\rm sv2} = \mathcal{U}_{\rm s} \times_1 \mathbf{J}_{\rm v2}. \tag{6.44}$$

By substituting (6.42) into (6.43), we have $\mathcal{U}_{\rm sv2} = \mathcal{U}_{\rm sv1} \times_3 \mathbf{\Psi}_{\rm v}$, where $\mathbf{\Psi}_{\rm v} \in \mathbb{C}^{K\times K}$ is a full rank matrix. To obtain the estimate of $\mathbf{\Psi}_{\rm v}$, we define $\mathbf{\Upsilon}_{\rm v} = [\mathbf{\Upsilon}_{\rm v1} \quad \mathbf{\Upsilon}_{\rm v2}] \in \mathbb{C}^{K\times 2K}$. We now generalize the matrix TLS problem formulation [119] to the tensor setting, as follows.

$$\begin{aligned} \hat{\mathbf{\Upsilon}}_{\rm v} &= \arg\min_{\mathbf{\Upsilon}_{\rm v}} \left\|\mathcal{U}_{\rm sv1} \times_3 \mathbf{\Upsilon}_{\rm v1} + \mathcal{U}_{\rm sv2} \times_3 \mathbf{\Upsilon}_{\rm v2}\right\|, \\ &\text{s.t.} \quad \mathbf{\Upsilon}_{\rm v}\mathbf{\Upsilon}_{\rm v}^H = \mathbf{I}_K, \end{aligned} \tag{6.45}$$

which finds a unitary matrix $\mathbf{\Upsilon}_{\rm v}$ with orthogonal submatrices to $\mathcal{U}_{\rm sv1}$ and $\mathcal{U}_{\rm sv2}$ in mode-3.

The mode-3 unfoldings of $\mathcal{U}_{\rm sv1}$ is given by

$$\mathbf{U}_{{\rm sv1}(3)} = \mathbf{U}_{{\rm s}(3)} \left(\mathbf{J}_{\rm v1} \otimes \mathbf{I}_{N_{\rm hdc}}\right)^T, \tag{6.46}$$

where $\mathbf{U}_{{\rm s}(3)} \in \mathbb{C}^{K\times M_{\rm vr}M_{\rm hr}M_{\rm f}}$ is the mode-3 unfolding of $\mathcal{U}_{\rm s}$. The mode-3 unfoldings of $\mathcal{U}_{\rm sv2}$ can be formulated in the same way. Since $\|\mathcal{A}\| = \left\|\mathbf{A}_{(n)}\right\|_{\rm F}$ $(n = 1, 2, \ldots, N)$ [76], we rewrite the tensor TLS problem (6.45) as

$$\begin{aligned} \hat{\mathbf{\Upsilon}}_{\rm v} &= \arg\min_{\mathbf{\Upsilon}_{\rm v}} \left\|\mathbf{\Upsilon}_{\rm v1}\mathbf{U}_{{\rm s}(3)} \left(\mathbf{J}_{\rm v1} \otimes \mathbf{I}_{N_{\rm hdc}}\right)^T + \mathbf{\Upsilon}_{\rm v2}\mathbf{U}_{{\rm s}(3)} \left(\mathbf{J}_{\rm v2} \otimes \mathbf{I}_{N_{\rm hdc}}\right)^T\right\|_{\rm F} \\ &= \arg\min_{\mathbf{\Upsilon}_{\rm v}} \left\|\mathbf{W}_{\rm v}\mathbf{\Upsilon}_{\rm v}^T\right\|_{\rm F}, \end{aligned} \tag{6.47}$$

where

$$\mathbf{W}_{\rm v} = \left[\left(\mathbf{J}_{\rm v1} \otimes \mathbf{I}_{N_{\rm hdc}}\right)\mathbf{U}_{{\rm s}(3)}^T \quad \left(\mathbf{J}_{\rm v2} \otimes \mathbf{I}_{N_{\rm hdc}}\right)\mathbf{U}_{{\rm s}(3)}^T\right] \in \mathbb{C}^{(N_{\rm ss}-1)N_{\rm hdc}\times 2K}. \tag{6.48}$$

The SVD of $\mathbf{W}_{\rm v}^H\mathbf{W}_{\rm v}$ is written as $\mathbf{W}_{\rm v}^H\mathbf{W}_{\rm v} = \dot{\mathbf{U}}_{\rm v}\dot{\mathbf{\Lambda}}_{\rm v}\dot{\mathbf{V}}_{\rm v}$, where $\dot{\mathbf{U}}_{\rm v} \in \mathbb{C}^{2K\times 2K}$ and $\dot{\mathbf{V}}_{\rm v} \in \mathbb{C}^{2K\times 2K}$ are the left and right singular matrices, respectively; and $\dot{\mathbf{\Lambda}}_{\rm v} \in \mathbb{C}^{2K\times 2K}$

contains the singular values. We partition $\dot{\mathbf{U}}_{\mathrm{v}}$ into four blocks:

$$\dot{\mathbf{U}}_{\mathrm{v}} = \begin{bmatrix} \dot{\mathbf{U}}_{\mathrm{v11}} & \dot{\mathbf{U}}_{\mathrm{v12}} \\ \dot{\mathbf{U}}_{\mathrm{v21}} & \dot{\mathbf{U}}_{\mathrm{v22}} \end{bmatrix} \in \mathbb{C}^{2K \times 2K}. \tag{6.49}$$

Let $\hat{\boldsymbol{\Upsilon}}_{\mathrm{v1}} = \dot{\mathbf{U}}_{\mathrm{v12}}^{T} \in \mathbb{C}^{K \times K}$ and $\hat{\boldsymbol{\Upsilon}}_{\mathrm{v2}} = \dot{\mathbf{U}}_{\mathrm{v22}}^{T} \in \mathbb{C}^{K \times K}$. According to the standard TLS [119], the estimate of $\boldsymbol{\Psi}_{\mathrm{v}}$ is given by $\hat{\boldsymbol{\Psi}}_{\mathrm{v}} = -\hat{\boldsymbol{\Upsilon}}_{\mathrm{v1}} \hat{\boldsymbol{\Upsilon}}_{\mathrm{v2}}^{-1}$, where the K eigenvalues of $\hat{\boldsymbol{\Psi}}_{\mathrm{v}}$, i.e., $\psi_{\mathrm{v},k}$, $k = 1, 2, \ldots, K$, are sorted in descending order. According to the array steering expression (6.2), the elevation angle of the k-th device can be estimated as

$$\hat{\theta}_k = \arccos\left(j\lambda \ln(\psi_{\mathrm{v},k})/(2\pi h)\right). \tag{6.50}$$

6.5.2 Estimation of Azimuth Angle

We use the tensor-MUSIC algorithm [78] to estimate the azimuth angle of each device. According to (6.40), we can discard the largest K singular values of the mode-n unfoldings of $\mathcal{Y}_{\mathrm{ss}}$ and obtain the noise subspace in mode-2, $\mathbf{U}_{\mathrm{dfh,n}}$. Then, we generalize the matrix-based MUSIC to the tensor, and the tensor MUSIC spectrum of the azimuth angle can be defined as

$$\mathrm{SP}_{\mathrm{MUSIC}}(\Phi) = \left\| \mathcal{A}_{\mathrm{ss}} \times_2 \mathbf{U}_{\mathrm{dfh,n}}^{H} \right\|^{-2}, \tag{6.51}$$

where $\Phi = [\phi_1, \phi_2, \ldots, \phi_K]$. The mode-2 unfolding of $\mathcal{A}_{\mathrm{ss}}$ can be expressed as

$$\mathbf{A}_{\mathrm{ss}(2)} = \mathbf{A}_{\mathrm{dfh}} \mathbf{Z}_{\mathrm{df}(2)} \left(\mathbf{I}_{M_{\mathrm{t}} N_{\mathrm{is}}} \otimes \mathbf{A}_{\mathrm{dfv0}} \right)^{T}, \tag{6.52}$$

where $\mathbf{Z}_{\mathrm{df}(2)}$ is the mode-2 unfolding of $\mathcal{Z}_{\mathrm{df}}$. According to a property of tensor multiplication and unfolding: $\|\mathcal{A}\| = \left\|\mathbf{A}_{(n)}\right\|_{\mathrm{F}}$, $n = 1, 2, \ldots, N$, we can rewrite the tensor MUSIC spectrum (6.51) as

$$\mathrm{SP}_{\mathrm{MUSIC}}(\Phi) = \left\| \mathbf{U}_{\mathrm{dfh,n}}^{H} \mathbf{A}_{\mathrm{dfh}} \mathbf{Z}_{\mathrm{df}(2)} \left(\mathbf{I}_{M_{\mathrm{t}} N_{\mathrm{is}}} \otimes \mathbf{A}_{\mathrm{dfv0}} \right)^{T} \right\|_{\mathrm{F}}^{-2}. \tag{6.53}$$

By substituting the estimated elevation angle of each device (6.50) into (6.53), the corresponding azimuth angle ϕ_k can be estimated by searching the prominent peaks

Algorithm 3 Tensor-based subspace estimation algorithm

- **Input**: The processed signal, $\mathcal{Y}_{\text{ss}}$, and the number of devices, K.
- **Output**: The estimated elevation and azimuth angles, $\hat{\theta}_k$ and $\hat{\phi}_k$, $k = 1, 2, \ldots, K$.
- Take the HOSVD of $\mathcal{Y}_{\text{ss}}$ and obtain $\mathcal{U}_{\text{s}}$, $\mathcal{U}_{\text{sv1}}$, and $\mathcal{U}_{\text{sv2}}$ according to (6.41) and (6.44).
- Estimate $\hat{\mathbf{\Psi}}_{\text{v}}$ by solving the tenosr TLS problem (6.45).
- Calculate the eigenvalues of $\hat{\mathbf{\Psi}}_{\text{v}}$, i.e., $\psi_{\text{v},k}$, $k = 1, 2, \ldots, K$, and estimate $\hat{\theta}_k$ by using (6.50).
- Calculate $\mathbf{U}_{\text{dfh,n}}$ and estimate $\hat{\phi}_k$ by searching the prominent peaks of (6.53).

of the tensor MUSIC spectrum (6.53). Algorithm 3 summarizes the procedure of the proposed tensor-based subspace estimation algorithm.

6.5.3 Complexity Analysis

We analyze the hardware and software complexity of the proposed tensor-based parameter estimation algorithm.

For the hardware complexity, the use of the proposed hybrid array reduces the hardware complexity to $O(M_{\text{rf}}) = O(N_{\text{vd}}N_{\text{hd}} + N_{\text{vs}}N_{\text{hs}})$, while a fully digital array using the same number of antennas would have a hardware complexity of $O(M_{\text{bs}})$. We compare the system power consumption between our system, and the systems using hourglass arrays [124] and OBAs [126][9]. According to [128], the power of a

[9]For a fair comparison, all these systems do not consider using spatial smoothing, and the periodicity of UCAs is considered here.

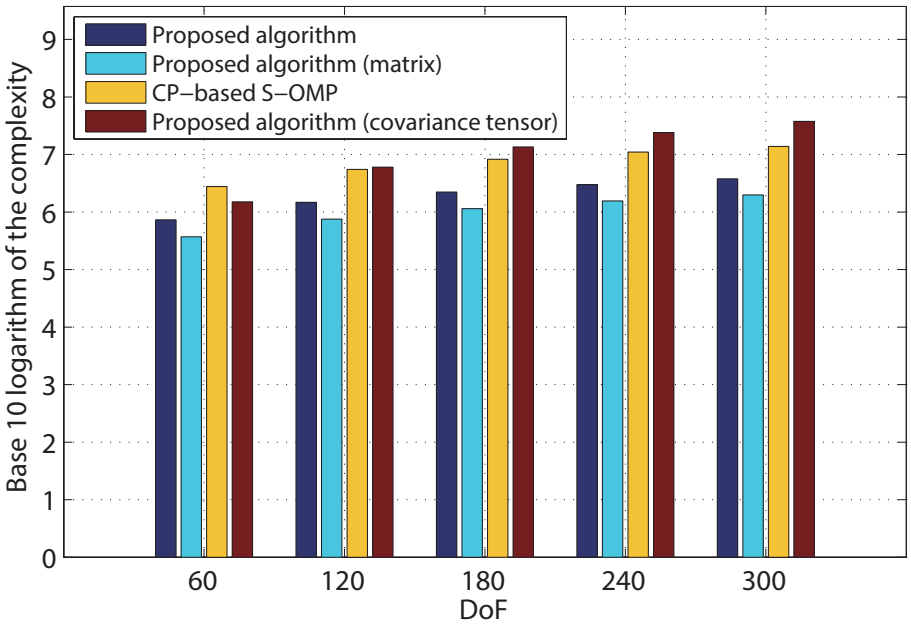


Figure 6.7 : Variation of the software complexity vs. DoF.

hybrid array is consumed by its RF chains, analog-to-digital converters (ADCs), local oscillators, power amplifiers, and phase shifters. Since the proposed method, and the methods using hourglass arrays and OBAs, are only different in terms of the design of RF connection matrices, the numbers of required phase shifters, local oscillators, and power amplifiers are the same across these three methods. As a result, the difference of system power consumption between the methods depends on the numbers of RF chains and ADSs. Also note that the number of RF chains is equal to the number of ADCs. Assume that the dimension of phase-shifter output ports is $M_{\rm hr} \times M_{\rm vr} = 29 \times 17$. In our system, only $M_{\rm rf} = N_{\rm vd}N_{\rm hd} + N_{\rm vs}N_{\rm hs} - 1 = 32$ RF chains (and ADSs) are required by solving (6.12). However, in the systems using hourglass arrays and OBAs, the numbers of required RF chains are 37 and 35, respectively.

