UNIVERSITY OF CALIFORNIA,
IRVINE

Beamforming in mmWave MIMO Systems

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical Engineering

by

Lisi Jiang

Dissertation Committee:
Professor Hamid Jafarkhani, Chair
Professor Lee Swindlehurst
Professor Ender Ayanoglu

2021

DEDICATION

To my family and my boyfriend.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF ALGORITHMS

ACKNOWLEDGMENTS

First, I would like to give my sincere gratitude to Prof. Hamid Jafarkhani for his guidance and support throughout my Ph.D. studies. This thesis would never have been possible without his advice. It is really a great honor to be his student, since he will always give me enough freedom to explore my own interest while keeping me on the right track. I truly appreciate his profound knowledge and keen insight. I will always keep in mind his valuable advice and guidance.

Besides my advisor, I would like to thank my thesis committee members, Prof. Ayanoglu and Prof. Swindlehurst for their insightful comments and feedback which help me to expand and deepen my thesis from various perspectives.

I would also like to thank other members in my research group over the past five years: Mojtaba Ahmadi Almasi, Biao He, Jun Guo, Xiaoyi Liu, Weiqi Li, Xun Zou, Saeed Karimi Bidhendi, Carles Diaz. It is a greart pleasure to work with them. Especially, I am truly honored to have the opportunity to collaborate with Mojtaba Ahmadi Almasi, from whom I acquired a great amount of knowledge and pleasure.

Many thanks to my friends, Yongce Chen, Ruijin Sun, Bimeng Gong, Kunchao Xu, Chen Yang and Shiyao Zhang. I will always remember the encouragement and happiness they gave me. It is such a fortune to have them around me.

Another gratitude goes to the National Science Foundation (NSF) Award Electrical, Communications and Cyber Systems Grant (ECCS) and Computing and Communication Foundations (CCF) for providing fundings for my research. I'm also grateful for the Department of Electrical Engineering and Computer Science (EECS) for offering me the opportunity to pursue a doctorate degree.

I would like to thank my mother, who has always pushed me to be disciplined and hard-working. I sincerely appreciate her for encouraging me to be stronger and better every day. Without her guidance, I would not have been able to have the achievements that I have today.

Last but not least, I want to express my gratitude and love to my boyfriend, Wen-Kai, for his companion and support throughout this journey. He is the one who always believe in me and he is my source of happiness. I am so lucky to have him in my life.

VITA

Lisi Jiang

EDUCATION

Doctor of Philosophy in Electrical Engineering **2021**
University of California, Irvine *Irvine, California*

Master of Science in Electrical Engineering **2016**
Beijing University of Posts and Telecommunications *Beijing, China*

Bachelor of Science in Electrical Engineering **2013**
Beijing University of Posts and Telecommunications *Beijing, China*

RESEARCH EXPERIENCE

Fellowship Scholar **2016–2018**
University of California, Irvine *Irvine, California*

Graduate Research Assistant **2018–2021**
University of California, Irvine *Irvine, California*

Intern-ship **June 2020–September 2020**
Facebook Connectivity *Menlo Park, California*

Reviewer-ship
IEEE JSAC, IEEE TWC, ETT, Globecom, ICC

TEACHING EXPERIENCE

Teaching Assistant **Jan. 2020– Mar. 2020**
University of California, Irvine *Irvine, California*

Teaching Assistant **Jan. 2021– Mar. 2021**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

Joint Beamwidth and Power Optimization in MmWave Hybrid Beamforming-NOMA Systems **2020**
IEEE Transactions on Wireless Communications

MmWave amplify-and-forward MIMO relay networks with hybrid precoding/combining design **2019**
IEEE Transactions on Wireless Communications

Multi-User Analog Beamforming in Millimeter Wave MIMO Systems Based on Path Angle Information **2018**
IEEE Transactions on Wireless Communications

A Two-Step mmWave-NOMA Beam Alignment Algorithm with Misalignment Probability Constraints **2021**
Submitted IEEE Transactions on Wireless Communications

Reconfigurable Intelligent Surface Assisted mmWave multi-UAV Wireless Cellular Networks **2021**
In preparation

REFEREED CONFERENCE PUBLICATIONS

Reconfigurable Intelligent Surface Assisted mmWave UAV Wireless Cellular Networks **June 2021**
IEEE International Conference on Communications (ICC)

Hybrid precoding/combining design in mmWave amplify-and-forward MIMO relay networks **May 2019**
IEEE International Conference on Communications (ICC)

Robust Multi-User Analog Beamforming in mmWave MIMO Systems **November 2018**
IEEE Global Conference on Signal and Information Processing (GlobalSIP)

ABSTRACT OF THE DISSERTATION

Beamforming in mmWave MIMO Systems

By

Lisi Jiang

Doctor of Philosophy in Electrical Engineering

University of California, Irvine, 2021

Professor Hamid Jafarkhani, Chair

Future wireless communication systems are expected to have much higher mobility and data rate. To achieve these goals, millimeter-wave (mmWave) communication has been considered as a key technology for the future wireless communication systems, because of the high data rates provided by the large bandwidth at the mmWave carrier frequency. However, obstacles such as the severe path-loss and the hardware complexity hinder the practical application of the mmWave communication. Incorporating beamforming into the mmWave communication systems is an effective way to combat the severe path-loss. In this dissertation, we explore the beamforming technique in mmWave communication systems. Four scenarios are investigated: multi-user networks, amplify-and-forward (AF) relay networks, mmWave non-orthogonal multiple access (NOMA) networks, and reconfigurable intelligent surface (RIS)-assisted mmWave unmanned aerial vehicles (UAV) networks.

For the multi-user networks, we design an analog-only beamforming scheme for downlink multi-user mmWave systems to optimize the beamforming gain and the inter-user interference at the same time. Traditional analog beamforming schemes, such as the beam selection method, use the array response vector corresponding to the strongest path of the channel to generate a beam pointing to the user. In multi-user systems, such schemes will lead to large inter-user interference, especially when the users are closely located. In this dissertation, we formulate a multi-objective problem

to strike a balance between the beamforming gain and the inter-user interference. Furthermore, to alleviate the effects of the channel estimation and feedback quantization errors, we design a robust beamforming scheme to provide robustness against imperfect channel information. We first develop a channel error model for the scattering clustered channel model, which can serve as a general channel error model for the mmWave channels. Then, we formulate a multi-objective problem using the stochastic approach to suppress the interference and enhance the beamforming gain at the same time.

For the AF relay networks, we consider the amplify-and-forward relay networks in mmWave systems and propose a hybrid precoder/combiner design approach. The phase-only RF precoding/combining matrices are first designed to support multi-stream transmission, where we compensate the phase for the eigenmodes of the channel. Then, the baseband precoders/combiners are performed to achieve the maximum mutual information. In addition, we also propose a robust joint transceiver design for imperfect channel state information.

For the mmWave-NOMA networks, we first take the limited channel coherence time into account for NOMA in mmWave hybrid beamforming systems. Due to the limited coherence time, the beamwidth of the hybrid beamformer affects the beam-training time, which in turn directly impacts the data transmission rate. To investigate this trade-off, we utilize a combined beam-training algorithm. Then, we formulate a sum-rate expression which considers the channel coherence time and beam-training time as well as users' power and other system parameters. Further, a joint power and beamwidth optimization problem is solved by iterating between the power allocation and the beamwidth optimization.

Further, We propose a new two-step beamwidth design and power allocation algorithm, in mmWave-NOMA systems, which takes the channel coherence time and users' locations into account. A joint beamwidth and power allocation optimization algorithm is proposed to maximize the sum-rate.

For the RIS-assisted mmWave UAV networks, we jointly optimize the deployment, user schedul-

ing, beamforming vector and RIS phases to maximize the sum-rate, with the constraints of the minimum rate, the UAV movement, the analog beamforming and the RIS phases.

Chapter 1

Introduction

1.1 Background

Communications over millimeter wave (mmWave) have received significant attention recently because of the high data rates provided by the large bandwidth at the mmWave carrier frequencies. Also, using large antenna arrays in mmWave communication systems is possible because the small wavelength allows integrating many antennas in a small area. Despite its advantages, the mmWave carrier frequencies suffer from relatively severe propagation losses. Meanwhile, the sparsity of the mmWave scattering environment usually results in rank-deficient channels [99].

To overcome the large path losses, large antenna arrays can be placed at both transmitters and receivers to guarantee sufficient received signal power [93, 60, 40, 117]. The large antenna arrays lead to a large number of radio frequency (RF) chains, which greatly increase the implementation cost and complexity. To reduce the number of RF chains, two approaches have been proposed : hybrid precoding [33], which connects analog phase shifters with a reduced number of RF chains, and analog beamforming [80, 53, 88], where a single RF chain is tied to the entire antenna array and the beamforming processing is performed with the RF analog components. The main advantage of

the hybrid precoding is that it can trade off between the low-complexity limited-performance analog phase shifters and the high-complexity good-performance digital precoding [47]. The analog beamforming, on the other hand, has the lowest hardware complexity with relative worse performance.

To further improve the throughput and coverage of the mmWave systems, techniques, such as multi-user [8, 89, 96, 78, 73, 94], relay [74, 138, 108, 118], non-orthogonal multiple access (NOMA) [27, 25, 147, 125, 3, 56, 139, 123, 9], reconfigurable intelligent surface (RIS) [17, 42, 35], and unmanned aerial vehicles (UAVs) [132, 131], are usually incorporated into the mmWave systems.

To implement beamforming in the aforementioned scenarios, the trading off between the performance and hardware complexity is the key issue. For example, for the multi-user system, how to cancel the interference among users while maximizing the beamforming gain for each user with limited number of RF chains is the crucial point to improve the system performance. In the relay networks, low-complexity hybrid precoding algorithm need to be proposed to achieve a good system performance. For the NOMA system, trade-off need to be made between the training complexity and alignment accuracy. In the UAV networks, where deployment and beamforming are jointly optimized, low-energy cost algorithms are usually adopted.

1.2 Contributions

In this dissertation, we explore the beamforming technique in mmWave communication systems. Four scenarios are investigated: multi-user networks, amplify-and-forward (AF) relay networks, mmWave NOMA networks, and RIS-UAV networks. For the multi-user networks, we design an analog-only beamforming scheme for downlink multi-user mmWave systems to optimize the beamforming gain and the inter-user interference at the same time. Traditional analog beamform-

ing schemes, such as the beam selection method, use the array response vector corresponding to the strongest path of the channel to generate a beam pointing to the user. In multi-user systems, such schemes will lead to large inter-user interference, especially when the users are closely located. In this paper, we formulate a multi-objective problem to strike a balance between the beamforming gain and the inter-user interference. To solve the problem, we first use the weighted-sum method to transform the multi-objective problem into a single-objective problem. Then, we use the semi-definite programming technique to make the analog beamforming with constant-magnitude constraints tractable. Furthermore, to alleviate the effects of the channel estimation and feedback quantization errors, we design a robust beamforming scheme to provide robustness against imperfect channel information. We first develop a channel error model for the scattering clustered channel model, which can serve as a general channel error model for the mmWave channels. Then, we formulate a multi-objective problem using the stochastic approach to suppress the interference and enhance the beamforming gain at the same time. The simulation results show that our proposed non-robust multi-user analog beamformer outperforms the traditional analog beamforming method when the SNR is high and our proposed robust beamformer can provide up to 109 % improvement in the sum-rate compared with the beam selection method.

For the AF relay networks, we consider the amplify-and-forward relay networks in mmWave systems and propose a hybrid precoder/combiner design approach. The phase-only RF precoding/combining matrices are first designed to support multi-stream transmission, where we compensate the phase for the eigenmodes of the channel. Then, the baseband precoders/combiners are performed to achieve the maximum mutual information. Based on the data processing inequality for the mutual information, we first jointly design the baseband source and relay nodes to maximize the mutual information before the destination baseband receiver. The proposed low-complexity iterative algorithm for the source and relay nodes is based on the equivalence between the mutual information maximization and the weighted minimum mean square error (MMSE). After we obtain the optimal precoder and combiner for the source and relay nodes, we implement the MMSE-SIC filter at the baseband receiver to keep the mutual information unchanged, thus obtain-

ing the optimal mutual information for the whole relay system. Simulation results show that our algorithm achieves better performance with lower complexity compared with other algorithms in the literature. In addition, we also propose a robust joint transceiver design for imperfect channel state information.

For the mmWave-NOMA networks, we first take the limited channel coherence time into account for NOMA in mmWave hybrid beamforming systems. Due to the limited coherence time, the beamwidth of the hybrid beamformer affects the beam-training time, which in turn directly impacts the data transmission rate. To investigate this trade-off, we utilize a combined beam-training algorithm. Then, we formulate a sum-rate expression which considers the channel coherence time and beam-training time as well as users' power and other system parameters. Further, a joint power and beamwidth optimization problem is solved by iterating between the power allocation and the beamwidth optimization. When allocating the power, we use the log-exponential reformulation and the sequential parametric convex approximation (SPCA) methods to solve the non-convex problem. Since beamwidth optimization involves too many variables, we propose an algorithm which iterates between clusters of users. Numerical results show that the optimized mmWave hybrid beamforming-NOMA system can achieve much higher sum-rates compared to NOMA with analog beamforming and traditional multiple access techniques.

Further, We propose a new two-step beamwidth design and power allocation algorithm, in mmWave-NOMA systems, which takes the channel coherence time and users' locations into account. A joint beamwidth and power allocation optimization algorithm is proposed to maximize the sum-rate. The log-exponential reformulation and the sequential parametric convex approximation method are used for the non-convex power allocation. Then, the optimal beamwidths are obtained by iterating between the two alignment stages. Further, because of the far users, the mmWave-NOMA systems are prone to beam misalignment. To this end, we impose extra constraints on the beamwidths to satisfy the misalignment probability requirements. An off-line search is used to find the appropriate ranges for the beamwidths. Then, the problem with misalignment analysis can be solved by

a similar joint optimization algorithm structure. Simulation results reveal that the proposed algorithm remarkably outperforms the traditional one-step approach in terms of the sum-rate. Also, compared to the one without beam misalignment, the sum-rate of the proposed algorithm with beam misalignment can lead to at most 10% decrease in sum-rate, which demonstrates that our algorithm can strike a good balance between misalignment and sum-rate.

For the RIS-assisted mmWave UAV networks, we jointly optimize the deployment, user scheduling, beamforming vector and RIS phases to maximize the sum-rate, with the constraints of the minimum rate, the UAV movement, the analog beamforming and the RIS phases. To solve this complex problem, we use an iterative method, in which when we optimize one variable, we fix the other three variables. When optimizing the deployment, we find the optimal position for the UAV by a sphere search. Then, we formulate an integer linear programming to find the best scheduling. We also design the analog beamforming vector by compensating the phases of the channel which combines the direct path and the RIS paths. When optimizing the RIS phases, we formulate a semi-definite programming to find the best phases. The proposed joint optimization outperforms the system without RIS assistance and the system without deployment optimization.

1.3 Dissertation Organization

This dissertation is organized as follows. In Chapter 2, a multi-user analog beamforming algorithm is proposed. Chapter 3 proposes a hybrid precoding algotithm for mmWave relay networks. In Chapter 4, a joint design of beamwidth and power allocation algorithm for hybrid beamforming in mmWave NOMA systems is proposed. Chapter 5 further proposes a two-step mmWave-NOMA beam alignment algorithm with misalignment probability constraints. In Chapter 6, joint optimization of deployment, scheduling, beamforming, and RIS phases adjustment in the RIS-assisted mmWave UAV network is investigated.

Notation: $\mathbb{C}^{m\times n}$ is the set of all $m \times n$ complex matrices with $\mathbb{C}^m \triangleq \mathbb{C}^{m\times 1}$ and $\mathbb{C} \triangleq \mathbb{C}^1$. $\mathbb{I}_m$ is the $m \times m$ identity matrix, and $\mathbf{0}_{m\times n}$ is the $m \times n$ all-zero matrix. $\mathbb{CN}(\boldsymbol{\mu}, \mathbf{K})$ is a circularly-symmetric complex Gaussian random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{K}$. Matrices $\mathbf{A}^T$ and $\mathbf{A}^H$ are the transpose and the Hermite transpose of matrix $\mathbf{A}$, respectively. Matrix $\mathbf{A} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_L]$ represents the concatenation of the L vectors $\boldsymbol{\alpha}_i$, and $\mathbf{B} = [\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_K]$ represents the concatenation of the K matrices $\mathbf{A}_i$.

Chapter 2

Multi-User Analog Beamforming in mmWave MIMO Systems Based on Path Angle Information

2.1 Introduction

Millimeter wave (mmWave) communication has been considered as a key technology for future wireless communication systems because of the high data rates provided by the large bandwidths at mmWave carrier frequencies. However, mmWave carrier frequencies suffer from relatively severe propagation losses, which reduce service coverage and impair communication performance [99]. Thus, large antenna arrays are usually proposed to be implemented at both transmitters and receivers to provide sufficient beamforming gain to mitigate the severe propagation attenuation [115, 86]. The large antenna arrays, however, lead to high system complexity for employing conventional full digital beamforming, where each antenna element is connected to a separate radio-frequency (RF) chain[60, 40, 117]. Therefore, analog beamforming [80, 53, 88], where a single RF

chain is tied to the entire antenna array and the beamforming processing is performed with the RF analog components, has been reignited in mmWave systems. To strike a balance between the system complexity and the beamforming precision, a hybrid architecture [33] that uses analog phase shifters in conjunction with a reduced number of RF chains is proposed for single-user mmWave multi-input multi-output (MIMO) systems.

To further improve the system throughput, multi-user systems, where a base station (BS) simultaneously serves a number of mobile stations, are often adopted. To cancel the interference among mobile stations, precoding is usually applied at the BS. For conventional multi-user systems, precoding is commonly done at the baseband, where each antenna element has a radio frequency (RF) chain[128, 106, 114, 102, 126, 39, 113, 60]. This kind of precoding is called fully digital beamforming. In [128, 106, 126], iterative beamforming algorithms that maximize the signal-to-interference-noise ratio (SINR) for all users were proposed. Unfortunately, these exists no closed-form solution for such iterative algorithms. Besides, the optimization problem is NP-complete, which means that they cannot be solved in reasonable time [39]. In [60] and [113], zero-forcing schemes for multi-user beamforming were proposed, which decoupled the multi-user beamforming problem and perfectly cancel the interference. In [114, 102, 39], the signal-to-leakage-and-noise ratio (SLNR) was chosen as the criterion of the optimization problem, which also leads to a decoupled optimization problem and provides an analytical closed-form solution. Also, in [39], a semidefinite-programming-based algorithm was proposed, which aims to minimize the total transmitted power with QoS requirements. However, similar to the single-user MIMO systems, all mentioned fully digital beamforming schemes are not practical for large antenna arrays in mmWave systems due to the high complexity and the large power consumption.

To address the difficulty of the limited number of RF chains in multi-user systems, two approaches have been proposed. One is the hybrid multi-user beamforming, in which the beamformer is constructed by the concatenation of a low-dimensional baseband (digital) beamformer and an RF (analog) beamformer [8, 89, 96]. The RF beamformer provides a high-dimensional phase-only control

and is usually used to enhance the array gain. The baseband beamformer, on the other hand, is usually used to cancel the interference. This method can achieve a performance close to a conventional digital beamformer [8, 89]. However, a two-stage feedback for both the RF beamforming and the baseband beamforming is needed. Such a two-stage feedback requires a tremendous overhead for large antenna arrays. This may become a limitation for mmWave MIMO systems and should be avoided if possible. The other approach is the analog multi-user beamforming, where the beamforming processing is only performed with RF analog components. Currently, many RF beamformers use discrete Fourier transform (DFT) vectors as the RF beamformer [78, 73, 94]. Since the DFT vectors have a form similar to that of the array response vectors of the arrays, such an RF beamformer will have the largest array gain [89, 78]. Additionally, in [124, 116, 57], an analog beam selection method for mmWave multi-user systems was proposed, where both the BS and the users are equipped with an analog beamforming codebook. The codebook consists of beamforming vectors which are the array response vectors with uniform spacing. The BS chooses the best beamforming vector, which maximizes the beamforming gain, from the codebook. This method performs well for line-of-sight (LOS) channels, because the selected beamformer is the match filters for the LOS channels. However, considering a non-LOS (NLOS) channel model, the performance of the beam selection method will be degraded due to the interference among different paths and different users. Besides, the beam selection method needs a training stage to find the best beam, whose overhead scales linearly with the number of users. All of the previous work either uses DFT vectors as the analog beamformers or uses the array response vectors as the analog beamformers. To the best of our knowledge, there is no structure design for the analog beams in the literature.

In this chapter, we aim to design an analog beamforming method in downlink multi-user systems which not only enhances the beamforming gain but also cancels the inter-user interference. In our method, the analog beams are not DFT structured vectors, but the objectives to be optimized. In the first part of the chapter, we propose an analog beamforming method based on perfect channel state information (CSI). A multi-objective problem (MOP) is first established to maximize the

beamforming gain and minimize the inter-user interference at the same time. We transform the MOP into a single-objective problem (SOP) by using the weighted-sum method. The semi-definite programming (SDP) technique is then introduced to deal with the constant-magnitude constraints for the analog beamforming. To further reduce the feedback overhead, we only use the angle of departures/angle of arrivals (AoD/AoA) of the channel instead of the full channel information.

Channel information is also critical for mmWave MIMO systems. Imperfect CSI will lead to severe performance degradation. Some papers, such as [29], [31] and [30], analyzed the performance of the imperfect CSI and proposed communication schemes for imperfect CSI in traditional MIMO systems, i.e., fully digital MIMO systems. However, to the best of our knowledge, there is no robust communication design for mmWave MIMO systems in the literature.

The second part of the chapter considers imperfect CSI caused by channel estimation and quantization in mmWave systems. We propose a robust design for the analog beamforming, which not only suppresses the interference and enhances the beamforming gain, but also provides robustness against imperfect CSI. We assume there exists angle errors in the AoD/AoA of the channel and simplify the error model into an additive error model by using Taylor expansion. Based on the statistical properties of the errors, a probabilistic objective similar to [28, 18, 21] is formulated. We maximize the average beamforming gain while keeping the probability of small leakage power as large as possible (i.e., we formulate a MOP to maximize the average array gain and the probability of small leakage power at the same time). The probabilistic objective is transformed into a deterministic one by applying Markov's inequality, and then we use the same technique as the proposed non-robust beamforming to deal with the MOP for robust beamforming.

The contributions of this chapter can be summarized as follows:

- We propose an analog beamforming scheme based on the path angle information in mmWave systems. The scheme strikes a balance between beamforming gain and inter-user interference only using partial channel information. Our proposed scheme outperforms the conven-

tional beam selection method due to the interference suppression.

- We consider the effects of the channel estimation and feedback quantization errors and develop a channel error model for the scattering clustered channel model, which can serve as a general channel error model for mmWave channels.

- We propose a robust analog beamforming scheme for mmWave systems to alleviate the effects of the channel estimation and feedback quantization errors. The proposed robust analog beamforming scheme brings about 109% improvement in sum-rate compared to the conventional beam selection method.

The remaining sections are organized as follows. In Section 2.2 , we describe the system model and the mmWave channel model. Section 2.3 formulates the proposed analog beamforming method based on perfect CSI. Section 2.4 presents the proposed robust analog beamforming design for imperfect CSI. Numerical examples are presented and discussed in Section 2.5. We provide concluding remarks in Section 2.6.

2.2 System model and problem formulation

2.2.1 System model and zero-forcing schemes

We consider a multi-user system including a BS with N_t antennas serving K single-antenna users as Fig. 2.1 depicts. The number of RF chains N_{RF} is set to be K to enable the multi-user transmission. In the case that the number of RF chains, N_{RF}, is less than the number of users, one can select N_{RF} out of K users and serve them using our algorithm. One such a user selection method is the proportional-fair (PF) method presented in [81]. Only analog beamforming is used for each user. The BS generates the analog beamforming vector for User i based on the estimated multi-path angles of the channels. We denote s_i as the transmitted symbol intended for User i with

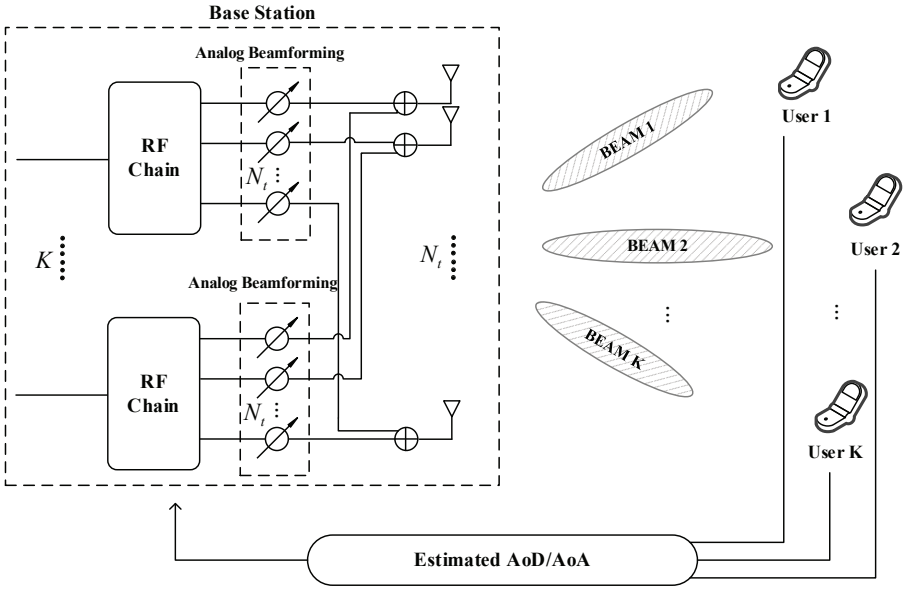

Figure 2.1: System model

$E[\|s_i\|^2] = 1$ and $\boldsymbol{w}_i \in \mathbb{C}^{N_t \times 1}$ as the beamforming vector for s_i. The channel between User i and the BS is denoted by $\boldsymbol{h}_i^H \in \mathbb{C}^{1 \times N_t}$. The received signal at User i can be expressed as

$$y_i = \boldsymbol{h}_i^H \boldsymbol{w}_i s_i + \sum_{k=1, k \neq i}^{K} \boldsymbol{h}_i^H \boldsymbol{w}_k s_k + n_i, \tag{2.1}$$

where n_i is the additive Gaussian noise with zero mean and σ^2 variance. The second term in (2.1) is called the co-channel interference (CCI) caused by other users.

Intuitively, the optimal multi-user system is the one that maximizes the signal-to-interference-plus-noise ratio (SINR) of every user. The SINR of User i is given by

$$SINR_i = \frac{|\boldsymbol{h}_i^H \boldsymbol{w}_i|^2}{\sigma^2 + \sum_{k=1, k \neq i}^{K} |\boldsymbol{h}_i^H \boldsymbol{w}_k|^2}. \tag{2.2}$$

However, using SINR as the optimization criterion generally results in a challenging optimization problem to deal with K coupled variables $\{\boldsymbol{w}_i\}_{i=1}^{K}$ [128, 106, 126].

One way to avoid solving the coupled problem is to focus on canceling the CCI by using zero-forcing (ZF) schemes [68]. As [60] and [113] did, the ZF schemes choose beamforming vectors

$\boldsymbol{w}_i$ by enforcing the conditions

$$\boldsymbol{h}_k^H \boldsymbol{w}_i = 0,\ \forall i,\ k = 1, ..., K, k \neq i. \tag{2.3}$$

This solution results in good performance since it completely cancels the CCI at every receiver. However, ZF does not optimize the beamforming gain, and thus is not optimal for SINR. Moreover, for analog beamforming, the elements of vector $\boldsymbol{w}_i$ have constant magnitudes, i.e., $|w_i^n| = constant$ where w_i^n denotes the n^{th} element in vector $\boldsymbol{w}_i$. We call these constraints, the constant-magnitude constraints. For a system with constant-magnitude constraints, the ZF conditions in (2.3) may not be feasible. To remedy these issues, we relax the ZF conditions in (2.3) and take the beamforming gain into account when choosing $\boldsymbol{w}_i$. The details will be explained in subsequent sections. In the next section, we will introduce the mmWave channel model which is very different from the traditional Rayleigh fading channel model. Based on the mmWave channel model, we will formulate our analog multi-user beamforming problem for mmWave systems.

2.2.2 Channel model

MmWave channels are expected to have limited scattering characteristic [98], which means the assumptions of a rich scattering environment become invalid. This is called sparsity in the literature and leads to the unreliability of traditional channel models, such as the Rayleigh fading channel model. To characterize the limited scattering feature, we adopt the clustered mmWave channel model in [33] and[32] with L_i scatters for the channel of User i. Each scatter is assumed to contribute a single propagation path between the BS and the user. For our single-antenna user system, the channel is modeled as a vector described by

$$\boldsymbol{h}_i^H = \sqrt{\frac{N_t}{L_i}} \sum_{l=1}^{L_i} (a_l^i)^* \boldsymbol{\alpha}_t (\theta_l^i)^H, \tag{2.4}$$

where $\boldsymbol{\alpha}_t(\theta_l^i)$ is the antenna array response vectors of the BS for path l with departure angle θ_l^i. Parameter $(a_l^i)^*$ is the complex path gain of path l modeled by a complex Gaussian distribution such as $\mathscr{CN}(0,1)$. While the algorithms and results in the chapter can be applied to arbitrary antenna arrays, we use uniform linear arrays (ULAs) in the simulations for simplicity. The array response vectors take the following form

$$\boldsymbol{\alpha}_t(\theta_l^i) = \frac{1}{\sqrt{N_t}}[1, e^{j\frac{2\pi}{\lambda}d\sin(\theta_l^i)}, ..., e^{j(N_t-1)\frac{2\pi}{\lambda}d\sin(\theta_l^i)}], \tag{2.5}$$

where λ is the signal wavelength, and d is the distance between antenna elements. The departure angle θ_l^i is assumed to have a uniform distribution over $[0, 2\pi]$.

To simplify the expression of the channels, we denote

$$\mathbf{A}_i = [\boldsymbol{\alpha}_t(\theta_1^i), \boldsymbol{\alpha}_t(\theta_2^i), ..., \boldsymbol{\alpha}_t(\theta_L^i)], \tag{2.6}$$

$$\tilde{\boldsymbol{h}}_i = [a_1^i, a_2^i, ..., a_L^i]^T. \tag{2.7}$$

Matrix $\mathbf{A}_i \in \mathbb{C}^{N_t \times L_i}$ contains all the array response vectors from the BS to User i and vector $\tilde{\boldsymbol{h}}_i \in \mathbb{C}^{L_i \times 1}$ contains the complex gain of all the paths from the BS to User i. The channel $\boldsymbol{h}_i^H$ can be expressed as the product of Hermitian $\tilde{\boldsymbol{h}}_i$ and Hermitian $\mathbf{A}_i$

$$\boldsymbol{h}_i^H = \tilde{\boldsymbol{h}}_i^H \mathbf{A}_i^H. \tag{2.8}$$

We call $\mathbf{A}_i$ the AoD matrix of User i. In fact, to estimate the mmWave channels is to estimate the AoDs and the complex gains. In this chapter, to further reduce the feedback overhead, we assume the BS only knows the AoD of the channels (i.e., the BS only knows $\mathbf{A}_i$).

2.2.3 Problem formulation

As we mentioned in the previous section, ZF schemes may not be effective for multi-user analog beamforming since the constant-magnitude constraints on $\boldsymbol{w}_i$ may cause the ZF conditions in (2.3) infeasible. To deal with this problem, we relax the equality condition in (2.3) and try to minimize the leakage interference. To be specific, we denote the leakage interference matrix as

$$\tilde{\mathbf{I}}_i = [\mathbf{A}_1, ..., \mathbf{A}_{i-1}, \mathbf{A}_{i+1}, ..., \mathbf{A}_K]^H, \tag{2.9}$$

where $\tilde{\mathbf{I}}_i \in \mathbb{C}^{\sum_{k=1,k\neq i}^{K} L_k \times N_t}$ is a matrix that contains the AoD matrices from the BS to all the users except User i. Originally, $\tilde{\mathbf{I}}_i$ should contain the channel vectors from the BS to all the users except User i. Since we assume we only know the AoDs of the channels, we use AoD matrices to represent the channels.

The traditional ZF schemes are basically forcing the $\boldsymbol{w}_i$ to lie in the null space of $\tilde{\mathbf{I}}_i$ so as to avoid the interference from User i to other users. The null space of $\tilde{\mathbf{I}}_i$ can be obtained through singular-value decomposition (SVD). We define the SVD of $\tilde{\mathbf{I}}_i$ as

$$\tilde{\mathbf{I}}_i = \tilde{\mathbf{U}}_i \tilde{\Sigma}_i [\tilde{\mathbf{V}}_i^{(1)} \;\; \tilde{\mathbf{V}}_i^{(0)}]^H, \tag{2.10}$$

where $\tilde{\mathbf{V}}_i^{(1)}$ holds the first $\sum_{k=1,k\neq i}^{K} L_k$ right singular vectors and $\tilde{\mathbf{V}}_i^{(0)}$ holds the last $N_t - \sum_{k=1,k\neq i}^{K} L_k$ right singular vectors. We assume that we implement a large antenna array at the BS, which means N_t is very large. Note that the mmWave channels have limited scattering characteristic, which means L_k is usually small. Normally, $N_t \gg L_k$. Therefore, we assume $N_t > \sum_{k=1,k\neq i}^{K} L_k$ to ensure that the null space of $\tilde{\mathbf{I}}_i$ (i.e., $\tilde{\mathbf{V}}_i^{(0)}$) exists.

For the multi-user analog beamforming, where ZF conditions may be infeasible, we want to minimize the leakage interference of $\boldsymbol{w}_i$. That is actually minimizing the projection from $\boldsymbol{w}_i$ to $\tilde{\mathbf{I}}_i$, which means $\boldsymbol{w}_i$ should have the largest projection on the null space of $\tilde{\mathbf{I}}_i$. In other words, we are trying

to find the $\boldsymbol{w}_i$ that has the largest projection on the null space of $\tilde{\mathbf{I}}_i$. The projected vector from $\boldsymbol{w}_i$ to the null space of $\tilde{\mathbf{I}}_i$ is

$$\boldsymbol{w}_i^p = \tilde{\mathbf{V}}_i^{(0)}(\tilde{\mathbf{V}}_i^{(0)})^H \boldsymbol{w}_i, \tag{2.11}$$

where $\tilde{\mathbf{V}}_i^{(0)}(\tilde{\mathbf{V}}_i^{(0)})^H$ is the projector matrix onto the the null space of $\tilde{\mathbf{I}}_i$. For simplicity, we maximize the square of the norm of $\boldsymbol{w}_i^p$ which is

$$\|\boldsymbol{w}_i^p\|^2 = \boldsymbol{w}_i^H \tilde{\mathbf{V}}_i^{(0)}(\tilde{\mathbf{V}}_i^{(0)})^H \tilde{\mathbf{V}}_i^{(0)}(\tilde{\mathbf{V}}_i^{(0)})^H \boldsymbol{w}_i = \boldsymbol{w}_i^H \tilde{\mathbf{V}}_i^{(0)}(\tilde{\mathbf{V}}_i^{(0)})^H \boldsymbol{w}_i. \tag{2.12}$$

To obtain an optimal SINR, only to minimize the leakage interference is not enough, since it ignores the beamforming gain. The beamforming gain refers to the improvement of the receive power which results from beamforming and we define the beamforming gain for $\boldsymbol{w}_i$ under our partial channel information assumption as

$$BG = \boldsymbol{w}_i^H \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{w}_i. \tag{2.13}$$

Taking both the beamforming gain and the leakage interference into account, we formulate a multi-objective optimization problem (MOP) as follows:

$$\begin{aligned} &\boldsymbol{w}_i^{opt} = \text{argmax}\{\boldsymbol{w}_i^H \tilde{\mathbf{V}}_i^{(0)}(\tilde{\mathbf{V}}_i^{(0)})^H \boldsymbol{w}_i, \boldsymbol{w}_i^H \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{w}_i\} \\ &s.t.\ \ \boldsymbol{w}_i \in \mathscr{W}, \end{aligned} \tag{2.14}$$

where $\mathscr{W}$ is the set of all constant-magnitude vectors with each element having a magnitude of $\frac{1}{\sqrt{N_t}}$. Problem (2.14) is a multi-objective problem with non-convex constraints, which is intractable. In the next section, we will solve this problem by the weighted sum method and the SDP technique.

2.3 Multi-user analog beamforming

To solve Problem (2.14), we are actually facing two challenges: 1) how to deal with the multi-objective problem; and 2) how to handle the non-convex constraints. For the MOP, we will use the simplest yet effective method (i.e., the weighted sum method), to transform it into an SOP. Then, through some algebraic transformation, we will transform the non-convex problem into an SDP and solve it using convex optimization.

2.3.1 Transforming the MOP into an SOP

The solution to an MOP may not exist because a single point that optimizes all objectives simultaneously usually does not exist. The idea of Pareto optimality is usually used to describe solutions for MOPs. A solution point is Pareto optimal if it is not possible to move from that point and improve at least one objective function without detriment to any other objective function. Alternatively, a point is weakly Pareto optimal if it is not possible to move from that point and improve all objective functions simultaneously. To solve an MOP, we need to ensure the necessary and/or sufficient condition for Pareto optimality. In other words, to solve an MOP is to find the Pareto optimal points. There are many ways to find the Pareto optimal points. Two general methods are visualization and scalarization. A scalarization method specifies a goal function $f : R^M \rightarrow R$ that for any conceivable operating point, it produces a scalar describing how preferable that point is (large value means high preference). To be specific, it means to solve the optimization problem:

$$\begin{aligned} &\max_{\boldsymbol{x}} f(g_1(\boldsymbol{x}), ..., g_M(\boldsymbol{x})) \\ s.t.\ \ &\boldsymbol{x} \in \mathscr{X} \end{aligned} \tag{2.15}$$

where $g_i(\boldsymbol{x})$ is the i^{th} objective function in the original MOP and $\mathscr{X}$ is the feasible region. For the weighted-sum method, function f is a weighted summation of the objective functions. As stated

in [85], the weighted sum method can provide a sufficient condition for Pareto optimality if all the weights are positive and the summation of the weights is equal to one. In other words, the solution to the SOP formulated by the weighted summation is Pareto optimal for the MOP.