As for the signal processing complexity, we compare the computational complexity of the proposed tensor-based algorithm with its matrix-based counterpart, which

formulates the signal model in the matrix form and uses matrix-based ESPRIT-MUSIC algorithm for DoA estimation. For matrix-based algorithms, the computational complexity of performing SVD to the measurement sample matrix and truncating its rank to K is $O(N_{\rm ss}N_{\rm hdc}M_{\rm t}N_{\rm is}K)$. The complexities of estimating the elevation and azimuth angles are $O(K^3 + N_{\rm ss}N_{\rm hdc})$ and $O(N_{\rm hdc}K^2 + N_{\rm hdc}^2KD)$, respectively. D is the size of search dimension. For the proposed tensor-based algorithm, the computational complexity of taking the HOSVD of the tensor model is $O(N_{\rm ss}N_{\rm hdc}M_{\rm t}N_{\rm is}K)$. The computational complexities of estimating elevation and azimuth angles are $O(N_{\rm ss}N_{\rm hdc}+K^3)$ and $O(N_{\rm ss}N_{\rm hdc}M_{\rm t}N_{\rm is}K+N_{\rm hdc}^2KD)$, respectively. The new tensor-based algorithm needs slightly more computations, but is in the same order as its matrix-based counterpart.

We also compare our algorithm with the CP-based simultaneous-orthogonal matching pursuit (S-OMP) algorithm [78]. The algorithm first applies CP decomposition to decompose the received signal tensor model, and then applies S-OMP to estimate the parameters. The complexities of the CP decomposition and S-OMP are $O(N_{\rm ss}N_{\rm hdc}M_{\rm t}N_{\rm is}K+N_{\rm ss}N_{\rm hdc}K^2+K^3)$ and $O(N_{\rm ss}N_{\rm hdc}M_{\rm t}N_{\rm is}(N_1+N_2))$, respectively, where $N_1 \gg K$ and $N_2 \gg K$ are the dimensions of the OMP grid. The complexity of CP-based subspace algorithm is much higher than that of our HOSVD-based algorithm.

Note that all the operations in our DoA estimation algorithm are on the signal data tensor model directly. If our algorithms operate on the signal covariance tensor, we need to calculate the signal covariance tensor model [2, 79, 94]

$$\mathcal{R}_{\rm ss} = \frac{1}{M_{\rm t}N_{\rm is}} \sum_{m=1}^{M_{\rm t}N_{\rm is}} \mathcal{Y}_{{\rm ss},m} \circ \mathcal{Y}_{{\rm ss},m}^{*} \in \mathbb{C}^{N_{\rm ss}\times N_{\rm hdc}\times N_{\rm ss}\times N_{\rm hdc}}, \tag{6.54}$$

where $\mathbf{Y}_{{\rm ss},m} \in \mathbb{C}^{N_{\rm ss}\times N_{\rm hdc}}$ is the m-th subtensor of $\mathcal{Y}_{\rm ss}$, $m = 1, 2, \ldots, M_{\rm t}N_{\rm is}$, and then, take the HOSVD of (6.54). The computational complexity of this process is $O(M_{\rm t}N_{\rm is}N_{\rm ss}^2N_{\rm hdc}^2 + N_{\rm ss}^2N_{\rm hdc}^2K)$, which needs much more computations than our

algorithms.

The results of the computational complexities of our algorithm, its matrix-based counterpart, CP-based S-OMP algorithm, and the proposed algorithm operating on the covariance tensor, as a function of the system DoF, $O(N_{\rm ss}N_{\rm hdc})$, are presented in Fig. 6.7, where $M_{\rm t} = 20$, $N_{\rm is} = 20$, $K = 15$, and $N_1 = N_2 = 50$. The figure shows that our proposed algorithm requires more computations than its matrix-based counterpart at the gain of significantly improved DoA estimation performance, as will be seen in Section 6.6. However, compared to the other two algorithms, the computational complexity of our algorithm is much lower.

6.6 Simulation Results

In this section, simulation results are provided to demonstrate the performance of the proposed algorithm in the mIoT networks. The system bandwidth is $B = 1$ GHz. The number of time frames is set to $M_{\rm t}$= 20. The reference radial frequency $f = 28$ GHz. The vertical spacing between adjacent receiving UCAs is $h = 0.5\lambda$ and the radius of the UCyA is $r = 2\lambda$, where $\lambda = c/f$ and c is the speed of light. The geometry parameters of the UCyA are $M_{\rm v} = 25$ and $M_{\rm h} = 30$. For the hybrid beamforming, we set $N_{\rm vd} = 5$, $N_{\rm hd} = 5$, $N_{\rm vs} = 5$, and $N_{\rm hs} = 6$, so there are $M_{\rm rf} = N_{\rm vd}N_{\rm hd} + N_{\rm vs}N_{\rm hs} - 1 = 54$ RF chains in our system.

Fig. 6.8 plots the root mean square errors (RMSEs) for the estimates of the azimuth and elevation angles versus the average received SNR, where the DoAs of $K = 50$ devices are estimated. By using the proposed nested sparse hybrid beamforming, we compare the proposed HOSVD-based ESPRIT-MUSIC (HB-H-EM) algorithm with its reduced version in the matrix form (HB-H-EM (M)), the CP-based S-OMP (HB-C-SO) algorithm [78], the HOSVD-based 2-D MUSIC (HB-H-2DM) algorithm [94], and the proposed algorithm but using OBA to design the RF connection matrix (HB-H-EM (OBA)). We also apply the proposed algorithm

for fully digital beamforming (DB-H-EM), and provide the CRLB [117] as a reference. We can see that all the estimated algorithms approach the CRLB, as the average received SNR increases. Fig. 6.8 also shows that our proposed tensor-based algorithm provides a better accuracy than its matrix-based counterparts. This is because the tensor-based algorithm can suppress the noise components in each mode of the signal tensor model, while the matrix-based algorithm can only suppress the noise in the time domain corresponding to the third mode in this chapter. By applying CP to decompose the signal tensor model, HB-C-SO achieves better estimation performance than other HOSVD-like algorithms. However, the performance improvement is limited because HB-C-SO uses S-OMP to estimate the parameters, generating quantized estimates only. We also observe that the precision of the angle estimation of our proposed algorithm is a bit lower than that of DB-H-EM. However, both DB-H-EM and HB-C-SO have a much higher complexity than our algorithm, as analyzed in Section 6.5.3. In addition, the estimation accuracy is nearly the same between the proposed HB-H-EM and HB-H-EM (OBA). This is because the DoA estimation accuracy depends on the dimension of the difference coarray, not the dimension of the RF chain network, while the constructed difference coarrays of the two methods are identical.

Fig. 6.9 shows the RMSEs for the estimates of DoAs versus the average received SNR. The number of devices is $K = 200$. By comparing Figs. 6.8 and 6.9, we see that for the fixed SNR and a fixed number of RF chains, the estimation accuracy of all the tested algorithms decreases as K increases. This is because as K grows, more signal components need to be estimated and distinguished. Compared with Fig. 6.8, Fig. 6.9 also shows that the performance degeneration of HB-H-2DM is larger than other algorithms. This is because HB-H-2DM uses signal covariance tensors for the 2-D DoA estimation, and its MUSIC spectrum is a product of multiple separable second-order mode-n spectra, which results in undesirable cross-terms [94]

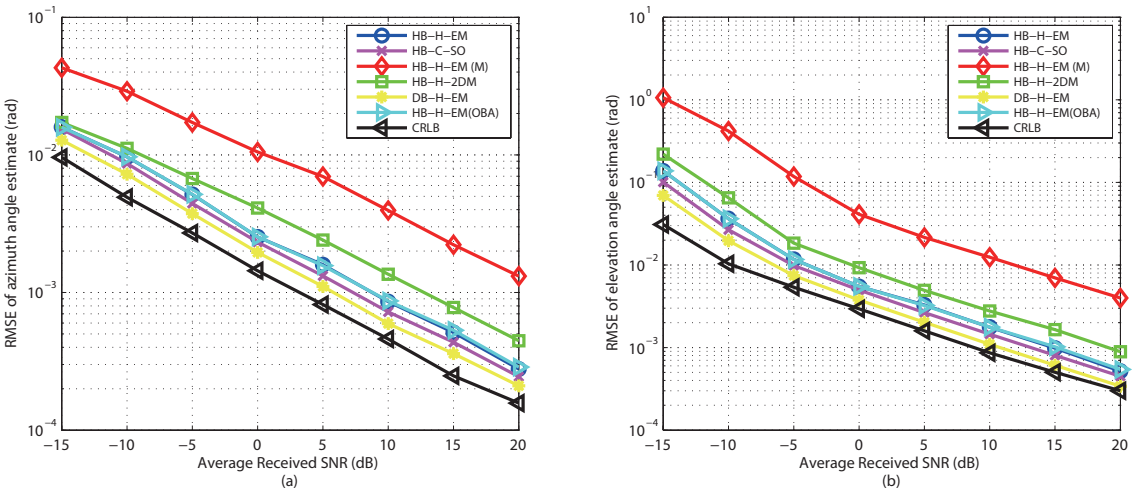


Figure 6.8 : The RMSE vs. the average received SNR for the estimation of DoAs for identifying $K = 50$ devices. (a) Azimuth angle; (b) Elevation angle.

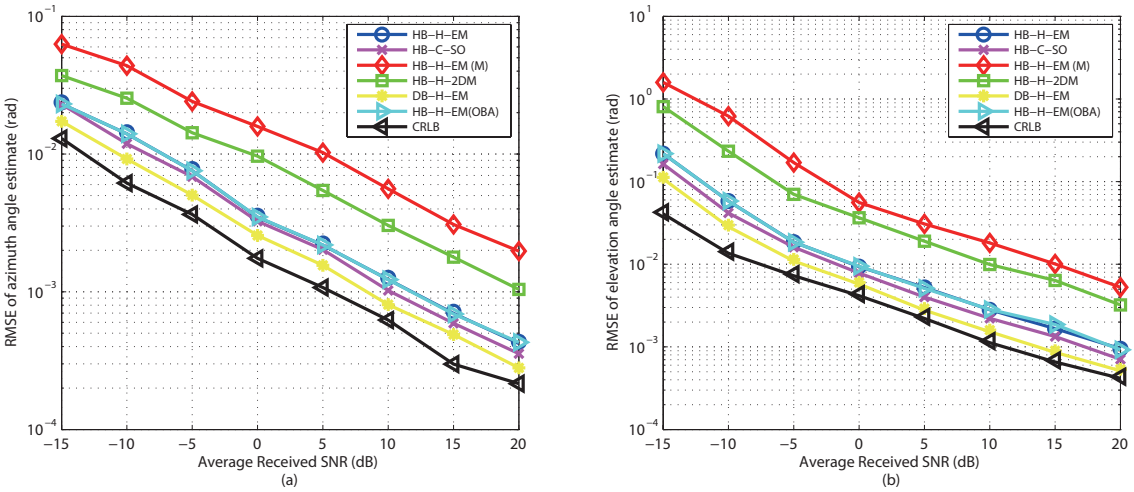


Figure 6.9 : The RMSE vs. the average received SNR for the estimation of DoAs for identifying $K = 200$ devices. (a) Azimuth angle; (b) Elevation angle.

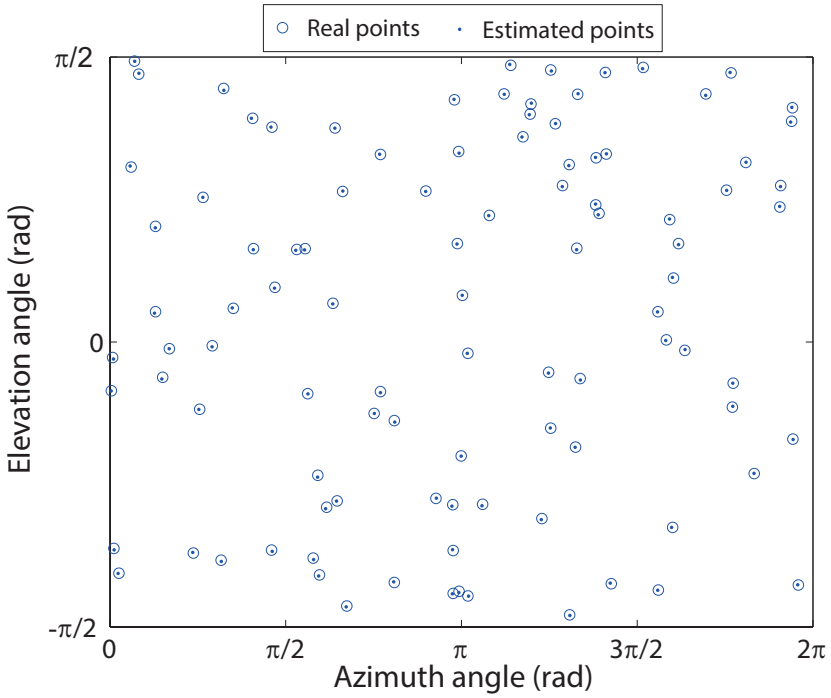


Figure 6.10 : 2-D DoA estimation by using the proposed algorithm for 100 devices.

and compromises the estimation accuracy.

Fig. 6.10 evaluates the performance of our proposed algorithm. Without loss of generality, here we estimate the 2-D DoAs of $K = 100$ devices, where SNR $= 5$ dB. As seen from the results, the proposed algorithm can accurately estimate the azimuth and elevation angles of 100 devices. All the estimates are well matched with the actual values, while we only use 54 RF chains in our system.

6.7 Summary

In this chapter, we have proposed a novel sparse nested hybrid UCyA for mIoT networks. A corresponded 2-D DoA estimation algorithm has also been presented, based on the second-order statistics of the received signals. We have shown that by exploiting the difference coarray technique and tailoring for the UCyA, the designed hybrid array only requires a small number of RF chains to achieve DoA estimation

for a massive number of IoT devices. We have also proposed a spatial smoothing-based method to enhance the n-ranks of the signal tensor model. By using the method, a large enough rank in each mode of the signal tensor model can be provided for the DoA estimation of K devices. By processing the signals in the tensor form and operating on the signal data tensor model directly, we have demonstrated that the proposed DoA estimation algorithm can significantly improve the estimation accuracy while reducing the computational complexity. We have also shown via simulations that, in the mIoT networks, the proposed hybrid array system can accurately estimate the 2-D DoAs of a large number of IoT devices.

Chapter 7

Conclusion and Outlook

In this thesis, we investigated channel parameter estimation methods for mmWave systems with large-scale UCyAs. To meet the different communication and deployment requirements of various scenarios, namely, indoor, UMi, RMa, and mIoT, both digital and hybrid beamformers were considered. Our goal was to find effective solutions which can improve the accuracy of the estimation methods while reducing the system complexity (in both hardware and software). We showed that both mathematics-based and mmWave property-based means can be employed to achieve the target. For example, we reduced the computational complexity of the estimation methods by transforming the signals into a low-dimensional beamspace. We also applied wideband signal-subspace methods to combine the signals across the wideband, so that the high temporal resolution offered by wideband mmWave systems can be exploited to improve delay estimation accuracy.

Chapter 3 was devoted to the channel parameter estimation for the indoor scenario with digital UCyAs. To reduce the computational complexity of the estimation method, we first exploited the sparsity and quasi-optical propagation mmWave and proposed a novel channel compression technique. We showed that by properly quantizing and selecting the received signals, the proposed technique can remain the necessary principal signal components, i.e., the LoS and single-bounce scattering paths, and mitigate the useless secondary components. As a result, the dimension of the received signal space can be reduced while improving the parameter estimation accuracy. We also developed a joint RSS-AoA estimation algorithm, which

exploits the different transmission power of paths to estimate the distance between the BS and the MS, and is based on the beamspace transformation to estimate the arrival angle of each paths. We showed that the beamspace transformation can transform signal vectors in the element space to a low-dimensional beamspace, so that the computational complexity of the parameter estimation can be further reduced. We finally described a novel indoor positioning approach based on the estimated parameters. We demonstrated that by exploiting the quasi-optical propagation at mmWave frequencies, only a single BS is required to implement spatial 3-D localization. This is particularly applicable for the mmWave indoor scenario, where high-accuracy parameter estimation and target localization are required in limited spaces.

Next, we extended the channel parameter estimation to outdoor mmWave communication scenarios, e.g., UMi and RMa, where hybrid UCyAs were considered. The hybrid UCyAs have lower hardware cost and power consumption than the UCyAs with digital front-end, as demonstrated in Chapter 4. We first designed a new two-step wideband hybrid beamforming strategy for the UMi scenario. The strategy reduces the dimension of received signals on the horizontal plane by exploiting the convergence of the Bessel function, and reduces the active beams in the vertical direction through preselection. We showed that by using this strategy, we can preserve the important recurrence relationship of the received signals needed for subspace-based angle and delay estimation, although a dramatically small number of RF chains is deployed. This is particularly useful for the UMi scenario, where large-scale UCyAs are deployed and high costs caused by the beamformers need to be reduced. Then, we generalized the conventional linear interpolation to reconstruct the received signals of the hybrid beamformer. We showed that this method can coherently combine the signals across the whole band and suppress the beam squint. As a result, the subspace-based algorithms can be applied to estimate the angles

and delays of the multi-path components. In Chapter 4, we also introduced a novel multi-parameter matching method. Compared to conventional exhaustive search methods, the proposed method only needs to add perturbation matrices to mitigate the mismatch of the estimated parameters, which has a much lower computational complexity.

In Chapter 5, we discussed the channel parameter estimation for hybrid directional UCyAs. We first designed a hybrid directional beamformer, which employs sweeping directional beamforming and exploits the convergence property of the Bessel function. We showed that the designed hybrid directional beamformer can deal with the severe free-space pathloss of mmWave links and maintain the angular resolution of the hybrid UCyA with a reduced number of RF chains. We then presented a new UCAMI to suppress the beam squint effect. We demonstrated that the proposed new UCAMI only needs to optimize the focusing matrices in the elevation angular domain, which has a much lower complexity than the conventional UCAMI [73], in which computationally expensive multi-dimensional optimization problems need to be solved. In Chapter 5, we also introduced the tensor-based signal processing method and employ it for the designed hybrid directional UCyAs. We showed that by employing tensor signal processing for channel parameter estimation, the receiver noises in all dimensions (time, frequency, and space) can be suppressed. As a result, the proposed tensor-based subspace estimation algorithm is able to achieve substantially higher estimation accuracy than existing matrix-based algorithms, with a comparable computational complexity in the RMa scenario.

In Chapter 6, we focused on the mIoT networks. We studied the sparse array-based hybrid beamformer and its corresponding 2-D angle estimation algorithm. We first introduced the sparse array technique, which is conventionally used to improve the system DoF with a limited number of antennas. We applied this technique to the hybrid beamforming design, and showed that the designed sparse hybrid beamform-

ing is capable of providing larger DoF than the fully connected beamforming array with much fewer RF chains. We also employed tensor signal processing in Chapter 6 and proposed a new spatial smoothing-based method to ensure that the signal and noise spaces in all modes of the signal tensor model can be decomposed. We evaluated the impact of the rank of the signal power matrix on the n-ranks of the signal tensor model, and verified that by using the proposed method, we can construct a signal tensor model that provides a large enough rank in each mode to perform the DoA estimation of all devices, even when the received signal powers of all the devices are equal. Tailored for the new hybrid UCyA array, we then proposed a new tensor-based 2-D DoA estimation algorithm, which is based on the second-order statistics of the signals and operates on the signal data model directly. Corroborated by the simulation results and the complexity analysis, we demonstrated that the algorithm can improve the estimation accuracy with an affordable computational complexity. This is particularly useful for the mIoT networks, where the high network access capability is required to support the connection of a large number of devices, and the requirement of accurate channel parameter estimation is also needed to be achieved while guaranteeing a low computational complexity.

Overall, this thesis provided practical design schemes for large-scale UCyAs and proposed the corresponding efficient channel parameter estimation algorithms. The presented hybrid UCyAs can significantly reduce the number of required RF chains with a marginal accuracy loss, and the proposed algorithms are capable of accurately estimating the channel parameters with low computational complexities. In this thesis, there are also some important scientific-technical findings. For example, we find that the number of required phase-mode vectors does not depend on the number of array response vectors, which is important for hardware complexity reduction (see Chapter 4). We also find that when emplying tensor signal processing, the ranks of signal spaces in all dimensions are dependent on the rank of the received signal.(see

Chapter 6).

The methods presented in this thesis have addressed the topics of high hardware and software complexities in mmWave large-scale antenna array systems, but several issues are yet to be considered and will be our future research directions. First of all, it is necessary to investigate the proposed channel estimation and positioning methods in some more realistic scenarios. In this thesis, we only provide the numerical simulations to demonstrate the proposed methods. We assume that the communication scenarios are simple, where all the transmitted signals can be received by the BSs. However, practical scenarios are complicated and the antenna arrays are not perfectly manufactured. In realistic scenarios, we cannot guarantee that the received noises are AWGNs, the transmitted signal power of the omnidirectional antennas is the same, the beams of antenna arrays can be precisely controlled, and so on. And regarding to NLOS paths, the single-bounce specular reflection components of mmWave signals would not be specular reflections. To improve and evaluate the practicalities of our work, we plan to construct testbeds and make experiments to demonstrate the proposed methods in the future. Specifically, we will construct a simple multi-antenna Universal Software Radio Peripheral (USRP) experimental system firstly, and then, develop it by mmWave large antennas. To obtain the explicit estimates of channel parameters and target positions in practical NLoS scenarios, some efficient and easy-implemented environment reconstruction methods will also be investigated.

Secondly, finding effective ways to address the source enumeration problem is important. The subspace-based channel parameter estimation algorithms, such as ESPRIT and MUSIC [15, 29–31], have a much lower computational complexity than the ML-based estimation algorithms [36], and can also achieve attractive parameter estimation performance. However, the high accurate estimation provided by the subspace-based channel parameter estimation algorithms relies on the correct

decomposition of the signal and noise subspaces. The process of subspace decomposition needs the a-priori knowledge of the dimensions of the signal and noise subspaces, which depends on the number of incoherent signals. Currently, many methods, based on Information Theoretic criteria, gerschgorin disks, matrix decomposition, etc., have been proposed to determine the number of incoherent sources, but these methods cannot achieve acceptable estimation performance in low SNR regions. In this sense, effective source enumeration methods has been a necessity.

Finally, we also plan to extend our work to antenna arrays with arbitrary architectures. In B5G and 6G wireless communication systems, cell-free and reconfigurable antenna arrays are promising for wide applications. To improve the practicability and applicability of the proposed methods, we will design universal applicable channel estimation methods adopted for B5G/6G mmWave communication systems with general applicability to a wide range of antenna array types and specifications.

Appendices

A Properties of Tensor Operation

The important properties of tensor operations used in this chapter are provided below.

Property 1. *The n-mode product satisfies the following properties:*

$$\mathcal{A} \times_n \mathbf{B} \times_n \mathbf{C} = \mathcal{A} \times_n (\mathbf{CB}); \tag{A.1}$$

$$\mathcal{A} \times_n \mathbf{B} \times_m \mathbf{D} = \mathcal{A} \times_m \mathbf{D} \times_n \mathbf{B}, \tag{A.2}$$

where $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, $\mathbf{B} \in \mathbb{C}^{J_n \times I_n}$, $\mathbf{C} \in \mathbb{C}^{K_n \times J_n}$, *and* $\mathbf{D} \in \mathbb{C}^{J_m \times I_m}$ $(n, m = 1, 2, \ldots, N$ *and* $n \neq m)$.