Applying the weighted sum method to our problem results in

$$\begin{aligned} \boldsymbol{w}_i^{opt} &= \text{argmax}\{\lambda_1 \boldsymbol{w}_i^H \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H \boldsymbol{w}_i + \lambda_2 \boldsymbol{w}_i^H \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{w}_i\} \\ s.t.\ \ \boldsymbol{w}_i &\in \mathscr{W} \end{aligned} \tag{2.16}$$

where $\lambda_1 + \lambda_2 = 1$ and $\lambda_i > 0, i \in \{1,2\}$. Parameter λ_i represents the importance of the i^{th} component in the objective function. Different values for λ_is will result in different solutions to the problem. If we want to obtain a smaller leakage interference, we will set a larger λ_1 for the first objective function. If we want to obtain a larger beamforming gain, we will set a larger λ_2 for the second objective function. We will evaluate the performance under different values of λ_is in Section V.

2.3.2 SDP formulation

Although we have transformed Problem (2.14) into Problem (2.16), it is still hard to solve because of the non-convex constraints. To make the problem tractable, we transform the problem into an SDP through some algebraic transformation. For the objective function, we have

$$\begin{aligned} &\lambda_1 \boldsymbol{w}_i^H \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H \boldsymbol{w}_i + \lambda_2 \boldsymbol{w}_i^H \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{w}_i \\ &= \text{Tr}(\boldsymbol{w}_i^H (\lambda_1 \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H + \lambda_2 \mathbf{A}_i \mathbf{A}_i^H) \boldsymbol{w}_i) \\ &= \text{Tr}((\lambda_1 \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H + \lambda_2 \mathbf{A}_i \mathbf{A}_i^H) \boldsymbol{w}_i \boldsymbol{w}_i^H). \end{aligned} \tag{2.17}$$

We denote $\boldsymbol{w}_i \boldsymbol{w}_i^H$ as $\mathbf{W}$. Matrix $\mathbf{W}$ is a symmetric semi-definite matrix with rank one.

The constant-magnitude constraints of Problem (2.16) are transformed into

$$\mathbf{W}_{ii} = \frac{1}{N_t},\ \forall i = 1, ..., N_t, \tag{2.18}$$

where $\mathbf{W}_{ii}$ represents the i^{th} diagonal element in $\mathbf{W}$.

Then, Problem (2.17) is transformed into an SDP as follows:

$$\begin{aligned} & SDP(\mathbf{W}^{opt}) = \text{argmax}\{\text{Tr}((\lambda_1 \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H + \lambda_2 \mathbf{A}_i \mathbf{A}_i^H)\mathbf{W})\} \\ s.t.\ & \mathbf{W}_{ii} = \frac{1}{N_t},\ \forall i = 1, ..., N_t; \\ & \mathbf{W} \succeq 0; \\ & \text{Rank}(\mathbf{W}) = 1. \end{aligned} \tag{2.19}$$

The rank-one constraint is still hard to deal with. In order to efficiently solve the optimization problem, we introduce semidefinite programming relaxation (SDR) by dropping the rank-one constraint in (2.19) to solve the optimization problem as (2.20) shows.

$$\begin{aligned} & SDR(\mathbf{W}^{opt}) = \text{argmax}\{\text{Tr}((\lambda_1 \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H + \lambda_2 \mathbf{A}_i \mathbf{A}_i^H)\mathbf{W})\} \\ s.t.\ & \mathbf{W}_{ii} = \frac{1}{N_t},\ \forall i = 1, ..., N_t; \\ & \mathbf{W} \succeq 0. \end{aligned} \tag{2.20}$$

Problem (2.20) is the relaxed version of Problem (2.19). Its solution will be an upper bound for the solution of the optimization problem in Problem (2.19).

2.3.3 Approximation

Problem (2.20) is a standard SDP. Its optimal solution $SDR(\mathbf{W}^{opt})$ can be found by standard tools of mathematical programming such as CVX [41]. Note that Problem (2.20) is the relaxed version

of Problem (2.19), which means we cannot guarantee $SDR(\mathbf{W}^{opt})$ is rank-one. In fact, according to our simulation results, $SDR(\mathbf{W}^{opt})$ is not necessarily rank-one. When the rank of $SDR(\mathbf{W}^{opt})$ is larger than one, we cannot recover $\boldsymbol{w}_i^{opt}$ from $SDR(\mathbf{W}^{opt})$ straightforwardly. In [122], a randomization technique is used to make an approximation. Its basic idea is to generate a set of candidate vectors $\{\boldsymbol{z}_m\}_{m=1}^{M}$ based on $SDR(\mathbf{W}^{opt})$ and choose the best candidate as the approximation of the $\boldsymbol{w}_i^{opt}$, where M is the size of the set of the candidate vectors.

To be specific, we first generate a set of complex Gaussian vectors $\mathscr{X} = \{\boldsymbol{x}_m \in \mathbb{C}^{N_t \times 1}\}_{m=1}^{M}$. For each $\boldsymbol{x}_m \in \mathscr{X}$, we generate the vector randomly using the distribution $\mathscr{N}(\mathbf{0}, SDR(\mathbf{W}^{opt}))$, where $\mathbf{0} \in \mathbb{C}^{N_t \times 1}$ is the mean vector and $SDR(\mathbf{W}^{opt})$ is the covariance matrix of the Gaussian random vector. In this way, we will have $E[\boldsymbol{x}_m \boldsymbol{x}_m^H] = SDR(\mathbf{W}^{opt})$. However, $\boldsymbol{x}_m$ may not have constant magnitudes for every element, thus it may not be a feasible solution for Problems (2.19) and (2.20). To deal with this issue, for each $\boldsymbol{x}_m$, we form $\boldsymbol{z_m} \in \mathbb{C}^{N_t \times 1}$ such that

$$\boldsymbol{z}_m^n = \frac{x_m^n}{\sqrt{N_t}|x_m^n|}, \ \forall n = 1, ..., N_t, \tag{2.21}$$

where $\boldsymbol{z}_m^n$ is the n^{th} element of vector $\boldsymbol{z}_m$ and x_m^n is the n^{th} element of vector $\boldsymbol{x}_m$. In this way, all the generated $\boldsymbol{z}_m$s satisfy the constant-magnitude constraints, and hence are feasible points for Problems (2.19) and (2.20). Therefore, we have

$$\mathrm{Tr}((\lambda_1 \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H + \lambda_2 \mathbf{A}_i \mathbf{A}_i^H) \boldsymbol{z}_m \boldsymbol{z}_m^H) \leq SDP(\mathbf{W}^{opt}) \leq SDR(\mathbf{W}^{opt}). \tag{2.22}$$

Since for all $\boldsymbol{z}_m s$, Inequality (2.22) will hold, this means

$$\mathrm{Tr}((\lambda_1 \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H + \lambda_2 \mathbf{A}_i \mathbf{A}_i^H) E[\boldsymbol{z}_m \boldsymbol{z}_m^H]) \leq SDP(\mathbf{W}^{opt}) \leq SDR(\mathbf{W}^{opt}). \tag{2.23}$$

Based on Inequality (2.23), we can choose the $\boldsymbol{z}_m$ from the candidate set $\mathscr{Z} = \{\boldsymbol{z}_m\}_{m=1}^{M}$ that maximizes $\boldsymbol{z}_m^H (\lambda_1 \tilde{\mathbf{V}}_i^{(0)} (\tilde{\mathbf{V}}_i^{(0)})^H + \lambda_2 \mathbf{A}_i \mathbf{A}_i^H) \boldsymbol{z}_m$ as the approximation for $\boldsymbol{w}_i^{opt}$.

Up to now, we have proposed a multi-user analog beamforming method based on the precisely estimated AoD. However, due to the limited feedback and imperfect estimation, we can not obtain the accurate AoD for every path, which makes the proposed beamforming scheme unstable. To deal with the uncertainty in the estimation of AoD, we further propose a robust beamforming scheme in the next section.

2.4 Robust Beamforming

To design the robust beamforming scheme, we first need to model the estimation errors. For traditional Rayleigh fading channel models, the estimation errors is simply modeled as a matrix consisting of i.i.d. complex Gaussian distributed entries, which is directly added to the presumed channel. However, in the clustered mmWave channel model, the errors cannot be simply modeled as the additive estimation errors, since the estimated angle errors appear in the index of the exponential function in the array response vectors. Therefore, we need to simplify the error model before designing a robust beamforming scheme.

2.4.1 Error model

In this section, we will develop an error model for the ULAs with an array response vector in (2.5). We assume that for the angle θ_l^i of the path l there exists an angle estimation/quantization error $\Delta\theta_l^i$ with mean 0 and variance σ_l^i. A Gaussian distribution $\mathcal{N}(0,\sigma_l^i)$ is a reasonable assumption, although we only use the first and second order statistics and do not need the distribution. Then, the array response vector with error $\Delta\theta_l^i$ can be expressed as

$$\boldsymbol{\alpha}(\theta_l^i+\Delta\theta_l^i)=\frac{1}{\sqrt{N_t}}[1,e^{j\frac{2\pi}{\lambda}d\sin(\theta_l^i+\Delta\theta_l^i)},...,e^{j(N_t-1)\frac{2\pi}{\lambda}d\sin(\theta_l^i+\Delta\theta_l^i)}]^T. \tag{2.24}$$

To extract the error out of the exponential function in (2.24), we expand the exponential function using the first-order Taylor expansion. To simplify the expression, we denote $\frac{2\pi}{\lambda}d$ as κ and the equation (2.24) is expanded as

$$e^{jn\kappa\sin(\theta_l^i+\Delta\theta_l^i)} \approx e^{jn\kappa\sin(\theta_l^i)} + jn\kappa\cos(\theta_l^i)\Delta\theta_l^i e^{jn\kappa\sin(\theta_l^i)}, \quad \forall n = 0,...,N_t-1. \tag{2.25}$$

We denoting $e_l^{i,n}$ as $jn\kappa\cos(\theta_l^i)\Delta\theta_l^i e^{jn\kappa(\sin(\theta_l^i)}$, which represents the error for the n^{th} element in the response vector of the l^{th} path of User i. Error $e_l^{i,n}$ can be written as:

$$e_l^{i,n} = jn\kappa\cos(\theta_l^i)\Delta\theta_l^i\cos(n\kappa(\sin(\theta_l^i))-n\kappa\cos(\theta_l^i)\Delta\theta_l^i\sin(n\kappa(\sin(\theta_l^i)),\forall n=0,...,N_t-1. \tag{2.26}$$

Defining the error vector $\boldsymbol{e}_l^i \triangleq \frac{1}{\sqrt{N_t}}[e_l^{i,0}, e_l^{i,1}, ..., e_l^{i,N_t-1}]^T$ as the error for the l^{th} path of User i, we now simplify the errors in the AoD into an additive random error as

$$\tilde{\boldsymbol{\alpha}}(\theta_l^i) = \boldsymbol{\alpha}(\theta_l^i+\Delta\theta_l^i) \approx \boldsymbol{\alpha}(\theta_l^i) + \boldsymbol{e}_l^i, \tag{2.27}$$

where we denote $\tilde{\boldsymbol{\alpha}}(\theta_l^i)$ as the imperfect array response vector of the l^{th} path of User i. Based on the mean and the variance of $\Delta\theta_l^i$, we can calculate the statistical characteristic of $\boldsymbol{e}_l^i$. First, we calculate the statistical characteristic of each element $e_l^{i,n}$ in the vector $\boldsymbol{e}_l^i$. Then, we calculate the cross-covariance between different elements (i.e., $e_l^{i,n}$ and $e_l^{i,m}$ where $m \neq n$) so as to calculate the covariance matrix of e_l^i. The mean and variance of each element $e_l^{i,n}$ are calculated by the following Eqs. (2.28) and (2.29).

$$E[e_l^{i,n}] = E[\Delta\theta_l^i](jn\kappa\cos(\theta_l^i)\cos(n\kappa(\sin(\theta_l^i))- n\kappa\cos(\theta_l^i)\sin(n\kappa(\sin(\theta_l^i)))) = 0, \forall n = 0,...,N_t-1. \tag{2.28}$$

$$
\begin{aligned}
D[e_l^{i,n}] = E[(e_l^{i,n})^* e_l^{i,n}] = E[(n\kappa\cos(\theta_l^i)\Delta\theta_l^i)^2 \\
(\cos^2(n\kappa(\sin(\theta_l^i)) + \sin^2(n\kappa(\sin(\theta_l^i))))] \\
= (n\kappa\cos(\theta_l^i))^2 E[(\Delta\theta_l^i)^2] = (n\kappa\cos(\theta_l^i)\sigma_l^i)^2, \\
\forall n = 0, ..., N_t - 1.
\end{aligned}
\tag{2.29}
$$

The cross-covariance between $e_l^{i,n}$ and $e_l^{i,m}$ is calculated as

$$
E[(e_l^{i,n})^* e_l^{i,m}] = nm\kappa^2\cos^2(\theta_l^i)(\sigma_l^i)^2, \forall m, n \; and \; m \neq n \tag{2.30}
$$

Based on Equations (2.29) and (2.30), we can calculate the covariance matrix of $\boldsymbol{e}_l^i$ as

$$
\mathbf{C}_l^i = \begin{bmatrix}
0 & 0 & \dots & 0 \\
0 & (\kappa\cos(\theta_l^i)\sigma_l^i)^2 & \dots & (N_t-1)\kappa^2\cos^2(\theta_l^i)(\sigma_l^i)^2 \\
\vdots & \vdots & \ddots & \vdots \\
0 & (N_t-1)\kappa^2\cos^2(\theta_l^i)(\sigma_l^i)^2 & \dots & (N_t-1)^2\kappa^2\cos^2(\theta_l^i)(\sigma_l^i)^2
\end{bmatrix}. \tag{2.31}
$$

Note that the first row and the first column of $\mathbf{C}_l^i$ are all zeros. This is because the first element of the array response vector (2.24) is always 1, which is independent of an error. In other words, the first element of the array response vector (2.24) is deterministic and this leads to the zeros in the first row and the first column of $\mathbf{C}_l^i$.

Since we have simplified the AoD error of each path for each user into an additive error, we can further model the errors for the whole AoD matrix as an additive error. Denoting the presumed AoD matrix of User i as $\mathbf{A}_i^p$, the AoD matrix of User i with errors can be modeled as

$$
\mathbf{A}_i = \mathbf{A}_i^p + \mathbf{E}_i, \tag{2.32}
$$

where $\mathbf{E}_i = [\boldsymbol{e}_1^i, \boldsymbol{e}_2^i, ..., \boldsymbol{e}_L^i] \in \mathbb{C}^{N_t \times L}$ is a matrix that contains all the error vectors for User i. Based on the assumption that the paths are independent [33][32], we can assume that the errors of different paths and users are independent, which means

$$E[\boldsymbol{e}_l^i (\boldsymbol{e}_q^i)^H] = \mathbf{0}_{N_t \times N_t}, \ \ \forall l \neq q. \tag{2.33}$$

In (2.33), $\mathbf{0}_{N_t \times N_t}$ represents the zero square matrix with dimension of N_t. Therefore, the covariance matrix of $\mathbf{E}_i$ is

$$\mathbf{C}_i = \sum_{l=1}^{L_i} \mathbf{C}_l^i. \tag{2.34}$$

The imperfect leakage interference matrix of User i could also be modeled in the same way as the imperfect AoD matrix. We denote the presumed leakage interference matrix of User i as $\tilde{\mathbf{I}}_i^p = [\mathbf{A}_1^p, ..., \mathbf{A}_{i-1}^p, \mathbf{A}_{i+1}^p, ..., \mathbf{A}_K^p]^T$. The imperfect leakage interference matrix of User i with errors can be modeled as

$$\tilde{\mathbf{I}}_i = \tilde{\mathbf{I}}_i^p + \tilde{\mathbf{E}}_i, \tag{2.35}$$

where $\tilde{\mathbf{E}}_i = [\mathbf{E}_1, ..., \mathbf{E}_{i-1}, \mathbf{E}_{i+1}, ..., \mathbf{E}_K]^T \in \mathbb{C}^{\sum_{k \neq i}^K L_k \times N_t}$ is a matrix that contains all the error matrices for all the users except User i. We assume the errors of different users are independent, which means

$$E[\mathbf{E}_i \mathbf{E}_j^H] = \mathbf{0}_{N_t \times N_t}, \ \ \forall i \neq j. \tag{2.36}$$

Therefore, the covariance matrix of $\tilde{\mathbf{E}}_i$ is

$$\tilde{\mathbf{C}}_i = \sum_{k \neq i}^{K} \mathbf{C}_k. \tag{2.37}$$

Now, we have simplified both the errors in the AoD matrix and the leakage interference matrix into the additive error. Based on this error model, we will propose a robust beamforming scheme to confront the uncertainty in the channel information.

2.4.2 Robust beamforming

The leakage interference matrix is random due to the uncertainty of errors. This means we cannot find a valid null space of $\tilde{\mathbf{I}}_i$. Therefore, the beamforming method proposed in the previous sections cannot be applied here. To deal with this problem, we use a probabilistic approach to restrict the leakage interference (i.e., we maximize the outage probability). The outage probability can be expressed as

$$P_{outage} = \Pr\{\boldsymbol{w}_i^H \tilde{\mathbf{I}}_i^H \tilde{\mathbf{I}}_i \boldsymbol{w}_i \leq \gamma_i\}, \tag{2.38}$$

where γ_i denotes a pre-specified leakage power level. Besides the leakage power, we also want to maximize the average beamforming gain of User i, which is defined as

$$BG_{avg} = E[\boldsymbol{w}_i^H \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{w}_i]. \tag{2.39}$$

We want to maximize the outage probability and the average beamforming gain at the same time. Therefore, a multi-objective optimization problem is constructed as

$$\begin{aligned} \boldsymbol{w}_i^{opt} &= \text{argmax } \{E[\boldsymbol{w}_i^H \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{w}_i], \Pr\{\boldsymbol{w}_i^H \tilde{\mathbf{I}}_i^H \tilde{\mathbf{I}}_i \boldsymbol{w}_i \leq \gamma_i\}\} \\ s.t.\ \boldsymbol{w}_i &\in \mathscr{W}, \end{aligned} \tag{2.40}$$

where $\mathscr{W}$ is the set of all constant-magnitude vectors with each element having a magnitude of $\frac{1}{\sqrt{N_t}}$. Problem (2.40) is an MOP with a constant-magnitude constraint and a probabilistic objective function. We first use Markov's inequality to transform the probabilistic objective into the expecta-

tion objective. Then, similar to the beamforming scheme we proposed in the previous part, we use the weighted-sum method and the SDP to deal with the multi-objective and constant-magnitude constraint, respectively.

Based on Markov's inequality, we have the following inequality as a foundation for our simplification of the probabilistic objective.

$$\Pr\{Z \leq \gamma\} = 1 - \Pr\{Z \geq \gamma\} \geq 1 - \frac{E[Z]}{\gamma}. \tag{2.41}$$

Then the probabilistic objective can be simplified as

$$\Pr\{\boldsymbol{w}_i^H \tilde{\mathbf{I}}_i^H \tilde{\mathbf{I}}_i \boldsymbol{w}_i \leq \gamma_i\} = \Pr\{\boldsymbol{w}_i^H (\tilde{\mathbf{I}}_i^p + \tilde{\mathbf{E}}_i)^H (\tilde{\mathbf{I}}_i^p + \tilde{\mathbf{E}}_i) \boldsymbol{w}_i \leq \gamma_i\} \tag{2.42a}$$

$$\geq 1 - \frac{E[\boldsymbol{w}_i^H (\tilde{\mathbf{I}}_i^p + \tilde{\mathbf{E}}_i)^H (\tilde{\mathbf{I}}_i^p + \tilde{\mathbf{E}}_i) \boldsymbol{w}_i]}{\gamma_i} \tag{2.42b}$$

$$= 1 - \frac{E[\mathrm{Tr}((\tilde{\mathbf{I}}_i^p + \tilde{\mathbf{E}}_i)^H (\tilde{\mathbf{I}}_i^p + \tilde{\mathbf{E}}_i) \boldsymbol{w}_i \boldsymbol{w}_i^H)]}{\gamma_i} \tag{2.42c}$$

$$= 1 - \frac{1}{\gamma_i} E[\mathrm{Tr}((\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{I}}_i^p \boldsymbol{w}_i \boldsymbol{w}_i^H + \tilde{\mathbf{E}}_i^H \tilde{\mathbf{I}}_i^p \boldsymbol{w}_i \boldsymbol{w}_i^H + \tag{2.42d}$$

$$(\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{E}}_i) \boldsymbol{w}_i \boldsymbol{w}_i^H + \tilde{\mathbf{E}}_i^H \tilde{\mathbf{E}}_i \boldsymbol{w}_i \boldsymbol{w}_i^H)] \tag{2.42e}$$

$$= 1 - \frac{1}{\gamma_i} \mathrm{Tr}(E[(\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{I}}_i^p \boldsymbol{w}_i \boldsymbol{w}_i^H] + E[\tilde{\mathbf{E}}_i^H \tilde{\mathbf{I}}_i^p] \boldsymbol{w}_i \boldsymbol{w}_i^H + \tag{2.42f}$$

$$E[(\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{E}}_i)] \boldsymbol{w}_i \boldsymbol{w}_i^H + E[\tilde{\mathbf{E}}_i^H \tilde{\mathbf{E}}_i] \boldsymbol{w}_i \boldsymbol{w}_i^H) \tag{2.42g}$$

$$= 1 - \frac{\mathrm{Tr}(((\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{I}}_i^p + \tilde{\mathbf{C}}_i) \boldsymbol{w}_i \boldsymbol{w}_i^H)}{\gamma_i} \tag{2.42h}$$

$$= 1 - \frac{\mathrm{Tr}(((\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{I}}_i^p + \tilde{\mathbf{C}}_i) \mathbf{W})}{\gamma_i}. \tag{2.42i}$$

In (2.42f), we exchange the operation order of the expectation and the Tr. Matrix $\mathbf{W} = \boldsymbol{w}_i \boldsymbol{w}_i^H$ is a symmetric semi-definite matrix with rank one. Equations in (2.42) transform the probabilistic objective function into a deterministic and convex function of $\mathbf{W}$.

The average beamforming gain for User i is an expectation over the instant beamforming gain,

which is not easy to deal with. To make the problem tractable, we perform some algebraic transformation and convert it into a deterministic and convex function of $\mathbf{W}$ as below.

$$\begin{aligned}
&E[\boldsymbol{w}_i^H \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{w}_i] \\
&= E[\boldsymbol{w}_i^H (\mathbf{A}_i^p + \mathbf{E}_i)(\mathbf{A}_i^p + \mathbf{E}_i)^H \boldsymbol{w}_i] \\
&= E[\mathrm{Tr}((\mathbf{A}_i^p + \mathbf{E}_i)(\mathbf{A}_i^p + \mathbf{E}_i)^H \boldsymbol{w}_i \boldsymbol{w}_i^H] \\
&= \mathrm{Tr}(E[(\mathbf{A}_i^p + \mathbf{E}_i)(\mathbf{A}_i^p + \mathbf{E}_i)^H \boldsymbol{w}_i \boldsymbol{w}_i^H]) \\
&= \mathrm{Tr}(E[\mathbf{A}_i^p (\mathbf{A}_i^p)^H \boldsymbol{w}_i \boldsymbol{w}_i^H] + \\
&\quad E[\mathbf{E}_i (\mathbf{A}_i^p)^H + \mathbf{A}_i^p \mathbf{E}_i^H] \boldsymbol{w}_i \boldsymbol{w}_i^H + E[\mathbf{E}_i \mathbf{E}_i^H] \boldsymbol{w}_i \boldsymbol{w}_i^H) \\
&= \mathrm{Tr}((\mathbf{A}_i^p (\mathbf{A}_i^p)^H + \mathbf{C}_i) \boldsymbol{w}_i \boldsymbol{w}_i^H) \\
&= \mathrm{Tr}((\mathbf{A}_i^p (\mathbf{A}_i^p)^H + \mathbf{C}_i) \mathbf{W}).
\end{aligned} \tag{2.43}$$

The introduction of matrix $\mathbf{W}$ will transform the non-convex constraints on $\boldsymbol{w}_i$ into

$$\mathbf{W}_{ii} = \frac{1}{N_t},\ \forall i = 1, ..., N_t, \tag{2.44}$$

where $\mathbf{W}_{ii}$ represents the i^{th} diagonal element in $\mathbf{W}$. This constraints are convex constraints and are easy to deal with.

Based on the above three simplifications, using the weighted-sum method, we can reformulate Problem (2.40) into

$$\begin{aligned}
&SDP_{robust}(\mathbf{W}^{opt}) = \mathrm{argmax}\{\lambda_1 \mathrm{Tr}((\mathbf{A}_i^p (\mathbf{A}_i^p)^H + \mathbf{C}_i)\mathbf{W}) + \\
&\lambda_2 \left(1 - \frac{\mathrm{Tr}(((\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{I}}_i^p + \tilde{\mathbf{C}}_i)\mathbf{W})}{\gamma_i}\right)\} \\
s.t.\ \ &\mathbf{W}_{ii} = \frac{1}{N_t},\ \forall i = 1, ..., N_t; \\
&\mathbf{W} \succeq 0; \\
&\mathrm{Rank}(\mathbf{W}) = 1,
\end{aligned} \tag{2.45}$$

where $\lambda_1 + \lambda_2 = 1$. Parameter λ_i represents the importance of the i^{th} component in the cost function.

To deal with the rank-one constraint, we introduce the SDR by dropping the rank constraint in (2.45). Therefore, an upper bound of Problem (2.45) can be achieved.

$$\begin{aligned} & SDR_{robust}(\mathbf{W}^{opt}) = \text{argmax}\{\lambda_1 \,\text{Tr}((\mathbf{A}_i^p(\mathbf{A}_i^p)^H + \mathbf{C}_i)\mathbf{W}) + \\ & \lambda_2 \left(1 - \frac{\text{Tr}(((\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{I}}_i^p + \tilde{\mathbf{C}}_i)\mathbf{W})}{\gamma_i}\right)\} \\ s.t.\ \ & \mathbf{W}_{ii} = \frac{1}{N_t},\ \forall i = 1,...,N_t; \\ & \mathbf{W} \succeq 0. \end{aligned} \tag{2.46}$$

The optimal solution $SDR_{robust}(\mathbf{W}^{opt})$ can be found by standard tools of mathematical programming [41]. Using the same approximation method as in the non-robust case, we can obtain the analog beamforming vector $\boldsymbol{w}_i^{opt}$.

2.5 Simulation Results

In this section, we evaluate the performance of the non-robust beamforming method in Section 2.3 and the robust beamforming method in Section 2.4. Note that our objective in this chapter is not to optimize the sum-rate due to the intractability of doing so. In fact, we strike a balance between maximizing the beamforming gain and minimizing the inter-user interference. Since the λ_i represents the importance of each term in the objective function of the MOP, we expect to find the best balance by evaluating different assignments of values of $\lambda_i s$. Therefore, we pick the combination of λ_1 and λ_2 that achieve the highest sum-rate. We also compare our multi-user analog beamforming with the beam selection method and other traditional fully-digital beamforming methods. For the robust beamforming method, we compare our robust analog beamforming with the non-robust ones to evaluate the improvement brought by our method.

2.5.1 Non-robust analog beamforming

In the simulation, we consider a multi-user MIMO system consisting of one BS equipped with a large antenna array and K single-antenna users. The channels are realized using Eq. (2.4). Due to the limited scattering characteristic of the mmWave channels, the number of paths should be small. Here, we assume each channel has $L = 6$ paths. The large antenna array at the BS is assumed to have $N_t = 64$ antennas, which is the same number of antennas in [1]. To leave enough dimension for the null space of the leakage interference, we assume the total number of users $K = 6$. This will leave $N_t - KL = 28$ dimensions for the null space. The θ_l^i of each path is assumed to be uniformly distributed in $[0, 2\pi]$. The results are averaged over $20,000$ channel realizations. The variance of AWGN noise per user is assumed to be the same for all users, i.e. $\sigma_1^2 = ... = \sigma_K^2 = \sigma^2$. And the large-scale fading path loss factor of all users are uniformly distributed in $[0.5, 1.5]$ dB. We have used these parameters in all figures unless we specifically mention otherwise.

Figs. 2.2 and 2.3 illustrate the sum-rate and the beamforming gain and interference of our analog beamforming under different λ_1 and λ_2 values, respectively. We, in general, evaluate 21 combinations of λ_1 and λ_2. To be specific, the λ_2 ranges from 0 to 1 with step-size 0.05 and $\lambda_1 = 1 - \lambda_2$. In Fig. 2.2, as λ_2 increases from 0 to 1, the sum-rate first increases and then decreases, which exhibits the tradeoff between the beamforming gain and the leakage interference. In Fig. 2.3, when $\lambda_2 = 0$, the leakage interference is minimized but the beamforming gain is quite low. When $\lambda_1 = 1$, the beamforming gain is maximized but the leakage interference is very high. Fig. 2.2 shows that there is a tradeoff between leakage interference and beamforming gain and as a result the MOP provides better performance than only optimizing beamforming gain or interference. Note that there is no need to set different weights for different users because we deal with a decoupled problem, i.e., the best λ_2 for each user should be the same. For example, when we fix the parameters for other users and change the weight the for User 1, the trend in the plot should be similar to Fig. 2.2. Therefore, we will set $\lambda_1 = 0.9$ and $\lambda_2 = 0.1$, which achieves the best sum-rate, for all users in the follow-up simulations.

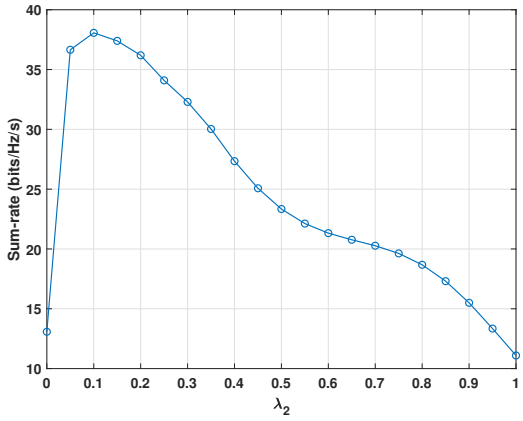

Figure 2.2: Sum-rate evaluation for different combinations of sum weights

To further illustrate the relationship between the two objectives in Problem (2.14), we plot Fig. 2.4 to show the tradeoff between beamforming gain and the norm of the projection of $\boldsymbol{w}_i$ onto the null space. As the beamforming gain increases, the norm of the projection decreases. As such, the use of multiple-objective optimization is justified.

In Fig. 2.5, the proposed analog non-robust beamforming method is compared with the analog beam selection method in [124], the digital ZF beamforming [113, 60] and the digital SLNR beamforming [114, 102, 39]. The analog beam selection method usually needs a hierarchical search to find the best beam, which results in a high training overhead. In our simulation, for simplicity, we directly use the array response vector of the strongest path of each user, which is the best beam for each user, as the solution for the analog beam selection. This is equivalent to a codebook with infinite vectors. As such, the performance of the beam selection that we report is better than what presented in [124, 116, 57]. Fig. 2.5 shows the empirical cumulative distribution function (CDF) of SINR when the $SNR = 25dB$ for different methods. For a fixed SINR value a, the corresponding CDF value p implies that $Pr(SINR \leq a) = p$. At $SINR = 0dB$, the corresponding CDFs of the digital SLNR beamformer, the digital ZF beamformer, our proposed analog ZF beamformer, and the beam selection method are around 0%, 2%, 2%, and 10%, respectively. At $SINR = 15dB$, the CDF numbers are around 0%, 10%, 30%, and 75%, respectively. Generally speaking, the fully

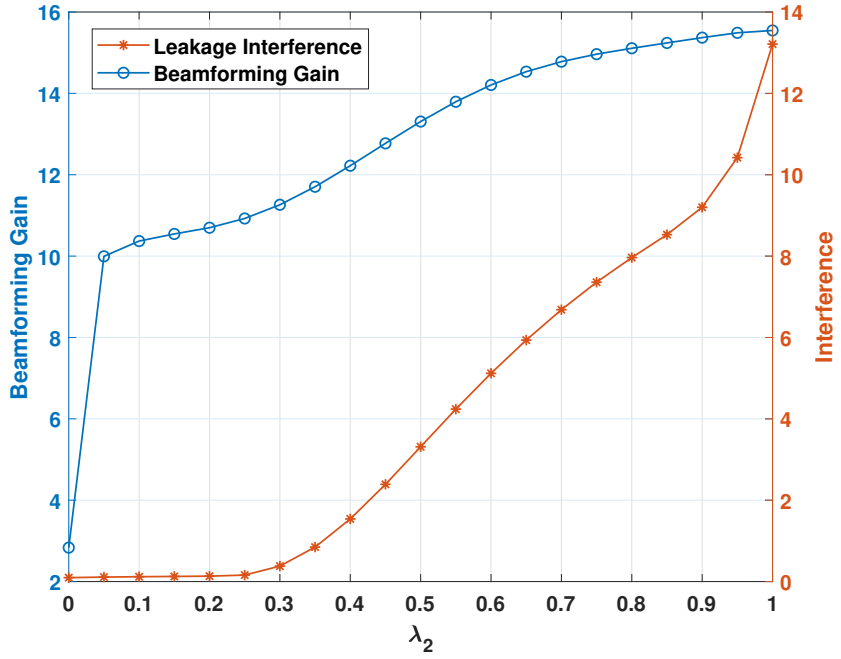

Figure 2.3: Beamforming Gain and Interference for different combinations of sum weights

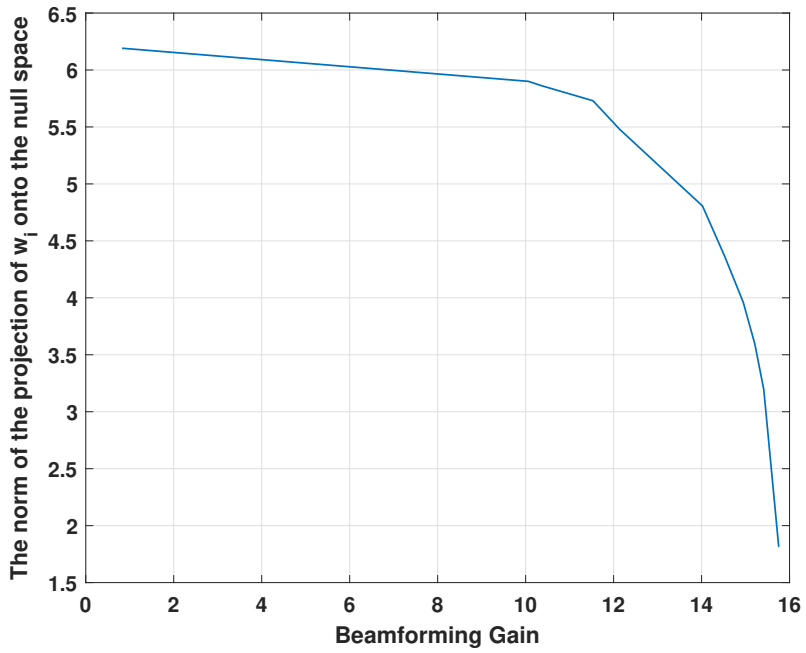

Figure 2.4: The tradeoff between the beamformining gain and the norm of the projection of $\boldsymbol{w}_i$ onto the null space

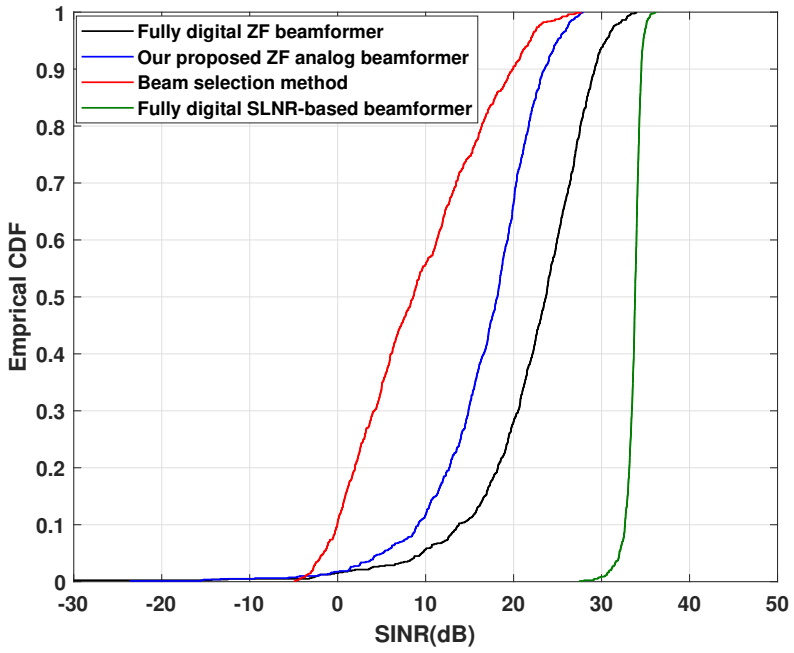

Figure 2.5: CDF of SINR when SNR=25dB

digital beamformers have CDF values that are lower than those of the analog beamformers at every SINR while requiring a higher cost of RF chains. Note that, the fully digital beamforming methods such as the zero-forcing method and the SLNR-based method are only used as the benchmark. The analog beamforming cannot outperform the fully digital beamforming. Besides, we only use the channel phase information, which will lead to performance degradation. The main method we are comparing with is the beam selection method, which is the main analog beamforming method adopted in mmWave systems and our proposed method has a much lower CDF value at most SINRs compared to the beam selection method. This is because we suppress the interference in our model while the beam selection method only maximizes the beamforming gain.