Property 2. *The Tucker decomposition decomposes a tensor* $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ *into a core tensor* $\mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times \cdots \times R_N}$ *multiplied by a factor matrix* $\mathbf{C}^{(n)} = [\mathbf{c}^{(n)}_{r_n=1}, \mathbf{c}^{(n)}_{r_n=2}, \ldots, \mathbf{c}^{(n)}_{r_n=R_n}] \in \mathbb{C}^{I_n \times R_n}$ $(\mathbf{c}^{(n)}_{r_n} \in \mathbb{C}^{I_n \times 1}$ *and* $n = 1, 2, \ldots, N)$ *in each mode, i.e.,*

$$\begin{aligned} \mathcal{A} &= \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} g_{r_1 r_2 \cdots r_N} \left(\mathbf{c}^{(1)}_{r_1} \circ \mathbf{c}^{(2)}_{r_2} \circ \cdots \mathbf{c}^{(N)}_{r_N} \right) \\ &= \left[\!\left[\mathcal{G}; \mathbf{C}^{(1)}, \mathbf{C}^{(2)}, \ldots, \mathbf{C}^{(N)} \right]\!\right]. \end{aligned} \tag{A.3}$$

The higher-order singular value decomposition (HOSVD) is a special case of the Tucker decomposition, where the core tensor is all-orthogonal [76], and the factor matrices are the unitary left singular matrices of the mode-n unfolding of $\mathcal{A}$.

Property 3. *The CANDECOMP/PARAFAC (CP) decomposition decomposes a tensor* $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ *into a sum of rank-one component tensors* $\mathbf{b}^{(n)}_r \in \mathbb{C}^{I_n}$, *as*

given by

$$\mathcal{A} = \sum_{r=1}^{R} \lambda_r \mathbf{b}_r^{(1)} \circ \mathbf{b}_r^{(2)} \circ \cdots \mathbf{b}_r^{(N)}, \tag{A.4}$$

where $R = Rank(\mathcal{A})$ *is the rank of* $\mathcal{A}$[1]. *Following [76], CP can be viewed as the special case of the Tucker decomposition, where the core tensor is superdiagonal. Thus, the CP model in* (A.4) *can be rewritten as a multilinear product:*

$$\mathcal{A} = \mathcal{D} \times_1 \mathbf{B}^{(1)} \times_2 \mathbf{B}^{(2)} \cdots \times_N \mathbf{B}^{(N)} = \left[\!\left[\mathcal{D}; \mathbf{B}^{(1)}, \mathbf{B}^{(2)}, \ldots, \mathbf{B}^{(N)}\right]\!\right], \tag{A.5}$$

where $\mathbf{B}^{(n)} = \left[\mathbf{b}_1^{(n)}, \mathbf{b}_2^{(n)}, \ldots, \mathbf{b}_R^{(n)}\right] \in \mathbb{C}^{J_n \times R}$ *is the factor matrix of* $\mathbf{b}_r^{(n)}$, *and* $\mathcal{D} \in \mathbb{C}^{R \times R \times \cdots \times R}$ *is a superdiagonal tensor*[2] *with* $d_{r,r,\cdots,r} = \lambda_r$.

Property 4. *The multilinear product of a tensor* $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ *with matrices* $\mathbf{B}^{(n)} \in \mathbb{C}^{J_n \times I_n}$, $n = 1, 2, \ldots, N$, *is a sequence of contractions, each being an n-mode product, i.e.,*

$$\mathcal{C} = \mathcal{A} \times_1 \mathbf{B}^{(1)} \times_2 \mathbf{B}^{(2)} \cdots \times_N \mathbf{B}^{(N)} \in \mathbb{C}^{J_1 \times J_2 \times \cdots \times J_N}. \tag{A.6}$$

Its mode-n unfolding is given by

$$\mathbf{C}_{(n)} = \mathbf{B}^{(n)} \mathbf{A}_{(n)} (\mathbf{B}^{(n+1)} \otimes \mathbf{B}^{(n+2)} \otimes \cdots \otimes \mathbf{B}^{(N)} \otimes \mathbf{B}^{(1)} \otimes \mathbf{B}^{(2)} \otimes \cdots \otimes \mathbf{B}^{(n-1)})^T. \tag{A.7}$$

[1]The rank of a tensor, $\mathcal{A}$, denoted Rank($\mathcal{A}$), is defined as the smallest number of rank-one tensors that yield $\mathcal{A}$ in a linear combination [76].

[2]A tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is diagonal if $a_{i_1 i_2 \cdots i_N} \neq 0$ only if $i_1 = i_2 = \cdots = i_N$. When $I_1 = I_2 = \cdots = I_N$, $\mathcal{A}$ is called as superdiagonal.

B Proofs and derivations for Chapter 3

B.1 Proof of the asymptotic property in Chapter 3 Section 3.4.2

Substituting (3.29)-(3.32) into (3.33), we have

$$f_{p,k}(\tilde{\phi}^{\mathrm{r}}_{l',\xi},\tilde{\theta}^{\mathrm{r}}_{l',\xi}) = \frac{1}{N}\sum_{n=1}^{N} e^{jp\varphi_n} e^{j\frac{2\pi}{\lambda}r\sin(\bar{\theta}^{\mathrm{r}}_{\xi})\cos(\bar{\phi}^{\mathrm{r}}_{\xi}-\varphi_n)} \times e^{-j\frac{2\pi}{\lambda}h(k-1)\cos(\tilde{\theta}^{\mathrm{r}}_{l',\xi})}. \tag{B.8}$$

When the number of antennas in each UCA tends to infinity, i.e., $N \to \infty$, (B.8) is reformulated as

$$\begin{aligned} f_{p,k}(\tilde{\phi}^{\mathrm{r}}_{l',\xi},\tilde{\theta}^{\mathrm{r}}_{l',\xi}) &= \frac{1}{2\pi}\int_{0}^{2\pi} e^{jp\varphi_n} e^{j\frac{2\pi}{\lambda}r\sin(\bar{\theta}^{\mathrm{r}}_{\xi})\cos(\bar{\phi}^{\mathrm{r}}_{\xi}-\varphi_n)} \\ &\quad \times e^{-j\frac{2\pi}{\lambda}h(k-1)\cos(\tilde{\theta}^{\mathrm{r}}_{l',\xi})}\mathrm{d}\varphi \\ &= j^p J_p\left(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right) e^{j\left[p\tilde{\phi}^{\mathrm{r}}_{l',\xi}-\frac{2\pi}{\lambda}h(k-1)\cos(\tilde{\theta}^{\mathrm{r}}_{l',\xi})\right]}. \end{aligned} \tag{B.9}$$

Compared (3.36) with (B.9), it can be found that the residual term $\varepsilon_p(\frac{2\pi}{\lambda}r\sin(\tilde{\theta}^{\mathrm{r}}_{l',\xi}), \tilde{\phi}^{\mathrm{r}}_{l',\xi}) \to 0$, when $N \to \infty$.

B.2 Illustration of two likely positions of each path in Chapter 3 Section 3.5.1

Without loss of generality, we use the Fig. B.1, which is a top view of one of the NLoS paths, for illustration. Based on the geometry between the positions of BS and MS shown in Fig. B.1, we can establish the following equations

$$\begin{cases} y^2 + x^2 = r'^2 \\ y = D\tan\phi + (D-x)\tan\phi \end{cases}. \tag{B.10}$$

It is obvious that $r' = \hat{d}\sin(\pi - \hat{\theta}^{\mathrm{r}})$ and $\phi = 2\pi - \hat{\phi}^{\mathrm{r}}$, where $\hat{\theta}^{\mathrm{r}}$, $\hat{\phi}^{\mathrm{r}}$, and $\hat{d}$ are the results of ranging and angle estimation. Because there is a quadratic equation, we

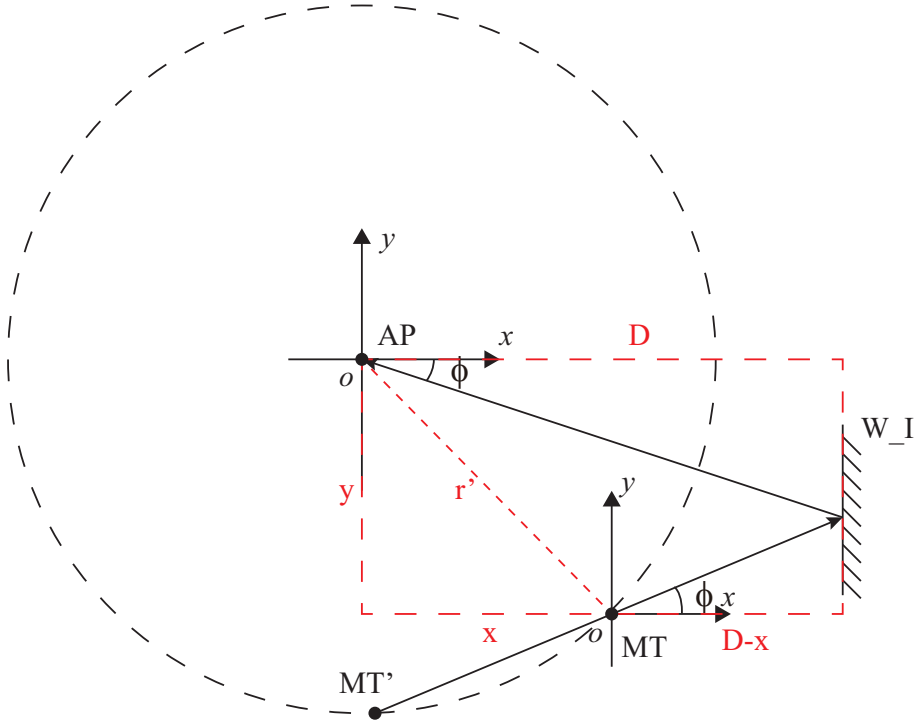


Figure B.1 : Top view of the MS and BS.

can have two solutions. From Fig. B.1, we find that both the MS and the MS' have the same $\hat{\theta}^{\mathrm{r}}$, $\hat{\phi}^{\mathrm{r}}$, and r', where r' is the horizontal distance between the MS and the BS. Thus, the solutions of the above equations are given by

$$\begin{cases} \hat{x}_{\mathrm{MT}} = \dfrac{2D\tan^2(\hat{\phi}^{\mathrm{r}}) + \sqrt{\hat{d}^2\sin^2(\hat{\theta}^{\mathrm{r}})(\tan^2(\hat{\phi}^{\mathrm{r}})+1) - 4D^2\tan^2(\hat{\phi}^{\mathrm{r}})}}{\tan^2(\hat{\phi}^{\mathrm{r}})+1} \\ \hat{y}_{\mathrm{MT}} = \dfrac{\left(2D - \sqrt{\hat{d}^2\sin^2(\hat{\theta}^{\mathrm{r}})(\tan^2(\hat{\phi}^{\mathrm{r}})+1) - 4D^2\tan^2(\hat{\phi}^{\mathrm{r}})}\right)\tan(\hat{\phi}^{\mathrm{r}})}{\tan^2(\hat{\phi}^{\mathrm{r}})+1} \end{cases} \tag{B.11}$$

and

$$\begin{cases} \hat{x}_{\mathrm{MT}'} = \dfrac{-2D\tan^2(\hat{\phi}^{\mathrm{r}}_1) + \sqrt{\hat{d}^2\sin^2(\hat{\theta}^{\mathrm{r}})(\tan^2(\hat{\phi}^{\mathrm{r}})+1) - 4D^2\tan^2(\hat{\phi}^{\mathrm{r}})}}{\tan^2(\hat{\phi}^{\mathrm{r}})+1} \\ \hat{y}_{\mathrm{MT}'} = \dfrac{\left(2D + \sqrt{\hat{d}^2\sin^2(\hat{\theta}^{\mathrm{r}})(\tan^2(\hat{\phi}^{\mathrm{r}})+1) - 4D^2\tan^2(\hat{\phi}^{\mathrm{r}})}\right)\tan(\hat{\phi}^{\mathrm{r}})}{\tan^2(\hat{\phi}^{\mathrm{r}})+1} \end{cases} . \tag{B.12}$$

It is obvious that we can obtain two likely positions, i.e., a "true" one and a "false" one. Due to estimation error, all "true" positions have different but adjacent coordinates. Finally, we use clustering algorithms to fuse the "true" positions.