To have an overall observation of the performance of the three methods, we plot the average sum-rate per user with SNR ranging from -15 dB to 30 dB in Fig. 2.6. The fully digital SLNR-based beamforming method has the best performance at every SNR. When the SNR is low, the beam selection method performs better compared to both the digital ZF beamformer and the proposed analog beamformer. For example, when SNR is 0 dB, the sum-rate of the beam selection is around

10, while the sum-rates of our proposed analog beamformer and the digital ZF beamformer are around 6 and 5, respectively. Although our proposed analog beamformer cannot beat the beam selection method at low SNR, it performs better than the digital ZF beamformer. The reason why our proposed analog beamformer cannot beat the beam selection method when SNR is low is because we only use partial channel information (AoD matrix), not the entire channel in (2.4), to maximize the beamforming gain. When the SNR is low, the power of interference can be ignored because the noise power is large, therefore, the method with the largest beamforming gain will have the largest SINR thus the best sum-rate. Note that, in our simulations, we directly use the array response vector of the strongest path as the beam selected for User i, which is the best performance for the beam selection, without considering any beam alignment loss. Therefore, as mentioned before, the beam selection method in our simulations would perform better than what presented in [124, 116, 57]. However, when the SNR is large, for example 25 dB, the performance of the beam selection is the worst among the four methods due to the severe interference. When the SNR is 25 dB, our proposed analog beamformer can achieve a sum-rate of about 35 while the sum-rate of the beam selection only reaches 21.

Although the beam selection method has a better performance in the low SNR region, it needs a multi-stage training process to obtain the precise beam, which will result in a waste of transmission resources. Our method, on the other hand, only needs one-stage feedback of the AoD, thus saving transmission resources. Moreover, to increase the precision of the beam direction for the beam selection method, one should increase the training overhead. However, we have proposed a robust beamforming method to confront the estimation/quantization error in the AoD. Our robust beamforming method can further reduce the feedback overhead since we do not need to know a completely precise AoD as confirmed by the simulations in the next section.

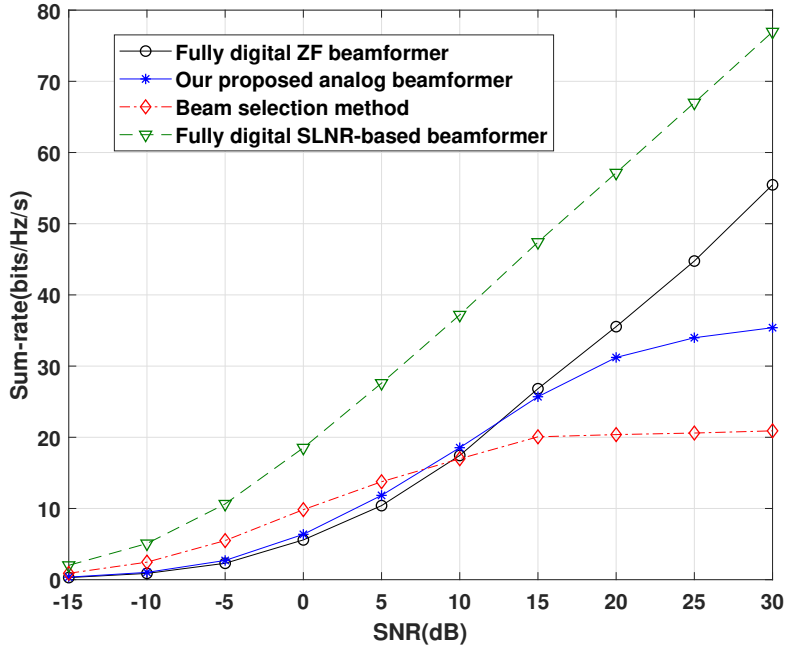

Figure 2.6: Sum-rate comparison

2.5.2 Robust analog beamforming

For the robust case, we evaluate the sum-rate performance with different levels of uncertainty, i.e. the error variance $\sigma^2_{error} = \{0.005, 0\}$. The leakage power level is set to be $\gamma_i = 0.1, \forall i = 1, ..., K$. We compare the performance of the proposed robust analog beamfomer, the non-robust digital ZF beamformer, the beam selection method, and the non-robust digital SLNR-based beamformer. In the previous section, we assumed an infinite codebook for the beam selection method, which avoids the quantization error. In this section, for the beam selection method, we assume there exists an error in the beam alignment angle and this error has the same statistical characteristic as the error in AoDs.

When the error variance $\sigma^2_{error} = 0$, we will have the same results as in the non-robust case. In fact, the ideas behind these two beamformers are the same. For the robust beamformer, when the error variance is 0, we actually minimize $\text{Tr}((\tilde{\mathbf{I}}_i^p)^H \tilde{\mathbf{I}}_i^p \mathbf{W})$, i.e., the leakage power of User i. This is the same as maximizing the projection of $\boldsymbol{w}_i$ onto the null space of $\tilde{\mathbf{I}}_i^p$, which is what we do for the

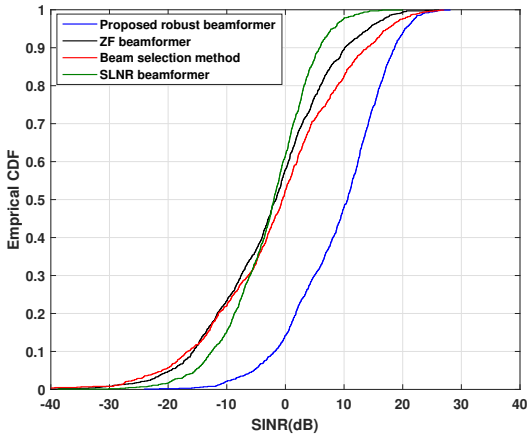

Figure 2.7: CDF of SINR when SNR=25dB. The variance of error is 0.005

non-robust beamformer.

Fig. 2.7 shows the SINR CDF of the three beamforming approaches when $\sigma^2_{error} = 0.005$. We set the $SNR = 25dB$. The proposed analog beamformer has a large improvement in performance compared with the beam selection method and the non-robust ZF beamformer. For example, when $SINR = 0dB$, the CDF values of the proposed robust analog beamformer, the beam selection method, the non-robust digital ZF beamformer and the non-robust SLNR-based beamformer are around 15%, 52%, 58% and 61%, respectively. When CDF is at 50%, the proposed robust analog beamformer can provide 11 dB ,12 dB and 13 dB improvement in SINR compared with the beam selection method, the non-robust ZF beamformer and the non-robust SLNR-based beamformer, respectively.

Fig. 2.8 plots the averaged sum-rate per user of the three beamforming methods when the SNR ranges from -15dB to 30dB with $\sigma^2_{error} = 0.005$. The proposed robust analog beamformer outperforms both the beam selection method and the non-robust ZF beamformer at every SNR. When SNR is 25dB, the proposed beamformer provides an improvement of 109%, 188% and 254% of the averaged sum-rate with respect to that of the beam selection method, the non-robust ZF beamformer and the non-robust SLNR-based beamformer, respectively.

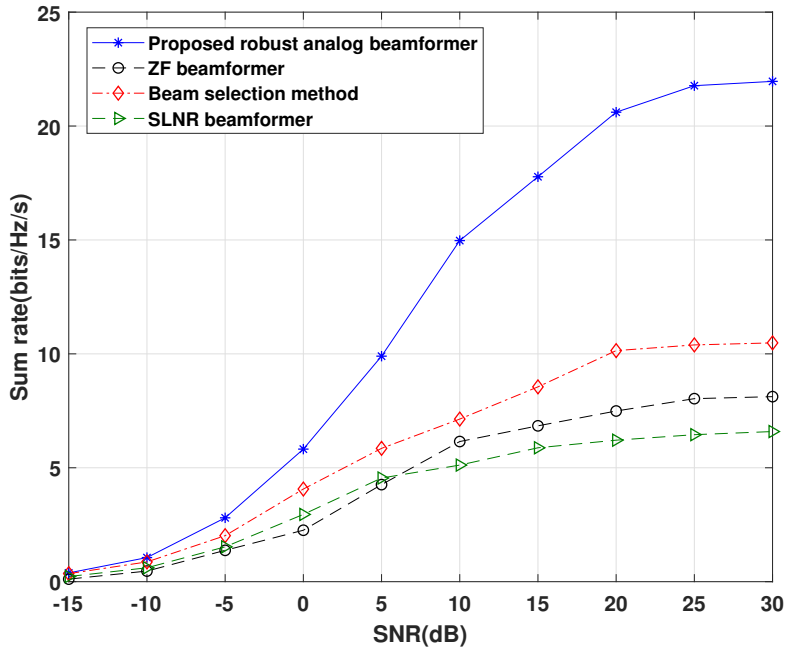

Figure 2.8: Averaged sum-rate per user. The variance of error is 0.005

Both Figs. 2.7 and 2.8 show that a slight estimation error will lead to severe system performance degradation. However, and our robust beamforming scheme can provide large improvement for imperfect CSI scenarios.

2.6 Conclusion

In this chapter, we proposed an analog beamforming scheme which strikes a balance between the beamforming gain and the inter-user interference. We formulated an MOP that maximizes the beamforming gain and minimizes the interference at the same time. The weighted-sum method was used to transform the MOP into an SOP and the SDP was adopted to make the constant-magnitude constraints for the analog beamforming tractable. Furthermore, to alleviate the effects of the channel estimation and feedback quantization errors, we designed a robust beamforming scheme to overcome the channel uncertainty. A probabilistic constraint was used and an MOP similar with the non-robust beamforming scheme was formulated. For the non-robust case, simulation results

showed that the proposed beamformer provides a better balance between the beamforming gain and the inter-user interference compared with other analog beamformers in the high SNR region. For the robust case, the simulation results demonstrated the highest robustness of our beamforming scheme against channel errors.

Chapter 3

MmWave Amplify-and-Forward MIMO Relay Networks with Hybrid Precoding/Combining Design

3.1 Introduction

Communications over millimeter wave (mmWave) have received significant attention recently because of the high data rates provided by the large bandwidth at the mmWave carrier frequencies. Also, using large antenna arrays in mmWave communication systems is possible because the small wavelength allows integrating many antennas in a small area. Despite its advantages, the mmWave carrier frequencies suffer from relatively severe propagation losses. Meanwhile, the sparsity of the mmWave scattering environment usually results in rank-deficient channels [99].

To overcome the large path losses, large antenna arrays can be placed at both transmitters and receivers to guarantee sufficient received signal power [93]. The large antenna arrays lead to a large number of radio frequency (RF) chains, which greatly increase the implementation cost

and complexity. To reduce the number of RF chains, hybrid analog/digital precoding has been proposed, which connects analog phase shifters with a reduced number of RF chains. The main advantage of the hybrid precoding is that it can trade off between the low-complexity limited-performance analog phase shifters and the high-complexity good-performance digital precoding [47].

Despite the help of large antenna arrays, the severe propagation losses still limit mmWave communications to take place within short ranges. Fortunately, the coverage can be greatly extended with the help of relay nodes [72]. Therefore, investigating the performance of hybrid precoding/combining in the relay scenario is important. For the conventional relay scenario, network beamforming in amplify-and-forward (AF) relay networks was studied in [66, 67].

For a mmWave relay scenario, large antenna arrays are usually implemented to mitigate the severe path loss. In addition, a hybrid precoding method is adopted. There are two typical hybrid precoding structures: (i) fully-connected structure (where each RF chain is connected to all antennas) [33], and (ii) sub-connected structure (where each RF chain is connected to a subset of antennas) [38]. For fully-connected mmWave networks with AF relay nodes, the authors in [74] designed hybrid precoding matrices using the orthogonal matching pursuit (OMP) algorithm. However, the performance of the OMP algorithm used in [74] depends on the orthogonality of the predetermined candidates for the analog precoders. In [138], a joint source and relay precoding design for mmWave AF relay network is proposed based on semidefinite programming (SDP). However, the proposed method in [138] is only suitable for one data stream scenario. In [108], to reduce the complexity, the RF and the baseband (BB) are separately designed and a minimum mean squared error (MMSE)-based design for the BB filters is proposed. Although the algorithm in [108] shows its advantage over the OMP algorithm in terms of sum spectral efficiency, it did not optimize the sum rate of the system. In fact, [108] can be seen as a special case of our proposed methods since we minimize the weighted mean squared error. In [118], an efficient algorithm is proposed via employing the so-called Alternating Direction Method of Multipliers (ADMM), which greatly re-

duces the distance between the hybrid precoder/combiner and the full-digital precoder/combiner. However, the ADMM algorithm has a high complexity and is sub-optimal in terms of the data rate for the system. For sub-connected structures, [137] proposes a MMSE-based relay hybrid precoding design. To make the problem tractable, [137] reformulates the original problem as three subproblems and proposes an iterative successive approximation (ISA) algorithm. The algorithm in [137] can also be extended to the fully-connected structure. Compared with the OMP algorithm, the ISA algorithm in [137] greatly improves the performance, however, the complexity of the ISA algorithm is high and it only optimizes the relay node.

In this chapter, we study the hybrid precoding for fully-connected mmWave AF relay networks in the domain of massive multiple-input and multiple-output (MIMO) systems. To reduce the complexity, we separate the RF and the BB. For the RF, we first design the phase-only RF precoding/combining matrices for multi-stream transmissions. We decompose the channel into parallel sub-channels through singular value decomposition (SVD) and compensate the phase of each sub-channel, i.e., each eigenmode of the channel. When the RF precoding and combining are performed, the digital baseband precoders/combiners are performed on the equivalent baseband channel to achieve the maximal mutual information. The problem of finding the optimal baseband precoders/combiners for the optimal mutual information is non-convex and intractable to solve by low-complexity methods. Based on the data processing inequality for the mutual information [43], we first jointly design the baseband source and relay nodes to maximize the mutual information before the destination baseband receiver. We propose a low-complexity iterative algorithm to design the precoder and combiner for the source and relay nodes, which is based on the equivalence between the mutual information maximization and the weighted MMSE [20]. After we obtain the optimal precoder and combiner for the source and relay nodes, we implement the MMSE successive interference cancellation (MMSE-SIC) filter [117] at the baseband receiver to keep the mutual information unchanged, thus obtaining the optimal mutual information for the whole relay system. Simulation results show that our algorithm outperforms the OMP in [74]. Moreover, our algorithm achieves better performance with lower complexity compared to the ISA algorithm in [137].

We also propose a robust hybrid precoding/combining approach considering the inevitable imperfect channel state information (CSI) in the second part of the chapter. Robust design for traditional relay systems has been well studied in papers, such as [134, 135, 101]. In [70, 82], the topic of imperfect channel state information in amplify-and-forward relay networks has been studied under amplify-and-forward relay networks with limited feedback. However, there is not much work on the effects of imperfect channel state information in mmWave relay networks. In [83], a robust OMP-based algorithm is proposed to maximize the receiving signal-to-noise ratio (SNR) at the destination node. Similar with the non-robust case, the performance of the OMP-based algorithm depends on the orthogonality of the predetermined candidates for the analog precoders. In this chapter, we adopt the well-known Kronecker model [143, 134] for the CSI mismatch. We first estimate the phase for RF precoding/combining to minimize the average estimation error. Then, we modify our proposed weighted MMSE approach for the perfect CSI to achieve a more robust performance for the baseband processing. Simulation results demonstrate the robustness of the proposed algorithm against CSI mismatch.

The contributions of this chapter can be summarized as follows:

- We propose a hybrid precoding/combining approach for perfect CSI in mmWave relay systems. The phase-only RF precoding/combining matrices are first designed to achieve large array gains and support multi-stream transmissions. Then, we design the baseband processing system to achieve maximal mutual information by transforming the highly complicated non-convex mutual information maximization problem into an easily tractable weighted MMSE problem. An iterative algorithm which decouples the joint design into four subproblems is developed.

- A robust design for the imperfect CSI is further proposed by modifying the non-robust precoding/combining design. To the best of our knowledge, except [83], this is the first mmWave relay system design that is robust to channel estimation error. Numerical results are provided to show the robustness of the proposed algorithm against CSI mismatch.

Compared with our conference version [65] which designs the hybrid filters for perfect CSI, we analyze the impact of imperfect CSI in this chapter, and further propose a robust design to strengthen the robustness of our proposed algorithm. The remaining sections are organized as follows. In Section 3.2 , we describe the system model and the mmWave channel model. Section 3.3 formulates the proposed hybrid precoding/combing approach for the perfect CSI. Section 3.4 presents the proposed robust hybrid design for the imperfect CSI. Numerical examples are presented and discussed in Section 3.5. We provide concluding remarks in Section 3.6.

3.2 System Model

In this section, we present the signal and channel model for a single user mmWave MIMO relay system with large antenna arrays and limited RF chains.

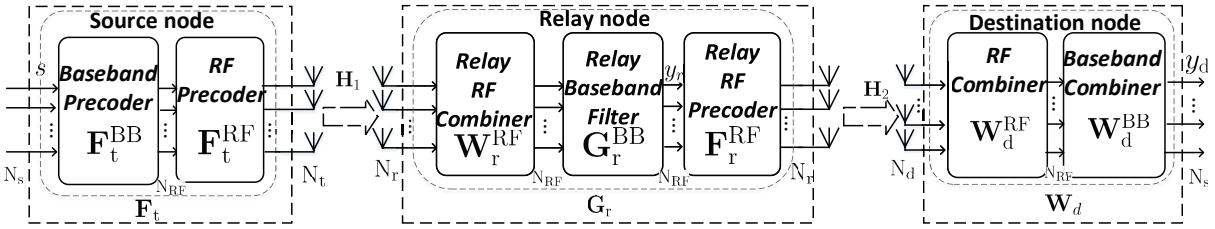

Figure 3.1: System model

3.2.1 System model

Consider a single-user mmWave MIMO relay system using hybrid precoding as illustrated in Fig. 1. The system consists of a source node with N_t transmission antennas, a relay node with N_r antennas for both transmitting and receiving signals, and a destination node with N_d antennas. Assuming N_s data streams are transmitted, the BS is equipped with N_{RF} RF chains such that $N_s \leq N_{RF} \leq N_t$. Using the N_{RF} transmit chains, an $N_{RF} \times N_s$ baseband precoder $\mathbf{F}_t^{BB}$ is applied. The RF precoder is an $N_t \times N_{RF}$ matrix $\mathbf{F}_t^{RF}$. Half duplex relaying is adopted. During the first time slot, the

BS transmits the N_s data streams to the relay through a MIMO channel $\mathbf{H}_1 \in \mathbb{C}^{N_r \times N_t}$. The relay receives the signal with an RF combiner $\mathbf{W}_r^{RF} \in \mathbb{C}^{N_{RF} \times N_r}$ and a baseband filter $\mathbf{G}_r^{BB} \in \mathbb{C}^{N_{RF} \times N_{RF}}$. During the second time slot, the relay transmits the data using one RF percoder $\mathbf{F}_r^{RF} \in \mathbb{C}^{N_r \times N_{RF}}$ through a MIMO channel $\mathbf{H}_2 \in \mathbb{C}^{N_d \times N_r}$ and the destination receives the data with one RF combiner $\mathbf{W}_d^{RF} \in \mathbb{C}^{N_{RF} \times N_d}$ and one baseband combiner $\mathbf{W}_d^{BB} \in \mathbb{C}^{N_s \times N_{RF}}$.

We assume the transmited signal is $\boldsymbol{s} = [s^1, s^2, ..., s^{N_s}]^T$ with $\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^H] = \mathbf{I}_{N_s} \in \mathbb{C}^{N_s \times N_s}$. During the first time slot, the received signal after the baseband filter at the relay can be expressed as

$$\boldsymbol{y}_r = \mathbf{G}_r^{BB}\mathbf{W}_r^{RF}\mathbf{H}_1\mathbf{F}_t^{RF}\mathbf{F}_t^{BB}\boldsymbol{s} + \mathbf{G}_r^{BB}\mathbf{W}_r^{RF}\boldsymbol{n}_1, \tag{3.1}$$

where $\boldsymbol{n}_1 \in \mathbb{C}^{N_r \times 1}$ is a zero-mean complex Gaussian noise vector at the relay node with covariance matrix $\mathbb{E}[\boldsymbol{n}_1\boldsymbol{n}_1^H] = \sigma_1^2\mathbf{I}_{N_r} \in \mathbb{C}^{N_r \times N_r}$. The power constraint at the source node is

$$\left\|\mathbf{F}_t^{RF}\mathbf{F}_t^{BB}\right\|_F^2 \leq E_t. \tag{3.2}$$

During the second time slot, the received signal after the combiners at the destination can be expressed as

$$\begin{aligned}\boldsymbol{y}_d = {} & \mathbf{W}_d^{BB}\mathbf{W}_d^{RF}\mathbf{H}_2\mathbf{F}_r^{RF}\mathbf{G}_r^{BB}\mathbf{W}_r^{RF}\mathbf{H}_1\mathbf{F}_t^{RF}\mathbf{F}_t^{BB}\boldsymbol{s} \\ & + \mathbf{W}_d^{BB}\mathbf{W}_d^{RF}\mathbf{H}_2\mathbf{F}_r^{RF}\mathbf{G}_r^{BB}\mathbf{W}_r^{RF}\boldsymbol{n}_1 + \mathbf{W}_d^{BB}\mathbf{W}_d^{RF}\boldsymbol{n}_2,\end{aligned} \tag{3.3}$$

where $\boldsymbol{n}_2 \in \mathbb{C}^{N_d \times 1}$ is a zero-mean complex Gaussian noise vector at the destination node with covariance matrix $\mathbb{E}[\boldsymbol{n}_2\boldsymbol{n}_2^H] = \sigma_2^2\mathbf{I}_{N_d} \in \mathbb{C}^{N_d \times N_d}$.

To simplify the expression, we define $\mathbf{F}_t = \mathbf{F}_t^{RF}\mathbf{F}_t^{BB} \in \mathbb{C}^{N_t \times N_s}$ as the hybrid precoding matrix at the transmitter, $\mathbf{G}_r = \mathbf{W}_r^{RF}\mathbf{G}_r^{BB}\mathbf{F}_r^{BB} \in \mathbb{C}^{N_r \times N_r}$ as the hybrid filter at the relay node, and $\mathbf{W}_d = \mathbf{W}_d^{BB}\mathbf{W}_d^{RF} \in \mathbb{C}^{N_s \times N_d}$ as the hybrid combiner at the destination node. Eq. (3.3) can be expressed as

$$\boldsymbol{y}_d = \mathbf{W}_d\mathbf{H}_2\mathbf{G}_r\mathbf{H}_1\mathbf{F}_t\boldsymbol{s} + \mathbf{W}_d\mathbf{H}_2\mathbf{G}_r\boldsymbol{n}_1 + \mathbf{W}_d\boldsymbol{n}_2. \tag{3.4}$$

The relay's power constraint is

$$\mathbb{E}[\|\mathbf{G}_{\mathrm{r}}\mathbf{H}_1\mathbf{F}_{\mathrm{t}}\boldsymbol{s}+\mathbf{G}_{\mathrm{r}}\boldsymbol{n}_1\|_F^2] \leq \mathrm{E}_{\mathrm{r}}. \tag{3.5}$$

Based on this hybrid precoding/combining system model, we can derive the achieved data rate for the system as

$$R = \frac{1}{2}\log_2 \det(\mathbf{I}_{\mathrm{N_s}} + \mathbf{W}_{\mathrm{d}}\mathbf{H}_2\mathbf{G}_{\mathrm{r}}\mathbf{H}_1\mathbf{F}_{\mathrm{t}}\mathbf{R}_{\mathrm{n}}^{-1}\mathbf{F}_{\mathrm{t}}^H\mathbf{H}_1^H\mathbf{G}_{\mathrm{r}}^H\mathbf{H}_2^H\mathbf{W}_{\mathrm{d}}^H), \tag{3.6}$$

where $\mathbf{R}_{\mathrm{n}} = \sigma_1^2\mathbf{W}_{\mathrm{d}}\mathbf{H}_2\mathbf{G}_{\mathrm{r}}\mathbf{G}_{\mathrm{r}}^H\mathbf{H}_2^H\mathbf{W}_{\mathrm{d}}^H + \sigma_2^2\mathbf{W}_{\mathrm{d}}\mathbf{W}_{\mathrm{d}}^H$ is the covariance matrix of the colored Gaussian noise at the output of the baseband combiner.

Generally, we want to jointly optimize the RF and baseband precoders/combiners to achieve the optimal data rate. However, finding the global optima for this problem (maxmizing R while imposing constant-amplitude on the RF analog precoder/combiners) is non-convex and intractable. Separated RF and baseband processing designs, as [89] did, are investigated to obtain satisfying performance. Therefore, we will separate the RF and baseband domain designs in this chapter.

3.2.2 Channel model

MmWave channels are expected to have limited scattering characteristic [98, 111, 136], which means the assumption of a rich scattering environment becomes invalid. This is called sparsity in the literature and leads to the unreliability of traditional channel models, such as the Rayleigh fading channel model. To characterize the limited scattering feature, we adopt the clustered mmWave channel model in [99, 111, 136, 105] with L scatters. Each scatter is assumed to contribute $\mathrm{N_{cl}}$

propagation paths to the channel matrix H. Then, the channel is given by

$$\mathbf{H} = \sqrt{\frac{\mathrm{N_t N_r}}{\mathrm{L N_{cl}}}} \sum_{l=1}^{\mathrm{L}} \sum_{n=1}^{\mathrm{N_{cl}}} \alpha_{l,n} \boldsymbol{a}_{\mathrm{r}}(\varphi_{l,n}^{\mathrm{r}}, \theta_{l,n}^{\mathrm{r}}) \boldsymbol{a}_{\mathrm{t}}^{H}(\varphi_{l,n}^{\mathrm{t}}, \theta_{l,n}^{\mathrm{t}}), \tag{3.7}$$

where $\alpha_{l,n}$ is the complex gain of the n^{th} path in the l^{th} scatter with the distribution $\mathbb{CN}(0,1)$, $\varphi_{l,n}^{\mathrm{r}}(\theta_{l,n}^{\mathrm{r}})$ and $\varphi_{l,n}^{\mathrm{t}}(\theta_{l,n}^{\mathrm{t}})$ are the random azimuth and elevation angles of arrival and departure. $\boldsymbol{a}_{\mathrm{r}}(\varphi_{l,n}^{\mathrm{r}}, \theta_{l,n}^{\mathrm{r}})$ and $\boldsymbol{a}_{\mathrm{t}}(\varphi_{l,n}^{\mathrm{t}}, \theta_{l,n}^{\mathrm{t}})$ are the receiving and transmitting antenna array response vectors, respectively. While the algorithms and results in the chapter can be applied to arbitrary antenna arrays, we use uniform linear arrays (ULAs) in the simulations for simplicity. The array response vectors take the following form [13]:

$$\boldsymbol{a}^{\mathrm{ULA}}(\varphi) = \frac{1}{\sqrt{\mathrm{N}}}[1, e^{jkdsin(\varphi)}, ..., e^{j(\mathrm{N}-1)kdsin(\varphi)}]^{T}, \tag{3.8}$$

where $k = \frac{2\pi}{\lambda}$. Parameter λ represents the wavelength of the carrier and d is the spacing between antenna elements. The angle φ is assumed to have a uniform distribution over $[0, 2\pi]$.

Since the channel in mmWave systems has limited scattering, we can further simplify the channel by assuming each scatter only contributes one path to the channel matrix. Then, the channel can be further expressed as

$$\mathbf{H} = \sqrt{\frac{\mathrm{N_t N_r}}{\mathrm{L}}} \sum_{l=1}^{\mathrm{L}} \alpha_l \boldsymbol{a}_{\mathrm{r}}(\varphi_l^{\mathrm{r}}, \theta_l^{\mathrm{r}}) \boldsymbol{a}_{\mathrm{t}}^{H}(\varphi_l^{\mathrm{t}}, \theta_l^{\mathrm{t}}). \tag{3.9}$$

The matrix formulation can be expressed as

$$\mathbf{H} = \sqrt{\frac{\mathrm{N_t N_r}}{\mathrm{L}}} \mathbf{A}_{\mathrm{r}} diag(\boldsymbol{\alpha}) \mathbf{A}_{\mathrm{t}}^{H}, \tag{3.10}$$

where $\mathbf{A}_{\mathrm{r}} = [\boldsymbol{a}_{\mathrm{r}}(\varphi_1^{\mathrm{r}}, \theta_1^{\mathrm{r}}), ..., \boldsymbol{a}_{\mathrm{r}}(\varphi_{\mathrm{L}}^{\mathrm{r}}, \theta_{\mathrm{L}}^{\mathrm{r}})]$ and $\mathbf{A}_{\mathrm{t}} = [\boldsymbol{a}_{\mathrm{t}}(\varphi_1^{\mathrm{t}}, \theta_1^{\mathrm{t}}), ..., \boldsymbol{a}_{\mathrm{t}}(\varphi_{\mathrm{L}}^{\mathrm{t}}, \theta_{\mathrm{L}}^{\mathrm{t}})]$ are matrices containing the receiving and transmitting array response vectors, and $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_{\mathrm{L}}]^{T}$.

3.3 Hybrid Precoder/Combiner Design

As discussed in Section 3.2, we use a hybrid design to reduce the number of RF chains. We first design the RF precoder/combiner. Then, based on the designed RF precoder/combiner, we design a low-complexity iterative algorithm for the baseband precoder/combiner to maximize the mutual information.

3.3.1 RF precoder/combiner design

Our goal for RF precoder/combiner is to make the channels decomposed into N_{RF} parallel sub-channels to support the multi-stream transmission. The main challenge is the constant-magnitude constraints on RF precoders and combiners. Without the constant-magnitude constraints, the optimal precoder/combiner should be the right/left singular matrix of the channel, which transmits the signals along the eigenmodes of the channel. Considering the constant-magnitude constraints, we cannot directly use the singular matrix to rotate the signals, but we can use the projection on each eigenmode as a criterion to choose RF precoder and combiner. For the i^{th} eigenmode, the best precoder should be the one that has the largest projection on that eigenmode, i.e., the one that casts the most energy along that eigenmode direction.

Using $\mathbf{H}_1$ as an example, we first perform the singular value decomposition (SVD) for the channel matrix.

$$\mathbf{H}_1 = \mathbf{U}_1 \Sigma_1 \mathbf{V}_1^H = \sum_{i=1}^{\mathrm{L}} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^H, \tag{3.11}$$

where $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the i^{th} vectors in matrices $\mathbf{U}_1$ and $\mathbf{V}_1$, respectively, which correspond to σ_i. The singular values σ_is are assumed to be in a descending order. L is the rank of the channel and is equal to the number of propagation paths for the mmWave scenario. Note that for mmWave systems, the channels usually have limited scattering characteristics, which means the number

of propagation paths is far less than $\min(\mathrm{N_t}, \mathrm{N_r})$. In such cases, the channel rank is equal to the number of propagation paths L. Eq. (3.11) indicates that channel $\mathbf{H}_1$ has L eigenmodes. We denote the i^{th} eigenmode by $\boldsymbol{u}_i \boldsymbol{v}_i^H$, and its gain by σ_i.

For our RF precoding/combining, we want to maximize the projection of the i^{th} data stream onto the i^{th} eigenmode, i.e., $\left|\boldsymbol{w}_i^H \boldsymbol{u}_i \boldsymbol{v}_i^H \boldsymbol{f}_i\right|$, where $\boldsymbol{f}_i$ and $\boldsymbol{w}_i$ are the i^{th} vector of precoder $\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}$ and combiner $\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}$, respectively. To approach the maximal projection, we have the following proposition.

Proposition 1. *The optimal phase-only vectors $\boldsymbol{f}_i$ and $\boldsymbol{w}_i$, which maximize the projection for the i^{th} data stream onto the i^{th} eigenmode of the channel, will satisfy the following conditions:*

$$phase(\boldsymbol{f}_i[m]) = phase(\boldsymbol{v}_i[m]) \ \forall m = 1, 2, ..., N_t, \tag{3.12}$$

$$phase(\boldsymbol{w}_i[n]) = phase(\boldsymbol{u}_i[n]) \ \forall n = 1, 2, ..., N_r, \tag{3.13}$$

where $\cdot[k]$ represents the k^{th} element of a vector.

Proof. First, we express the vectors in polar coordinates. Due to the magnitude-constant constraints, vectors $\boldsymbol{f}_i$ and $\boldsymbol{w}_i$ are expressed as $\boldsymbol{f}_i = \frac{1}{\sqrt{\mathrm{N_t}}}[e^{j\theta_1^i}, e^{j\theta_2^i}, ..., e^{j\theta_{\mathrm{N_t}}^i}]^T$ and $\boldsymbol{w}_i = \frac{1}{\sqrt{\mathrm{N_r}}}[e^{j\varphi_1^i}, e^{j\varphi_2^i}, ..., e^{j\varphi_{\mathrm{N_r}}^i}]^T$. Since there are no constant-magnitude constraints for $\boldsymbol{v}_i$ and $\boldsymbol{u}_i$, each element in the vector has its own magnitude. The polar forms of $\boldsymbol{v}_i$ and $\boldsymbol{u}_i$ are $\boldsymbol{v}_i = [r_1^i e^{j\alpha_1^i}, r_2^i e^{j\alpha_2^i}, ..., r_{\mathrm{N_t}}^i e^{j\alpha_{\mathrm{N_t}}^i}]^T$ and $\boldsymbol{u}_i = [\rho_1^i e^{j\beta_1^i}, \rho_2^i e^{j\beta_2^i}, ..., \rho_{\mathrm{N_r}}^i e^{j\beta_{\mathrm{N_r}}^i}]^T$, respectively. Then, the projection can be calculated as

$$\left|\boldsymbol{w}_i^H \boldsymbol{u}_i \boldsymbol{v}_i^H \boldsymbol{f}_i\right| = \left|\frac{1}{\sqrt{\mathrm{N_r}}} \sum_{n=1}^{\mathrm{N_r}} \rho_n^i e^{j(\varphi_n^i - \beta_n^i)}\right| \left|\frac{1}{\sqrt{\mathrm{N_t}}} \sum_{m=1}^{\mathrm{N_t}} r_m^i e^{j(\alpha_m^i - \theta_m^i)}\right|. \tag{3.14}$$

According to the Cauchy-Schwartz inequality, we have

$$\begin{aligned}&\left|\frac{1}{\sqrt{\mathrm{N_t}}}\sum_{m=1}^{\mathrm{N_t}} r_m^i e^{j(\alpha_m^i-\theta_m^i)}\right|^2 \\ &\leq \frac{1}{\mathrm{N_t}}\sum_{m=1}^{\mathrm{N_t}}\left|r_m^i\right|^2\sum_{m=1}^{\mathrm{N_t}}\left|e^{j(\alpha_m^i-\theta_m^i)}\right|^2=\frac{1}{\mathrm{N_t}}\sum_{m=1}^{\mathrm{N_t}}\left|r_m^i\right|^2.\end{aligned} \tag{3.15}$$

Equality can be achieved in (3.15) if and only if $\theta_m^i=\alpha_m^i\ \forall m=1,2,...,\mathrm{N_t}$. This means the maximal $\left|\boldsymbol{v}_i^H\boldsymbol{f}_i\right|$ is achieved when $\theta_m^i=\alpha_m^i\ \forall m=1,2,...,\mathrm{N_t}$. Similarly, the maximal $\left|\boldsymbol{w}_i^H\boldsymbol{u}_i\right|$ is achieved when $\varphi_n^i=\beta_n^i\ \forall n=1,2,...,\mathrm{N_r}$. Therefore, we have the conclusion that the optimal phase-only vectors $\boldsymbol{f}_i$ and $\boldsymbol{w}_i$, which maximize $|\boldsymbol{w}_i^H\boldsymbol{u}_i\boldsymbol{v}_i^H\boldsymbol{f}_i|$, will satisfy the conditions in (3.12) and (3.13). □

Our RF precoders and combiners are actually compensating the phase of each sub-channel. Note that when the number of antennas is large enough, the response vectors $\boldsymbol{a}_{\mathrm{r}}(\varphi_l^{\mathrm{r}},\theta_l^{\mathrm{r}})$s and $\boldsymbol{a}_{\mathrm{t}}(\varphi_l^{\mathrm{t}},\theta_l^{\mathrm{t}})$s become orthogonal to each other. $\mathbf{A}_{\mathrm{t}}$ and $\mathbf{A}_{\mathrm{r}}$ become the left and right singular matrices of the channel and they directly become our RF precoder and combiner. In this case, we can perfectly decompose the channel into independent parallel sub-channels. The equivalent channel after RF processing is diagonal, which makes it easier for baseband processing.

3.3.2 Baseband system

In this section, we focus on designing the baseband precoding/combining matrices. First, we define the equivalent baseband channels for $\mathbf{H}_1$ and $\mathbf{H}_2$ as

$$\tilde{\mathbf{H}}_1=\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{H}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}, \tag{3.16}$$

$$\tilde{\mathbf{H}}_2=\mathbf{W}_{\mathrm{d}}^{\mathrm{RF}}\mathbf{H}_2\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}. \tag{3.17}$$

Based on the equivalent channels, we simplify our system model as shown in Fig. 3.2.

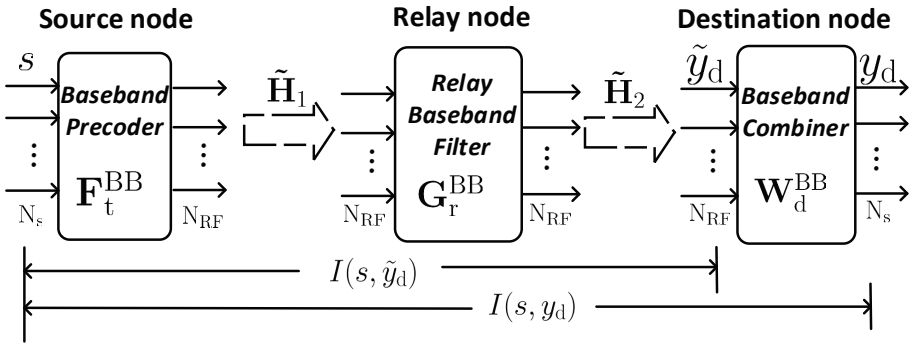

Figure 3.2: Baseband System model

Using the equivalent channels (3.16) and (3.17), we rewrite the received signals at the destination node as

$$\tilde{\boldsymbol{y}}_{\mathrm{d}} = \tilde{\mathbf{H}}_2 \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}} \boldsymbol{s} + \tilde{\mathbf{H}}_2 \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\boldsymbol{n}}_1 + \tilde{\boldsymbol{n}}_2, \tag{3.18}$$

$$\boldsymbol{y}_{\mathrm{d}} = \mathbf{W}_{\mathrm{d}}^{\mathrm{BB}} \tilde{\mathbf{H}}_2 \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}} \boldsymbol{s} + \mathbf{W}_{\mathrm{d}}^{\mathrm{BB}} \tilde{\mathbf{H}}_2 \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\boldsymbol{n}}_1 + \mathbf{W}_{\mathrm{d}}^{\mathrm{BB}} \tilde{\boldsymbol{n}}_2, \tag{3.19}$$

where $\tilde{\boldsymbol{n}}_1 = \mathbf{W}_{\mathrm{r}}^{\mathrm{RF}} \boldsymbol{n}_1$ and $\tilde{\boldsymbol{n}}_2 = \mathbf{W}_{\mathrm{d}}^{\mathrm{RF}} \boldsymbol{n}_2$.

Our ultimate goal for the baseband design is to maximize the mutual information $I(\boldsymbol{s}, \boldsymbol{y}_{\mathrm{d}})$. However, directly optimizing $I(\boldsymbol{s}, \boldsymbol{y}_{\mathrm{d}})$ is intractable. According to the data processing inequality [43], $I(\boldsymbol{s}, \boldsymbol{y}_{\mathrm{d}}) \leq I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})$. We first design $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ to maximize the mutual information $I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})$. After we get the maximum $I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})$, we implement the MMSE-SIC for $\mathbf{W}_{\mathrm{d}}^{\mathrm{BB}}$, which according to [117] is information lossless. In this way, we make $I(\boldsymbol{s}, \boldsymbol{y}_{\mathrm{d}}) = I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})$. Since $I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})$ is maximized, $I(\boldsymbol{s}, \boldsymbol{y}_{\mathrm{d}})$ is also maximized because of the data processing inequality and the independence of $I(\boldsymbol{s}; \tilde{\boldsymbol{y}}_{\mathrm{d}})$ from $\mathbf{W}_{\mathrm{d}}^{\mathrm{BB}}$.

3.3.3 $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ design

In this section, we jointly design $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ to maximize $I(\boldsymbol{s},\tilde{\boldsymbol{y}}_{\mathrm{d}})$. According to [20], there exists a relationship between $I(\boldsymbol{s},\tilde{\boldsymbol{y}}_{\mathrm{d}})$ and the MSE matrix $\mathbf{E}_{\mathrm{MMSE}}$, i.e.,

$$I(\boldsymbol{s},\tilde{\boldsymbol{y}}_{\mathrm{d}}) = \log_2 \det(\mathbf{E}_{\mathrm{MMSE}}^{-1}), \tag{3.20}$$

where the MSE matrix $\mathbf{E}_{\mathrm{MMSE}}$ is defined as the mean square error covariance matrix given the MMSE receiver. The detailed proof can be found in [20]. We give a brief derivation procedure below.

The MMSE receiver is defined as

$$\begin{aligned}\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}} &= \operatorname{argmin} \mathbb{E}[\|\mathbf{W}_{\mathrm{d}}^{\mathrm{BB}}\tilde{\boldsymbol{y}}_{\mathrm{d}} - \boldsymbol{s}\|^2] = (\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H \\ &\quad (\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}(\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H + \mathbf{R}_{\tilde{\mathrm{n}}})^{-1},\end{aligned} \tag{3.21}$$

where $\mathbf{R}_{\tilde{\mathrm{n}}} = \sigma_1^2\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}(\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}})^H + \sigma_2^2\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}(\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}})^H$.

The MMSE matrix $\mathbf{E}_{\mathrm{MMSE}}$ can be calculated by

$$\begin{aligned}\mathbf{E}_{\mathrm{MMSE}} &= \mathbb{E}[(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\boldsymbol{y}}_{\mathrm{d}} - \boldsymbol{s})(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\boldsymbol{y}}_{\mathrm{d}} - \boldsymbol{s})^H] \\ &= (\mathbf{I}_{\mathrm{N_s}} - \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}) \\ &\quad (\mathbf{I}_{\mathrm{N_s}} - \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H + \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\mathbf{R}_{\tilde{\mathrm{n}}}(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}})^H.\end{aligned} \tag{3.22}$$

Substituting (3.21) into (3.22), we can express $\mathbf{E}_{\mathrm{MMSE}}$ as

$$\mathbf{E}_{\mathrm{MMSE}} = (\mathbf{I}_{\mathrm{N_s}} + (\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H\mathbf{R}_{\tilde{\mathrm{n}}}^{-1}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^{-1}. \tag{3.23}$$

From (3.23), we can obtain (3.20).

Based on (3.20), we can establish the equivalence between the $I(\boldsymbol{s},\tilde{\boldsymbol{y}}_{\mathrm{d}})$ maximization problem and

a WMMSE problem as [20] did.

The $I(\boldsymbol{s},\tilde{\mathbf{y}}_{\mathrm{d}})$ maximization problem is formulated as

$$\begin{aligned} &\min_{\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}},\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}} -I(\boldsymbol{s},\tilde{\mathbf{y}}_{\mathrm{d}}) \\ s.t.\ &\left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2 \le \mathrm{E}_{\mathrm{t}}, \\ &\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\boldsymbol{s}+\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] \le \mathrm{E}_{\mathrm{r}}. \end{aligned} \tag{3.24}$$

The WMMSE problem is formulated as

$$\begin{aligned} &\min_{\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}},\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}},\mathbf{V}} \mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}}) \\ s.t.\ &\left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2 \le \mathrm{E}_{\mathrm{t}}, \\ &\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\boldsymbol{s}+\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] \le \mathrm{E}_{\mathrm{r}}, \end{aligned} \tag{3.25}$$

where $\mathbf{V}$ is a constant weight matrix.

We will show that Problems (3.24) and (3.25) have the same optimum solution, i.e., the points that satisfy the KKT conditions for (3.24) and (3.25) are the same. Same as [20], we set the partial derivatives of the Lagrange functions of (3.24) and (3.25) to zero. Note that the power constraints of (3.24) and (3.25) are the same. To prove the equivalence, we only need to calculate the partial derivatives of $-I(\boldsymbol{s},\tilde{\mathbf{y}}_{\mathrm{d}})$ and $\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}})$ w.r.t $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$. Note that $\partial \log\det(\mathbf{X}) = \mathrm{Tr}(\mathbf{X}^{-1}\partial\mathbf{X})$. Taking $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ as an example, for $I(\boldsymbol{s},\tilde{\mathbf{y}}_{\mathrm{d}})$, we have

$$\frac{\partial -I(\boldsymbol{s},\tilde{\mathbf{y}}_{\mathrm{d}})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}} = -\frac{\partial \log_2\det(\mathbf{E}_{\mathrm{MMSE}})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}} = -\frac{\mathrm{Tr}(\mathbf{E}_{\mathrm{MMSE}}^{-1}\partial\mathbf{E}_{\mathrm{MMSE}})}{(\log 2)\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}}, \tag{3.26}$$

Note that $\partial \mathbf{X}^{-1} = -\mathbf{X}^{-1}(\partial \mathbf{X})\mathbf{X}^{-1}$ and $\partial(\mathbf{AX}) = \partial(\mathbf{X})\mathbf{A} + \partial(\mathbf{A})\mathbf{X}$. For $\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}})$, we have

$$\begin{aligned}
&\frac{\partial \,\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}} = -\frac{\mathrm{Tr}(\partial(\mathbf{V}(\mathbf{E}_{\mathrm{MMSE}}^{-1})^{-1}))}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}} \\
&= -\frac{\mathrm{Tr}(\mathbf{E}_{\mathrm{MMSE}}\partial(\mathbf{E}_{\mathrm{MMSE}}^{-1})\mathbf{E}_{\mathrm{MMSE}}\mathbf{V} + \partial(\mathbf{V})\mathbf{E}_{\mathrm{MMSE}})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}} \\
&= -\frac{\mathrm{Tr}(\mathbf{E}_{\mathrm{MMSE}}\partial(\mathbf{E}_{\mathrm{MMSE}}^{-1})\mathbf{E}_{\mathrm{MMSE}}\mathbf{V})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}}.
\end{aligned} \tag{3.27}$$

If we set the constant weight matrix $\mathbf{V} = \frac{\mathbf{E}_{\mathrm{MMSE}}^{-1}}{\log 2}$, then we have

$$\frac{\partial - I(\boldsymbol{s}, \tilde{\mathbf{y}}_{\mathrm{d}})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}} = \frac{\partial \,\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}}. \tag{3.28}$$

Similarly, we can derive

$$\frac{\partial - I(\boldsymbol{s}, \tilde{\mathbf{y}}_{\mathrm{d}})}{\partial \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}} = \frac{\partial \,\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}})}{\partial \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}}. \tag{3.29}$$

From Eqs. (3.28) and (3.29), we can conclude that the KKT-conditions of (3.24) and (3.25) can be satisfied simultaneously, which suggests that it is possible to solve the mutual information maximization problem through the use of WMMSE by choosing an appropriate weight, i.e., $\mathbf{V}$. To maximize $I(\boldsymbol{s}, \tilde{\mathbf{y}}_d)$, we propose an iterative algorithm based on the WMMSE problem (3.25). Note that in the proposed algorithm, we also iteratively search for the appropriate weight matrix. When the algorithm converges, we will obtain the desired weight matrix as well as the optimal filters that maximize $I(\boldsymbol{s}, \tilde{\mathbf{y}}_{\mathrm{d}})$. The detailed algorithm is described as follows:

1. Calculate the MMSE receiver $\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}$ in Eq. (3.21) and the MSE matrix $\mathbf{E}_{\mathrm{MMSE}}$ in Eq. (3.22).

2. Update $\mathbf{V}$ by setting $\mathbf{V} = \frac{\mathbf{E}_{\mathrm{MMSE}}^{-1}}{\log 2}$.

3. Fixing $\mathbf{V}$ and $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$, then we find $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ that minimizes $\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}}) = \mathrm{Tr}(\mathbf{V}((\mathbf{I}_{\mathrm{N_s}} - \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})(\mathbf{I}_{\mathrm{N}}$ $\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H + \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\mathbf{R}_{\tilde{\mathrm{n}}}(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}})^H))$ under the power constraints, i.e.,

$$\begin{aligned}\hat{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} = \text{argmin}\ \ & \mathrm{Tr}(\mathbf{V}((\mathbf{I}_{\mathrm{N_s}} - \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}) \\ & (\mathbf{I}_{\mathrm{N_s}} - \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H + \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\mathbf{R}_{\tilde{\mathrm{n}}}(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}})^H)) \\ s.t.\ & \mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\boldsymbol{s} + \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] \leq \mathrm{E_r}.\end{aligned} \tag{3.30}$$

Problem (3.30) is a convex optimization for $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ and we can solve it using the KKT condition. Denoting the Lagrange function of Problem (3.30) as $L^r(\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}, \lambda^{\mathrm{r}}) = \mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}}) + \lambda^{\mathrm{r}}(\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\boldsymbol{s} + \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2 - E_{\mathrm{r}})$, the KKT conditions are

$$\frac{\partial L^{\mathrm{r}}(\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}, \lambda^{\mathrm{r}})}{\partial \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}} = 0, \tag{3.31}$$

$$\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\boldsymbol{s} + \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] - \mathrm{E_r} \leq 0, \tag{3.32}$$

$$\lambda^{\mathrm{r}}(\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\boldsymbol{s} + \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] - \mathrm{E_r}) = 0, \tag{3.33}$$

$$\lambda^{\mathrm{r}} \geq 0. \tag{3.34}$$

Solving (3.31), we have

$$\begin{aligned}\hat{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} = &((\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2)^H\mathbf{V}\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2 + \lambda^{\mathrm{r}}(\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}})^H\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}})^{-1} \\ &(\tilde{\mathbf{H}}_2)^H(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}})^H\mathbf{V}(\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H(\tilde{\mathbf{H}}_1)^H \\ &(\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}(\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H + \sigma_1^2\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}(\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}})^H)^{-1}.\end{aligned} \tag{3.35}$$

Based on (3.32) and (3.33), we can obtain the Lagrange multiplier λ^{r} as follows. First, we calculate $\hat{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$ by setting $\lambda^{\mathrm{r}} = 0$. If the power constraint is satisfied, then we set $\lambda^{\mathrm{r}} = 0$. If the power constraint is not satisfied, then, we initialize λ^{r} with a pre-defined value and substitute it into (3.32) and start a bisection search for λ^{r} until the power constraint is satisfied.

4. Fixing $\mathbf{V}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$, then we find the $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ to minimize $\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}}) = \mathrm{Tr}(\mathbf{V}((\mathbf{I}_{\mathrm{N_s}} - \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})(\mathbf{I}$ $\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H + \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\mathbf{R}_{\tilde{\mathrm{n}}}(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}})^H))$ under the power constraints, i.e.,

$$\begin{aligned}&\hat{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}} = \mathrm{argmin}\ \ \mathrm{Tr}(\mathbf{V}((\mathbf{I}_{\mathrm{N_s}} - \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})(\mathbf{I}_{\mathrm{N_s}} \\ &- \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}})^H + \mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\mathbf{R}_{\tilde{\mathrm{n}}}(\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}})^H)) \\ &s.t.\ \ \left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2 \le \mathrm{E_t}, \\ &\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}s + \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] \le \mathrm{E_r}.\end{aligned} \tag{3.36}$$

Problem (3.36) is a convex optimization for $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ and, similar with Problem (3.30), we can solve Problem (3.36) using the KKT conditions. Denoting the Lagrange function of Problem (3.36) as $L^{\mathrm{t}}(\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}, \lambda_1^{\mathrm{t}}, \lambda_2^{\mathrm{t}})$, the KKT conditions are

$$\frac{\partial L^{\mathrm{t}}(\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}, \lambda_1^{\mathrm{t}}, \lambda_2^{\mathrm{t}})}{\partial \mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}} = 0, \tag{3.37}$$

$$\left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2 - \mathrm{E_t} \le 0, \tag{3.38}$$

$$\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}s+\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2]-\mathrm{E}_{\mathrm{r}}\leq 0, \tag{3.39}$$

$$\lambda_1^{\mathrm{t}}(\left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2-\mathrm{E}_{\mathrm{t}})=0, \tag{3.40}$$

$$\lambda_2^{\mathrm{t}}(\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}s+\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2]-\mathrm{E}_{\mathrm{r}})=0, \tag{3.41}$$

$$\lambda_1^{\mathrm{t}},\lambda_2^{\mathrm{t}}\geq 0. \tag{3.42}$$

The optimal solution for $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ can be expressed as

$$\begin{aligned}\hat{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}=&((\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1)^H\mathbf{V}\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\\&+\lambda_1^{\mathrm{t}}(\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}})^H\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}+\lambda_2^{\mathrm{t}}(\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1)^H\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1)^{-1}\\&(\mathbf{V}\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}}\tilde{\mathbf{H}}_2\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1)^H,\end{aligned} \tag{3.43}$$

where λ_1^{t} and λ_2^{t} are the non-negative Lagrange multipliers corresponding to the power constraints. Similar with Problem (3.30), to obtain λ_1^{t} and λ_2^{t}, we consider four cases: i) if the power constraints are satisfied when $\lambda_1^{\mathrm{t}}=0$ and $\lambda_2^{\mathrm{t}}=0$, we will set λ_1^{t} and λ_2^{t} equal to 0; ii) if case (i) is not satisfied, we then set $\lambda_1^{\mathrm{t}}=0$ and start a bisection search for λ_2^{t} until the KKT condition (3.41) and the power constraint (3.38) are satisfied; iii) if (3.41) and (3.38) cannot be satisfied simultaneously through the bisection search for λ_2^{t}, we then set $\lambda_2^{\mathrm{t}}=0$ and start a bisection search for λ_1^{t} until (3.39) and (3.40) are satisfied; iv) if we cannot find

the appropriate λ_1^t to satisfy (3.39) and (3.40) at the same time, we can obtain λ_1^t and λ_2^t through a two-layer bisection search. The search algorithm is described in Algorithm 1.

Algorithm 1: Two-layer bisection search for λ_1^t and λ_2^t

initialize $\lambda_{1,\min}^t = \lambda_{2,\min}^t = 0$, $\lambda_{1,\max}^t$,$\lambda_{2,\max}^t$;
while $\lambda_{1,\max}^t - \lambda_{1,\min}^t > \varepsilon_1$ **do**
  setting $\lambda_1^t = \frac{\lambda_{1,\min}^t + \lambda_{1,\max}^t}{2}$;
  while $\lambda_{2,\max}^t - \lambda_{2,\min}^t > \varepsilon_2$ **do**
    setting $\lambda_2^t = \frac{\lambda_{2,\min}^t + \lambda_{2,\max}^t}{2}$;
    calculate $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ according to (3.43);
    if $\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}s + \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] \leq \mathrm{E_r}$ **then**
      $\lambda_{2,\max}^t = \lambda_2^t$;
    end if
    if $\mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}s + \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] \geq \mathrm{E_r}$ **then**
      $\lambda_{2,\min}^t = \lambda_2^t$;
    end if
  end while
  calculate $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ according to (3.43);
  if $\left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2 \leq \mathrm{E_t}$ **then**
    $\lambda_{1,\max}^t = \lambda_1^t$;
  end if
  if $\left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2 \geq \mathrm{E_t}$ **then**
    $\lambda_{1,\min}^t = \lambda_1^t$;
  end if
end while

The entire design procedures for $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ are summarized in Algorithm 2.

3.3.4 Convergence analysis

Since the constant weight matrix $\mathbf{V}$ changes at each iteration, it does not generate a monotonic decreasing sequence, which means we cannot directly prove the convergence of the proposed algorithm. Fortunately, according to [20], the iteration procedure to maximize the mutual information through minimizing WMMSE is the same optimization procedure for an equivalent optimization

Algorithm 2: Design for $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$

1: Initialize $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}(0)}$ and $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}(0)}$;
2: Set $k=1$;
3: **while** $|I(\boldsymbol{s},\tilde{\boldsymbol{y}}_{\mathrm{d}})^{(k)} - I(\boldsymbol{s},\tilde{\boldsymbol{y}}_{\mathrm{d}})^{(k-1)}| > \varepsilon$ **do**
4: Calculate the MMSE receiver $\mathbf{W}_{\mathrm{d}}^{\mathrm{MMSE}(k)}$ according to (3.21) and the MSE matrix $\mathbf{E}_{\mathrm{MMSE}}^{(k)}$ in Eq. (3.22);
5: Update $\mathbf{V}$ by setting $\mathbf{V} = \frac{(\mathbf{E}_{\mathrm{MMSE}}^{(k)})^{-1}}{\log 2}$;
6: Calculate $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}(k)}$ as Step III illustrates;
7: Calculate $\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}(k)}$ as Step IV illustrates;
8: $k = k+1$;
9: **end while**

problem as below

$$\begin{aligned} \min_{\substack{\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}},\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}},\\ \mathbf{V},\mathbf{W}_{\mathrm{d}}^{\mathrm{BB}}}} \ & \mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}}) - \log\det(\mathbf{V}) \\ s.t. \ & \left\|\mathbf{F}_{\mathrm{t}}^{\mathrm{RF}}\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\right\|_F^2 \le \mathrm{E}_{\mathrm{t}}, \\ & \mathbb{E}[\left\|\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\tilde{\mathbf{H}}_1\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}\boldsymbol{s} + \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}\mathbf{W}_{\mathrm{r}}^{\mathrm{RF}}\boldsymbol{n}_1\right\|_F^2] \le \mathrm{E}_{\mathrm{r}}. \end{aligned} \tag{3.44}$$

The proof of the equivalence is similar to the proof in [20] and we omit the detailed prove for brevity. The main idea is that the alternating minimization of the objective in (3.44) corresponds to Steps 1-4 in our proposed algorithm. For example, when $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$,$\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ and $\mathbf{V}$ are fixed, optimizing (3.44) w.r.t. $\mathbf{W}_{\mathrm{d}}^{\mathrm{BB}}$ becomes minimizing MMSE, which gives the same result as Step 1. When $\mathbf{F}_{\mathrm{t}}^{\mathrm{BB}}$,$\mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}$ and $\mathbf{W}_{\mathrm{d}}^{\mathrm{BB}}$ are fixed, the optimal solution for $\mathbf{V}$ which minimizes the objective $\mathrm{Tr}(\mathbf{V}\mathbf{E}_{\mathrm{MMSE}}) - \log\det(\mathbf{V})$ in (3.44) is the same as Step 2.

Based on this equivalence, we can prove the convergence of the proposed algorithm by proving the monotonic convergence of Problem (3.44). According to [20], the objective in Problem (3.44) has a lower bound, which is the negative of the maximum mutual information. Due to the alternating minimization process, the objective in Problem (3.44) decreases monotonically. Since a sequence of monotonically decreasing numbers with a lower bound converges, Problem (3.44) converges

and so does our proposed algorithm.

3.3.5 Complexity analysis

Since we provide closed-form solutions for each iteration, the main complexity lies in the search for the appropriate Lagrange multipliers. Let us define the search accuracy as ε. This is a relative measure for the search interval. For example, if the length of our search interval is l, then the threshold for the search termination is set to be εl. Based on the accuracy ε, the number of iterations in the bisection search in Step 3 is bounded by $\mathscr{O}(\log_2 \frac{1}{\varepsilon})$. In Step 4, we use a two-layer bisection search, whose number of iteration is bounded by $\mathscr{O}(\log_2^2 \frac{1}{\varepsilon})$. So, for each outer iteration, the total number of inner iterations is $\mathscr{O}(\log_2 \frac{1}{\varepsilon}) + \mathscr{O}(\log_2^2 \frac{1}{\varepsilon})$. Compared with the algorithm in [137], for which the number of inner iterations is $\mathscr{O}(2(2\mathrm{N_{RF}N_r})^{2.5} \log \frac{1}{\varepsilon})$ for each outer iteration, the complexity of our algorithm is much lower especially for large antenna arrays.

3.3.6 $\mathbf{W_d^{BB}}$ design

Since $I(\boldsymbol{s}, \boldsymbol{y}_\mathrm{d}) \leq I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_\mathrm{d})$ [43], after we find the maximum $I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_\mathrm{d})$, the optimal $I(\boldsymbol{s}, \boldsymbol{y}_\mathrm{d})$ will be obtained if the destination node baseband processing does not cause any information loss. According to [117], MMSE-SIC is information lossless. Therefore, we use MMSE-SIC for our destination baseband design. To simplify the expression, let us define

$$\tilde{\boldsymbol{y}}_\mathrm{d} = \tilde{\mathbf{H}}_2 \mathbf{G}_\mathrm{r}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 \mathbf{F}_\mathrm{t}^{\mathrm{BB}} \boldsymbol{s} + \tilde{\mathbf{H}}_2 \mathbf{G}_\mathrm{r}^{\mathrm{BB}} \tilde{\boldsymbol{n}}_1 + \tilde{\boldsymbol{n}}_2 = \mathbf{G}\boldsymbol{s} + \bar{\boldsymbol{v}}, \tag{3.45}$$

where $\mathbf{G} = [\boldsymbol{g}_1, ..., \boldsymbol{g}_{\mathrm{N_s}}] \in \mathbf{C}^{\mathrm{N_{RF}} \times \mathrm{N_s}}$, $\bar{\boldsymbol{v}}$ is the colored noise with covariance matrix $\mathbf{R}_{\tilde{\mathrm{n}}}$.

To implement the MMSE-SIC for the k^{th} stream, we subtract the effect of the first $k-1$ streams

from the output and obtain

$$\tilde{\boldsymbol{y}}_{\mathrm{d}'} = \tilde{\boldsymbol{y}}_{\mathrm{d}} - \sum_{i=1}^{k-1} \boldsymbol{g}_i s_i + \bar{\boldsymbol{v}} = \boldsymbol{g}_k s_k + \sum_{j=k+1}^{\mathrm{N_s}} \boldsymbol{g}_j s_j + \bar{\boldsymbol{v}}. \tag{3.46}$$

Define $\mathbf{W}_{\mathrm{d}}^{\mathrm{BB}} = [\boldsymbol{w}_1, ..., \boldsymbol{w}_{\mathrm{N_s}}]^H$, the baseband filter for the k^{th} stream is derived as the MMSE filter:

$$\boldsymbol{w}_k^H = \boldsymbol{g}_k^H \Big(\sum_{j=k+1}^{\mathrm{N_s}} \boldsymbol{g}_j \boldsymbol{g}_j^H + \mathbf{R}_{\tilde{\mathrm{n}}} \Big)^{-1}. \tag{3.47}$$

3.4 Robust Design

So far, we have designed the mmWave relay precoders/combiners under the perfect channel information. However, it is hard to avoid estimation/quantization errors while obtaining the channel information. To study the effects of imperfect channel estimation, we adopt the model provided in [143, 101, 134]. In this model, the relationship between the channel values and the corresponding estimated channel values are:

$$\mathbf{H}_1 = \bar{\mathbf{H}}_1 + \Phi_1^{\frac{1}{2}} \Delta_1 \Theta_1^{\frac{1}{2}}, \tag{3.48}$$

$$\mathbf{H}_2 = \bar{\mathbf{H}}_2 + \Phi_2^{\frac{1}{2}} \Delta_2 \Theta_2^{\frac{1}{2}}, \tag{3.49}$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are the actual channel matrices, i.e., the channels that we cannot precisely estimate, and $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$ are the estimated channels. The transmitting covariance matrix of the channel estimation error at the source node and the relay node are denoted by $\Theta_1 \in \mathbb{C}^{\mathrm{N_t} \times \mathrm{N_t}}$ and $\Theta_2 \in \mathbb{C}^{\mathrm{N_r} \times \mathrm{N_r}}$, respectively. The receiving covariance matrix of the channel estimation error at the relay node and the destination node are denoted by $\Phi_1 \in \mathbb{C}^{\mathrm{N_r} \times \mathrm{N_r}}$ and $\Phi_2 \in \mathbb{C}^{\mathrm{N_d} \times \mathrm{N_d}}$, respectively.

Δ_1 and Δ_2 are Gaussian random matrices with independent and identically distributed (i.i.d.) zero mean and unit variance entries and are the unknown parts of the CSI mismatch. we adopt the exponential model[134, 101] for the channel estimation error covariance matrices Φ_1, Θ_1, Φ_2 and Θ_2. To be specific, the entries of the matrices are given as $\Phi_1(i,j) = \sigma_{e,1}^2 \beta_1^{|i-j|}$, $\Theta_1(i,j) = \alpha_1^{|i-j|}$, $\Phi_2(i,j) = \sigma_{e,2}^2 \beta_2^{|i-j|}$ and $\Theta_2(i,j) = \alpha_2^{|i-j|}$, where α_1, β_1, α_2 and β_2 are the correlation coefficients and $\sigma_{e,1}^2$ and $\sigma_{e,2}^2$ denote the estimation error covariance. For simplicity, we assume $\alpha_1 = \alpha_2 = \alpha$, $\beta_1 = \beta_2 = \beta$ and $\sigma_{e,1}^2 = \sigma_{e,2}^2 = \sigma_e^2$.

As shown in Section 3.5, the imperfect channel information will result in severe performance degradation. For example, the achievable data rate of [74, 137] can be decreased to half of what it is for the perfect CSI.

To further increase the robustness of our proposed algorithm, in this section, we will propose a robust precoding/combining design for the mmWave relay system based on our proposed WMMSE algorithm.

3.4.1 RF design

Recall that our RF precoding/combining is actually phase compensation for each eigenmode. The eigenmodes are obtained through SVD. When considering the imperfect CSI, the phase of each eigenmode cannot be precisely acquired. Let us take the actual channel $\mathbf{H}_1$ and the estimated $\bar{\mathbf{H}}_1$ as an example. The left singular matrices of $\mathbf{H}_1$ and $\bar{\mathbf{H}}_1$ are denoted by $\mathbf{U}_1$ and $\bar{\mathbf{U}}_1$. We denote the phase of each entry in $\mathbf{U}_1$ and $\bar{\mathbf{U}}_1$ by $\theta_{i,j}$ and $\bar{\theta}_{i,j}$, respectively. The phase difference in each entry can be calculated as $\Delta\theta_{i,j} = \theta_{i,j} - \bar{\theta}_{i,j}$. Let us assume that $\Delta\theta_{i,j}$ has a distribution $p(\Delta\theta)$. We want to make an estimation on $\Delta\theta_{i,j}$ to minimize the mean square estimation error $\mathbb{E}[(\Delta\hat{\theta}_{i,j} - \Delta\theta_{i,j})^2]$. The estimation $\Delta\hat{\theta}_{i,j} = \mathbb{E}[\Delta\theta_{i,j}]$ can be calculated based on the distribution $p(\Delta\theta)$. Note that we can only obtain the estimated channel. Once we calculate $\Delta\hat{\theta}_{i,j}$, we can calibrate the phase of each entry in $\bar{\mathbf{U}}_1$ as $\hat{\theta}_{i,j} = \bar{\theta}_{i,j} + \Delta\hat{\theta}_{i,j}$. Using the same approach, we can calibrate the phase of singular

matrices of $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$. Then, based on (3.12) and (3.13), we can calculate the RF precoders and combiners $\bar{\mathbf{F}}_t^{\mathrm{RF}}$, $\bar{\mathbf{W}}_r^{\mathrm{RF}}$, $\bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}}$ and $\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{RF}}$ based on the calibrated singular matrices of $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$.

As we analyzed above, to calculate the RF precoders and combiners, we must know the distribution of $\Delta\theta_{i,j}$ to make the estimation $\Delta\hat{\theta}_{i,j} = \mathbb{E}[\Delta\theta_{i,j}]$. However, the theoretical analysis for the phase distribution is intractable. To obtain the phase distribution, we simulate 100 channel realizations based on the imperfect channel models in (3.48), where we set $\mathrm{N_r} = 32$, $\mathrm{N_t} = 48$ and $L = 20$. We adopt the correlation model from [143, 101, 134] where the entries of the correlation matrices are selected as $\Phi_1(i,j) = \sigma_{e,1}^2 \beta_1^{|i-j|}$, $\Theta_1(i,j) = \alpha_1^{|i-j|}$. In the simulation, we set $\alpha_1 = 0$, $\beta_1 = 0$ and $\sigma_{e,1}^2 = 0.1$.

We collect the phase difference in each matrix entry from 60 simulations. In Fig. 3.3, we plot the simulated probability density function (PDF) of the $\Delta\theta$ in solid line. We use a generalized normal distribution [87] to approximate the distribution. The PDF of a generalized normal distribution is expressed as $f(x) = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} e^{-(|x-\mu|/\alpha)^\beta}$. We can see the approaching effect of different value of the shaping parameter β in Fig. 3.3. We use Kullback-Leibler distance as a performance measure for the approximation, which is calculated by $D_{\mathrm{KL}}(\boldsymbol{Y}||\boldsymbol{X}) = \sum_{i=1}^{N} \log(\frac{Y_i}{X_i}) Y_i$ where $\boldsymbol{Y}$ and $\boldsymbol{X}$ are the probability distributions. The lower the Kullback-Leibler distance, the closer the two distributions are. Note that the absolute value of KL-distance varies when the number of total points (i.e., N) changes. In our simulation, the best approximation comes with the one with $\beta = 2$, since it has the lowest Kullback-Leibler distance. When $\beta = 2$, the generalized normal distribution in Fig. 3.3 is a Gaussian distribution with 0 mean, which means we can estimate $\Delta\hat{\theta}_{i,j} = 0$.

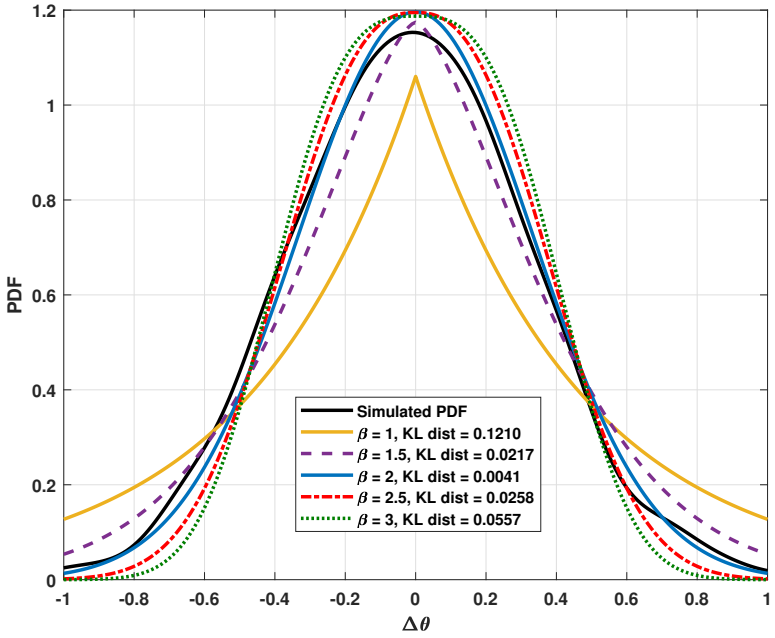

Figure 3.3: Approximation of the Simulated PDF

3.4.2 Baseband design

Based on the RF precoders and combiners $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}}$, $\bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}}$, $\bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}}$ and $\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{RF}}$ designed in the last subsection, the equivalent baseband channels after the RF processing are

$$\begin{aligned}\tilde{\mathbf{H}}_1 &= \bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}}\mathbf{H}_1\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}} = \bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}}\bar{\mathbf{H}}_1\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}} + \bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}}\Phi_1^{\frac{1}{2}}\Delta_1\Theta_1^{\frac{1}{2}}\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}} \\ &= \tilde{\bar{\mathbf{H}}}_1 + \tilde{\Phi}_1^{\frac{1}{2}}\Delta_1\tilde{\Theta}_1^{\frac{1}{2}},\end{aligned} \tag{3.50}$$

$$\begin{aligned}\tilde{\mathbf{H}}_2 &= \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{RF}}\mathbf{H}_2\bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}} = \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{RF}}\bar{\mathbf{H}}_2\bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}} + \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{RF}}\Phi_2^{\frac{1}{2}}\Delta_2\Theta_2^{\frac{1}{2}}\bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}} \\ &= \tilde{\bar{\mathbf{H}}}_2 + \tilde{\Phi}_2^{\frac{1}{2}}\Delta_2\tilde{\Theta}_2^{\frac{1}{2}},\end{aligned} \tag{3.51}$$

where we denote the true equivalent baseband channels by $\tilde{\mathbf{H}}_1$ and $\tilde{\mathbf{H}}_2$. The estimated equivalent baseband channels are denoted by $\tilde{\bar{\mathbf{H}}}_1$ and $\tilde{\bar{\mathbf{H}}}_2$, which are the channels we obtain at the source node. We define $\tilde{\Phi}_1 := \bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}}\Phi_1(\bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}})^H$, $\tilde{\Phi}_2 := \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{RF}}\Phi_2(\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{RF}})^H$, $\tilde{\Theta}_1 := (\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}})^H\Theta_1\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}}$ and $\tilde{\Theta}_2 :=$

$(\bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}})^H \Theta_2 \bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}}$.

For (3.50) and (3.51) , we have the following properties [46] (Using $\tilde{\mathbf{H}}_1$ as an example)

$$\mathbb{E}_{\Delta_1}[\tilde{\mathbf{H}}_1 \mathbf{C} \tilde{\mathbf{H}}_1^H] = \tilde{\bar{\mathbf{H}}}_1 \mathbf{C} \tilde{\bar{\mathbf{H}}}_1^H + \mathrm{Tr}(\mathbf{C}\tilde{\Theta}_1)\tilde{\Phi}_1, \tag{3.52}$$

$$\mathbb{E}_{\Delta_1}[\tilde{\mathbf{H}}_1^H \mathbf{C} \tilde{\mathbf{H}}_1] = \tilde{\bar{\mathbf{H}}}_1^H \mathbf{C} \tilde{\bar{\mathbf{H}}}_1 + \mathrm{Tr}(\tilde{\Phi}_1 \mathbf{C})\tilde{\Theta}_1. \tag{3.53}$$

To design a robust baseband system, we need to redesign the algorithm in Section 3.3.3 based on the imperfect channel models (3.50) and (3.51). The main idea is similar, i.e., that we first optimize the baseband filters $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}$ and $\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$ to maximize the average $\mathbb{E}_{\Delta_1,\Delta_2}[I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})]$, and then we use MMSE-SIC for $\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{BB}}$. Note that we denote the baseband precoder/combiner based on the estimated equivalent baseband channels $\tilde{\bar{\mathbf{H}}}_1$ and $\tilde{\bar{\mathbf{H}}}_2$ by $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}$, $\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$ and $\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{BB}}$. The main challenge here is if there still exists an equivalent relationship between the average mutual information maximization and the WMMSE minimization.