B.3 Derivation of the CRLB in Chapter 3 Section 3.6

The i-th array response vector of the k-th UCA in (3.29) can be approximated as [129]

$$\begin{aligned} a_i^{\mathrm{r}}(\tilde{\phi}_{l',\xi}^{\mathrm{r}}, \tilde{\theta}_{l',\xi}^{\mathrm{r}}) \approx\ & e^{j\frac{2\pi}{\lambda}\left[r\sin(\bar{\theta}_\xi^{\mathrm{r}})\cos(\bar{\phi}_\xi^{\mathrm{r}}-\varphi_n)-h(k-1)\cos(\bar{\theta}_\xi^{\mathrm{r}})\right]} \\ & \times e^{j\frac{2\pi}{\lambda}\Delta_{\tilde{\theta}_{l',\xi}^{\mathrm{r}}}\left[r\cos(\bar{\theta}_\xi^{\mathrm{r}})\cos(\bar{\phi}_\xi^{\mathrm{r}}-\varphi_n)+h(k-1)\sin(\bar{\theta}_\xi^{\mathrm{r}})\right]} \\ & \times e^{j\frac{2\pi}{\lambda}\Delta_{\tilde{\phi}_{l',\xi}^{\mathrm{r}}}\left[-r\sin(\bar{\theta}_\xi^{\mathrm{r}})\sin(\bar{\phi}_\xi^{\mathrm{r}}-\varphi_n)\right]}. \end{aligned} \tag{B.13}$$

Using the above approximation, the covariance matrix of the received signal vector in (3.27) can be calculated as

$$\mathbf{R}_{\dot{\mathbf{y}}} = \mathbb{E}\left\{\dot{\mathbf{y}}(t)\dot{\mathbf{y}}^H(t)\right\} \approx \sum_{\xi=1}^{4}\sigma_\xi^2\Xi_\xi + \sigma_{\mathrm{n}}^2\mathbf{I}_{N_{\mathrm{r}}}, \tag{B.14}$$

where $\sigma_\xi^2 = \sigma_{M_\xi}^2\sigma_{\beta_\xi}^2$. Ξ_ξ can be written as

$$\Xi_\xi = \left(\tilde{\mathbf{a}}^{\mathrm{r}}(\bar{\phi}_\xi^{\mathrm{r}}, \bar{\theta}_\xi^{\mathrm{r}})\tilde{\mathbf{a}}^{\mathrm{r}H}(\bar{\phi}_\xi^{\mathrm{r}}, \bar{\theta}_\xi^{\mathrm{r}})\right) \odot B_\xi = D_\xi B_\xi D_\xi^H, \tag{B.15}$$

where $D_\xi = \mathrm{diag}\left(\tilde{\mathbf{a}}^{\mathrm{r}}(\bar{\phi}_\xi^{\mathrm{r}}, \bar{\theta}_\xi^{\mathrm{r}})\right)$. Each element of B_ξ is

$$[B_\xi]_{n_{\mathrm{r}1},n_{\mathrm{r}2}} = \exp\left(-\frac{2\pi^2}{\lambda^2}\left(\varrho_{\xi,n_{\mathrm{r}1},n_{\mathrm{r}2}}^2 + \Upsilon_{\xi,n_{\mathrm{r}1},n_{\mathrm{r}2}}^2\right)\right), \tag{B.16}$$

where

$$\begin{aligned} \varrho_{\xi,n_{\mathrm{r}1},n_{\mathrm{r}2}} = \sigma_{\tilde{\theta}_\xi^{\mathrm{r}}}\big[r\cos(\bar{\theta}_\xi^{\mathrm{r}})\big(&\cos(\bar{\phi}_\xi^{\mathrm{r}}-\varphi_{n_{\mathrm{r}1}}) \\ &-\cos(\bar{\phi}_\xi^{\mathrm{r}}-\varphi_{n_{\mathrm{r}2}})\big) + h(k_1-k_2)\sin(\bar{\theta}_\xi^{\mathrm{r}})\big], \end{aligned} \tag{B.17}$$

$\Upsilon_{\xi,n_{\mathrm{r}1},n_{\mathrm{r}2}} = \sigma_{\tilde{\phi}_\xi^{\mathrm{r}}} r\sin(\bar{\theta}_\xi^{\mathrm{r}})\left(\sin(\bar{\phi}_\xi^{\mathrm{r}}-\varphi_{n_{\mathrm{r}1}}) - \sin(\bar{\phi}_\xi^{\mathrm{r}}-\varphi_{n_{\mathrm{r}2}})\right)$, $n_{\mathrm{r}1} = N(k_1-1)+n_1$, and $n_{\mathrm{r}2} = N(k_2-1)+n_2$. We have $\sigma_{\tilde{\theta}_\xi^{\mathrm{r}}}^2 = \frac{1}{N_{\mathrm{p}}'}\sum_{l'=1}^{N_{\mathrm{p}}'}\Delta_{\tilde{\theta}_{l',\xi}^{\mathrm{r}}}^2$, $\sigma_{\tilde{\phi}_\xi^{\mathrm{r}}}^2 = \frac{1}{N_{\mathrm{p}}'}\sum_{l'=1}^{N_{\mathrm{p}}'}\Delta_{\tilde{\phi}_{l',\xi}^{\mathrm{r}}}^2$ $n_1 = 1,2,\ldots,N$, $n_2 = 1,2,\ldots,N$, $k_1 = 1,2,\ldots,K$ known as the azimuth and elevation angular spreads. By defining $\mathbf{u} \triangleq \left[\mathbf{u}_{\bar{\phi}^{\mathrm{r}}}^T, \mathbf{u}_{\bar{\theta}^{\mathrm{r}}}^T\right]^T$, $\mathbf{v} = [\sigma_1^2,\ldots,\sigma_4^2,\sigma_{\mathrm{n}}^2]$, where $\mathbf{u}_{\bar{\phi}^{\mathrm{r}}} = \left[\bar{\phi}_1^{\mathrm{r}}, \bar{\phi}_2^{\mathrm{r}}, \bar{\phi}_3^{\mathrm{r}}, \bar{\phi}_4^{\mathrm{r}}\right]^T$and $\mathbf{u}_{\bar{\theta}^{\mathrm{r}}} = \left[\bar{\theta}_1^{\mathrm{r}}, \bar{\theta}_2^{\mathrm{r}}, \bar{\theta}_3^{\mathrm{r}}, \bar{\theta}_4^{\mathrm{r}}\right]^T$, the Fisher information matrix can be expressed as

$$[\mathbf{J}_{\chi,\chi}]_{f,f'} = T_s\mathrm{tr}\left(\mathbf{R}_{\dot{\mathbf{y}}}^{-1}\frac{\partial\mathbf{R}_{\dot{\mathbf{y}}}}{\partial[\chi]_f}\mathbf{R}_{\dot{\mathbf{y}}}^{-1}\frac{\partial\mathbf{R}_{\dot{\mathbf{y}}}}{\partial[\chi]_{f'}}\right), \tag{B.18}$$

where f and $f' = 1, 2, \ldots, 21$. The partial derivatives in (B.18), i.e., $\partial \mathbf{R}_{\dot{\mathbf{y}}}/\partial \bar{\phi}_{\xi}^{\mathrm{r}}$, $\partial \mathbf{R}_{\dot{\mathbf{y}}}/\partial \bar{\theta}_{\xi}^{\mathrm{r}}$, $\partial \mathbf{R}_{\dot{\mathbf{y}}}/\partial \sigma_{\xi}^{2}$, and $\partial \mathbf{R}_{\dot{\mathbf{y}}}/\partial \sigma_{\mathrm{n}}^{2}$, can be obtained based on (B.14) and (B.15). $\mathbf{J}_{\mathbf{u},\mathbf{u}}$, $\mathbf{J}_{\mathbf{u},\mathbf{v}}$, and $\mathbf{J}_{\mathbf{v},\mathbf{v}}$ can be defined similarly to $\mathbf{J}_{\chi,\chi}$ as shown in (B.18) and they are related as

$$\mathbf{J}_{\chi,\chi} = \begin{bmatrix} \mathbf{J}_{\mathbf{u},\mathbf{u}} & \mathbf{J}_{\mathbf{u},\mathbf{v}} \\ \mathbf{J}_{\mathbf{u},\mathbf{v}}^{T} & \mathbf{J}_{\mathbf{v},\mathbf{v}} \end{bmatrix}. \tag{B.19}$$

Invoking the block matrix inversion lemma, the CRLB concerning the covariance matrix of the error of the estimated signal parameter vector $\mathbf{u}$ is obtained as

$$\mathbf{C} = \left(\mathbf{J}_{\mathbf{u},\mathbf{u}} - \mathbf{J}_{\mathbf{u},\mathbf{v}}\mathbf{J}_{\mathbf{v},\mathbf{v}}^{-1}\mathbf{J}_{\mathbf{u},\mathbf{v}}^{T}\right)^{-1}, \tag{B.20}$$

which implies

$$\mathbb{E}\left\{(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^{T}\right\} \geq \mathbf{C}. \tag{B.21}$$

C Proofs and derivations for Chapter 4

C.1 Proof of Lemma 1 in Chapter 4 Section 4.3.2

According to the property of Bessel function, i.e., $J_{-v}(x) = (-1)^v J_v(x)$, we have $|J_{-v}(x)| = |J_v(x)|$, so here we only use $J_v(x)$ with $v \in \mathbb{Z}^+$ for illustration convenience. Let $x = v\rho, \rho \in (0, 1]$. The Bessel function, $J_v(x)$, whose order v exceeds its argument, x, can be written in the form [109]

$$J_v(v\rho) = \frac{1}{\pi} \int_0^\pi \exp\left(-vF(\vartheta, \rho)\right) d\vartheta, \tag{C.22}$$

where

$$\begin{aligned} F(\vartheta, \rho) = &\log\left(\frac{\vartheta + \sqrt{\vartheta^2 - \rho^2 \sin^2 \vartheta}}{\rho \sin \vartheta}\right) \\ &- \cot \vartheta \sqrt{\vartheta^2 - \rho^2 \sin^2 \vartheta}. \end{aligned} \tag{C.23}$$

The partial derivative of (C.22) with respect to ρ is calculated as

$$\begin{aligned} \frac{\partial}{\partial \rho} J_v(v\rho) &= -\frac{v}{\pi} \int_0^\pi \frac{\partial F(\vartheta, \rho)}{\partial \rho} \exp\left(-vF(\vartheta, \rho)\right) d\vartheta \\ &= \frac{v}{\pi\rho} \int_0^\pi g(\vartheta, \rho) \exp\left(-vF(\vartheta, \rho)\right) d\vartheta, \end{aligned} \tag{C.24}$$

where $g(\vartheta, \rho) = \left(\vartheta - \rho^2 \sin \vartheta \cos \vartheta\right) / \sqrt{\vartheta^2 - \rho^2 \sin^2 \vartheta}$. Considering that

$$\begin{aligned} g(\vartheta, \rho) &= \frac{\vartheta - \rho^2 \sin \vartheta \cos \vartheta}{\sqrt{\vartheta^2 - \rho^2 \sin^2 \vartheta}} \geq \frac{\vartheta - \sin \vartheta \cos \vartheta}{\sqrt{\vartheta^2 - \rho^2 \sin^2 \vartheta}} \\ &\geq \frac{\vartheta - \sin \vartheta}{\sqrt{\vartheta^2 - \rho^2 \sin^2 \vartheta}} \geq 0, \end{aligned} \tag{C.25}$$

we have $\partial J_v(v\rho)/\partial \rho > 0$, and conclude that $J_v(v\rho)$ is a positive increasing function of ρ. Thus, $J_v(v\rho) < J_v(v)$.

On the other hand, the partial derivative of (C.22) with respect to v is calculated as

$$\frac{\partial}{\partial v} J_v(v\rho) = -\frac{1}{\pi} \int_0^\pi F(\vartheta, \rho) \exp\left(-vF(\vartheta, \rho)\right) d\vartheta. \tag{C.26}$$

Because

$$\frac{\partial}{\partial\vartheta}F(\vartheta,\rho)=\frac{(1-\rho\cot\vartheta)^2}{\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}}+\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}\geq 0 \tag{C.27}$$

and $\partial F(0,\rho)/\partial\rho=-\sqrt{1-\rho^2}/\rho\leq 0$, we have $F(\vartheta,\rho)\geq F(0,\rho)\geq F(0,1)=0$, and hence $\partial J_v(v\rho)/\partial v<0$. This means that $J_v(v\rho)$ is a positive decreasing function of v, i.e., $J_v(v\rho)\leq J_1(\rho)$. Therefore, we have $J_v(v\rho)<J_v(v)\leq J_1(1)\approx 0.44$ with $\rho\in(0,1]$ and $v\in\mathbb{Z}^+$. When $|v|>|x|$, $|J_v(x)|\approx 0$, $v\in\mathbb{Z}$.

C.2 Proof of Theorem 1 in Chapter 4 Section 4.3.2

According to Lemma 1, we observe that $J_p(\varpi_{m,l})$ cannot be omitted if $|p|\leq|\varpi_{m,l}|=|2\pi f_m r\sin(\theta_{\mathrm{R},l})/c|\leq 2\pi f_m r/c$. Because $f_0\leq f_m$ and $p\in\mathbb{Z}$, we set the highest order $P=\max(|p|)=\lfloor 2\pi f_0 r/c\rfloor$.