According to (3.23), the average EMMSE $\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}]$ can be expressed as

$$\begin{aligned} &\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}] \\ &= \mathbb{E}_{\Delta_1,\Delta_2}[(\mathbf{I}_{\mathrm{N_s}} + (\tilde{\mathbf{H}}_2 \bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 \bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}})^H \mathbf{R}_{\tilde{\mathrm{n}}}^{-1} \tilde{\mathbf{H}}_2 \bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 \bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}})^{-1}] \end{aligned} \tag{3.54}$$

Then, there exists an equivalent relationship the an upper bound for the average mutual information of $\mathbb{E}_{\Delta_1,\Delta_2}[I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})]$ and the average EMMSE $\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}]$:

$$\begin{aligned} &\mathbb{E}^{\mathrm{UB}}_{\Delta_1,\Delta_2}[I(\boldsymbol{s}, \tilde{\boldsymbol{y}}_{\mathrm{d}})] = \\ &\log_2 \det(\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{I}_{\mathrm{N_s}} + (\tilde{\mathbf{H}}_2 \bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 \bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}})^H \mathbf{R}_{\tilde{\mathrm{n}}}^{-1} \tilde{\mathbf{H}}_2 \bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 \bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}]) = \\ &\log_2(\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}]^{-1}) \end{aligned} \tag{3.55}$$

Equation (3.55) implies that we can maximize the upper bound of the average mutual information through the WMMSE minimization as discussed in Section 3.3.3.

The expression of $\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\text{MMSE}}]$ in (3.55) includes a matrix inverse operator, which complicates the following calculation for $\bar{\mathbf{F}}_{\text{t}}^{\text{BB}}$ and $\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}$. To derive a simpler expression for $\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\text{MMSE}}]$, we first calculate the average MSE matrix $\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\text{MSE}}]$. The MSE matrix is given by

$$\begin{aligned}\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\text{MSE}}] = \bar{\mathbf{W}}_{\text{d}}^{\text{BB}}(\mathbf{A}+\mathbf{R}_{\tilde{\tilde{\text{n}}}})(\bar{\mathbf{W}}_{\text{d}}^{\text{BB}})^H - (\bar{\mathbf{W}}_{\text{d}}^{\text{BB}})^H\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}} \\ - (\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}})^H\bar{\mathbf{W}}_{\text{d}}^{\text{BB}} + \mathbf{I}_{\text{N}_\text{s}},\end{aligned} \tag{3.56}$$

where

$$\mathbf{A} := \tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\mathbf{A}_1(\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}})^H + \text{Tr}(\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\mathbf{A}_1(\bar{\mathbf{G}}_{\text{r}}^{\text{BB}})^H\tilde{\Theta}_2^H)\tilde{\Phi}_2,$$

$$\mathbf{A}_1 := \tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}}(\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}})^H + \text{Tr}(\bar{\mathbf{F}}_{\text{t}}^{\text{BB}}(\bar{\mathbf{F}}_{\text{t}}^{\text{BB}})^H\tilde{\Theta}_1^H)\tilde{\Phi}_1.$$

Based on (3.56), we can derive the $\bar{\mathbf{W}}_{\text{d}}^{\text{MMSE}}$, which minimizes $\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\text{MSE}}]$, as

$$\bar{\mathbf{W}}_{\text{d}}^{\text{MMSE}} = (\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}})^H(\mathbf{A}+\mathbf{R}_{\tilde{\tilde{\text{n}}}})^{-1}. \tag{3.57}$$

Substituting (3.57) into (3.56), we have

$$\begin{aligned}&\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\text{MMSE}}] = \bar{\mathbf{W}}_{\text{d}}^{\text{MMSE}}(\mathbf{A}+\mathbf{R}_{\tilde{\tilde{\text{n}}}})(\bar{\mathbf{W}}_{\text{d}}^{\text{MMSE}})^H \\ &- (\bar{\mathbf{W}}_{\text{d}}^{\text{MMSE}})^H\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}} - (\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}})^H\bar{\mathbf{W}}_{\text{d}}^{\text{BB}} + \mathbf{I}_{\text{N}_\text{s}} \\ &= \mathbf{I}_{\text{N}_\text{s}} - (\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}})^H(\mathbf{A}+\mathbf{R}_{\tilde{\tilde{\text{n}}}})^{-1}\tilde{\bar{\mathbf{H}}}_2\bar{\mathbf{G}}_{\text{r}}^{\text{BB}}\tilde{\bar{\mathbf{H}}}_1\bar{\mathbf{F}}_{\text{t}}^{\text{BB}}.\end{aligned} \tag{3.58}$$

Based on (3.58), we can amend our results based on the imperfect channel model, using the same

procedure as Section 3.3.3. For $\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$, the amended expression is

$$\begin{aligned}\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} = &(\mathbf{K}_1 + \lambda^r (\bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}})^H \bar{\mathbf{F}}_{\mathrm{r}}^{\mathrm{RF}})^{-1} (\tilde{\bar{\mathbf{H}}}_2)^H (\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}})^H \bar{\mathbf{V}} (\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}})^H \\ &(\mathbf{K}_2 + \sigma_{\mathrm{I}}^2 \bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}} (\bar{\mathbf{W}}_{\mathrm{r}}^{\mathrm{RF}})^H)^{-1},\end{aligned} \tag{3.59}$$

where

$$\bar{\mathbf{V}} = \frac{\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}]^{-1}}{\log 2},$$

$$\mathbf{K}_1 := (\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}} \tilde{\bar{\mathbf{H}}}_2)^H \bar{\mathbf{V}} \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}} \tilde{\bar{\mathbf{H}}}_2 + \mathrm{Tr}(\tilde{\Phi}_2 (\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}})^H \bar{\mathbf{V}} \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}}) \tilde{\Theta}_2,$$

$$\mathbf{K}_2 := \tilde{\bar{\mathbf{H}}}_1 \bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}} (\tilde{\bar{\mathbf{H}}}_1 \bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}})^H + \mathrm{Tr}(\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}} (\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}})^H \tilde{\Theta}_1) \tilde{\Phi}_1.$$

For $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}$, the amended expression is

$$\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}} = (\mathbf{T}_1 + \lambda_{\mathrm{I}}^t (\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}})^H \bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{RF}} + \lambda_2^t \mathbf{T}_2)^{-1} (\bar{\mathbf{V}} \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}} \tilde{\bar{\mathbf{H}}}_2 \bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\bar{H}}_1)^H, \tag{3.60}$$

where

$$\mathbf{T}_1 := \tilde{\bar{\mathbf{H}}}_1^H (\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}})^H \mathbf{B} \bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\bar{\mathbf{H}}}_1 + \mathrm{Tr}(\tilde{\Phi}_1 (\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}})^H \mathbf{B} \bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}) \tilde{\Theta}_1,$$

$$\mathbf{T}_2 := (\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}} \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1)^H \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}} \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}} \tilde{\mathbf{H}}_1 + \mathrm{Tr}(\tilde{\Phi}_1 (\mathbf{F}_{\mathrm{r}}^{\mathrm{RF}} \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}})^H \mathbf{F}_{\mathrm{r}}^{\mathrm{RF}} \mathbf{G}_{\mathrm{r}}^{\mathrm{BB}}) \tilde{\Theta}_1,$$

$$\mathbf{B} := \tilde{\bar{\mathbf{H}}}_2^H (\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}})^H \mathbf{V} \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}} \tilde{\bar{\mathbf{H}}}_2 + \mathrm{Tr}(\tilde{\Phi}_2 (\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}})^H \mathbf{V} \bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}}) \tilde{\Theta}_2.$$

Based on the above modifications, our robust baseband design for $\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$ and $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}$ is as follows:

1. Calculate the MMSE receiver $\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{MMSE}}$ in Eq. (3.57) and the MMSE matrix $\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}]$ in Eq. (3.58).
2. Update $\bar{\mathbf{V}}$ by setting $\bar{\mathbf{V}} = \frac{\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}]^{-1}}{\log 2}$.

3. Fix $\bar{\mathbf{V}}$ and $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}$, then we find $\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$ that minimizes $\mathrm{Tr}(\bar{\mathbf{V}}\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}])$ under the power constraints. The solution is given by Eq.(3.59).

4. Fix $\bar{\mathbf{V}}$ and $\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$, then we find $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}$ that minimizes $\mathrm{Tr}(\bar{\mathbf{V}}\mathbb{E}_{\Delta_1,\Delta_2}[\mathbf{E}_{\mathrm{MMSE}}])$ under the power constraints. The solution is given by Eq.(3.60).

After we obtain $\bar{\mathbf{G}}_{\mathrm{r}}^{\mathrm{BB}}$ and $\bar{\mathbf{F}}_{\mathrm{t}}^{\mathrm{BB}}$, we will use MMSE-SIC to design $\bar{\mathbf{W}}_{\mathrm{d}}^{\mathrm{BB}}$, which is the same as what we did in Section 3.3.3.

3.5 Simulation Results

3.5.1 Non-robust case

In this section, we consider a relay MIMO system consisting of one source node equipped with a $\mathrm{N_t} = 64$ antenna array, a relay node with an $\mathrm{N_r} = 32$ antenna array and a destination node with a $\mathrm{N_d} = 48$ antenna array unless other number of antennas are specifically mentioned. The number of antennas is chosen from [137] for the purpose of the comparison. For simplicity, we use the channel model in Eq. (3.9) for channel realization. Due to the limited scattering characteristic of the mmWave channels, the number of paths should be less than the number of relay antennas. Here, we assume each channel has $\mathrm{L} = 20$ paths. The φ_l of each path is assumed to be uniformly distributed in $[0, 2\pi]$. The results are averaged over 2000 channel realizations. The SNR of the source-to-relay link and the relay-to-destination are assumed to be the same. In the simulation, we calculate the variances of AWGN noises σ_1 and σ_2 according to the source power and the relay power to maintain the same SNR.

In Fig. 3.4, we equally set the power of source node and the relay node, all to be $\mathrm{N_s}$. We compare our algorithm with the ADMM in [118], the ISA in [137] and the OMP in [74] in terms of the achievable data rate. We use three scenarios: i) the number of data streams is $\mathrm{N_s} = 4$ and the

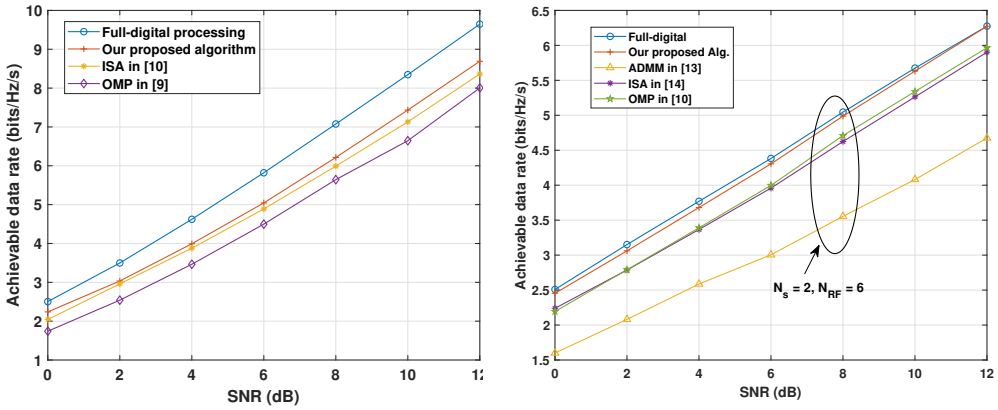

Figure 3.4: Achievable rate comparison with $64 \times 32 \times 48$ when $E_s = E_r = N_s$

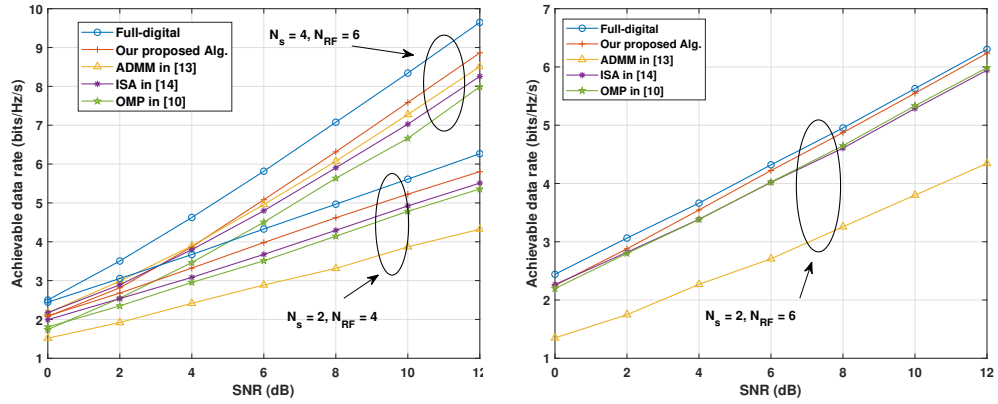

Figure 3.5: Achievable rate comparison with $64 \times 32 \times 48$ when $E_s = 2E_r = 2N_s$

number of RF chains is $N_{RF} = 6$; ii) the number of data streams is $N_s = 2$ and the number of RF chains is $N_{RF} = 4$; iii) the number of data streams is $N_s = 2$ and the number of RF chains is $N_{RF} = 6$. The full-digital method is used as a benchmark, where we use the singular matrices of $\mathbf{H}_1$ and $\mathbf{H}_2$ as the precoding/combing matrices. When $N_s = 4$, our algorithm outperforms ADMM by 2%, ISA by 4% and OMP by 9% at SNR $= 12$ dB. When $N_s = 2$ and $N_{RF} = 4$, our algorithm outperforms ADMM by 32%, ISA by 6% and OMP by 9% at SNR $= 12$ dB. When $N_s = 2$ and $N_{RF} = 6$, our algorithm outperforms ADMM by 34%, ISA by 6 % and OMP by 5% at SNR $= 12$ dB.

In Fig. 3.5, we set $E_s = 2E_r = 2N_s$. Our proposed algorithm outperforms the other three methods in three scenarios. When $N_s = 4$, our algorithm can provide 4%, 7% and 11% gains over ADMM, ISA and OMP, respectively, at SNR $= 12$ *dB*. When $N_s = 2$ and $N_{RF} = 4$, our algorithm can provide a gain of 34% over ADMM, 5% over ISA and 8% over OMP at SNR $= 12$ dB. When $N_s = 2$ and $N_{RF} = 6$, our algorithm can provide a gain of 44% over ADMM, 5% over ISA and 4% over OMP at SNR $= 12$ dB.

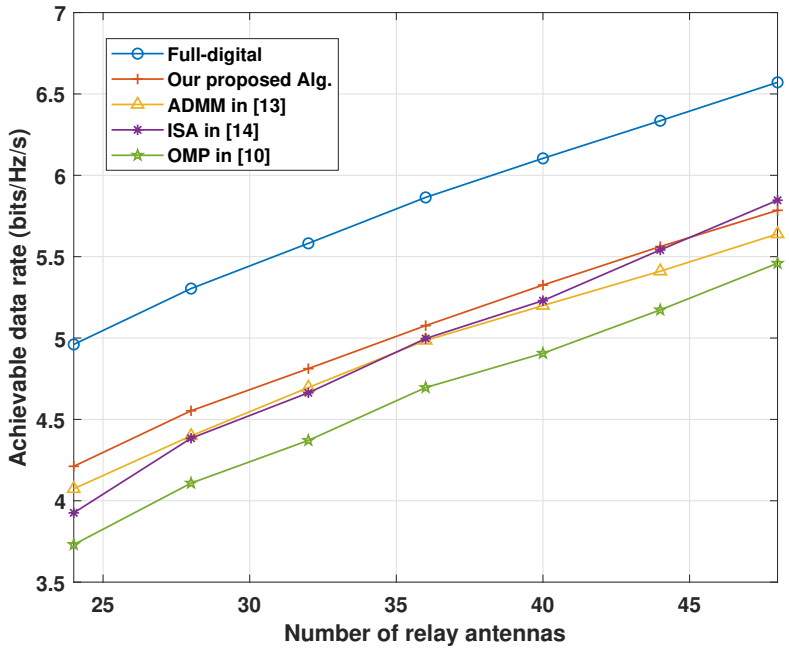

Figure 3.6: Achievable rate comparison with different relay antennas when $N_s = 4$, $N_{RF} = 6$ and SNR $= 5$dB

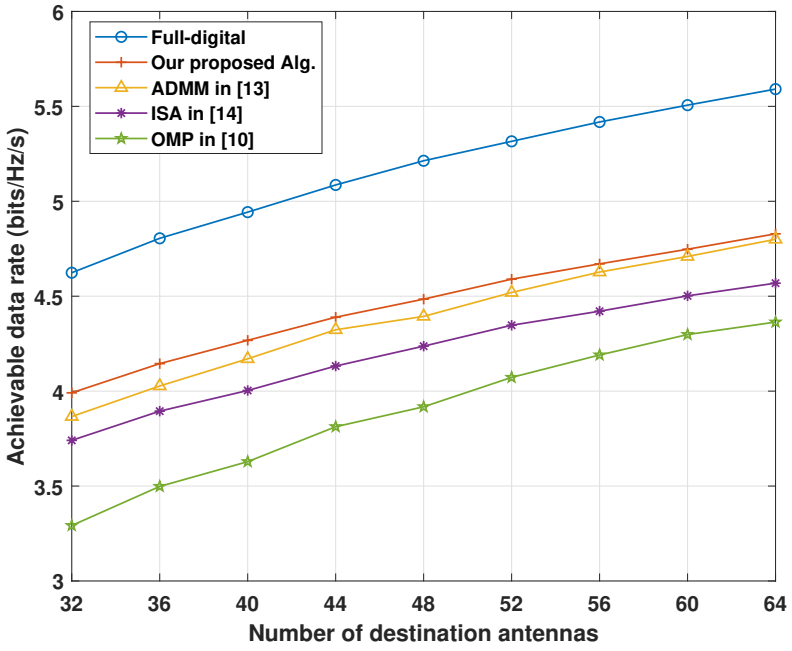

Figure 3.7: Achievable rate comparison with different destination antennas when $N_s = 4$, $N_{RF} = 6$ and SNR $= 5$ dB

Fig. 3.6 compares the achievable rate of different algorithms for different number of relay antennas when $N_s = 4$, $N_{RF} = 6$ and SNR $= 5$ dB. The full-digital method is used as a benchmark. As expected, when the number of antennas at the relay node increases, the performance of all different algorithms improves because of the additional antenna gain. Our proposed method has the best achievable rate performance among the four methods except for $N_r = 48$. When $N_r = 48$, ISA has the highest achievable rate among the four methods. However, as the number of antennas at the relay node increases, the complexity of the ISA increases greatly, which will lead to a high power consumption.

Fig. 3.7 compares the achievable rate for different number of antennas at the destination node when $N_s = 4$, $N_{RF} = 6$ and SNR $= 5$ dB. Similar to Fig. 3.6, when the number of antennas at the destination node increases, the performance of all different algorithms improves because of the additional antenna gain. Our proposed method has the best achievable rate performance among the four methods.

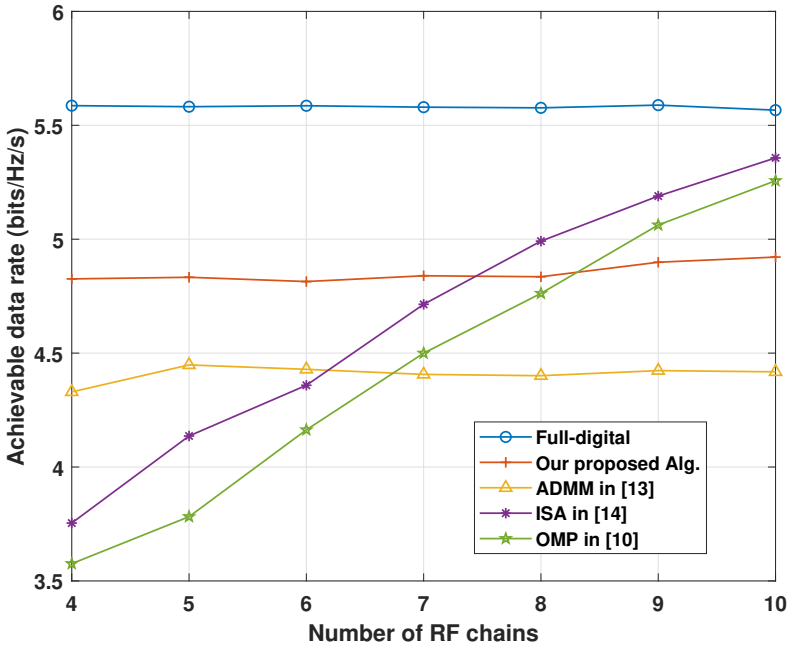

Figure 3.8: Achievable rate comparison with different RF chains when $N_s = 4$ and SNR $= 5$ dB using channel model (3.9)

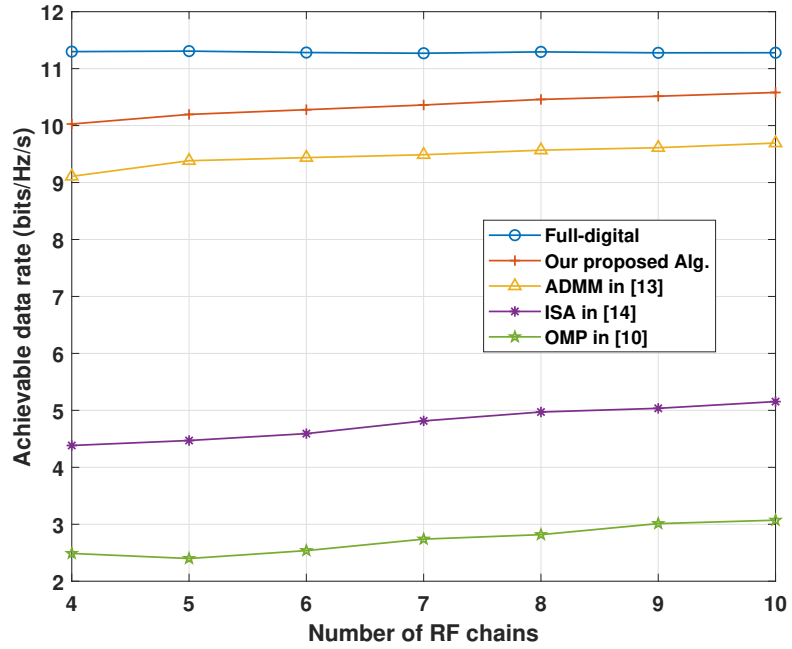

Figure 3.9: Achievable rate comparison with different RF chains when $N_s = 4$ and SNR $= 5$ dB using channel model (3.7)

Fig. 3.8 compares the achievable rate among the four methods for different number of RF chains when $N_s = 4$ and SNR $= 5$ dB. Since our proposed method is designed to maximize the mutual information between the destination node and the source node after RF precoding/combining, the gap between our method and the full-digital method is more-or-less fixed, which is caused by the analog processing. However, ISA and OMP are approximation algorithms jointly iterating between the RF and the baseband. Therefore, as the number of RF chains increases, the performance improves. When the number of RF chains is larger than 8, ISA and OMP will outperform our proposed algorithm. However, larger number of RF chains leads to higher complexity and more power consumption. Also, the performance of the approximation algorithms depends on the limited scattering characteristic of the channel. The more sparse the channel is, the better performance the approximation algorithms achieve. In Fig. 3.8, we use the highly limited scattering channel model in (3.9), where each scatter only contributes to one path, thus the approximation algorithms have good performance. If we use the general channel model in (3.7), the performance of approximation algorithms degrades greatly as shown in Fig. 3.9. In Fig. 3.9, we set the number of propagation paths N_{cl} in each scatter to be 2 and the number of scatters L to be 20. In this case, the performance of ISA and OMP falls far behind our proposed algorithm.

Fig. 3.10 shows the convergence performance of different algorithms with respect to the number of iterations. In our algorithm, we update the WMMSE matrix, the digital relay matrix and the digital precoding matrix sequentially in each iteration. In ISA, the digital relay filter, the analog relay receiver and the analog relay precoder are updated sequentially in each iteration. In ADMM, the source node, the relay node and the destination node are optimized alternatively in each iteration. In Fig. 3.10, our algorithm has the fastest convergence rate while ADMM has the slowest convergence rate. Moreover, our algorithm has much lower complexity in each iteration compared with ISA. ISA needs to solve three optimization sub-problems, and in each sub-problem it needs to solve an optimization problem through an iterative method. In our algorithm, we have closed-form solutions for each step. In addition, since we preform the baseband processing after the RF processing, the matrix dimensions are greatly reduced compared to ISA.

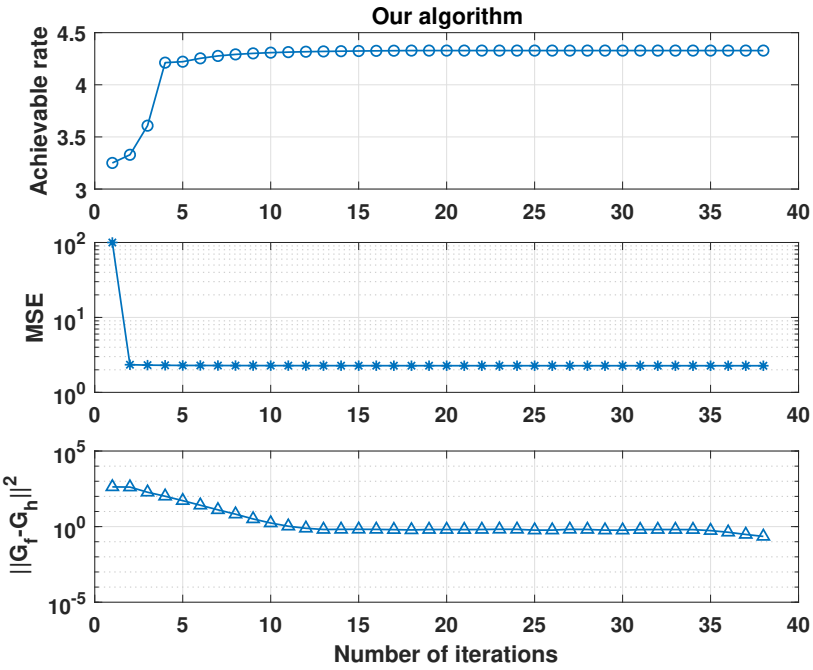

Figure 3.10: Convergence rate comparison

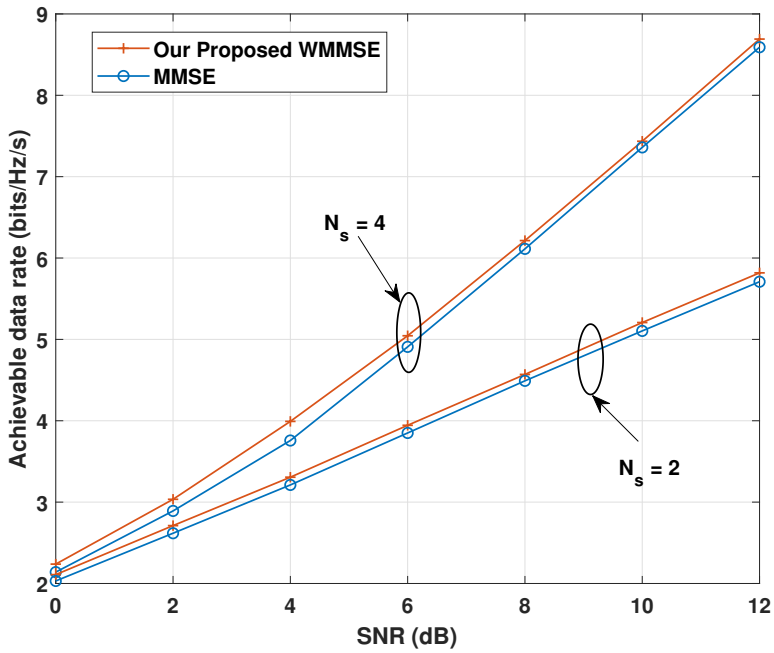

Figure 3.11: Baseband algorithms comparison

Fig. 3.11 compares the baseband processing algorithms. Note that we apply the MMSE algorithm only on the baseband, i.e., on the $\tilde{\mathbf{H}}_1$ and $\tilde{\mathbf{H}}_2$. Our proposed WMMSE algorithm outperforms the MMSE algorithm in terms of the achievable data rate since we optimized the mutual information $I(\mathbf{s}, \mathbf{y_d})$. In fact, if we set our weight matrix to be the identity matrix, our algorithm degenerates to the MMSE algorithm. Therefore, the MMSE algorithm can be considered as a special case of our proposed WMMSE algorithm and our algorithm strictly performs better than MMSE.

3.5.2 Robust case

As we described in Section 3.4, we adopt the channel estimation error model from [143, 101, 134] where the entries of the correlation matrices are selected as $\Phi_1(i,j) = \sigma_{e,1}^2 \beta_1^{|i-j|}$, $\Theta_1(i,j) = \alpha_1^{|i-j|}$, $\Phi_2(i,j) = \sigma_{e,2}^2 \beta_2^{|i-j|}$ and $\Theta_2(i,j) = \alpha_2^{|i-j|}$. Parameters α_1, β_1, α_2 and β_2 are the correlation coefficients and $\sigma_{e,1}^2$ and $\sigma_{e,2}^2$ denote the estimation error covariance. For simplicity, we assume $\alpha_1 = \alpha_2 = \alpha$, $\beta_1 = \beta_2 = \beta$ and $\sigma_{e,1}^2 = \sigma_{e,2}^2 = \sigma_e^2$. The antenna settings are the same as the non-robust part and the number of scatters is set to be 20. The actual channels $\mathbf{H}_1$ and $\mathbf{H}_2$ are generated based on sparse channel model (3.9) and the estimated channels are generated by $\bar{\mathbf{H}}_1 = \mathbf{H}_1 - \Phi_1^{\frac{1}{2}} \Delta_1 \Theta_1^{\frac{1}{2}}$ and $\bar{\mathbf{H}}_2 = \mathbf{H}_2 - \Phi_2^{\frac{1}{2}} \Delta_2 \Theta_2^{\frac{1}{2}}$.

Fig. 3.12 shows the effects of the channel estimation error. We provide the performance of our algorithm and those of [74, 137]. For this simulation, we have chosen $\sigma^2 = 0.1$, $\alpha = 0.6$ and $\beta = 0.4$. As shown in Fig. 3.12, the imperfect channel information will result in severe performance degradation. The achievable data rate of [74, 137] can be decreased to half of what it is for the perfect CSI.

The achievable data rate performances of the proposed robust scheme with various antenna covariance values are depicted in Figs. 3.13 and 3.14. When SNR is low, the estimation error can be neglected compared to the noise, therefore the non-robust algorithm achieves good performance which can be even better than that of the robust algorithm. When SNR goes up, the performance of

the non-robust algorithm starts to degrade. In Fig. 3.14, the performance becomes worse than that of the low SNR region for large σ_e^2. Meanwhile, the proposed robust design offers significant gain considering various σ_e^2, which demonstrates the effectiveness of the modified robust transceiver optimization.

In Fig. 3.15, we compare our robust algorithm with the OMP algorithm in [83]. We set $\alpha = 0$ and $\beta = 0$ for simplicity. The proposed robust design provides a large gain over the algorithm in [83] in all three σ_e^2 settings, showing the advantage of our algorithm.

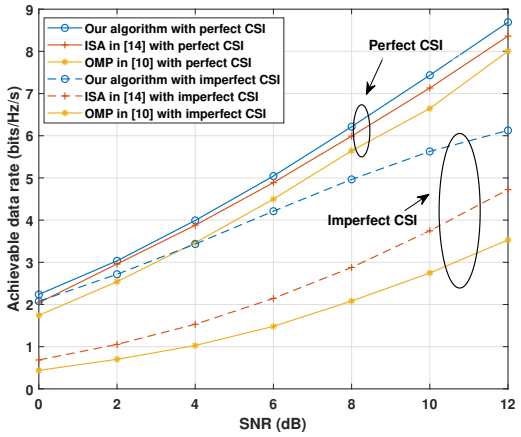

Figure 3.12: The effect of imperfect CSI when the error covariance $\sigma_e^2 = 0.1$, $\alpha = 0.6$ and $\beta = 0.4$

3.6 Conclusion

In this chapter, we considered mmWave AF relay networks in the domain of massive MIMO. We designed the hybrid precoding/combining matrices for the source node, the relay node, and the destination node. We first performed the RF processing to decompose the channel into parallel sub-channels by compensating the phase of each eigenmode of the channel. Given the RF processing matrices, we designed the baseband matrices to maximize the mutual information. The baseband processing is divided into two parts. We first jointly designed the source node and the relay node by making use of the equivalence between maximizing the mutual information and the

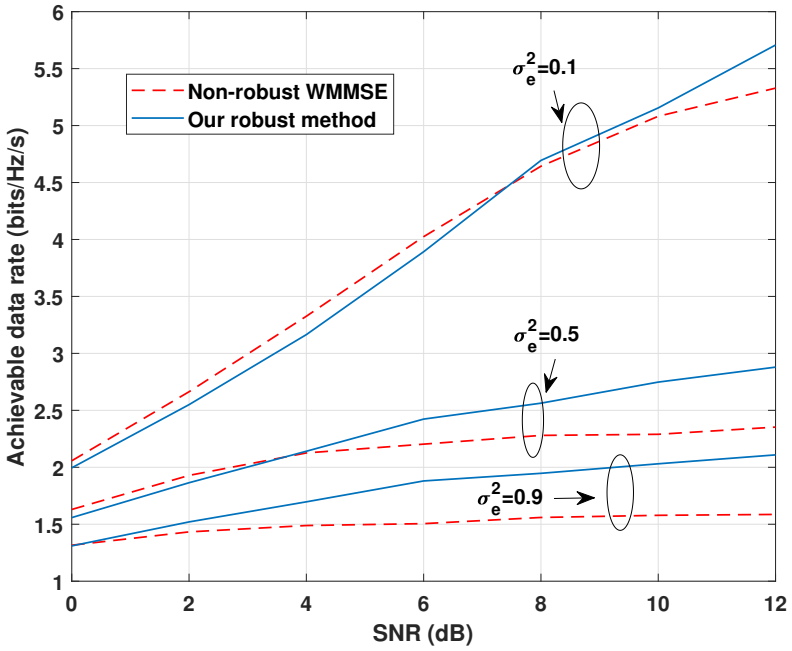

Figure 3.13: Achievable rate comparison when $\alpha = 0$, $\beta = 0$

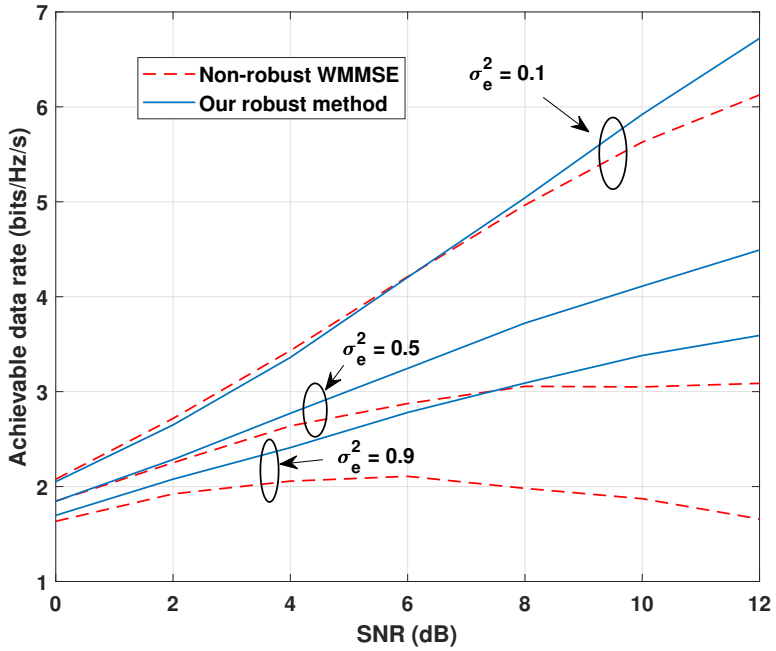

Figure 3.14: Achievable rate comparison when $\alpha = 0.5$, $\beta = 0.5$

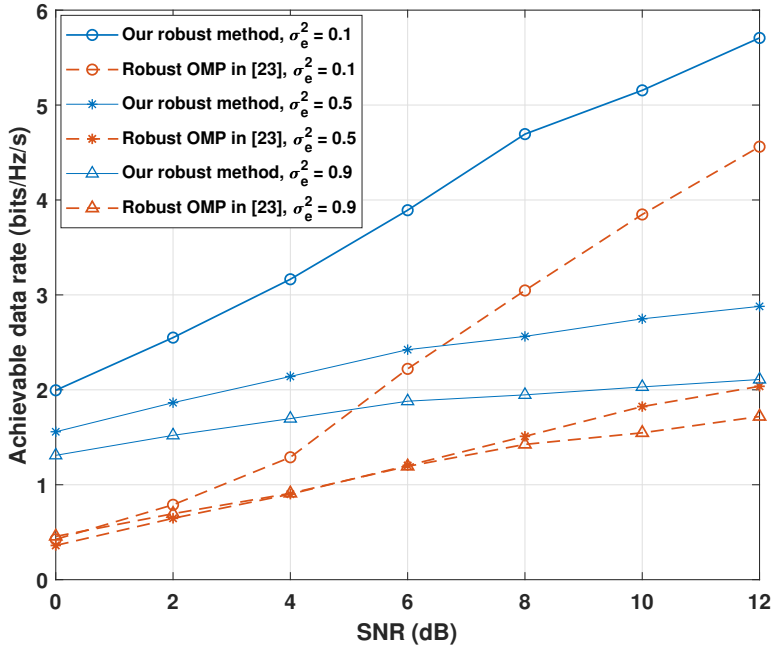

Figure 3.15: Achievable rate comparison with the OMP algorithm in [83] when $\alpha = 0$, $\beta = 0$

WMMSE. Given the optimal baseband source and relay filters, we implemented MMSE-SIC for baseband destination node to obtain the maximal mutual information. In addition, a robust hybrid precoding/combining design was proposed for the imperfect CSI. Simulation results show that our algorithm achieves better performance with lower complexity compared with other algorithms in the literature.

Chapter 4

Joint Beamwidth and Power Optimization in MmWave Hybrid Beamforming-NOMA Systems

4.1 Introduction

Current wireless communication networks operating under 6 GHz are restrained by limited spectral resources. Subsequently, it is necessary to use the millimeter-Wave (mmWave) band ranging from 30 to 300 GHz to increase the available spectrum [99]. Short wavelength and a large path loss are key characteristics of mmWave communication systems. Due to the short wavelength, a large number of antennas can be packed in a small area in mmWave devices. This feature combined with beamforming can be used to tackle severe path loss. Since a fully-digital beamforming may not be practical, various architectures have been proposed for mmWave outdoor communications, i.e., analog beamforming with multiple RF chains [62], hybrid beamforming [12, 112, 64], beamspace multiple-input multiple-output (MIMO) [15], and reconfigurable antenna-based MIMO [11, 51].

The fully-digital architecture needs one radio frequency (RF) chain per antenna. Hence, power consumption by large number of RF chains and hardware complexity are the main obstacles in implementing the fully-digital architecture. Although multi-user analog beamforming reduces the hardware complexity of the system and uses only the angle of arrival (AoA) information for beam alignment, it may not completely direct the total energy of a beam toward a desired receiver [62]. Accordingly, alternative methods such as beamspace MIMO [15] and reconfigurable antenna-based MIMO [11, 51] architectures reduce the number of required RF chains by dedicating one RF chain to each channel path instead of each antenna. However, these architectures are not able to change their beamwidth, which seems to be necessary and desired in the mmWave networks [125]. This is because in lens-based architectures, the lens operates like a passive phase-shifter network. Hence, it may not be possible to adjust the beamwidth. In contrast, not only does the hybrid beamforming architecture reduce the number of required RF chains, but also, thanks to the use of phase-shifters, it can adjust the transmission beamwidth. Hence, in this chapter, we adopt the hybrid beamforming architecture which is a feasible solution to meet the demands in mmWave networks.

Non-orthogonal multiple access (NOMA) aims to improve the spectral efficiency and simultaneously serve more than one user at the same frequency/time/code in single-carrier and multi-carrier systems [121, 36]. Especially, NOMA transmits the users' signal at the same time slot and frequency band by using superposition coding (SC) and decodes the desired signal by exploiting successive interference cancellation (SIC) at the receiver [103]. In this chapter, we leverage power-domain NOMA in which each user has a different level of power.

Recently, NOMA has been incorporated into the mmWave communication, termed mmWave-NOMA, to enhance spectral efficiency and connectivity of the network. Here, we review the work on mmWave-NOMA networks in the downlink transmission with a single transmitter [27, 25, 123, 48, 133, 147, 125, 3, 9]. In [27], a random beamforming method is studied for mmWave directional transmission. In [25], two NOMA users with different directions are assigned the same

beamforming codeword using phase-shifters with finite resolution. NOMA is combined with lens-based beamspace MIMO in [123], and a power allocation algorithm is proposed. Energy-efficiency of mmWave-NOMA networks is evaluated in [48]. A joint power allocation optimization to design beamforming vectors for mmWave-NOMA networks is presented in [133]. The coverage and rate of mmWave-NOMA networks for analog beamforming in the presence of misalignment between the transmit and receive beams is analyzed in [147]. The impact of beamwidth on user pairing in mmWave-NOMA is studied in [125]. Also, [3] evaluates the effect of beam misalignment on the sum-rate performance of mmWave-NOMA networks with hybrid beamforming. Further, NOMA is utilized in lens-based mmWave reconfigurable antennas to increase the number of served users and improve the sum-rate in [9]. What is common among the above works is the assumption that there is sufficient time to train the beams.

In practical scenarios, neglecting the effect of beam-training duration may cast doubt on the performance of the mmWave-NOMA networks. Especially, since the channel coherence time in mmWave bands is limited [97], the beam-training duration should be adequately small. Thus, on one hand, a small beam-training duration results in a wide beamwidth, i.e., low beamforming gain, and noisy channel estimation. On the other hand, a long beam-training time provides robust beamforming and accurate channel estimation but imposes a delay in data transmission. This may not be desirable in delay-sensitive systems as it leaves less time for data transmission and leads to low sum-rates. There is a rich literature on fast beam-training algorithms [57, 7, 109, 92, 145, 90, 34, 95, 50, 59]. This issue is very crucial in mmWave-NOMA networks in which more users are trained at each frequency/time resource. Beamwidth control and sum-rate trade-off in the mmWave analog beamforming-NOMA network for two users are evaluated in [49]. The impact of beam-training duration on the sum-rate of the system is determined and then an optimization problem that maximizes the sum-rate subject to the training duration and allocated power for each user is investigated. However, due to the inter-cluster interference, an extension of this architecture to the mmWave hybrid beamforming-NOMA network is quite challenging. In this chapter, motivated by [49], we study the beamwidth control and sum-rate trade-off for the mmWave hybrid

beamforming-NOMA network. There are two major differences between [49] and our work. First, we consider a hybrid beamforming system which produces side lobes and as a result inter-cluster interference. Second, we do not allow receivers to have a beamwidth wider than that of their intended transmit beam. Otherwise, the receiver cannot catch the entire transmission energy. Neither the first case nor the second case is considered in [49]. The contributions of this chapter are listed below:

1. We consider the well-studied mmWave hybrid beamforming combined with NOMA for limited coherence time scenarios. The system can control the beamwidth, using the phase-shifters deployed in the hybrid beamformer, and allocate power to NOMA users. To this end, a tone-based beam-training algorithm [145] compatible to our mmWave-NOMA system is utilized. The algorithm combines the exhaustive search [57] and tone-based beam-training [145] algorithms.

2. Unlike the existing multi-beam mmWave-NOMA systems, we take the channel coherence time into account. The limited coherence time leads to a trade-off between the beamwidth resolution and the data transmission rate. We also formulate a new sum-rate expression for optimization.

3. A joint power and beamwidth optimization algorithm is proposed which iterates between the power allocation and beamwidth optimization.

4. The numerical results verify the effectiveness of the joint optimization algorithm. Also, three significant results are revealed. First, at low signal-to-noise ratios (SNRs), both power allocation and beamwidth control play a major role in the sum-rate while at high SNRs, beamwidth is the only important parameter. Second, for very short channel coherence times and high SNRs the optimization is not required and predefined fixed values can be used instead. Third, a bottleneck for achieving high sum-rates is a small number of antennas, which results in a low resolution beamwidth, especially at large coherence time and low

SNRs.

The rest of the chapter is organized as follows. In Section 4.2, the system model is described. Section 4.3 formulates the optimization problem. In Section 4.4, the allocated power and beamwidth are determined through the proposed optimization algorithm and its convergence analysis is provided. Numerical results are presented in Section 4.5. Section 4.6 concludes this chapter.

4.2 System Model

We assume a narrow band mmWave downlink system composed of a single cell with a base station (BS) and M_{UE} user devices. The BS is equipped with N_{RF} RF chains and N_{BS} antennas whereas each user has one RF chain and N_{UE} antennas. Each RF chain is connected to the antennas through phase-shifters. The architecture of the BS and a typical user is shown in Fig. 4.1. Due to the hybrid beamforming structure at the BS, the number of antennas is larger than the number of RF chains, $N_{\mathrm{BS}} > N_{\mathrm{RF}}$, and due to the analog beamforming at the users, we have $N_{\mathrm{UE}} > 1$. Further, the BS transmits N_s streams simultaneously by steering N_{B} beams toward the users. To implement hybrid beamforming, the condition $N_{\mathrm{B}} \leq N_{\mathrm{RF}}$ should be satisfied. In this chapter, however, we assume $N_{\mathrm{B}} = N_{\mathrm{RF}}$ to reduce the complexity and cost of the system. Indeed, if we consider sending one stream via one beam, it results in $N_s = N_{\mathrm{RF}}$. On the other hand, to establish better connectivity by increasing the number of simultaneously served users in a dense area and further improve spectral efficiency, we use NOMA in the proposed mmWave hybrid beamforming network. Hence, each beam can serve more than one user. That is, the transmitter simultaneously sends N_{RF} streams toward $M_{\mathrm{UE}} > N_{\mathrm{RF}}$ users which are grouped into N_{RF} clusters, i.e., $M_{\mathrm{UE}} = \sum_{n=1}^{N_{\mathrm{RF}}} K_n$, where K_n denotes the number of users in the nth cluster. Note that NOMA requires the number of users in each cluster to be more than one, which should be satisfied by $K_n > 1$. Hereafter, the mth user equipment in the nth cluster is represented by $\mathrm{UE}_{n,m}$.

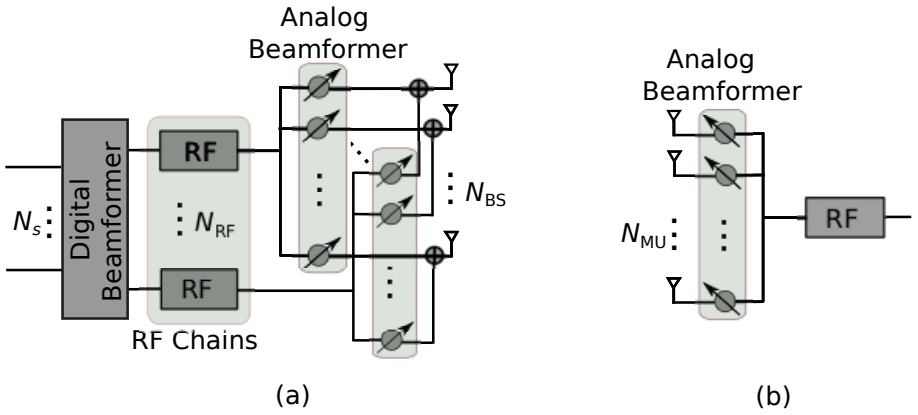

Figure 4.1: Schematic of (a) the BS with hybrid beamforming structure and (b) a user equipment with analog beamforming structure.

4.2.1 Channel Model

We use the widely adopted extended Saleh-Valenzuela model as a multi-path channel (MPC) model in our mmWave hybrid beamforming-NOMA system [12, 105]. In this model, each LoS and NLoS path is defined by a channel gain and an array steering vector at the transmitter and an array response vector at the receiver. Hence, the channel matrix between the BS and $\text{UE}_{n,m}$ in downlink is given by

$$\mathbf{H}_{n,m} = \frac{1}{\sqrt{L_{n,m}+1}} \left(\beta_{n,m,0}\mathbf{G}_{n,m,0} + \sum_{l=1}^{L_{n,m}} \beta_{n,m,l}\mathbf{G}_{n,m,l} \right), \tag{4.1}$$

where $\beta_{n,m,0}$ and $\beta_{n,m,l}$ denote the channel gain of LoS and NLoS channels, respectively. $\mathbf{G}_{n,m,0} \in \mathbb{C}^{N_{\text{UE}} \times N_{\text{BS}}}$ is the LoS channel matrix and $\mathbf{G}_{n,m,l}$ is the lth NLoS channel matrix. In particular, $\mathbf{G}_{n,m,l}$, $0 \leq l \leq L_{n,m}$, is given by

$$\mathbf{G}_{n,m,l} = \mathbf{a}_{\text{UE}}(\theta^{\text{az}}_{n,m,l}, \theta^{\text{el}}_{n,m,l})\mathbf{a}^{\dagger}_{\text{BS}}(\phi^{\text{az}}_{n,m,l}, \phi^{\text{el}}_{n,m,l}), \tag{4.2}$$

where $\theta^{\text{az}}_{n,m,l}$ ($\theta^{\text{el}}_{n,m,l}$) and $\phi^{\text{az}}_{n,m,l}$ ($\phi^{\text{el}}_{n,m,l}$) are normalized azimuth (elevation) AoA and angle of departure (AoD), respectively. Also, $\mathbf{a}_{\text{BS}} \in \mathbb{C}^{N_{\text{BS}} \times 1}$ and $\mathbf{a}_{\text{UE}} \in \mathbb{C}^{N_{\text{UE}} \times 1}$ are the antenna array steering vector and array response vector of the BS and $\text{UE}_{n,m}$, respectively. In mmWave outdoor commu-

nications, to further reduce the interference, sectorized BSs can be employed. Mostly, each sector in the azimuth domain is much wider than that of the elevation domain [62, 112]. Reasonably, we assume that the BS separates the clusters in the azimuth domain and considers fixed elevation angles. Further, we assume that the sector-level beamwidth for the BS is defined by ω_{BS} and for each user is defined by ω_{UE}. Hence, the BS implements only azimuth beamforming and neglects elevation beamforming. In this case, the antenna configuration is a uniform linear array (ULA) and the superscript "el" is dropped. For a ULA, the steering vector is defined as

$$\mathbf{a}_{\mathrm{BS}}(\phi_{n,m,l}) = \left[1, e^{-j\pi\phi_{n,m,l}}, \ldots, e^{-j\pi(N_{\mathrm{BS}}-1)\phi_{n,m,l}}\right]^T, \tag{4.3}$$

where $\phi_{n,m,l} \in [-1,1]$ is related to the AoD $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ as $\phi_{n,m,l} = \frac{2D\sin(\varphi)}{\lambda}$. Note that D denotes the antenna spacing and λ denotes the wavelength of the propagation. The antenna array response vector for $\mathbf{a}_{\mathrm{UE}}(\theta_{n,m,l})$ can be written in a similar fashion. AoD/AoA variations over the coherence time are trivial and can be ignored [140]. Let T and T_b denote the coherence time and the time duration over which AoD/AoA remain unchanged, respectively. In [140], it is shown that the coherence time duration is far less than T_b, i.e., $T \ll T_b$, which ensures that AoD/AoA do not change over the coherence time. In this chapter, the channel gain captures path loss and shadow fading effects. The assumption on AoD/AoA and the channel gain state that the channel model in (4.1) represents a long term channel which is widely adopted in the literature [109, 59, 140]. Ignoring AoD/AoA variations and instantaneous channel fluctuations are valid assumptions since the power allocation and the beamwidth control are done over the coherence time. This follows from the fact that the long term channel model can be effectively used in long term resource allocation [140].

It is demonstrated that in dense urban environments, with high probability, the mmWave channels contain only one or two paths, with the dominant one that carries most of the signal energy [5]. Therefore, with a single path assumption, the MPC model described in (4.1) is converted to a single

path channel model given by

$$\mathbf{H}_{n,m} = \beta_{n,m}\mathbf{a}_{\mathrm{UE}}(\theta_{n,m})\mathbf{a}^{\dagger}_{\mathrm{BS}}(\phi_{n,m}). \tag{4.4}$$

Hence, the BS communicates to the users through a single path channel. It is worth mentioning that the users are ordered based on their channel gain, i.e., $\beta_{n,1} \geq, \ldots, \geq \beta_{n,M}$ where $\beta_{n,m}$ is captured through channel quality indicator (CQI) [79]. Although it is assumed that the channel is single path, in some rare cases there might be more than one dominant path. To mitigate the multipath issue, rake receivers or orthogonal frequency-division multiplexing (OFDM) can be used. It should also be mentioned that due to the availability of large bandwidth, in mmWave systems, wide band transmission is preferred. For this case, the considered narrow band system should be combined with OFDM. In general, the extension of our narrow band system to the wide band is straightforward and studied in the literature. For instance, the OFDM-based NOMA has been considered in [103] and other similar work.

4.2.2 Beam-Training

Each transmission frame in mmWave directional communications depends on the channel coherence time and consists of two parts: (i) beam-training and (ii) data transmission as depicted in Fig. 4.2. At the first step, the channel parameters AoDs, AoAs, and effective channel are estimated by channel estimation algorithms. In this chapter, we assume that the channel parameters are perfectly estimated [109, 49]. In particular, the estimation of AoDs and AoAs is performed using beam alignment algorithms and takes much more time compared to the effective channel estimation. The beam alignment algorithms should be fast, accurate, and energy-efficient. At the second step, during the remaining time, the data is transmitted. Recently, a few codebook-based beam-training algorithms have been proposed for mmWave hybrid beamforming systems [57, 7, 92, 145, 79]. Even the current fastest algorithms take a considerable portion of the coherence time that leaves a

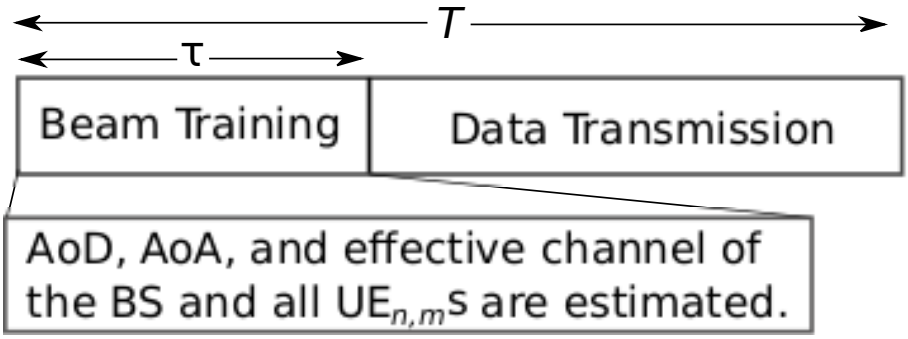

Figure 4.2: Schematic of the transmission frame in the mmWave-NOMA system.

short time for data transmission and can diminish the achievable rate of a user [24]. On the other hand, a smaller beam-training duration means wider beamwidth, which supplies lower beamforming gain. Consequently, in the mmWave systems, there exists a trade-off between the training duration and data transmission duration. This trade-off becomes more notable in the mmWave-NOMA networks in which more channels should be estimated. Motivated by this, finding an optimal beam-training duration and user power allocation for data transmission to increase the sum-rate of the mmWave-NOMA system will be the subject of this chapter.

As mentioned before, the main part of a beam-training algorithm consists of beam alignment followed by an effective channel estimation. In general, there are two different search algorithms for beam alignment, exhaustive search [57] and hierarchical search [7]. The former algorithm examines all beam pairs in the codebook for BS and UE and determines the best pair that maximizes the beamforming gain. The training time for this algorithm is proportional to the size of the beam's search space which is given by

$$\tau = \left\lceil \frac{\omega_{\mathrm{BS}}}{\eta} \right\rceil \left\lceil \frac{\omega_{\mathrm{UE}}}{\mu} \right\rceil T_p, \tag{4.5}$$

where η and μ denote the beamwidths of the BS and the UE, respectively. Further, T_p is the time for pilot transmission. On the other hand, the hierarchical search algorithm is designed based on multi-level codebook designs and uses bisection beam search. At the first level, the algorithm chooses a wider beam with a low resolution which has a small beam search space. The algo-

rithm refines the search iteratively using the next-level codebook within the subspace defined in the wider-level. At each level, the algorithm performs an exhaustive search to find the best pair. Compared to the exhaustive search, the hierarchical algorithm takes less training time with the same beam resolution and length of the pilot sequence at the cost of the higher probability of misalignment [79].

The exhaustive and hierarchical algorithms are designed only for single-user and multi-user scenarios. In particular, multi-user beam alignment algorithms assume that each user has a distinct AoD and might not be efficient for NOMA systems in which users are allowed to have the same AoD. Particularly, the hierarchical algorithm has a higher probability of beam misalignment at the low SNR regime [79]. This can be a major barrier in realizing the hierarchical algorithm in mmWave-NOMA networks, where the users with low SNR are paired with the users with high SNR. It seems that the exhaustive algorithm is a proper candidate for the beam alignment in the mmWave-NOMA system since it works better at the low SNR regime [79]. In the exhaustive search algorithm, all beams are aligned with the same resolution. That is, the beamwidth of the beams at the BS is equal. In some scenarios, this may impose a limitation on designing an optimal mmWave-NOMA system. To overcome this issue, we adopt a multi-user tone-based beam-training algorithm proposed in [145]. The algorithm consists of three steps summarized as follows. At the first step, each user transmits a pilot using one omni-directional antenna with a unique frequency tone in the uplink. Given a predefined resolution $\frac{\omega_{\text{BS}}}{\eta}$ for each user, the BS searches for the best AoD that maximizes the beamforming gain. It is worth mentioning that the BS can estimate the AoDs with different predefined resolutions. At the second step, using the estimated AoDs, the BS simultaneously transmits a pilot for each user over a unique frequency tone in the downlink. Each user estimates the AoA with a predefined resolution $\frac{\omega_{\text{UE}}}{\mu}$. Finally, each user transmits an orthogonal pilot sequence to the BS, and the BS estimates the channel. We note that using a unique tone for each user requires more hardware complexity compared to the search algorithm. There are two main differences between the tone-based algorithm and the exhaustive search algorithm. First, due to using unique frequency tones, the beam alignment for each user is done indepen-

dently. Hence, the BS can select different beamwidth values for different users and each user can also have a distinct beamwidth value. Second, the beam-training time is shorter than those of the exhaustive search algorithm. That is, the total training time for the tone-based algorithm is $\tau = \max\{(\lceil \frac{\omega_{\text{BS}}}{\eta_n} \rceil + \lceil \frac{\omega_{\text{UE}}}{\mu_{n,m}} \rceil) T_p\}$, where η_n and $\mu_{n,m}$ denote the beamwidth of the nth beam of the BS and the $\text{UE}_{n,m}$, respectively. It is clear when η_n and $\mu_{n,m}$ are the same as those of the exhaustive search algorithm, the training time for the tone-based algorithm is smaller than (4.5).

Although the algorithm in [145] is applicable to the mmWave-NOMA structure and can remarkably reduce the training time, similar to the hierarchical algorithm, it may result in a higher probability of misalignment. This is due to the use of omni-directional antennas at the first step which does not provide enough beamforming gain, especially for low-SNR users. To tackle this challenge, we modify the algorithm at the cost of sacrificing the speed of beam-training. We assume that each user steers directional beams with the predefined beamwidth μ. Then, we combine the first and second steps and perform an exhaustive beam search to find the best beam pair that achieves the highest beamforming gain. Note that the BS communicates with each user via a unique frequency tone. Further, the third step remains unchanged. Therefore, the training time becomes $\tau = \max\{\lceil \frac{\omega_{\text{BS}}}{\eta_n} \rceil \lceil \frac{\omega_{\text{UE}}}{\mu_{n,m}} \rceil T_p\}$. When beamwidth for the BS and users are the same as those of (4.5), the beam-training time for the modified tone-based algorithm is similar to that of the exhaustive search algorithm. In summary, we adopt the tone-based beam alignment algorithm in [145] and instead of hierarchical search we use exhaustive search.

4.2.3 Data Transmission

In mmWave-NOMA systems, during the data transmission, the transmit symbols are superposition coded at the BS. Then, at the user side, unintended symbols are removed via SIC. More details on these two processes are provided as follows. Let $\mathbf{s} \in \mathbb{C}^{N_{\text{RF}} \times 1}$ denote the information signal vector such that its nth element s_n satisfies $\mathbb{E}[s_n s_n^*] = \frac{1}{N_{\text{RF}}}$ for $n = 1, 2, \ldots, N_{\text{RF}}$. At the baseband of

the BS, the superposition coded signal of the nth stream is given by $s_n = \sum_{m=1}^{K_n} \sqrt{P_{n,m}} z_{n,m}$ where $P_{n,m}$ and $z_{n,m}$ are the allocated power and transmit symbol for the mth user in the nth cluster, respectively. Then, the hybrid beamforming is done in digital and analog precoding stages. The BS applies the digital precoder $\mathbf{F}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{RF}} \times N_{\mathrm{RF}}}$ using RF chains, and then applies the analog precoder $\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{N_{\mathrm{BS}} \times N_{\mathrm{RF}}}$ using phase-shifters. Thus, the transmit signal vector after superposition coding and beamforming, $\mathbf{x} \in \mathbb{C}^{N_{\mathrm{BS}} \times 1}$, is expressed as

$$\mathbf{x}^T = \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{s}^T. \tag{4.6}$$

Each element of all beamforming vectors has a constant magnitude of $\frac{1}{\sqrt{N_{\mathrm{BS}}}}$. Further, the total power of the hybrid beamforming is constrained to $\left\| \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \right\|_F^2 = N_{\mathrm{RF}}$. On the other hand, the received signal by $\mathrm{UE}_{n,m}$, $\mathbf{r}_{n,m} \in \mathbb{C}^{N_{\mathrm{UE}} \times 1}$, is given by

$$\mathbf{r}_{n,m} = \mathbf{H}_{n,m} \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{s} + \mathbf{n}_{n,m}, \tag{4.7}$$

where $\mathbf{n}_{n,m} \in \mathbb{C}^{N_{\mathrm{UE}} \times 1}$ is the additive white Gaussian noise vector with zero-mean and σ^2 variance for each element, i.e., $\mathscr{CN}(0, \sigma^2)$. Then, the received vector at $\mathrm{UE}_{n,m}$ followed by the analog combiner $\mathbf{w}_{n,m} \in \mathbb{C}^{N_{\mathrm{UE}} \times 1}$ is obtained as

$$\begin{aligned} y_{n,m} = & \underbrace{\sqrt{P_{n,m}} \mathbf{w}_{n,m}^{\dagger} \mathbf{H}_{n,m} \mathbf{F}_{\mathrm{RF}} \mathbf{f}_{\mathrm{BB}}^{n} z_{n,m}}_{\text{desired signal}} \\ & + \underbrace{\sum_{k \neq m}^{K_n} \sqrt{P_{n,k}} \mathbf{w}_{n,m}^{\dagger} \mathbf{H}_{n,m} \mathbf{F}_{\mathrm{RF}} \mathbf{f}_{\mathrm{BB}}^{n} z_{n,k}}_{\text{intra-cluster interference}} \\ & + \underbrace{\sum_{q \neq n}^{N_{\mathrm{RF}}} \sum_{\ell=1}^{K_q} \sqrt{P_{q,\ell}} \mathbf{w}_{n,m}^{\dagger} \mathbf{H}_{n,m} \mathbf{F}_{\mathrm{RF}} \mathbf{f}_{\mathrm{BB}}^{\ell} z_{q,\ell}}_{\text{inter-cluster interference}} + \underbrace{\mathbf{w}_{n,m}^{\dagger} \mathbf{n}_{n,m}}_{\text{noise}}. \end{aligned} \tag{4.8}$$

Each user decodes the intended signal by using SIC. As such, after applying SIC, the received signal at $\mathrm{UE}_{n,1}$ is given by

$$\begin{aligned} y_{n,1} =& \sqrt{P_{n,1}}\mathbf{w}_{n,1}^{\dagger}\mathbf{H}_{n,1}\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}^{n}z_{n,1} \\ &+ \sum_{q\neq n}^{N_{\mathrm{RF}}}\sum_{\ell=1}^{K_q}\sqrt{P_{q,\ell}}\mathbf{w}_{n,1}^{\dagger}\mathbf{H}_{n,1}\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}^{\ell}z_{q,\ell} + \mathbf{w}_{n,1}^{\dagger}\mathbf{n}_{n,1}, \end{aligned} \tag{4.9}$$

and the received signal at $\mathrm{UE}_{n,m}$, for $m > 1$, is given by

$$\begin{aligned} y_{n,m} =& \sqrt{P_{n,m}}\mathbf{w}_{n,m}^{\dagger}\mathbf{H}_{n,m}\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}^{n}z_{n,m} \\ &+ \underbrace{\sum_{k=1}^{m-1}\sqrt{P_{n,k}}\mathbf{w}_{n,m}^{\dagger}\mathbf{H}_{n,m}\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}^{n}z_{n,k}}_{\text{residual intra-cluster interference}} \\ &+ \sum_{q\neq n}^{N_{\mathrm{RF}}}\sum_{\ell=1}^{K_q}\sqrt{P_{q,\ell}}\mathbf{w}_{n,m}^{\dagger}\mathbf{H}_{n,m}\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}^{\ell}z_{q,\ell} + \mathbf{w}_{n,m}^{\dagger}\mathbf{n}_{n,m}. \end{aligned} \tag{4.10}$$

One can observe that the desired signal of the first user in (4.9) is contaminated by the inter-cluster interference and noise, whereas the desired signal of the other users represented by (4.10) is contaminated by the residual intra-cluster and inter-cluster interference and noise.

4.2.4 Clustering

In this section, we describe a simple, yet effective clustering method for two NOMA users per cluster which is the case in our system model. The reason for choosing two users per cluster will be explained in the next sections. Before proceeding, we define cluster-head and far UE terms. In a cluster, we call the closer UE to the BS the cluster-head and the other UE the far UE. A clustering algorithm for two NOMA users per cluster, mainly designed based on the following two key points, has been proposed in [6, Algorithm 1]: (i) A key point to maximize the sum-rate in NOMA is to ensure that the high channel gain users are selected as the cluster-heads. (ii) The channel gain difference between the cluster-head and the far UE should be sufficiently high.

Before applying the clustering method in [6], we select $2N_{\mathrm{RF}}$ users and divide them into two groups. The first group consists of the N_{RF} users with the highest channel gains denoted by $\mathrm{UE}_{n,1}$ for $n = 1, 2, \cdots, N_{\mathrm{RF}}$. The second group includes the remaining N_{RF} users denoted by $\mathrm{UE}_{n,2}$ for $n = 1, 2, \cdots, N_{\mathrm{RF}}$. Further, the users of the first group are called the cluster-heads and the users of the second group are called the far users. The following conditions result in good performance.

Condition 1: The cluster-heads are located in distinctive directions.

Condition 2: The far users $\mathrm{UE}_{n,2}$ have the lowest channel gains among all the users and are paired with $\mathrm{UE}_{1,1}$, $\mathrm{UE}_{2,1}$, $\cdots$, $\mathrm{UE}_{N_{\mathrm{RF}},1}$, respectively.

Then, we use the clustering algorithm proposed in [6] in our mmWave-NOMA network. To make sure that Conditions 1 and 2 hold, we replace the users that violate them. Since the probability of existing high channel gain users in mmWave cells is almost one, new cluster-heads that will not violate Condition 1 are always available. Since the sum-rate is mainly determined by the channel gain of the cluster-heads, replacing the users that violate Condition 2 will not affect the sum-rate dramatically. Therefore, to ease the calculations, for the rest of the chapter, it is assumed that $\mathrm{UE}_{n,1}$ and $\mathrm{UE}_{n,2}$ are clustered together.

4.2.5 Hybrid Beamforming Gain and SINR

After the clustering is performed, an efficient beamforming is used to reduce/eliminate the inter-cluster interference. We use the zero-forcing beamforming (ZFBF) method which is widely adopted in the literature [3, 123, 6, 22, 107]. This method is low-complex and highly efficient. In fact, it is shown that when the channels of the users inside a cluster are highly correlated, ZFBF can significantly suppress the inter-cluster interference. In ideal cases, i.e., a perfect correlation, ZFBF is able to completely eliminate the inter-cluster interference. First, we describe an ideal beamforming gain which is the same as that of an ideal ZFBF. Then, to take the practical issues into account, we describe a non-ideal beamforming gain which reflects the impact of the imperfect channel corre-

lation in ZFBF. We note that when the channels between the users are not highly correlated, the singular value decomposition (SVD) method is used to design the beamforming matrix [123].

Let us define $\mathbf{f}_n = \mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}^n$ as the hybrid beamforming vector of the nth beam at the BS. An ideal hybrid beamformer leads to $|\mathbf{a}_{\mathrm{BS}}^{\dagger}(\phi_{n,m})\mathbf{f}_n| = \sqrt{G_{\mathrm{BS}}^{\mathrm{id}}(\phi'_{n,m},\eta_n)}$ in which $G_{\mathrm{BS}}^{\mathrm{id}}$ is the beamforming gain of the ideal beamformer at the BS and $\phi'_{n,m}$ and η_n denote the angle off the boresight direction at the BS and the beamwidth of the nth beam, respectively. It is worth mentioning that in this chapter the parameter $G_{\mathrm{BS}}^{\mathrm{id}}$ is irrespective of how the hybrid beamforming is designed. Essentially, the value of $G_{\mathrm{BS}}^{\mathrm{id}}$ depends on the beamformer $\mathbf{f}_n$, where $||\mathbf{f}_n|| = 1$, and the size of the transmit antenna array. Also, note that an ideal beamforming vector is obtained when there is no channel estimation error and perfect beam alignment is done while considering an infinite resolution for the phase-shifters. Further, the beamwidth depends on the design of the analog beamformer and the digital beamformer. In particular, the beamforming gain is defined as

$$G_{\mathrm{BS}}^{\mathrm{id}}(\phi'_{n,m},\eta_n) = \begin{cases} \frac{2\pi}{\eta_n}, & \text{if } |\phi'_{n,m}| \leq \frac{\eta_n}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{4.11}$$

Further, the beamforming gain of the ideal analog beamformer at $\mathrm{UE}_{n,m}$ is assumed to be $|\mathbf{w}_{n,m}^{\dagger}\mathbf{a}_{\mathrm{UE}}(\theta_{n,m})| = \sqrt{G_{\mathrm{UE}}^{\mathrm{id}}(\theta'_{n,m},\mu_{n,m})}$ in which $G_{\mathrm{UE}}^{\mathrm{id}}$ is the gain of the ideal analog beamformer and $\theta'_{n,m}$ and $\mu_{n,m}$ respectively denote the angle off the boresight direction at $\mathrm{UE}_{n,m}$ and the beamwidth of $\mathrm{UE}_{n,m}$. Similar to $G_{\mathrm{BS}}^{\mathrm{id}}$, the ideal beamforming gain is defined as

$$G_{\mathrm{UE}}^{\mathrm{id}}(\theta'_{n,m},\mu_{n,m}) = \begin{cases} \frac{2\pi}{\mu_{n,m}}, & \text{if } |\theta'_{n,m}| \leq \frac{\mu_{n,m}}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{4.12}$$

Note that ideal beamforming at the BS and users results in the complete cancellation of the inter-cluster interference represented in (4.8).

In practice, achieving the ideal beamforming gain may not be possible because of the application

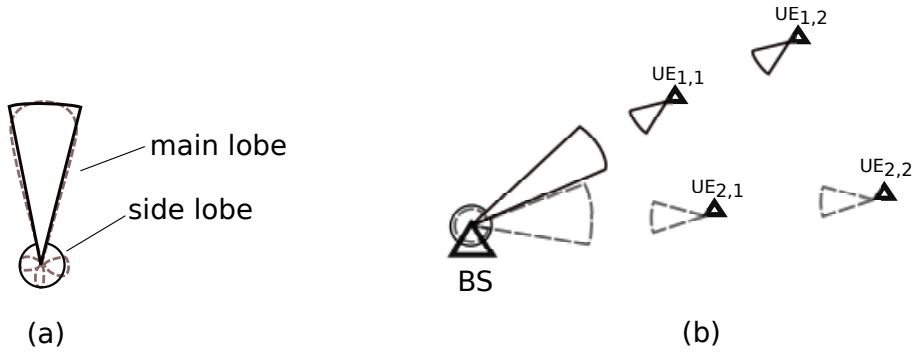

Figure 4.3: (a) A non-ideal beam is modeled with a constant main lobe gain and side lobe gain, (b) The impact of the side lobe gain of each beam on the UEs located in the other cluster.

of suboptimal solutions for the analog beamformer, finite resolution of the phase-shifters, channel estimation error, and beam misalignment. These problems reduce the gain in the main lobe and introduce a side lobe for each beam. Hence, the beamforming model should take these issues into account [7, 147]. A more practical model for the beamforming gain of the nth beam is given by

$$G_{\mathrm{BS}}(\phi'_{n,m},\eta_n)=\begin{cases}\frac{2\pi-(2\pi-\eta_n)\xi}{\eta_n}, & \text{if } |\phi'_{n,m}|\leq\frac{\eta_n}{2},\\ \xi, & \text{otherwise,}\end{cases} \tag{4.13}$$

where $0\leq\xi<1$ with $\xi\ll 1$ for narrow beams, which is widely adopted in the literature [92, 95, 127]. Note that there is another common model for the beamforming gain with side lobe level varying with the beamwidth [147]. In this chapter, to make the analysis tractable, we use the model described above that satisfies the total power of 1, i.e., $\int_0^{2\pi} G_{\mathrm{BS}}(\phi'_{n,m},\eta_n)\mathrm{d}\phi'_{n,m}=\frac{\eta_n}{2\pi}\frac{2\pi-(2\pi-\eta_n)\xi}{\eta_n}+\frac{2\pi-\eta_n}{2\pi}\xi=1$. Similarly, the model for the beamforming gain of $\mathrm{UE}_{n,m}$ is given by

$$G_{\mathrm{UE}}(\theta'_{n,m},\mu_{n,m})=\begin{cases}\frac{2\pi-(2\pi-\mu_{n,m})\xi}{\mu_{n,m}}, & \text{if } |\theta'_{n,m}|\leq\frac{\mu_{n,m}}{2},\\ \xi, & \text{otherwise.}\end{cases} \tag{4.14}$$

In the above equations, the main lobe's gain is distributed uniformly in the entire beamwidth and the side lobe's gain is assumed to be constant [145, 90] as demonstrated in Fig. 4.3.(a). However,

$$\gamma_{n,m} = \frac{P_{n,m}\beta_{n,m}^2 G_{\mathrm{BS}}(\phi'_{n,m},\eta_n) G_{\mathrm{UE}}(\theta'_{n,m},\mu_{n,m})}{\sum_{k=1}^{m-1} P_{n,k}\beta_{n,m}^2 G_{\mathrm{BS}}(\phi'_{n,m},\eta_n) G_{\mathrm{UE}}(\theta'_{n,m},\mu_{n,m}) + \sum_{q\neq n}^{N_{\mathrm{RF}}}\sum_{\ell=1}^{K_q} P_{q,\ell}\beta_{n,m}^2 G_{\mathrm{UE}}(\phi'_{n,m},\mu_{n,m})\xi + \sigma^2}. \quad (4.16)$$

in reality, the main lobe's gain changes over the beamwidth and the side lobe's gain depends on the size of the beamwidth. For example, for a narrower beam, the side lobe's gain is higher [13]. Further, the side lobe results in interference that impacts the UEs located in other clusters as shown in Fig. 4.3.(b). In our formulation, this interference is modeled by the inter-cluster interference term in (4.8).