On the other hand, in the case of $Q\neq 0$, because $p\in[-P,P]\cap\mathbb{Z}$ and $N_{\mathrm{H}}\geq 2P$, we have $|p-QN_{\mathrm{H}}|\geq|\varpi_{m,l}|$. According to Lemma 1, we obtain

$$\begin{aligned}
&|\varepsilon_{p,Q}(\varpi_{m,l},\phi_{\mathrm{R},l})|\\
&=\left|j^{(QN_{\mathrm{H}}-p)}J_{(QN_{\mathrm{H}}-p)}(\varpi_{m,l})\exp\left(j(QN_{\mathrm{H}}-p)\phi_{\mathrm{R},l}\right)\right|\\
&=\left|J_{(p-QN_{\mathrm{H}})}(\varpi_{m,l})\right|\approx 0.
\end{aligned} \tag{C.28}$$

In this case, (4.17) can be approximated by

$$\begin{aligned}
A_{\mathrm{PM},p}&=\sqrt{N_{\mathrm{H}}}\left[j^pJ_p(\varpi_{m,l})e^{-jp\phi_{\mathrm{R},l}}\right.\\
&\quad\left.+\sum_{Q=-\infty,Q\neq 0}^{\infty}\varepsilon_{p,Q}(\varpi_{m,l},\phi_{\mathrm{R},l})\right]\\
&\approx\sqrt{N_{\mathrm{H}}}j^pJ_p(\varpi_{m,l})e^{-jp\phi_{\mathrm{R},l}}.
\end{aligned} \tag{C.29}$$

This concludes the proof of Theorem 1.

D Proofs and derivations for Chapter 5

D.1 Proof of Theorem 2 in Chapter 5 Section 5.3

Let $\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}) = \frac{2\pi}{c} f_{m_\mathrm{f}} r \sin(\theta_{k_{m_\mathrm{b}}})$. The Q-DFT of $a_{\mathrm{h},m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}, \phi_{k_{m_\mathrm{b}}})$ can be expressed as

$$\begin{aligned}
a_{\mathrm{QDFT},p,m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}, \phi_{k_{m_\mathrm{b}}}) &= \sum_{m_\mathrm{h}=1}^{M_\mathrm{h}} \left(\frac{1}{\sqrt{M_\mathrm{h}}} e^{j\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})\cos(\phi_{k_{m_\mathrm{b}}} - \varphi_{m_\mathrm{h}})} \right) e^{-j\frac{2\pi(m_\mathrm{h}-1)}{M_\mathrm{h}}p} \\
&\overset{\text{(a)}}{=} \sum_{m_\mathrm{h}=1}^{M_\mathrm{h}} \left(\frac{1}{\sqrt{M_\mathrm{h}}} \sum_{q=-\infty}^{\infty} j^q J_q(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) e^{jq(\phi_{k_{m_\mathrm{b}}} - \varphi_{m_\mathrm{h}})} \right) e^{-j\frac{2\pi(m_\mathrm{h}-1)}{M_\mathrm{h}}p} \\
&\overset{\text{(b)}}{=} \frac{1}{\sqrt{M_\mathrm{h}}} \sum_{Q=-\infty}^{\infty} M_\mathrm{h} j^{(QM_\mathrm{h}-p)} J_{(QM_\mathrm{h}-p)}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) e^{j(QM_\mathrm{h}-p)\phi_{k_{m_\mathrm{b}}}} \\
&\overset{\text{(c)}}{=} \sqrt{M_\mathrm{h}} [j^p J_p(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) e^{-jp\phi_{k_{m_\mathrm{b}}}} + \sum_{Q=-\infty, Q\neq 0}^{\infty} \varepsilon_{p,Q}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}), \phi_{k_{m_\mathrm{b}}})]
\end{aligned} \tag{D.30}$$

where

$$\varepsilon_{p,Q}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}), \phi_{k_{m_\mathrm{b}}}) = j^{(QM_\mathrm{h}-p)} J_{(QM_\mathrm{h}-p)}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) e^{j(QM_\mathrm{h}-p)\phi_{k_{m_\mathrm{b}}}}. \tag{D.31}$$

In (D.30), (a) and (c) follow the important properties of the Bessel function, i.e., $e^{jx\cos y} = \sum_{v=-\infty}^{\infty} j^v J_v(x) e^{jvy}$ and $J_{-v}(x) = (-1)^v J_v(x)$, respectively; (b) is obtained by letting $p + q = QM_\mathrm{h}$; and (c) stems from the property of the Bessel function $J_{-v}(x) = (-1)^v J_v(x)$ [109].

Consider that the number of antennas per UCA, M_h, is large, i.e, $M_\mathrm{h} \gg P$. Let $M_\mathrm{h} = \alpha P$ and $\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}) = \beta P$, where $\alpha \gg 1$ and $0 < \beta < 1$. According to [109], we have $J_v(v\rho) < J_v(v)$ and $J_{v_1}(v_1\rho) < J_{v_2}(v_2\rho)$, where $v_1 > v_2$ and $\rho \in (0, 1)$. Since $P \geq \lfloor 2\pi f_{m_\mathrm{f}} r/c \rfloor$, we have $J_{(QM_\mathrm{h}-p)}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) < J_{(\alpha-1)P}(\beta P)$ and $J_P(\beta P) \leq J_p(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}))$. Set $\alpha = 3$ and $\beta = 0.5$ for an example. In general, $P > 3$. Hence,$J_p(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) \geq J_3(1.5) \approx 0.06$ and

$$\begin{aligned}
&J_{(Q_2M_\mathrm{h}-p)}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) < J_{(Q_1M_\mathrm{h}-p)}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) \\
&< J_{(M_\mathrm{h}-p)}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) < J_6(1.5) \approx 0.0002,
\end{aligned} \tag{D.32}$$

where $Q_2 > Q_1 > 1$. Compared with $J_p(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}))$, the amplitude of $J_{(QM_\mathrm{h}-p)}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}))$ is so small and can be omitted. We suppress $\varepsilon_{p,Q}(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}}), \phi_{k_{m_\mathrm{b}}})$ and approximate (D.30) as

$$a_{\mathrm{QDFT},p,m_\mathrm{f},m_\mathrm{b}}(\theta_{k_{m_\mathrm{b}}}, \phi_{k_{m_\mathrm{b}}}) \approx \sqrt{M_\mathrm{h}} j^p J_p(\gamma_{m_\mathrm{f}}(\theta_{k_{m_\mathrm{b}}})) \exp(-jp\phi_{k_{m_\mathrm{b}}}). \tag{D.33}$$

This concludes the proof of Theorem 2.

E Proofs and derivations for Chapter 6

E.1 Proof of Theorem 4 in Chapter 6 Section 6.3.2

Let $\gamma(\theta_k) = 2\pi r \sin(\theta_k)/\lambda$. The phase-space transformation of $\mathbf{a}_{\mathrm{h}}(\theta_k, \phi_k)$ can be expressed as

$$
\begin{aligned}
a_{\mathrm{hps},p}(\theta_k, \phi_k) &= \sum_{m_{\mathrm{h}}=1}^{M_{\mathrm{h}}} \left(a_{\mathrm{h},m_{\mathrm{h}}}(\theta_k, \phi_k)\right) e^{-j\frac{2\pi(m_{\mathrm{h}}-1)}{M_{\mathrm{h}}}p} \\
&= \sum_{m_{\mathrm{h}}=1}^{M_{\mathrm{h}}} \left(\frac{1}{\sqrt{M_{\mathrm{h}}}} e^{j\gamma(\theta_k)\cos(\phi_k - \varphi_{m_{\mathrm{h}}})}\right) e^{-j\frac{2\pi(m_{\mathrm{h}}-1)}{M_{\mathrm{h}}}p} \\
&\overset{\text{(a)}}{=} \sum_{m_{\mathrm{h}}=1}^{M_{\mathrm{h}}} \left(\frac{1}{\sqrt{M_{\mathrm{h}}}} \sum_{q=-\infty}^{\infty} j^q J_q\left(\gamma(\theta_k)\right) e^{jq(\phi_k - \varphi_{m_{\mathrm{h}}})}\right) \\
&\qquad \times e^{-j\frac{2\pi(m_{\mathrm{h}}-1)}{M_{\mathrm{h}}}p} \\
&\overset{\text{(b)}}{=} \frac{1}{\sqrt{M_{\mathrm{h}}}} \sum_{Q=-\infty}^{\infty} M_{\mathrm{h}} j^{(QM_{\mathrm{h}}-p)} J_{(QM_{\mathrm{h}}-p)}\left(\gamma(\theta_k)\right) \\
&\qquad \times e^{j(QM_{\mathrm{h}}-p)\phi_k} \\
&\overset{\text{(c)}}{=} \sqrt{M_{\mathrm{h}}} \Bigg[j^p J_p\left(\gamma(\theta_k)\right) e^{-jp\phi_k} \\
&\qquad + \sum_{Q=-\infty, Q\neq 0}^{\infty} \varepsilon_{p,Q}\left(\gamma(\theta_k), \phi_k\right) \Bigg]
\end{aligned} \tag{E.34}
$$

where

$$
\varepsilon_{p,Q}\left(\gamma(\theta_k), \phi_k\right) = j^{(QM_{\mathrm{h}}-p)} J_{(QM_{\mathrm{h}}-p)}\left(\gamma(\theta_k)\right) e^{j(QM_{\mathrm{h}}-p)\phi_k}. \tag{E.35}
$$

In (E.34), (a) and (c) follow the important properties of the Bessel function, i.e., $e^{jx\cos y} = \sum_{v=-\infty}^{\infty} j^v J_v(x) e^{jvy}$ and $J_{-v}(x) = (-1)^v J_v(x)$, respectively. (b) is obtained by letting $p + q = QM_{\mathrm{h}}$ [109].

Let $x = v\rho, \rho \in (0, 1)$ and $v \in \mathbb{Z}^+$. The Bessel function, $J_v(x)$, whose order v exceeds its argument, x, can be written in the following form [109]

$$
J_v(v\rho) = \frac{1}{\pi} \int_0^{\pi} \exp\left(-vF(\vartheta, \rho)\right) d\vartheta, \tag{E.36}
$$

where

$$F(\vartheta,\rho)=\log\left(\frac{\vartheta+\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}}{\rho\sin\vartheta}\right)-\cot\vartheta\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}. \tag{E.37}$$

The partial derivative of (E.36) with respect to ρ is given by

$$\begin{aligned}\frac{\partial}{\partial\rho}J_v(v\rho)&=-\frac{v}{\pi}\int_0^\pi\frac{\partial F(\vartheta,\rho)}{\partial\rho}\exp\left(-vF(\vartheta,\rho)\right)d\vartheta\\&=\frac{v}{\pi\rho}\int_0^\pi g(\vartheta,\rho)\exp\left(-vF(\vartheta,\rho)\right)d\vartheta,\end{aligned} \tag{E.38}$$

where $g(\vartheta,\rho)=\left(\vartheta-\rho^2\sin\vartheta\cos\vartheta\right)/\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}$. Given that

$$\begin{aligned}g(\vartheta,\rho)&=\frac{\vartheta-\rho^2\sin\vartheta\cos\vartheta}{\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}}\geq\frac{\vartheta-\sin\vartheta\cos\vartheta}{\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}}\\&\geq\frac{\vartheta-\sin\vartheta}{\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}}\geq 0,\end{aligned} \tag{E.39}$$

we have $\partial J_v(v\rho)/\partial\rho>0$, and conclude that $J_v(v\rho)$ is an increasing function of ρ. Thus, $J_v(v\rho)<J_v(v)$.

On the other hand, the partial derivative of (E.36) with respect to v is given by

$$\frac{\partial}{\partial v}J_v(v\rho)=-\frac{1}{\pi}\int_0^\pi F(\vartheta,\rho)\exp\left(-vF(\vartheta,\rho)\right)d\vartheta. \tag{E.40}$$

Because

$$\frac{\partial}{\partial\vartheta}F(\vartheta,\rho)=\frac{(1-\rho\cot\vartheta)^2}{\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}}+\sqrt{\vartheta^2-\rho^2\sin^2\vartheta}\geq 0 \tag{E.41}$$

and $\partial F(0,\rho)/\partial\rho=-\sqrt{1-\rho^2}/\rho\leq 0$, we have $F(\vartheta,\rho)\geq F(0,\rho)\geq F(0,1)=0$ and hence, $\partial J_v(v\rho)/\partial v<0$. This means that $J_v(v\rho)$ is a decreasing function of v, i.e., $J_v(v\rho)<J_1(\rho)$. Therefore, we have $J_v(v\rho)<J_v(v)<J_1(1)\approx 0.4$ with $\rho\in(0,1)$ and $v\in\mathbb{Z}^+$. For $|v|>|x|$, $|J_v(x)|\approx 0$ with $v\in\mathbb{Z}^+$. Based on this property, , both $\varepsilon_{p,Q}\left(\gamma(\theta_k),\phi_k\right)$ and $J_p\left(\gamma(\theta_k)\right)$ in (E.34) can be suppressed in the case of $|p|>P>\gamma(\theta_k)$, since $P\geq\lfloor 2\pi r/\lambda\rfloor>2$ and $M_{\mathrm{h}}\geq\lfloor 4\pi r/\lambda\rfloor>2$. When $|p|\leq P$, we can only ignore $\varepsilon_{p,Q}\left(\gamma(\theta_k),\phi_k\right)$. Thus, (E.34) can be approximated by (6.7). This concludes the proof.

E.2 Proof of Theorem 5 in Chapter 6 Section 6.4

Define $\mathcal{Y}_{\text{dfs}} = [\![\mathcal{Z}_{\text{df}}; \mathbf{A}_{\text{dfv}}, \mathbf{A}_{\text{dfh}}, \mathbf{D}]\!]$, which is the noise-free model of $\mathcal{Y}_{\text{df}}$. Thus, $\mathcal{Y}_{\text{dfs}}$ consists of all the signal space components.

Because $\mathbf{A}_{\text{dfv}} \in \mathbb{C}^{N_{\text{vdc}} \times K}$ and $\mathbf{A}_{\text{dfh}} \in \mathbb{C}^{N_{\text{hdc}} \times K}$ are Vandermonde matrices, and in our system, we have $N_{\text{vdc}} \geq K$ and $N_{\text{hdc}} \geq K$, according to uniqueness condition of the CP decomposition, the n-ranks of $\mathcal{Y}_{\text{dfs}}$ depends on the rank of $\mathbf{D}$.

On the other hand, the SVD of the mode-n unfolding of $\mathcal{Y}_{\text{dfs}}$, $\mathbf{Y}_{\text{dfs}(n)}$, can be written as $\mathbf{Y}_{\text{dfs}(n)} = \mathbf{U}_{\text{vs},n}\mathbf{\Sigma}_{\text{vs},n}\mathbf{V}_{\text{vs},n}^{H}$, where $n = 1, 2, 3$, and we have $\text{Rank}(\mathbf{Y}_{\text{dfs}(n)}) = \text{Rank}(\mathbf{U}_{\text{vs},n}) = \text{Rank}(\mathbf{\Sigma}_{\text{vs},n}) = \text{Rank}(\mathbf{V}_{\text{vs},n})$. When $\text{Rank}(\mathbf{D}) < K$, we have $\text{Rank}(\mathbf{Y}_{\text{dfs}(n)}) < K$, and thus $\text{Rank}(\mathbf{U}_{\text{v},n}) < K$.

This concludes the proof of Theorem 5.

Bibliography

[1] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3198–3213, July 2008.

[2] P. Forster and G. Ginolhac, "Derivation of the theoretical performance of a tensor MUSIC algorithm," *Signal Process.*, vol. 129, no. 1, pp. 97–105, Dec. 2016.

[3] J. Steinwandt, F. Roemer, and M. Haardt, "Performance analysis of ESPRIT-type algorithms for non-circular sources," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*. Vancouver, BC, Canada: IEEE, May 2013, pp. 3986–3990.

[4] Z. Lin, T. Lv, J. A. Zhang *et al.*, "Tensor-based high-accuracy position estimation for 5G mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*. Dublin, Ireland: IEEE, Jun. 2020, pp. 1–6.

[5] Ericsson. (2019, Jun.) Ericsson mobility report. [Online]. Available: https://www.ericsson.com/en/mobility-report/reports

[6] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[7] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[8] O. Andrisano, V. Tralli, and R. Verdone, "Millimeter waves for short-range multimedia communication systems," *Proc. IEEE*, vol. 86, no. 7, pp. 1383–1401, Jul. 1998.

[9] T. Bogale and L. Le, "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges," *IEEE Vehic. Techno. Mag.*, vol. 11, no. 1, pp. 64–75, Mar. 2016.

[10] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[11] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[12] H. Q. Ngo and E. G. Larsson, "Large-scale multipair two-way relay networks with distributed AF beamforming," *IEEE Commun. Lett.*, vol. 17, no. 12, pp. 1–4, Dec. 2013.

[13] T. M. Nguyen, V. N. Ha, and L. B. Le, "Resource allocation optimization in multi-user multi-cell massive MIMO networks considering pilot contamination," *IEEE Access*, vol. 3, pp. 1272–1287, 2015.

[14] Z. Lin, T. Lv, and P. T. Mathiopoulos, "3-D indoor positioning for millimeter-Wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2472–2486, Jun. 2018.

[15] T. Lv, F. Tan, H. Gao *et al.*, "A beamspace approach for 2-D localization of incoherently distributed sources in massive MIMO systems," *Signal Process.*, vol. 121, pp. 30–45, Apr. 2016.

[16] B. T. Sieskul, "An asymptotic maximum likelihood for joint estimation of nominal angles and angular spreads of multiple spatially distributed sources," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1534–1538, Jan. 2010.

[17] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, pp. 7112–7139, Nov. 2014.

[18] S. Haghighatshoar and G. Caire, "Massive MIMO channel subspace estimation from low-dimensional projections," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 303–318, Jan. 2017.

[19] J. Romme and B. Kull, "On the relation between bandwidth and robustness of indoor UWB communication," in *Proc. IEEE Conf. Ultra Wideband Systems and Technologies*. Reston, VA: IEEE, Nov. 2003, pp. 255–259.

[20] T. Gigl, G. Janssen, V. Dizdarevic *et al.*, "Analysis of a UWB indoor positioning system based on received signal strength," in *Proc. Workshop Positioning, Navigation Commun.* Hannover, Germany: IEEE, Mar. 2007, pp. 97–101.

[21] N. Garcia, H. Wymeersch, E. G. Larsson *et al.*, "Direct localization for massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2475–2487, May 2017.

[22] A. Shahmansoori, G. E. Garcia, and G. Destino, "5G position and orientation estimation through millimeter wave MIMO," in *Proc. IEEE Int.*

Global Commun. Workshops (GC Wkshps). San Diego, CA: IEEE, Dec. 2015, pp. 1–6.

[23] T. Rappaport, S. Sun, R. Mayzus *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[24] Y. Zeng and R. Zhang, "Millimeter wave MIMO with lens antenna array: a new path division multiplexing paradigm," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1557–1571, Apr. 2016.

[25] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE Journal Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, May 2014.

[26] J. Zhang, X. Ge, Q. Li, M. Guizani, and Y. Zhang, "5G millimeter-wave antenna array: Design and challenges," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 106–112, Oct. 2017.

[27] L. Zhu and J. Zhu, "Optimal design of uniform circular antenna array in mmWave LOS MIMO channel," *IEEE Access*, vol. 6, no. c, pp. 61 022–61 029, Sep. 2018.

[28] P. Wang, Y. Li, and B. Vucetic, "Millimeter wave communications with symmetric uniform circular antenna arrays," *IEEE Commun. Letters*, vol. 18, no. 8, pp. 1307–1310, Aug. 2014.

[29] X. Guo, Q. Wan, X. Shen *et al.*, "Low-complexity parameters estimator for multiple 2D domain incoherently distributed sources," *Turk. J. Elect. Eng. Comput. Sci.*, vol. 3, no. 19, pp. 445–462, May 2011.

[30] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT-a subspace rotation approach

to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 34, no. 5, pp. 1340–1342, Oct. 1986.

[31] J. Li and X. Zhang, "Closed-form blind 2D-DOD and 2D-DOA estimation for MIMO radar with arbitrary arrays," *Wirel. Pers. Commun*, vol. 69, no. 3, pp. 175–186, 2013.

[32] T. Lv, Z. Lin, P. Huang *et al.*, "Optimization of the energy-efficient relay-based massive IoT network," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 3043–3058, Aug. 2018.

[33] A. Alkhateeb, O. E. Ayach, G. Leus *et al.*, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[34] J. A. Zhang, X. Huang, V. Dyadyuk *et al.*, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 79–87, Feb. 2015.

[35] R. W. Heath Jr., N. G. Prelcic, S. Rangan *et al.*, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[36] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," *Signal Process.*, vol. 50, no. 1-2, pp. 57–69, Apr. 1996.

[37] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Jul. 1986.

[38] A. F. Molisch, V. V. Ratnam, S. Han *et al.*, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.

[39] J. H. Brady and A. M. Sayeed, "Wideband communication with high-dimensional arrays: New results and transceiver architectures," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*. London, U.K.: IEEE, Jun. 2015, pp. 1042–1047.

[40] ICT-317669 METIS Project. (Apr.) Metis channel models. D1.1.2 V1.2.

[41] ICT-671650 mmMAGIC Project. (2017, May) Measurement results and final mmmagic channel models. Deliv. D2.2. [Online]. Available: https://5g-mmmagic.eu/results/deliverables

[42] ICT-608637 MiWEBA Project. (June) Channel modeling and characterization. Deliv. D5.1.

[43] ICT-671650 mmMAGIC Project. (2016, Mar.) Measurement campaigns and initial channel models for preferred suitable frequency ranges. Deliv. D2.1. [Online]. Available: https://5g-mmmagic.eu/results/deliverables

[44] R. Tian, Y. Liang, X. Tan *et al.*, "Overlapping user grouping in IoT oriented massive MIMO systems," *IEEE Access*, vol. 5, pp. 14 177–14 186, Jul. 2017.

[45] M.-T. Martinez-Ingles, D. Gaillot, J. Pascual-Garcia *et al.*, "Deterministic and experimental indoor mmW channel modeling," *IEEE Ant. Wireless Prop. Lett.*, vol. 13, pp. 1047–1050, May 2014.

[46] A. Sayeed and V. Raghavan, "Maximizing MIMO capacity in sparse multipath with reconfigurable antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 156–166, Jun. 2007.

[47] A. Guerra, F. Guidi, and D. Dardari, "Position and orientation error bound for wideband massive antenna arrays," in *Proc. IEEE Int. Conf. on Commun. Workshops (ICCW)*. London, UK: IEEE, Jun. 2015, pp. 1–6.

[48] P. Pal and P. P. Vaidyanathan, "Nested arrays in two dimensions, Part I: Geometrical considerations," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4694–4705, Sep. 2012.

[49] C. Pielli, A. Biason, A. Zanella, and M. Zorzi, "Joint optimization of energy efficiency and data compression in TDMA-based medium access control for the IoT," in *Proc. IEEE Globecom Workshops (GC Workshops).* Washington, DC, USA: IEEE, Dec. 2016, pp. 1–6.

[50] H. El-Sayed, G. Athanasiou, and C. Fischione, "Evaluation of localization methods in millimeter-wave wireless systems," in *Proc. IEEE 19th Int. Workshop Comput. Aided Mod. and Design of Commun. Links and Networks (CAMAD).* Athens, Greece: IEEE, Dec. 2014, pp. 345–349.

[51] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, "Toward massive machine type cellular communications," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 120–128, Feb. 2017.

[52] Z. Lin, T. Lv, W. Ni, J. A. Zhang, and R. P. Liu, "Tensor-based multi-dimensional wideband channel estimation for mmWave hybrid cylindrical arrays," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7608–7622, Dec. 2020.

[53] M. Vari and D. Cassioli, "mmWaves RSSI indoor network localization," in *Proc. IEEE Int. Conf. on Commun. Workshop (ICCW).* Sydney, Australia: IEEE, Jun. 2014, pp. 127–132.

[54] Z. Lin, T. Lv, W. Ni, J. A. Zhang, J. Zeng, and R. P. Liu, "Joint estimation of multipath angles and delays for millimeter-Wave cylindrical arrays with hybrid front-ends," *IEEE Trans. Wireless Commun.*, pp. 1–15, Mar. 2021.

[55] Z. Lin, T. Lv, W. Ni, J. A. Zhang, and R. P. Liu, "Nested hybrid cylindrical array design and DoA estimation for massive IoT networks," *IEEE J. Sel. Areas Commun.*, pp. 1–15, Aug. 2020.

[56] "Estimation of nominal direction of arrival and angular spread using an array of sensors," *Signal Processing*, vol. 50, no. 1, pp. 57–69, Mar. 1996.

[57] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digit. Signal Process.*, vol. 8, no. 3, pp. 185–210, Nov. 1998.

[58] H. Boujemaa, "Extension of COMET algorithm to multiple diffuse source localization in azimuth and elevation," *Eur. Trans. Telecommun.*, vol. 16, no. 6, pp. 557–566, Nov. 2005.

[59] S. Shahbazpanahi, S. Valaee, and A. B. Gershman, "A covariance fitting approach to parametric localization of multiple incoherently distributed sources," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 592–600, Feb. 2004.

[60] A. Paulraj, R. Roy, and T. Kailath, "A subspace rotation approach to signal parameter estimation," *Proc. IEEE*, vol. 7, no. 74, pp. 1044–1046, Jul. 1986.