Hence, using (4.9), (4.13), and (4.14), the signal-to-interference-plus-noise ratio (SINR) of $\mathrm{UE}_{n,1}$ in the nth beam is expressed as

$$\gamma_{n,1} = \frac{P_{n,1}\beta_{n,1}^2 G_{\mathrm{BS}}(\phi'_{n,1},\eta_n) G_{\mathrm{UE}}(\theta'_{n,1},\mu_{n,1})}{\sum_{q\neq n}^{N_{\mathrm{RF}}}\sum_{\ell=1}^{K_q} P_{q,\ell}\beta_{n,1}^2 G_{\mathrm{UE}}(\theta'_{n,1},\mu_{n,1})\xi + \sigma^2}, \quad (4.15)$$

and, using (4.10), (4.13), and (4.14), the SINR of $\mathrm{UE}_{n,m}$, $m > 1$, is given by (4.16) on the top of next page.

Due to the single cell assumption, we can conclude that the users do not receive any interference from the side lobe and only receive signal from the main lobe. Further, it is assumed that codebooks for a specific level (beam resolution) are designed efficiently such that the steered beams by the BS do not overlap [57]. Hence, each user receives the desired signal and intra-cluster interference sent through the main lobe of the desired beam and the inter-cluster interference sent through the side lobe of the other beams as visualized in Fig. 4.3.(b). Further, in the modified beam-training algorithm described in Section 4.2.2 and adopted in this section, the beams directed by the BS can have different beamwidth values, i.e., the beams have different resolutions. In this case, during the data transmission, beam overlap may occur, which can impose severe inter-cluster interference.

To avoid this, we assume that there is a proper angle gap between the two neighboring beams. In mmWave hybrid beamforming, a limited number of RF chains is used, i.e., the number of beams is limited [3]. Therefore, the direction of each beam is selected to satisfy the required angle gap between the beams. Further, each cluster's users are served via a common beam directed by the BS. Therefore, the training time for $\mathrm{UE}_{n,m}$ is given by $\tau_{n,m} = \lceil \frac{\omega_{\mathrm{BS}}}{\eta_n} \rceil \lceil \frac{\omega_{\mathrm{UE}}}{\mu_{n,m}} \rceil T_p$ as explained in Section 4.2.2. Accordingly, the achievable rate for $\mathrm{UE}_{n,m}$ can be calculated as

$$R_{n,m} = \left(1 - \frac{\tau}{T}\right) \log_2(1 + \gamma_{n,m}), \tag{4.17}$$

where T denotes the channel coherence time as indicated in Fig. 4.2. It is worth mentioning that the chosen frame duration is smaller than the channel coherence time.

4.3 Problem Formulation

Here, NOMA is performed for two UEs per cluster which is compatible with the multi-user superposition transmission schemes recently adopted by 3GPP [23, 2]. Further, the BS is assumed to generate only two beams. Extension to more than two clusters will be addressed in future work.

To optimize the sum-rate performance, η_n, $\mu_{n,m}$, and $P_{n,m}$ should be optimized according to

$$\underset{\boldsymbol{\eta},\boldsymbol{\mu},\mathbf{P}}{\text{maximize}} \quad \sum_{n=1}^{2}\sum_{m=1}^{2} R_{n,m} \tag{4.18a}$$

$$\text{subject to} \quad \eta_{\min} \leq \eta_n \leq \omega_{\text{BS}}, \tag{4.18b}$$

$$\mu_{\min} \leq \mu_{n,m} \leq \min\{\omega_{\text{UE}}, \eta_n\}, \tag{4.18c}$$

$$\tau \leq T, \tag{4.18d}$$

$$\tau \geq \frac{\omega_{\text{BS}}}{\eta_n}\frac{\omega_{\text{UE}}}{\mu_{n,m}} T_p, \tag{4.18e}$$

$$R_{n,m} \geq R_{\min}, \tag{4.18f}$$

$$\sum_{n=1}^{2}\sum_{m=1}^{2} P_{n,m} \leq P_{\text{tot}}, \tag{4.18g}$$

$$P_{n,m} > 0, \tag{4.18h}$$

where $\boldsymbol{\eta} = [\eta_1, \eta_2]$, $\boldsymbol{\mu} = [\mu_{1,1}, \mu_{1,2}, \mu_{2,1}, \mu_{2,2}]$, $\mathbf{P} = [P_{1,1}, P_{1,2}, P_{2,1}, P_{2,2}]$, and P_{tot} denotes the total power of the BS. The smallest beamwidth resolutions for the BS and UE are denoted by $\eta_{\min}$ and $\mu_{\min}$, respectively. Here, we assume $\eta_{\min} = \mu_{\min}$. The beamwidth resolution relates to the number of antennas. Usually, the number of antennas at a BS is larger than those of UEs. Thus, the BS can generate narrower beams. However, we assume that the minimum beamwidths of the BS and UEs are identical. For the sake of simplicity, we relax $\tau \geq \lceil \frac{\omega_{\text{BS}}}{\eta_n} \rceil \lceil \frac{\omega_{\text{UE}}}{\mu_{n,m}} \rceil T_p$ to $\tau \geq \frac{\omega_{\text{BS}}}{\eta_n}\frac{\omega_{\text{UE}}}{\mu_{n,m}} T_p$ in (4.18e). After we obtain the optimal $\boldsymbol{\eta}$ and $\boldsymbol{\mu}$, we can recalculate $\tau = \max\{\lceil \frac{\omega_{\text{BS}}}{\eta_n} \rceil \lceil \frac{\omega_{\text{UE}}}{\mu_{n,m}} \rceil T_p\}$.

4.4 Joint beamwidth control and power allocation

Problem (4.18) is an intractable non-convex optimization problem and needs to be decomposed. We propose an algorithm which iterates between the power allocation and the beamwidth optimization. When allocating the power, we fix the beamwidths and when optimizing the beamwidths, we keep the powers fixed. We assume that the BS and the users are aligned after the training pro-

cess, which means $|\phi'_{n,m}| \leq \frac{\eta_n}{2}$ and $|\theta'_{n,m}| \leq \frac{\mu_n}{2}$. We also assume the users within the same cluster have the same beamwidth, i.e., $\mu_{n,1} = \mu_{n,2} = \mu_n$. However, the users in different clusters do not necessarily have the same beamwidth.

4.4.1 Power allocation

When the beamwidth and the training time are fixed, the beamforming gains are also fixed. Then, problem (4.18) is simplified to:

$$\underset{\mathbf{P}}{\text{maximize}} \quad \sum_{n=1}^{2}\sum_{m=1}^{2} R_{n,m} \tag{4.19a}$$

$$\text{subject to} \quad (4.18f)-(4.18h). \tag{4.19b}$$

Although Problem (4.19) has been greatly simplified compared to Problem (4.18), its objective is still complicated and non-convex. To transform Problem (4.19) into a tractable form, we use the log-exponential reformation idea in [77]. Introducing slack variables $\boldsymbol{S} = [x_{n,m}, d_{n,m}], n = 1,2, m = 1,2$, we can transform the objective in Problem (4.19) into a linear form by $\sum_{n=1}^{2}\sum_{m=1}^{2} \log_2 \frac{2^{x_{n,m}}}{2^{d_{n,m}}}$. For the sake of brevity, we denote $G_{\text{BS}}(\eta_n)G_{\text{UE}}(\mu_n)$ by G_n and $G_{\text{UE}}(\mu_n)$ by G^n_{UE} and make the following definitions:

$$\text{SI}_{n,1} \triangleq P_{n,1}\beta^2_{n,1}G_n + \sum_{\ell=1,q\neq n}^{2} P_{q,\ell}\beta^2_{n,1}G^n_{\text{UE}}\xi + \sigma^2, \tag{4.20}$$

$$\text{SI}_{n,2} \triangleq \sum_{\ell=1}^{2} P_{n,\ell}\beta^2_{n,2}G_n + \sum_{\ell=1,q\neq n}^{2} P_{q,\ell}\beta^2_{n,2}G^n_{\text{UE}}\xi + \sigma^2, \tag{4.21}$$

$$\mathrm{I}_{n,1} \triangleq \sum_{\ell=1,q\neq n}^{2} P_{q,\ell}\beta_{n,1}^2 G_{\mathrm{UE}}^n \xi + \sigma^2, \tag{4.22}$$

$$\mathrm{I}_{n,2} \triangleq P_{n,1}\beta_{n,2}^2 G_n + \sum_{\ell=1,q\neq n}^{2} P_{q,\ell}\beta_{n,2}^2 G_{\mathrm{UE}}^n \xi + \sigma^2. \tag{4.23}$$

Then, Problem (4.19) can be rewritten as

$$\begin{aligned}
\underset{\boldsymbol{S},\mathbf{P}}{\text{maximize}} \quad & (1-\frac{\tau}{T})\sum_{n=1}^{2}\sum_{m=1}^{2}(x_{n,m}-d_{n,m}) && \text{(4.24a)}\\
\text{subject to} \quad & 2^{x_{n,1}} \leq \mathrm{SI}_{n,1}, && \text{(4.24b)}\\
& 2^{x_{n,2}} \leq \mathrm{SI}_{n,2}, && \text{(4.24c)}\\
& 2^{d_{n,1}} \geq \mathrm{I}_{n,1}, && \text{(4.24d)}\\
& 2^{d_{n,2}} \geq \mathrm{I}_{n,2}, && \text{(4.24e)}\\
& (1-\frac{\tau}{T})(x_{n,m}-d_{n,m}) \geq R_{\min}, && \text{(4.24f)}\\
& (4.18g)-(4.18h). && \text{(4.24g)}
\end{aligned}$$

In Problem (4.24), the optimum is achieved when the constraints (4.24b)-(4.24e) satisfy with equality. Let us use (4.24b) as an example to show that the equality should be satisfied at the optimum. Assuming the opposite, we can increase $x_{n,1}$ while keeping other variables fixed. This results in increasing the cost function and contradicts the optimality assumption. Since constraints (4.24b)-(4.24e) achieve equality at the optimum, the non-convex objective of Problem (4.19) is equivalently decomposed into (4.24a) and constraints (4.24b)-(4.24e).

Unfortunately, constraints (4.24d) and (4.24e) are still non-convex. To relax the non-convex constraint to convex constraints, we use a sequential parametric convex approximation method (SPCA)

[14]. In this method, the non-convex feasible set is sequentially approximated by an inner convex approximation. Using (4.24d) as an example, at Iteration k, since function $2^{d_{n,1}}$ is a convex function, i.e., $2^y - 2^x \geq 2^x \log 2(y-x)$, we have a lower bound of $2^{d_{n,1}}$ as:

$$2^{d^*_{n,1}[k-1]} \log 2(d_{n,1} - d^*_{n,1}[k-1]) + 2^{d^*_{n,1}[k-1]} \leq 2^{d_{n,1}}, \tag{4.25}$$

where $d^*_{n,1}[k-1]$ is the optimal solution at Iteration $k-1$. Based on (4.25), we can relax (4.24d) into a convex constraint as

$$2^{d^*_{n,1}[k-1]} \log 2(d_{n,1} - d^*_{n,1}[k-1]) + 2^{d^*_{n,1}[k-1]} \geq \sum_{\ell=1, q\neq n}^{2} P_{q,\ell} \beta^2_{n,1} G^n_{\mathrm{UE}} \xi + \sigma^2. \tag{4.26}$$

Using the same method for (4.24e), we have

$$2^{d^*_{n,2}[k-1]} \log 2(d_{n,2} - d^*_{n,2}[k-1]) + 2^{d^*_{n,2}[k-1]} \geq P_{n,1} \beta^2_{n,2} G_n + \sum_{\ell=1, q\neq n}^{2} P_{q,\ell} \beta^2_{n,2} G^n_{\mathrm{UE}} \xi + \sigma^2. \tag{4.27}$$

At each iteration, we can relax Problem (4.24) into the following convex problem:

$$\underset{\boldsymbol{S}, \mathbf{P}}{\text{maximize}} \quad (1 - \frac{\tau}{T}) \sum_{n=1}^{2} \sum_{m=1}^{2} (x_{n,m} - d_{n,m}) \tag{4.28a}$$

$$\text{subject to} \quad (4.24b), \tag{4.28b}$$

$$(4.24c), \tag{4.28c}$$

$$(4.26) - (4.27), \tag{4.28d}$$

$$(4.18g) - (4.18h). \tag{4.28e}$$

This is a convex problem, which can be efficiently solved by off-the-shelf solutions, such as CVX [41].

By relaxing Problem (4.24) to Problem (4.28) in each iteration, we can propose an iterative algorithm to provide an approximation solution for Problem (4.24). Detailed steps are presented in Alg. 3. According to [14], Alg. 3 converges.

Algorithm 3: Power allocation

Set the sum-rate $R_{\text{sum}}[-1] \leftarrow 0$, the maximal iteration number $k_{max} \leftarrow 1000$ and the convergence threshold $\varepsilon \leftarrow 10^{-3}$;
repeat
  Choose a feasible start point $\mathbf{P}^*[0]$;
  $x^*_{n,m}[0] \leftarrow \log_2(P^*_{n,m}[0]\beta^2_{n,m}G_n)$, $d^*_{n,1}[0] \leftarrow \log_2(\sum^2_{\ell=1,q\neq n} P^*_{q,\ell}[0]\beta^2_{n,1}G^n_{\text{UE}}\xi + \sigma^2)$;
  $d^*_{n,2}[0] \leftarrow \log_2(P^*_{n,1}[0]\beta^2_{n,2}G_n + \sum^2_{\ell=1,q\neq n} P^*_{q,\ell}[0]\beta^2_{n,2}G^n_{\text{UE}}\xi + \sigma^2)$;
until (4.24f) is satisfied
$k \leftarrow 0$;
while $R_{\text{sum}}[k] - R_{\text{sum}}[k-1] \geq \varepsilon R_{\text{sum}}[k-1]$ and $k \leq k_{max}$ **do**
  $k \leftarrow k+1$;
  Solve (4.28) to obtain $\boldsymbol{S}^*[k]$ and $\mathbf{P}^*[k]$;
end while
Return $\mathbf{P}^*[k]$.

4.4.2 Beamwidth optimization

When the powers are fixed, the problem to optimize the beamwidth can be rewritten as

$$\underset{\boldsymbol{\eta},\boldsymbol{\mu}}{\text{maximize}} \quad \sum_{n=1}^{2}\sum_{m=1}^{2} R_{n,m} \tag{4.29a}$$

$$\text{subject to} \quad (4.18b)-(4.18f). \tag{4.29b}$$

Problem (4.29) is a very complicated problem with non-convex objective and constraints. To simplify the problem, we first perform the following variable substitutions:

$$\mu_n = \frac{A}{f_n}, \tag{4.30}$$

$$\eta_n = \frac{A}{h_n}, \tag{4.31}$$

where $f_n \triangleq G_{\mathrm{UE}}(\mu_n) - \xi$, $h_n \triangleq G_{\mathrm{BS}}(\eta_n) - \xi$ and $A = 2\pi - 2\pi\xi$ is a constant. By assuming $G_{\mathrm{BS}}(\eta_n)G_{\mathrm{UE}}(\mu_n) \approx \frac{(2\pi-2\pi\xi)^2}{\eta_n\mu_n} + \xi^2$, we have $\eta_n\mu_n = \frac{A^2}{h_n f_n}$ and $G_{\mathrm{BS}}(\eta_n)G_{\mathrm{UE}}(\mu_n) = h_n f_n + \xi^2$. Then, the SINR for $\mathrm{UE}_{n,1}$ and $\mathrm{UE}_{n,2}$ can be rewritten as

$$\gamma_{n,1} = \frac{P_{n,1}\beta_{n,1}^2(h_n f_n + \xi^2)}{\sum_{\ell=1,q\neq n}^{2} P_{q,\ell}\beta_{n,1}^2(f_n + \xi)\xi + \sigma^2}, \tag{4.32}$$

$$\gamma_{n,2} = \frac{P_{n,2}\beta_{n,2}^2(h_n f_n + \xi^2)}{P_{n,1}\beta_{n,2}^2(h_n f_n + \xi^2) + \sum_{\ell=1,q\neq n}^{2} P_{q,\ell}\beta_{n,2}^2(f_n + \xi)\xi + \sigma^2}. \tag{4.33}$$

Instead of finding the optimal beamwidth, we find the optimal f_n and h_n. We can rewrite Problem (4.29) as

$$\underset{\boldsymbol{h},\boldsymbol{f}}{\text{maximize}} \quad \sum_{n=1}^{2}\sum_{m=1}^{2} R_{n,m} \tag{4.34a}$$

$$\text{subject to} \quad \frac{A}{\omega_{\mathrm{BS}}} \leq h_n \leq \frac{A}{\eta_{\min}}, \tag{4.34b}$$

$$\frac{A}{\omega_{\mathrm{UE}}} \leq f_n \leq \frac{A}{\mu_{\min}}, \tag{4.34c}$$

$$h_n \leq f_n, \tag{4.34d}$$

$$R_{n,m} \geq R_{\min}, \tag{4.34e}$$

$$\tau \leq T, \tag{4.34f}$$

$$\tau = \max\{\frac{\omega_{\mathrm{BS}}\omega_{\mathrm{UE}}h_n f_n}{A^2}T_p\}, \tag{4.34g}$$

where $\boldsymbol{h} = [h_1, h_2]$ and $\boldsymbol{f} = [f_1, f_2]$. Problem (4.34) is still intractable with a non-convex objective.

To further decompose the problem, we will iterate between the two clusters, i.e., we first fix Cluster 2 to optimize the beamwidths in Cluster 1 and then fix Cluster 1 to optimize the beamwidths in Cluster 2.

4.4.3 Optimal beamwidth search for each cluster

Let us assume the beamwidths of Cluster 2 are fixed and take the optimization of the beamwidths in Cluster 1 as an example. The optimization for Cluster 2 is similar. In this case, the SINRs of $\mathrm{UE}_{2,1}$ and $\mathrm{UE}_{2,2}$ are fixed. We denote them by $\gamma_{2,1}^{\mathrm{fix}}$ and $\gamma_{2,2}^{\mathrm{fix}}$. We also denote the corresponding variables f_2 and h_2 as f_2^{fix} and h_2^{fix}. Then, the beamwidth optimization problem for Cluster 1 is as follows:

$$\underset{h_1, f_1}{\text{maximize}} \quad \sum_{m=1}^{2} R_{1,m} + (1-\frac{\tau}{T}) \sum_{m=1}^{2} \log_2(1+\gamma_{2,m}^{\mathrm{fix}}) \tag{4.35a}$$

$$\text{subject to} \quad \frac{A}{\omega_{\mathrm{BS}}} \le h_1 \le \frac{A}{\eta_{\min}}, \tag{4.35b}$$

$$\frac{A}{\omega_{\mathrm{UE}}} \le f_1 \le \frac{A}{\mu_{\min}}, \tag{4.35c}$$

$$h_1 \le f_1, \tag{4.35d}$$

$$R_{1,m} \ge R_{\min}, \tag{4.35e}$$

$$(1-\frac{\tau}{T}) \log_2(1+\gamma_{2,m}^{\mathrm{fix}}) \ge R_{\min}, \tag{4.35f}$$

$$\tau \le T, \tag{4.35g}$$

$$\tau \ge \frac{\omega_{\mathrm{BS}} \omega_{\mathrm{UE}} h_1 f_1 T_p}{A^2}, \tag{4.35h}$$

$$\tau \ge \frac{\omega_{\mathrm{BS}} \omega_{\mathrm{UE}} h_2^{\mathrm{fix}} f_2^{\mathrm{fix}} T_p}{A^2}. \tag{4.35i}$$

To simplify Problem (4.35), we discuss how to pick the optimal value for τ and remove it from the objective function. There are two cases for the optimal τ:

- Case 1: $h_1 f_1 < h_2^{\text{fix}} f_2^{\text{fix}}$. In this case, the τ should be set to $\tau^* = \frac{\omega_{\text{BS}} \omega_{\text{UE}} h_2^{\text{fix}} f_2^{\text{fix}}}{A^2} T_p$. Then, the objective function should be $(1 - \frac{\tau^*}{T}) \sum_{n=1}^{2} \sum_{m=1}^{2} \log_2(1 + \gamma_{n,m})$.

- Case 2: $h_1 f_1 \geq h_2^{\text{fix}} f_2^{\text{fix}}$. In this case, the τ should be set according to the value of $h_1 f_1$, which is $\tau = \frac{\omega_{\text{BS}} \omega_{\text{UE}} h_1 f_1}{A^2}$. Then, the objective function should be $(1 - \frac{\omega_{\text{BS}} \omega_{\text{UE}} h_1 f_1 T_p}{A^2 T}) \sum_{n=1}^{2} \sum_{m=1}^{2} \log_2(1 + \gamma_{n,m})$.

Since the solution for the two cases are different and the objective function may change, the search for the optimal beamwidths is complicated and needs to be simplified. First, we introduce a variable $g_n = h_n f_n$. Then, the SINR $\text{UE}_{n,1}$ and $\text{UE}_{n,2}$ can be rewritten as

$$\gamma_{n,1} = \frac{P_{n,1} \beta_{n,1}^2 (g_n + \xi^2)}{\sum_{\ell=1, q \neq n}^{2} P_{q,\ell} \beta_{n,1}^2 (f_n + \xi) \xi + \sigma^2}, \tag{4.36}$$

$$\gamma_{n,2} = \frac{P_{n,2} \beta_{n,2}^2 (g_n + \xi^2)}{P_{n,1} \beta_{n,2}^2 (g_n + \xi^2) + \sum_{\ell=1, q \neq n}^{2} P_{q,\ell} \beta_{n,2}^2 (f_n + \xi) \xi + \sigma^2}. \tag{4.37}$$

Since we fix the parameters for Cluster 2, $\log_2(1 + \gamma_{2,1}) + \log_2(1 + \gamma_{2,2})$ is a constant, which we denote by C. Then, we define a function $F(g_1, f_1) \triangleq \log_2(1 + \gamma_{1,1}) + \log_2(1 + \gamma_{1,2}) + C$ which has the following property:

Proposition 2. *For $F(g_1, f_1)$ with its domain defined by $(f_1, g_1) \in [lb_f, ub_f] \times [lb_g, \min\{f_1^2, ub_g\}]$, $0 < lb_g \leq lb_f^2$ and $0 < ub_g \leq ub_f^2$, the maximum point lies on the boundary $g_1 = f_1^2$, $f_1 \in [lb_f, \sqrt{ub_g}]$.*

Proof. See Appendix A.1. □

Proposition 2 implies that if we want to find the maximum point of $F(g_1, f_1)$, we only need to search on the boundary $g_1 = f_1^2, f_1 \in [lb_f, ub_f]$. This simplifies $F(g_1, f_1)$ to $F(f_1^2, f_1)$. We further

define $F_b(f_1) \triangleq F(f_1^2, f_1)$. Then, to find the maximum point of $F(g_1, f_1)$, we can perform a line search for $F_b(f_1)$ on $f_1 \in [lb_f, ub_f]$.

Next, we define the function $G(g_1, f_1) \triangleq (1 - \frac{\omega_{\text{BS}}\omega_{\text{UE}}g_1}{A^2}T_p)F(g_1, f_1)$. Function $G(g_1, f_1)$ has the following property:

Proposition 3. *For $G(g_1, f_1)$ with its domain defined by $(f_1, g_1) \in [\sqrt{lb_g}, ub_f] \times [lb_g, \min\{f_1^2, ub_g\}]$, $ub_f, lb_g, ub_g > 0$, $ub_g < ub_f^2$, the maximum point lies on the boundary $g_1 = f_1^2, f_1 \in [\sqrt{lb_g}, \sqrt{ub_g}]$.*

Proof. See Appendix A.2. □

Proposition 3 implies that if we want to find the maximum point of $G(g_1, f_1)$, we only need to search on the boundary $g_1 = f_1^2, f_1 \in [lb_f, ub_f]$. This simplifies $G(g_1, f_1)$ into $G(f_1^2, f_1)$. We define $G_b(f_1) \triangleq G(f_1^2, f_1)$. Then, to find the maximum point of $G(g_1, f_1)$, we can perform the line search for $G_b(f_1)$ on $f_1 \in [lb_f, ub_f]$.

To find the maximum point for Problem (4.35), we plot its feasible region with boundaries colored in green and blue in Figs. 4.4 and 4.5. According to Propositions 2 and 3, the maximum point lies on the blue boundary and we only need to search on the blue boundary. However, the objective function varies along the blue boundary. To conduct an effective search, we need to divide the blue boundary into two different subsets. Moreover, different initial conditions lead to different division strategies. There are two cases:

- Case 1: $g_1^{(0)} < g_2^{\text{fix}}$, where $g_1^{(0)}$ is the initial point. In this case, along the blue boundary, when we increase g_1 from $g_1^{(0)}$ to g_2^{fix}, the objective function is $(1 - \frac{\tau^*}{T})F_b(f_1)$. If we continue to increase g_1, the objective function changes to $G_b(f_1)$. Then, the blue boundary is divided as shown in Fig. 4.4. On Subset a, we perform a line search over $(1 - \frac{\tau^*}{T})F_b(f_1)$ to find a maximum point $(f_1^{(F)}, (f_1^{(F)})^2)$. On Subset b, we perform a line search over $G_b(f_1)$ to find the maximum point $(f_1^{(G)}, (f_1^{(G)})^2)$. Then, we compare the values of $(1 - \frac{\tau^*}{T})F_b(f_1^{(F)})$ and $G_b(f_1^{(G)})$, to pick the larger one as the optimal solution (f_1^*, g_1^*).

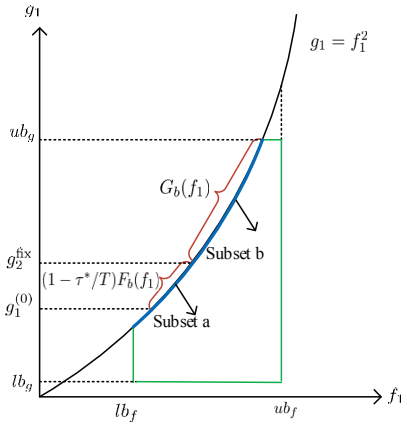

Figure 4.4: The search region division for Case 1.

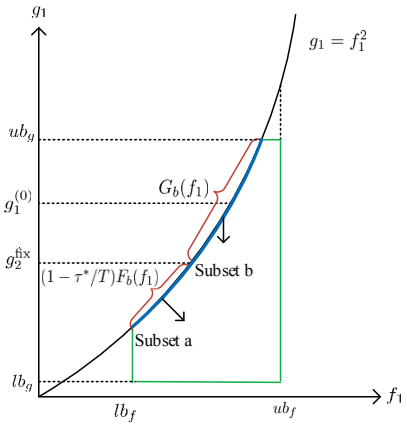

Figure 4.5: The search region division for Case 2.

- Case 2: $g_1^{(0)} \geq g_2^{\text{fix}}$, where $g_1^{(0)}$ is the initial point. In this case, when we decrease g_1 to be less than g_2^{fix}, the objective function changes from $G_b(f_1)$ to $(1-\frac{\tau^*}{T})F_b(f_1)$. Then, the blue boundary is divided as shown in Fig. 4.5. On Subset a, we perform a line search over $(1-\frac{\tau^*}{T})F_b(f_1)$ to find a maximum point $(f_1^{(F)}, (f_1^{(F)})^2)$. On Subset b, we perform a line search over $G_b(f_1)$ to find the maximum point $(f_1^{(G)}, (f_1^{(G)})^2)$. Then, we compare the values of $(1-\frac{\tau^*}{T})F_b(f_1^{(F)})$ and $G_b(f_1^{(G)})$, to pick the larger one as the optimal solution (f_1^*, g_1^*).

While performing the line search, we also need to consider the minimum rate constraint. The details of line search for $(1-\frac{\tau^*}{T})F_b(f_1)$ and $G_b(f_1)$ are described in Alg. 4 and Alg. 5, respectively.

Algorithm 4: Line search over $(1-\frac{\tau^*}{T})F_b(f_1)$

Input τ^*, the search interval $[f_{\min}, f_{\max}]$ and the step size Δ;
Initialize $R_{\max} \leftarrow 0$ and $f_1^{(F)} \leftarrow f_{\min}$;
for $f_1 = f_{\min} : \Delta : f_{\max}$ **do**
    if $(1-\frac{\tau^*}{T})F_b(f_1) > R_{\max}$ and $(1-\frac{\tau^*}{T})\log_2(1+\gamma_{1,1}) \geq R_{\min}$ and $(1-\frac{\tau^*}{T})\log_2(1+\gamma_{1,2}) \geq R_{\min}$ **then**
        $R_{\max} \leftarrow (1-\frac{\tau^*}{T})F_b(f_1)$;
        $f_1^{(F)} \leftarrow f_1$;
    end if
end for
Return $R_{\max}$, $f_1^{(F)}$.

Algorithm 5: Line search over $G_b(f_1)$

Input the search interval $[f_{\min}, f_{\max}]$ and the step size Δ;
Initialize $R_{\max} \leftarrow 0$ and $f_1^{(G)} \leftarrow f_{\min}$;
for $f_1 = f_{\min} : \Delta : f_{\max}$ **do**
    $\tau \leftarrow \frac{\omega_{\mathrm{BS}}\omega_{\mathrm{UE}}f_1^2}{A^2}T_p$
    if $G_b(f_1) > R_{\max}$ and $(1-\frac{\tau}{T})\log_2(1+\gamma_{1,m}) \geq R_{\min}$ and $(1-\frac{\tau}{T})\log_2(1+\gamma_{2,m}^{\mathrm{fix}}) \geq R_{\min}$ **then**
        $R_{\max} \leftarrow G_b(f_1)$;
        $f_1^{(G)} \leftarrow f_1$;
    end if
end for
Return $R_{\max}$, $f_1^{(G)}$.

4.4.4 Joint beamwidth optimization

For a fixed power allocation, our final beamwidth search algorithm, iterating between the beamwidth search for the two clusters, is presented in Alg. 6.

Algorithm 6 Joint Beamwidth Optimization

Set the sum-rate $R_{sum}[-1] \leftarrow 0$, the maximal iteration number $k_{max} \leftarrow 1000$ and the convergence threshold $\varepsilon \leftarrow 10^{-3}$;
Input: $R_{sum}[0]$, μ_2,η_2,$\mu_1^{(0)}$,$\eta_1(0)$, ω_{BS}, ω_{UE}, $\mu_{\min}$ and $\eta_{\min}$;
Calculate $f_1^{(0)}$,$g_1^{(0)}$,$f_2^{(0)}$ and $g_2^{(0)}$ according to (4.30), (4.31) and $g_n = f_n h_n$;
Calculate lb_f, ub_f, lb_g and ub_g based on ω_{BS}, ω_{UE}, $\mu_{\min}$ and $\eta_{\min}$;
$k \leftarrow 0$;
while $R_{sum}[k] - R_{sum}[k-1] \geq \varepsilon R_{sum}[k-1]$ and $k \leq k_{max}$ **do**
    $k \leftarrow k+1$;
    if $g_1^{(k-1)} < g_2^{(k-1)}$ **then**
        $\tau^* \leftarrow \frac{\omega_{\mathrm{BS}}\omega_{\mathrm{UE}}g_2^{(k-1)}T_p}{A^2}$;
        Do line search for $(1-\frac{\tau^*}{T})F_b(f_1)$ on interval $f_1 \in [\sqrt{g_1^{(k-1)}}, \sqrt{g_2^{(k-1)}}]$ using Alg. 4 to get the maximum point $(f_1^{(F)}, (f_1^{(F)})^2)$;

Do line search for $G_b(f_1)$ on interval $f_1 \in [\sqrt{g_2^{(k-1)}}, \sqrt{ub_g}]$ using Alg. 5 to get the maximum point $(f_1^{(G)}, (f_1^{(G)})^2)$;
if $F_b(f_1^{(F)}) == 0$ and $G_b(f_1^{(G)}) == 0$ **then**
  $(f_1^{(k)}, g_1^{(k)}) \leftarrow (f_1^{(k-1)}, g_1^{(k-1)})$;
end if
Compare the value of $(1 - \frac{\tau^*}{T})F_b(f_1^{(F)})$ and $G_b(f_1^{(G)})$, and pick the larger one as the optimal solution $(f_1^{(k)}, g_1^{(k)})$;
else if $g_1^{(k-1)} \geq g_2^{(k-1)}$ **then**
Do line search for $G_b(f_1)$ on interval $f_1 \in [\sqrt{g_2^{(k-1)}}, \sqrt{ub_g}]$ using Alg. 5 to get the maximum point $(f_1^{(G)}, (f_1^{(G)})^2)$;
$\tau^* \leftarrow \frac{\omega_{\text{BS}}\omega_{\text{UE}}g_2^{(k-1)}T_p}{A^2}$;
Do line search for $(1 - \frac{\tau^*}{T})F_b(f_1)$ on interval $f_1 \in [lb_f, \sqrt{g_2^{(k-1)}}]$ using Alg. 4 to get the maximum point $(f_1^{(F)}, (f_1^{(F)})^2)$;
if $F_b(f_1^{(F)}) == 0$ and $G_b(f_1^{(G)}) == 0$ **then**
  $(f_1^{(k)}, g_1^{(k)}) \leftarrow (f_1^{(k-1)}, g_1^{(k-1)})$;
end if
Compare the value of $(1 - \frac{\tau^*}{T})F_b(f_1^{(F)})$ and $G_b(f_1^{(G)})$, and pick the larger one as the optimal solution $(f_1^{(k)}, g_1^{(k)})$;
end if
if $g_2^{(k-1)} < g_1^{(k)}$ **then**
$\tau^* \leftarrow \frac{\omega_{\text{BS}}\omega_{\text{UE}}g_1^{(k)}T_p}{A^2}$;
Do line search for $(1 - \frac{\tau^*}{T})F_b(f_2)$ on interval $f_2 \in [\sqrt{g_2^{(k-1)}}, \sqrt{g_1^{(k)}}]$ using Alg. 4 to get the maximum point $(f_2^{(F)}, (f_2^{(F)})^2)$;
Do line search for $G_b(f_2)$ on interval $f_2 \in [\sqrt{g_1^{(k)}}, \sqrt{ub_g}]$ using Alg. 5 to get the maximum point $(f_2^{(G)}, (f_2^{(G)})^2)$;
if $F_b(f_2^{(F)}) == 0$ and $G_b(f_2^{(G)}) == 0$ **then**
  $(f_2^{(k)}, g_2^{(k)}) \leftarrow (f_2^{(k-1)}, g_2^{(k-1)})$;

end if

Compare the value of $(1-\frac{\tau^*}{T})F_b(f_2^{(F)})$ and $G_b(f_2^{(G)})$, and pick the larger one as the optimal solution $(f_2^{(k)}, g_2^{(k)})$;

else if $g_2^{(k-1)} \geq g_1^{(k)}$ **then**

Do line search for $G_b(f_2)$ on interval $f_2 \in [\sqrt{g_1^{(k)}}, \sqrt{ub_g}]$ using Alg. 5 to get the maximum point $(f_2^{(G)}, (f_2^{(G)})^2)$;

$\tau^* \leftarrow \frac{\omega_{\mathrm{BS}}\omega_{\mathrm{UE}}g_1^{(k)}T_p}{A^2}$;

Do line search for $(1-\frac{\tau^*}{T})F_b(f_2)$ on interval $f_1 \in [lb_f, \sqrt{g_1^{(k)}}]$ using Alg. 4 to get the maximum point $(f_2^{(F)}, (f_2^{(F)})^2)$;

if $F_b(f_2^{(F)}) == 0$ and $G_b(f_2^{(G)}) == 0$ **then**

$(f_2^{(k)}, g_2^{(k)}) \leftarrow (f_2^{(k-1)}, g_2^{(k-1)})$;

end if

Compare the value of $(1-\frac{\tau^*}{T})F_b(f_2^{(F)})$ and $G_b(f_2^{(G)})$, and pick the larger one as the optimal solution $(f_2^{(k)}, g_2^{(k)})$;

end if

Calculate $R_{sum}[k]$;

end while

Calculate μ_n and η_n based on f_n, h_n, where $h_n = g_n/f_n$;

Return μ_n^* and η_n^*.

4.4.5 The joint algorithm

Based on Alg. 3 and Alg. 6, we can propose a joint optimization algorithm, which iterates between the power allocation and the beamwidth optimization. The details of the algorithm are described in Alg. 7.

Algorithm 7: Joint Power and Beamwidth Optimization

1: Set the sum-rate $R_{sum}[-1] \leftarrow 0$, the maximal iteration number $n_{max} \leftarrow 1000$ and the convergence threshold $\varepsilon \leftarrow 10^{-3}$;
2: Choose a feasible start point $\mathbf{P}^*[0]$, $\boldsymbol{\mu}^*[0]$, $\boldsymbol{\eta}^*[0]$ and $\boldsymbol{\phi}^*[0]$;
3: $n \leftarrow 0$;
4: **while** $R_{sum}[n] - R_{sum}[n-1] \geq \varepsilon R_{sum}[n-1]$ and $n \leq n_{max}$ **do**
5: $n \leftarrow n+1$;
6: Search the optimal beamwidth using Alg. 6 with $\mathbf{P}^*[n]$ to obtain $\boldsymbol{\mu}^*[n]$, $\boldsymbol{\eta}^*[n]$ and $\boldsymbol{\phi}^*[n]$;
7: Solve Problem (4.19) using Alg. 3 with $\boldsymbol{\mu}^*[n-1]$, $\boldsymbol{\eta}^*[n-1]$, $\boldsymbol{\phi}^*[n-1]$ to obtain $\mathbf{P}^*[n]$;
8: **end while**
9: Return $\mathbf{P}^*[n]$, $\boldsymbol{\mu}^*[n]$, $\boldsymbol{\eta}^*[n]$.