[61] A. M. Sayeed and N. Behdad, "Continuous aperture phased MIMO: A new architecture for optimum line-of-sight links," in *Proc. IEEE Int. Symp. Ant. Propag. (APSURSI)*. Washington, USA: IEEE, Jul. 2011, pp. 293–296.

[62] D. Nion and N. D. Sidiropoulos, "A PARAFAC-based technique for detection and localization of multiple targets in a MIMO radar system," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*. Taipei, Taiwan: IEEE, Apr. 2009, pp. 2077–2080.

[63] A. Hu, T. Lv, H. Gao *et al.*, "An ESPRIT-based approach for 2-D

localization of incoherently distributed sources in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 996–1011, Oct. 2014.

[64] A. Shahmansoori, G. E. Garcia, G. Destino *et al.*, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Comm.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.

[65] Z. Lin, T. Lv, J. A. Zhang, and R. P. Liu, "3D wideband mmWave localization for 5G massive MIMO systems," in *Proc. IEEE Int. Global Commun. (Globecom).* Waikoloa, HI, USA: IEEE, Dec. 2019, pp. 1033–1038.

[66] X. Gao, L. Dai, S. Han, C. Lin, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, Sep. 2017.

[67] A. Liao, Z. Gao, Y. Wu *et al.*, "2D unitary ESPRIT based super-resolution channel estimation for millimeter-wave massive MIMO with hybrid precoding," *IEEE Access*, vol. 5, pp. 24 747–24 757, Nov. 2017.

[68] N. Amiot, T. Pedersen, M. Laaraiedh *et al.*, "A hybrid positioning method based on hypothesis testing," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 348–351, Aug. 2012.

[69] F. Sellone, "Robust auto-focusing wideband DOA estimation," *Signal Process.*, vol. 86, no. 17–37, pp. 3198–3213, Jan. 2006.

[70] B. D. V. Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust. Speech Sig. Proc. Mag.*, vol. 5, no. 5, pp. 4–24, Apr. 1988.

[71] F. Raimondi, P. Comon, and O. Michel, "Wideband multilinear array processing through tensor decomposition," in *Proc. IEEE Int. Conf. Acoust.*

Speech Signal Process. (ICASSP). Shanghai, China: IEEE, Mar. 2016, pp. 2951–2955.

[72] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 33, no. 4, pp. 823–831, Aug. 1985.

[73] G. Bienvenu, P. Fuerxer, G. Vezzosi *et al.*, "Coherent wide band high resolution processing for linear array," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, vol. 4. Glasgow, UK: IEEE, May 1989, pp. 2799–2802.

[74] F. E. D. Raimondi, R. C. Farias, O. J. Michel, and P. Comon, "Wideband multiple diversity tensor array processing," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5334–5346, Oct. 2017.

[75] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Sep. 2009.

[76] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, Mar. 2000.

[77] A. Cichocki, D. Mandic, L. D. Lathauwer *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.

[78] Z. Zhou, J. Fang, L. Yang *et al.*, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, Jul. 2017.

[79] F. Wen, N. Garcia, J. Kulmer *et al.*, "Tensor decomposition based beamspace ESPRIT for millimeter wave MIMO channel estimation," in *Proc. IEEE Int. Global Commun. (GLOBECOM)*. Abu Dhabi, United Arab Emirates: IEEE, Dec. 2018, pp. 1–7.

[80] A. Gharaibeh, M. A. Salahuddin, S. J. Hussini, A. Khreishah, I. Khalil, M. Guizani, and A. Al-Fuqaha, "Smart cities: A survey on data management, security, and enabling technologies," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 4, pp. 2456–2501, Aug. 2017.

[81] Mckinsey Global Institute, "The internet of things: Mapping the value beyond the hype," Tech. Rep., June 2015, tech. Rep.

[82] J. A. Stankovic, "Research directions for the internet of things," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 3–9, Feb. 2014.

[83] A. Whitmore, A. Agarwal, and L. D. Xu, "The internet of things-A survey of topics and trends," *L. Inf Syst Front*, vol. 17, no. 4, pp. 261–274, Apr. 2015.

[84] J. Zou, H. Yu, W. Miao, and C. Jiang, "Packet-Based preamble design for random access in massive IoT communication systems," *IEEE Access*, vol. 5, pp. 11 759–11 767, June 2017.

[85] G. Hattab and D. Cabric, "Energy-Efficient massive cellular IoT shared spectrum access via mobile data aggregators," in *Proc. IEEE Int. Conf. Pervasive Computing and Comm. (PERCOM)*. Rome, Italy: IEEE, Oct. 2017, pp. 1–6.

[86] T. Lv, Y. Ma, J. Zeng, and P. T. Mathiopoulos, "Millimeter-wave NOMA transmission in cellular M2M communications for internet of things," *IEEE Internet of Things J.*, vol. 5, no. 3, pp. 1989–2000, Mar. 2018.

[87] P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, Jan. 2010.

[88] A. T. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propag.*, vol. 16, no. 2, pp. 172–175, Mar. 1968.

[89] E. Vertatschitsch and S. Haykin, "Nonredundant arrays," *Proc. IEEE*, vol. 74, no. 1, pp. 217–217, Jan. 1986.

[90] P. Pal and P. P. Vaidyanathan, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 1053–1058, Feb. 2011.

[91] T. Basikolo, K. Ichige, and H. Arai, "Direction of arrival estimation for quasi-stationary signals using nested circular array," in *Proc. Int. Workshop Compressed Sens. Theory Appl. Radar, Sonar Remote Sens. (CoSeRa)*. Aachen, Germany: IEEE, Sep. 2016, pp. 193–196.

[92] ——, "Nested circular array and its concentric extension for underdetermined direction of arrival estimation," *IEICE Trans. Commun.*, vol. 101, no. 4, pp. 1084–1096, Apr. 2018.

[93] P. Pal and P. P. Vaidyanathan, "Nested arrays in two dimensions, Part II: Application in two dimensional array processing," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4706–4718, Sep. 2012.

[94] C. L. Liu and P. P.Vaidyanathan, "Tensor music in multidimensional sparse arrays," in *Proc. IEEE Asil. Conf. on Sig., Sys., and Comp.* Pacific Grove, CA, USA: IEEE, Nov. 2015, pp. 1783–1787.

[95] Y. Zhao, Y. Liu, T. He *et al.*, "FREDI: Robust RSS-based ranging with

multipath effect and radio interference," in *Proc. IEEE INFOCOM.* Turin, Italy: IEEE, April 2013, pp. 505–509.

[96] Y. Shen and M. Win, "On the accuracy of localization systems using wideband antenna arrays," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 1–8, Jan. 2010.

[97] A. Olivier, G. Bielsa, I. Tejado *et al.*, "Lightweight indoor localization for 60-GHz millimeter wave systems," in *Proc. IEEE Int. Symp. Ant. Propag. (APSURSI).* London, UK: IEEE, Jun. 2016, pp. 1–9.

[98] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[99] G. R. Maccartney, T. S. Rappaport, S. Sun, and S. Deng, "Indoor office wideband millimeter-wave propagation measurements and models at 28 GHz and 73 GHz for ultra-dense 5G wireless networks (invited paper)," *IEEE Access*, vol. 3, pp. 2388–2424, Oct. 2015.

[100] J. Järveläinen, S. Nguyen, K. Haneda, R. Naderpour, and U. Virk, "Evaluation of millimeter-wave line-of-sight probability with point cloud data," *IEEE Wireless Commun. Lett.*, vol. 5, no. 3, pp. 228–231, Jun. 2016.

[101] O. Landron, M. J. Feuerstein, and T. S. Rappaport, "A comparison of theoretical and empirical reflection coefficients for typical exterior wall surfaces in a mobile radio environment," *IEEE Trans. Antennas Propagat.*, vol. 44, no. 3, pp. 341–351, Mar. 1996.

[102] Y. Wang, Z. Tian, S. Feng *et al.*, "Efficient channel statistics estimation for millimeter-wave MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP).* Shanghai, China: IEEE, Mar. 2016, pp. 3411–3415.

[103] C. P. Mathews and M. D. Zoltowski, "Eigenstructure techniques for 2-D angle estimation with uniform circular arrays," *IEEE Trans. Signal Process*, vol. 42, no. 9, pp. 2395–2407, Sep. 1994.

[104] M. Ester, H.-P. Kriegel, and J. Sander, "Spatial data mining: A database approach," in *Proc. 5th Int. Symp. Large Spatial Databases (SSD)*. Berlin, Germany: IEEE, Jul. 1997, pp. 47–66.

[105] C. Chen and X. Zhang, "A low-complexity joint 2D-DOD and 2D-DOA estimation algorithm for MIMO radar with arbitrary arrays," *Int. J. Electron.*, vol. 100, no. 10, pp. 1455–1469, 2013.

[106] 5GCMSIG White Paper. (2016, Oct.) 5G channel model for bands up to 100 GHz. Ver.2.1. [Online]. Available: http://www.5gworkshops.com/5GCM.html

[107] X. Gao, L. Dai, S. Han *et al.*, "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, April 2016.

[108] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 7, pp. 2207–2225, July 2016.

[109] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed. Cambridge, UK: Cambridge Univ. Press, 1952.

[110] R. J. Mailloux, *Phased Array Antenna Handbook*, 2nd ed. USA: Artech House, 2005.

[111] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[112] J. Zhang, I. Podkurkov, M. Haardt *et al.*, "Efficient multidimensional parameter estimation for joint wideband radar and communication systems based on OFDM," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*. New Orleans, LA, USA: IEEE, Mar. 2017, pp. 3091–3100.

[113] M. C. Vanderveen, A. J. van der Veen, and A. Paulraj, "Estimation of multipath parameters in wireless communications," *IEEE Trans. Signal Process.*, vol. 46, no. 3, pp. 682–690, Mar. 1998.

[114] A. J. v. d. Veen, P. Ober, and E. Deprettere, "Azimuth and elevation computation in high resolution DOA estimation," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1828–1832, July 1992.

[115] A. L. Swindlehurst and T. Kailath, "Azimuth/elevation direction finding using regular array geometrics," *IEEE Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 145–156, Jan. 1993.

[116] B. Steingrimsson, Z.-Q. Luo, and K. M. Wong, "Soft quasi-maximum-likelihood detection for multiple antenna wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2710–2718, Nov. 2003.

[117] D. Wang, M. Fattouche, and X. Zhan, "Pursuance of mm-level accuracy: Ranging and positioning in mmWave systems," *IEEE Systems J.*, vol. 13, no. 2, pp. 1169–1180, Jun. 2019.

[118] F. Roemer, M. Haardt, and G. Del Galdo, "Analytical performance assessment of multi-dimensional matrix- and tensor-based ESPRIT-type algorithms," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2611–2625, May 2014.

[119] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, USA: NJ: Prentice-Hall, 1993.

[120] K. Abed-Meraim and Y. Hua, "A least-squares approach to joint Schur decomposition," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, vol. 4. Seattle, WA, USA: IEEE, May 1998, pp. 2541–2544.

[121] *Study on channel model for frequency spectrum above 6 GHz (Rel. 14)*, Std., 2016.

[122] L. Zhao, G. Geraci, T. Yang *et al.*, "A tone-based AoA estimation and multiuser precoding for millimeter wave massive MIMO," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5209–5225, Dec. 2017.

[123] X. Shen and A. I. Zayed, *Multiscale signal analysis and modeling.* New York, USA: Springer, 2013.

[124] C. L. Liu and P. P.Vaidyanathan, "Hourglass arrays and other novel 2-D sparse arrays with reduced mutual coupling," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3369–3383, Jul. 2017.

[125] S. Han, C. I, Z. Xu *et al.*, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[126] R. T. Hoctor and S. A. Kassam, "The unifying role of the coarray in aperture synthesis for coherent and incoherent imaging," *Proc. IEEE*, vol. 78, no. 4, pp. 735–752, Apr. 1990.

[127] W.-K. Ma, T.-H. Hsieh, and C.-Y. Chi, "DOA estimation of quasistationary signals via Khatri-Rao subspace," in *Proc. Int. Conf. Acoust., Speech Signal Process. (ICASSP).* Taipei, Taiwan: IEEE, Apr. 2009, pp. 2165–2168.

[128] L. N. Ribeiro, S. Schwarz, M. Rupp *et al.*, "Energy efficiency of mmwave massive MIMO precoding with low-resolution DACs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 298–312, May 2018.

[129] H. Boujemaa, "Extension of COMET algorithm to multiple diffuse source localization in azimuth and elevation," *Eur. Trans. Telecomm.*, vol. 16, no. 6, pp. 557–566, Nov./Dec. 2014.