4.4.6 Convergence and complexity analysis

To prove the convergence of Alg. 7, we first need to prove the convergence of Alg. 3 and Alg. 6. The convergence of Alg. 3 has been proved in [14]. In Alg. 6, to maximize the sum-rate, we optimize the beamwidth for one cluster while keeping the other cluster fixed. Such a step cannot decrease the sum-rate and generates a non-decreasing sequence of sum-rate values. Therefore, the convergence of Alg. 6 is guaranteed because the algorithm generates a sequence of non-decreasing sum-rates with an upper bound (the maximum sum-rate).

In Alg. 7, when allocating the power, we increase the sum-rate while keeping the beams in the feasible region. When optimizing the beamwidth, we search the feasible region for the beams to find the maximum sum-rate while guaranteeing the minimum rate constraint and keeping the powers in the feasible region. By doing so, we generate a monotonically increasing sequence with an upper bound (the maximum sum-rate), which proves the convergence.

Here, we provide the complexity analysis of the proposed algorithm. In our algorithm, we iteratively optimize the power allocation and beamwidth. In the power allocation algorithm, we use SPCA to gradually convexify the original non-convex problem. In each iteration, the complexity mainly lies in solving Problem (4.28). We use an off-the-shelf solution, i.e., CVX to solve Problem (4.28), which uses the interior-point method. The computational complexity of CVX is

$\mathcal{O}((3M_{\text{UE}})^{3.5})$, where M_{UE} is the total number of users and $3M_{\text{UE}}$ is the number of variables in Problem (4.28). In the beamwidth optimization, we iteratively optimize the beamwidth for each sector. In each iteration, the main complexity lies in the line-search algorithm, with the complexity $\mathcal{O}(\frac{ub_f - lb_f}{\Delta})$, ub_f and lb_f are the upper bound and lower bound of variable f in Problem (4.35), respectively, and Δ is the stepsize of the line-search algorithm.

4.5 Simulation Results

In this section, we present the simulation results of the joint power and beamwidth optimization algorithm. Four UEs are considered which are divided in two clusters each with two UEs. It is assumed that the UEs inside each cluster have different distances from the BS. Four multiple access techniques are investigated. The first technique is OMA in which UEs are served in different time slots. The second technique is a combination of OMA and NOMA called NOMA-OMA. In NOMA-OMA, UEs that belong to the same cluster are supported by a fixed-power NOMA and each cluster is supported by OMA at each time slot. The third technique is Fixed-NOMA in which all UEs are served by a fixed-power NOMA at one time slot. Finally, the fourth technique is the jointly optimized power and beamwidth NOMA system presented in Section 4.4, called Optimized-NOMA. For all techniques, first, the beams are trained and then the data transmission is done.

To evaluate the performance of the Optimized-NOMA, the parameters are set as follows. The minimum rate for all UEs is assumed to be $R_{\min} = 0.1$ bits/s/Hz. Further, for the Fixed-NOMA, we allocate $\frac{1}{5}$ of the total power to the cluster-head and $\frac{4}{5}$ of the total power to the far UE as done in [26]. Also, the power is equally divided between the two clusters. The SNR used in the simulations indicates the transmit SNR, i.e., SNR=$\frac{P_{\text{tot}}}{\sigma^2}$, $\sigma^2 = 1$. In the first cluster, the channel gains of the near and far UEs from the BS are $\beta_{1,1}^2 = -17$dB and $\beta_{1,2}^2 = -26.5$dB. In the second cluster, the channel gains of the near and far UEs are $\beta_{2,1}^2 = -19$dB and $\beta_{2,2}^2 = -25$dB. Also, the side lobe

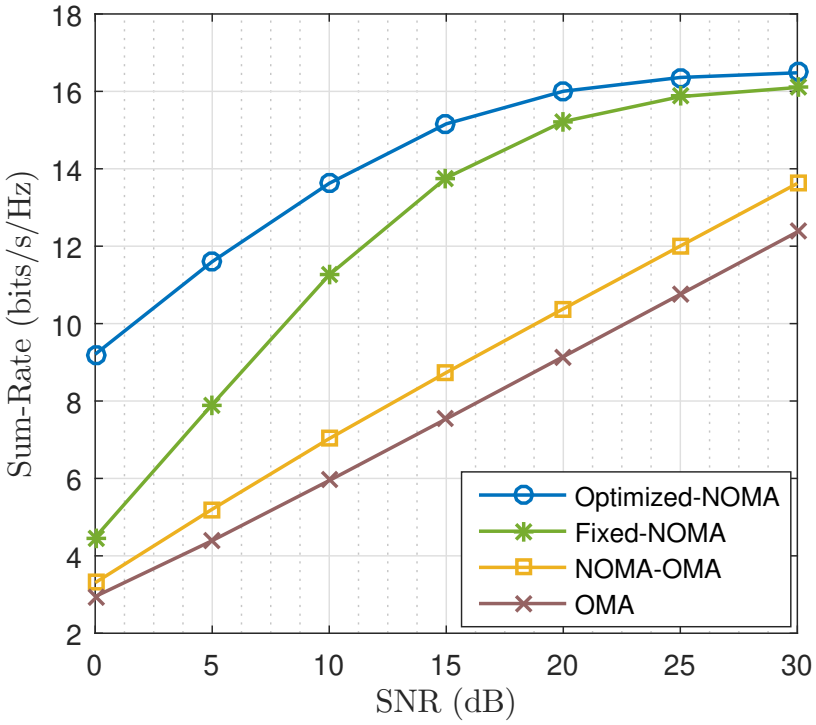

Figure 4.6: Performance of the sum-rate versus SNR for a large channel coherence time, i.e., $T = 5 \times 10^3 T_p$.

level is constant and is given as $\xi = 0.1$. For the Optimized-NOMA, we use $\omega_{\text{BS}} = \omega_{\text{UE}} = 120^\circ$ and $\eta_{\min} = \mu_{\min} = 5^\circ$ unless it is mentioned otherwise. Further, the convergence threshold is set to $\varepsilon = 10^{-3}$.

Fig. 4.6 demonstrates the performance of the sum-rate versus SNR. It is assumed that $T = 5 \times 10^3 T_p$ which indicates a large channel coherence time and $\eta = \mu_1 = \mu_2 = 10^\circ$. For all SNRs, by increasing SNR the sum-rate increases. The Optimized-NOMA achieves the highest sum-rate. Especially, at low SNRs, the performance gap is larger. For instance, at SNR $= 0$dB the gap between the Optimized-NOMA and Fixed-NOMA is more than 5 bits/s/Hz which reveals our joint optimization algorithm designs the powers and beamwidths very efficiently. As SNR increases, the gap decreases which is due to the fact that in the Fixed-NOMA the BS and UEs steer strong beams even if the powers and beamwidths are not optimized. The Fixed-NOMA technique preforms better than the NOMA-OMA and OMA techniques. The reason is that the Fixed-NOMA serves

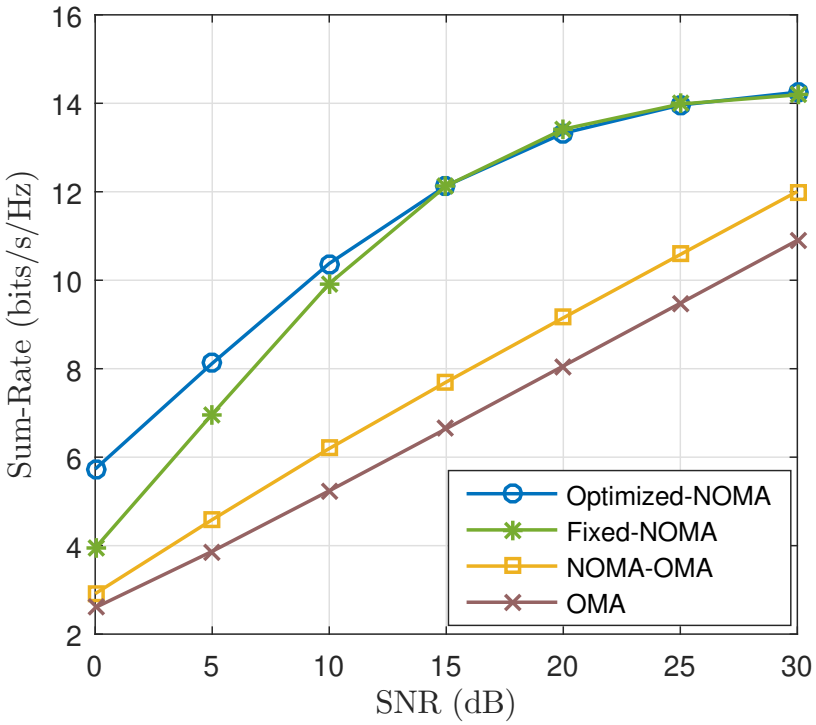

Figure 4.7: Performance of the sum-rate versus SNR for a short channel coherence time $T = 1 \times 10^3 T_p$.

all the users at the same time and takes the advantages of the spectrum sharing among UEs.

In Fig. 4.7, we repeat the same simulation as in Fig. 4.6 for a relatively short channel coherence time, i.e., $T = 1 \times 10^3 T_p$. Similarly, by increasing SNR, the sum-rate increases for all the techniques. However, compared to Fig. 4.6, at low SNRs, the rate gap between the Optimized-NOMA and Fixed-NOMA is small. Moreover, at high SNR regions, these two techniques achieve identical sum-rates. This is because when the channel coherence time is short, the optimization algorithm does not allocate a large portion of T to the beam-training, e.g., τ is small. Thus, the optimized beamwidths are not narrow enough to provide higher gain. Also, at high SNRs, the optimized powers have trivial effects on the sum-rate compared to the predefined fixed values which is an interesting observation. This observation indicates that for a short channel coherence time like $T = 1 \times 10^3 T_p$ and high SNR, the optimization is not required and fixed-NOMA can be used instead. For a smaller coherence time, the optimized-NOMA shows better performance only at low SNRs. Nevertheless, severe path loss and shadowing in mmWave bands makes the low SNR regime very crucial. Especially, NOMA is supposed to consider near and far users, where the far users likely receive signal through NLoS low SNR channels [5]. We emphasize that at high SNRs, by increasing the coherence time, the rate gap between the optimized-NOMA and the fixed-NOMA becomes larger (See Fig. 4.6).

Fig. 4.8 shows the sum-rate performance versus the normalized channel coherence time, i.e., T/T_p for the moderate SNR= 20dB. In this simulation, two sets of beamwidths are considered for the first three techniques: (i) $\eta = \mu = 5^\circ$ (narrow beamwidth) and (ii) $\eta = \mu = 10^\circ$ (relatively wide beamwidth). The Optimized-NOMA outperforms the other techniques for both sets of the beamwidths. For the Fixed-NOMA with $\eta = \mu = 10^\circ$ and short normalized channel coherence times, the performance is very close to that of the Optimized-NOMA. This is expected as we explained before. However, for $\eta = \mu = 5^\circ$ and short normalized channel coherence times, the sum-rate of the Fixed-NOMA is much smaller than that of the Optimized-NOMA. This is because at small T/T_p, the Fixed-NOMA assigns more time to the beam-training and leaves less time for

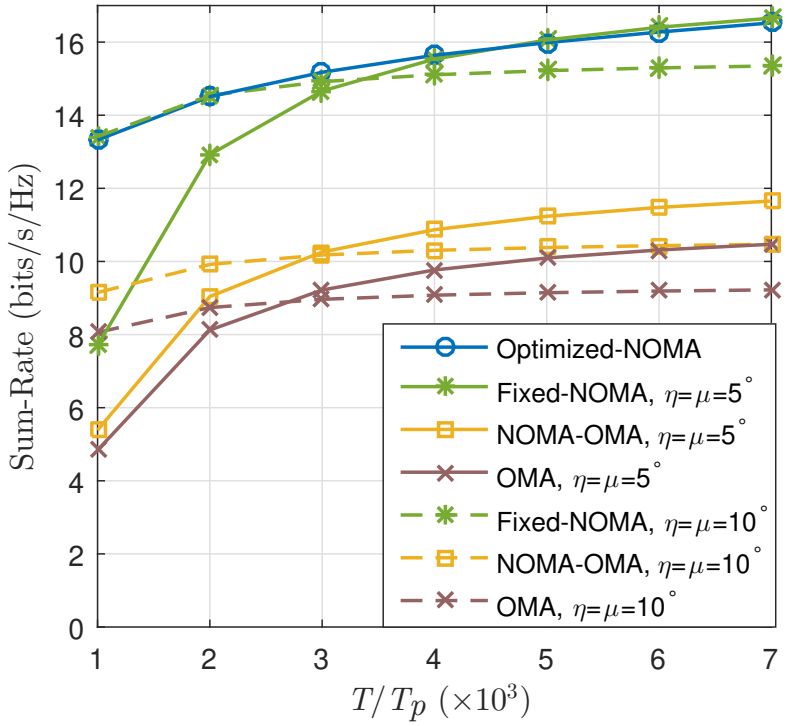

Figure 4.8: Performance of the sum-rate versus T/T_p for SNR=20dB.

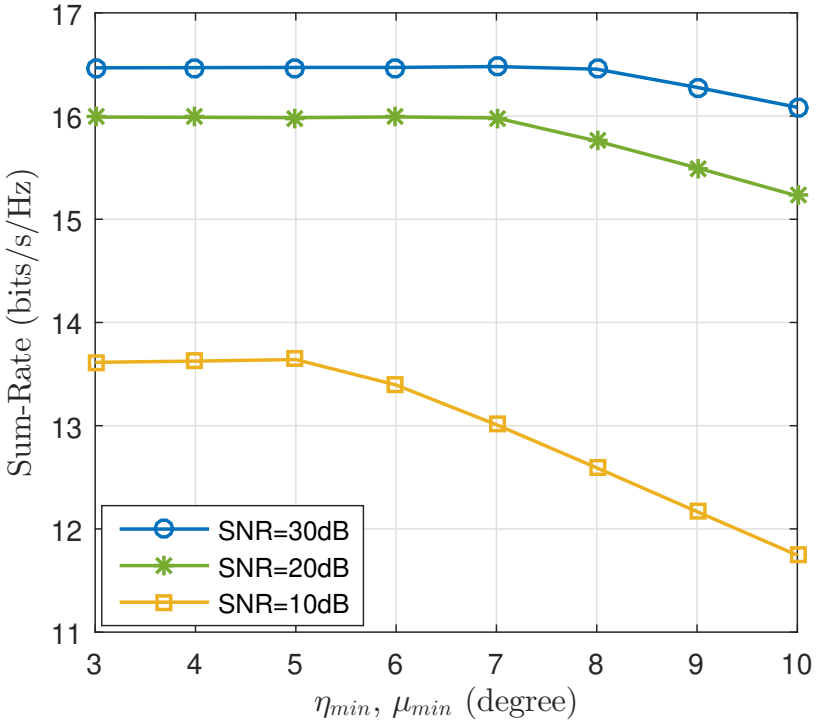

Figure 4.9: Performance of the sum-rate of the Optimized-NOMA versus minimum BS and UE beamwidth (η_{min} and μ_{min}) for $T = 5 \times 10^3 T_p$ and various SNRs.

the data transmission. As the normalized time goes up, more time is available for data transmission, and the narrow beam provides a higher sum-rate. This statement also is supported by Fig. 4.8 where at large channel coherence times, the Optimized-NOMA selects the minimum beamwidth. Hence, the Optimized-NOMA and the Fixed-NOMA with $\eta = \mu = 5^\circ$ achieve identical sum-rates at large normalized channel coherence times. Using Optimized-NOMA, a wide beamwidth is preferred for short T/T_p while a narrow beamwidth is preferred for large T/T_p.

In Fig. 4.9, we simulate the performance of the sum-rate versus the minimum beamwidth of BS and UEs. The simulation is done for SNRs 10dB, 20dB, and 30dB and $T = 5 \times 10^3 T_p$. In practice, the number of antennas at the BS and UEs is limited and even for large T/T_p, a narrow beamwidth may not be generated. At high SNRs, e.g. 30dB, increasing the minimum beamwidth does not affect the sum-rate severely. As such, compared to $\eta_{\min} = \mu_{\min} = 3^\circ$, at 10°, the sum-rate is reduced only by 0.5 bits/s/Hz. At SNR= 20dB, the sum-rate drops about 0.8 bits/s/Hz which is larger than the drop at SNR= 30dB. The minimum beamwidth has a major effect at SNR= 10dB. When SNR is low, narrow beams can still provide high gains to compensate for the low SNR. As the minimum beamwidth increases and the SNR is low, the optimization algorithm cannot select narrow beams. As a result, the sum-rate dramatically decreases. In this case, simulation results indicate that decreasing the beamwidth from 3° to 10° decreases the sum-rate by about 2 bits/s/Hz.

4.6 Conclusions

In this chapter, NOMA is incorporated into mmWave hybrid beamforming systems. We also consider the beam-training time because of the limited channel coherence time in mmWave directional communications. By combining the exhaustive search and tone-based beam-training algorithms, a new beam-training algorithm is employed. The formulated sum-rate expression consists of the channel coherence time and beam-training time. To maximize the sum-rate, a joint power allocation and beamwidth control optimization problem is solved by an algorithm which iterates between

the power allocation and the beamwidth optimization. The non-convex power allocation is solved by the log-reformulation and SPCA. The beamwidth optimization is solved by iterating between the two clusters. A boundary-search algorithm is proposed to reduce the search complexity for the beamwidth in each cluster. The numerical results demonstrate that an efficient power allocation and beam-training time can lead to higher sum-rates compared to the conventional mmWave-NOMA without optimized parameters, NOMA-OMA, and OMA. The only exception is that for a short channel coherence time and high SNR, the optimized-NOMA and the fixed-NOMA have identical sum-rate performance. Also, at low SNRs, the size of the antenna array is a major obstacle in achieving higher sum-rates.

Chapter 5

A Two-Step mmWave-NOMA Beam Alignment Algorithm with Misalignment Probability Constraints

5.1 Introduction

Because available spectral in under 6 GHz range is limited, the use of millimeter-Wave (mmWave) bands is a promising solution for the next generation of wireless communications called 5G [99]. The transmission in mmWave bands suffers from severe path loss which constrains the communication range to short distances. On the other hand, because of the short wavelength, a huge number of antennas can be utilized to provide a large beamforming gain. Fully-digital beamforming can achieve near maximum beamforming gain, but requires one radio frequency (RF) chain per antenna. This makes a fully-digital beamforming system expensive, power-hungry, and not practical. To resolve the issue, two practical beamforming systems are analog beamforming and hybrid beamforming. In analog beamforming systems, one RF chain is deployed and beamform-

ing is done via phase-shifters [63]. In contrast, the hybrid beamforming systems are equipped with several RF chains and the beamforming is done in digital and analog stages [12, 112, 64]. Alternatively, dedicating RF chains to channel paths instead of antennas, for example using beamspace MIMO [16] or reconfigurable antenna-based MIMO [11, 51] architectures, reduces the number of required RF chains as well.

Non-orthogonal multiple access (NOMA) as another promising solution for 5G aims at enhancing the spectral efficiency through serving more than one user at the same frequency/time/code [121, 148, 37]. In particular, signals in power-domain NOMA are transmitted using superposition coding (SC) and decoded by exploiting successive interference cancellation (SIC) at the receiver. Incorporation of power domain-NOMA in the mmWave systems, i.e., mmWave-NOMA, is vastly studied in the literature [27, 25, 147, 125, 3, 56, 139, 123, 9]. Ref. [27] proposes a random beamforming method for the mmWave-NOMA directional transmission. Ref. [25] evaluates sum-rate for a two-user NOMA with different directions but the same beamforming codeword using finite-resolution phase-shifters. In [147], the effect of beam misalignment on the coverage and rate of the analog beamforming systems is evaluated. Ref. [125] uses beamwidth control to properly pair the mmWave-NOMA users to maximize the sum-rate. The impact of beam alignment on the sum-rate and users' order in hybrid beamforming systems is analyzed in [3]. Further, the physical security issue of the mmWave-NOMA is investigated in [56]. Ref. [139] implements the mmWave-NOMA with one-bit quantized angle information. To serve more than one user per beam, the combination of NOMA with lens-based beamspace MIMO is performed in [123]. Finally, [9] incorporates NOMA in mmWave reconfigurable antennas to increase the number of served users and improve the sum-rate.

In all aforementioned works, it is assumed that the channel coherence time is sufficiently large to align the beams with high resolution and the beam alignment overhead is ignored. To fully exploit the potential of the large beamforming gain in mmWave systems, transmitter and receiver beams must be aligned. Such an alignment significantly increases the signal-to-noise ratio (SNR)

and consequently leads to higher throughput. However, finding the best direction increases the alignment overhead. Hence, there is a tradeoff between alignment and throughput. A narrower beamwidth leads to higher transmission rate due to the higher directivity gain (i.e., higher SNR) but imposes considerable alignment overhead since many directions should be examined. A larger beamwidth requires less alignment overhead although it reduces the transmission rate. In the mmWave-NOMA, the alignment becomes more critical, where one transmit beam and at least two receive beams, with a large channel gain difference, need alignment. A narrower beamwidth is in favor of the low channel gain receivers since they need sufficiently high SNR to achieve a required rate. In contrast, high channel gain receivers might be interested in a wider beamwidth as less time is allocated to alignment. These issues result in a new alignment-throughput tradeoff in the mmWave-NOMA. To address this tradeoff, we propose a joint formulation of a two-step beamwidth design and power allocation in the mmWave-NOMA, and study the impact of the network parameters on the throughput.

5.1.1 Related Work

Since channel estimation is a challenging task in mmWave directional communications due to the large-scale antenna array, beam alignment is adopted as a promising solution [75]. Recall that in the directional communications, there exists an alignment-throughout trade-off. In the literature, several works have investigated this trade-off. In [109], a joint beamwidth selection and scheduling algorithm was designed for multiple concurrent transmissions. Ref. [130] investigated a joint energy harvest-ratio and beamwidth selection algorithm in the mmWave communications in which a frame is divided into three segments, energy harvesting, beam searching, and data transmission. Then, [34] studied a fair beamwidth selection, user association, and power allocation algorithm in the mmWave bands in order to maximize the minimum user throughput. Further, [69] proposed a hierarchical beam search method for mmWave multi-user communications. In these works, each beam is dedicated to serve only one user and users located in the same direction are

served by time division multiple access or space division multiple access. Hence, the works [109, 69, 130, 34] may not be properly utilized in the mmWave-NOMA communications. To resolve this issue, [49] and [10] have recently developed a joint beamwidth selection and power allocation algorithm to maximize the sum-rate for NOMA users for the analog and hybrid mmWave-NOMA, respectively. Our work in [10] considers the simultaneous beamwidth design of the far and near users while taking the limited channel coherence time for the mmWave-NOMA into account. For NOMA to be more effective, the channel gain difference between the users must be large, i.e., the near and far users have high and low channel gains, respectively. In our one-step algorithm in [10] and other similar beam alignment algorithms, not leveraging this important property leads to a slow beam alignment. We use the channel gain difference in the this chapter by utilizing a two-step algorithm. In the first step, we align a fairly wide beam to the near user. Then, in the second step, a narrower beam is aligned to the near and far users. Such a two-step beam alignment algorithm takes advantage of the NOMA user locations and is much faster than the one-step algorithm. We also include additional constraints on the beamwidths to satisfy given misalignment probability requirements.

5.1.2 Contributions

The main contributions of this chapter are as follows:

- We propose a new two-step beamwidth selection algorithm compatible with mmWave-NOMA systems.
- We formulate an optimization problem that brings together beamwidth selection and power allocation. We solve the optimization problem by a joint optimization algorithm that iterates between the power allocation and the beamwidth.
- We further formulate another optimization problem that includes additional constraints on

the beamwidths to satisfy given misalignment probability requirements and solve it.

The reminder of the chapter is organized as follows. In Section 5.2, the system model is explained. In Section 5.3, the proposed two-step beam alignment is described. Section 5.4 formulates the optimization problem. In Section 5.5, the allocated power and beamwidth are determined through the proposed optimization algorithm and its convergence analysis is provided. Section 5.6 addresses the beam misalignment probability. In this section, the power coefficients and beamwidths are optimized to include the beam misalignment. Numerical results are presented in Section 5.7. Section 5.8 concludes this chapter.

5.2 System Model

We consider a mmWave downlink system which consists of a single cell and a BS located at the center. Also, we consider M user equipments (UEs) distributed uniformly in the cell. The system model is depicted in Fig. 5.1.(a). The BS with N_{BS} antennas and each UE with N_{UE} antennas are equipped with mmWave analog beamforming systems. The beamforming system is shown in Fig. 5.1.(b). A similar system is assumed for every UE. In order to apply the power-domain NOMA to the described system, two UEs located at the same direction but different distances from the BS should be detected and paired. Throughout this chapter, the user closer to the BS is called the *near user* ($\mathrm{UE_n}$) and the further user is called the *far user* ($\mathrm{UE_f}$). To determine $\mathrm{UE_n}$ and $\mathrm{UE_f}$, we assume the BS knows the long term channel for each user, which includes the path-loss and large-scale fading effects. The long term channel can be estimated through channel quality indicator (CQI) [79] or reference signals receive power (RSRP) [4], which can be obtained before beam alignment. We divide the users according to the quality of the channels. To this end, the cell is divided into *inner cell* and *outer cell* in which the $\mathrm{UE_n}$ and the $\mathrm{UE_f}$ are located, respectively. Without loss of generality, let us assume there are M_{n} UEs in the inner cell and M_{f} UEs in the outer cell such that $M = M_{\mathrm{n}} + M_{\mathrm{f}}$.

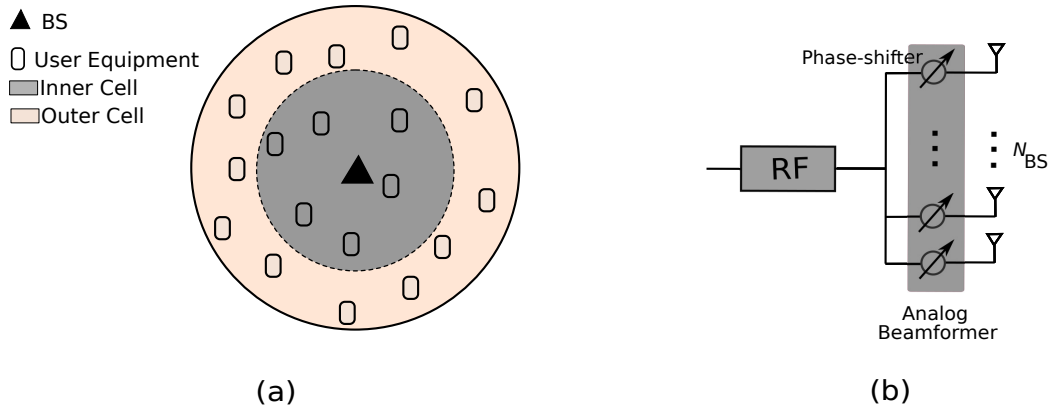

Figure 5.1: (a) A single cell with the BS at the middle. The users are located either in the inner cell or the outer cell. (b) The mmWave analog beamforming system designed for the BS.

5.2.1 Channel Model

The channel matrix between the BS and a typical UE in downlink is assumed to follow the widely adopted the Saleh-Valenzuela model [104] which is an aggregation of L scatterers. The expression of the channel matrix is given by

$$\mathbf{H} = \frac{1}{\sqrt{L}} \sum_{l=1}^{L} h_l \mathbf{a}_{\mathrm{UE}}(\theta_l) \mathbf{a}_{\mathrm{BS}}^{\dagger}(\phi_l), \tag{5.1}$$

where h_l denotes the gain of the l-th path. Further, the variables $\phi_l \in [0, 2\pi]$ and $\theta_l \in [0, 2\pi]$ are the azimuth angle of departure (AoD) and azimuth angle of arrival (AoA) of the l-th path, respectively. Also, the vectors $\mathbf{a}_{\mathrm{BS}} \in \mathbb{C}^{N_{\mathrm{BS}} \times 1}$ and $\mathbf{a}_{\mathrm{UE}} \in \mathbb{C}^{N_{\mathrm{UE}} \times 1}$ are antenna array response vectors at the BS and UE, respectively. The antenna array vector of the l-th path at the BS is given by

$$\mathbf{a}_{\mathrm{BS}}(\phi_l) = \left[1, e^{j\frac{2\pi}{\lambda} d_0 \sin(\phi_l)}, \ldots, e^{j(N_{\mathrm{BS}}-1)\frac{2\pi}{\lambda} d_0 \sin(\phi_l)}\right]^T, \tag{5.2}$$

where d_0 denotes the antenna spacing and λ denotes the wavelength of the propagation. The antenna array response vector for $\mathbf{a}_{\mathrm{UE}}(\theta_l)$ can be written in a similar fashion.

It is worth mentioning that the channel model in (5.1) indicates a 2-dimensional beamforming system. Hence, in this chapter, we focus on designing beamwidths in azimuth. In the future work,

we will consider performing 3-dimensional beamforming that allows us to control the beamwidth in azimuth and elevation. We should also note that, in this chapter, a long term resource allocation is assumed [109, 59, 140, 119, 120, 10]. To this end, the channel is defined as a long term channel model where the following assumptions are taken into account: (i) the variables AoD and AoA remain unchanged during the channel coherence time and (ii) the channel gain h_l contains the path loss and large-scale fading effects and does not change over the coherence time. The former assumption is due to the fact that the coherence time is much less than the time duration over which AoD and AoA do not change [140]. The later implies that the instantaneous channel fluctuations are ignored during the coherence time which is a valid assumption due to the long term power allocation and the beamwidth control [140, 109, 10, 5].

5.2.2 Signal Model and Beamforming Gain

Assume that the BS wants to transmit the symbols s_{n} and s_{f} to $\mathrm{UE_n}$ and $\mathrm{UE_f}$, respectively, using NOMA. We let $s \in \mathbb{C}$ denote the information signal vector such that $\mathbb{E}[ss^*] = 1$. Recall that NOMA superimposes the signals such that the transmit NOMA signal becomes $x = \sqrt{P_{\mathrm{n}}}s_{\mathrm{n}} + \sqrt{P_{\mathrm{f}}}s_{\mathrm{f}}$ where P_{n} and P_{f} are the allocated powers to $\mathrm{UE_n}$ and $\mathrm{UE_f}$, respectively. This stage is followed by the beamforming stage where the analog precoder $\mathbf{f} \in \mathbb{C}^{N_{\mathrm{BS}} \times 1}$ is performed. Each element of the precoder has a constant magnitude of $\frac{1}{\sqrt{N_{\mathrm{BS}}}}$ and the total power of the analog beamforming is limited to $||\mathbf{f}||^2 = 1$. After applying the analog combiner $\mathbf{w}^{\mathrm{i}} \in \mathbb{C}^{N_{\mathrm{UE}} \times 1}$ $\mathrm{i} \in \{\mathrm{n}, \mathrm{f}\}$, the received signal by $\mathrm{UE_i}$ is expressed as

$$\begin{aligned}\mathbf{y}^{\mathrm{i}} &= \mathbf{w}^{\mathrm{i}\dagger}\mathbf{H}^{\mathrm{i}}\mathbf{f}x + \mathbf{n}^{\mathrm{i}} \\ &= \sqrt{P_{\mathrm{n}}}\mathbf{w}^{\mathrm{i}\dagger}\mathbf{H}^{\mathrm{i}}\mathbf{f}s_{\mathrm{n}} + \sqrt{P_{\mathrm{f}}}\mathbf{w}^{\mathrm{i}\dagger}\mathbf{H}^{\mathrm{i}}\mathbf{f}s_{\mathrm{f}} + \mathbf{n}^{\mathrm{i}},\end{aligned} \tag{5.3}$$

where $\mathbf{n}^{\mathrm{i}} \in \mathbb{C}^{N_{\mathrm{UE}} \times 1}$ is the additive, white Gaussian noise vector with zero-mean and σ^2 variance for each element, i.e., $\mathscr{CN}(0, \sigma^2)$. Suppose the users are ordered based on their effective channel

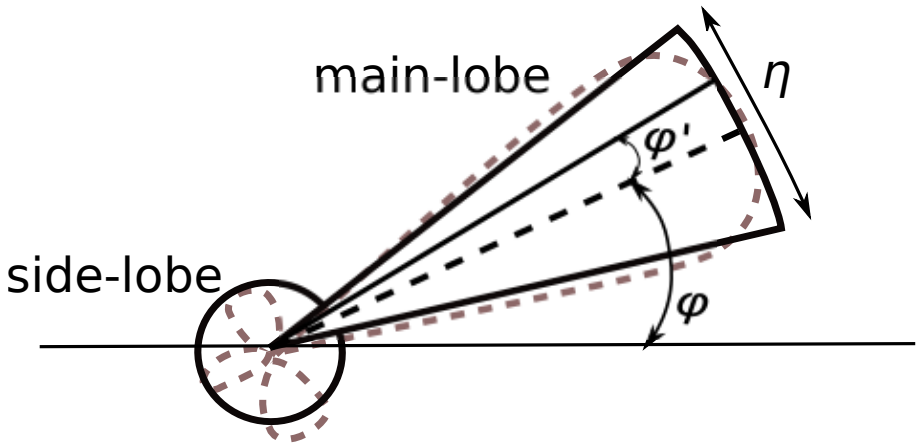

Figure 5.2: A realistic beamforming model (dashed line) with variable main-lobe and side-lobe gain and an estimated beamforming model (solid line) with a constant main-lobe and side-lobe gain. η, ϕ, ϕ' denote the beamwidth, AoD of the BS, and the angle off the boresight direction, respectively.

gain such that $|\mathbf{w}^{\mathrm{n}\dagger}\mathbf{H}^{\mathrm{n}}\mathbf{f}|^2 \geq |\mathbf{w}^{\mathrm{f}\dagger}\mathbf{H}^{\mathrm{f}}\mathbf{f}|^2$ which leads to $P_{\mathrm{n}} \leq P_{\mathrm{f}}$. Due to the different power levels, $\mathrm{UE_n}$ applies SIC and the remain is $\mathrm{UE_n}$'s signal contaminated only by the noise. Further, $\mathrm{UE_f}$ treats the $\mathrm{UE_n}$'s signal as the inter-user interference. Hence, the intended signal is contaminated by the interference plus noise.

Note that in our analog beamforming system, we transmit one stream using beam steering with an analog beamformer. If the channel is single-path, i.e., $L = 1$, the beam should be directed in that single path through beam alignment. If the channel is multi-path, i.e., $L > 1$, since all paths can be found by beam alignment, the path with best quality can be selected and the beam is directed in the direction of the best path during the data transmission. Therefore, only one path is used for the transmission. In other words, applying the beam alignment procedure reduces the multi-path MIMO channel to a single-path SISO channel and the gain of the SISO channel is given by $|\mathbf{w}^{\dagger}\mathbf{H}\mathbf{f}|^2$ [5]. In fact, this procedure can be explained by the channel model below:

$$\mathbf{H} = \sqrt{\frac{N_{\mathrm{BS}}N_{\mathrm{UE}}}{L}}\mathbf{A}_{\mathrm{BS}}\mathbf{V}\mathbf{A}_{\mathrm{UE}}, \tag{5.4}$$

where $\mathbf{A}_{\mathrm{BS}} = [\mathbf{a}_{\mathrm{BS}}(\bar{\phi}_1), ..., \mathbf{a}_{\mathrm{BS}}(\bar{\phi}_{N_{\mathrm{BS}}})]$ and $\mathbf{A}_{\mathrm{UE}} = [\mathbf{a}_{\mathrm{UE}}(\bar{\theta}_1), ..., \mathbf{a}_{\mathrm{UE}}(\bar{\theta}_{N_{\mathrm{UE}}})]$ are two DFT matrices.

$\{\bar{\phi}_m = \frac{2\pi(m-1)}{N_{\text{BS}}} - \pi\}_{m=1}^{N_{\text{BS}}}$ and $\{\bar{\theta}_n = \frac{2\pi(n-1)}{N_{\text{UE}}} - \pi\}_{n=1}^{N_{\text{UE}}}$ divide the entire angular space $[-\pi, \pi]$ into N_{UE} and N_{BS} directions, respectively. Matix $\mathbf{V}$ is a sparse matrix with only L nonzero entries whose positions are directly determined by the AoAs and AoDs. Based on (5.4), the beamalignment procedure can be expressed as

$$(\mathbf{w}^{\text{opt}}, \mathbf{f}^{\text{opt}}) = \arg \max_{\mathbf{w} \in \mathbf{A}_{\text{UE}}, \mathbf{f} \in \mathbf{A}_{\text{BS}}} |\mathbf{w}^{\dagger} \mathbf{H} \mathbf{f}|^2. \tag{5.5}$$

The solution to (5.5) can be found as $\mathbf{w}^{\text{opt}} = \mathbf{a}_{\text{UE}}(\bar{\theta}_i)$ and $\mathbf{f}^{\text{opt}} = \mathbf{a}_{\text{BS}}(\bar{\phi}_j)$, where i and j are , respectively, the row index and column index of the item with largest magnitude in $\mathbf{V}$. This solution shows that beam alignment finds the strongest path in $\mathbf{H}$. Since beam alignment practically reduces the multi-path channel to a single-path channel, in our system, the beamforming gain can be defined as the product of the beamforming gain of the BS and the beamforming gain of the UE, i.e., $G_{\text{BS}} G_{\text{UE}}$ such that $G_{\text{BS}} = |\mathbf{a}_{\text{BS}}^{\dagger} \mathbf{f}|^2$ and $G_{\text{UE}} = |\mathbf{w}^{\dagger} \mathbf{a}_{\text{UE}}|^2$, where $\mathbf{a}_{\text{BS}}$ and $\mathbf{a}_{\text{UE}}$ are the AoD and AoA steering vectors of the stongest path [7]. The beamforming gain at the BS is modeled by

$$G_{\text{BS}}(\phi', \eta) = \begin{cases} \frac{2\pi - (2\pi - \eta)\alpha}{\eta}, & \text{if } |\phi'| \leq \frac{\eta}{2}, \\ \alpha, & \text{otherwise}, \end{cases} \tag{5.6}$$

where η and ϕ' denote the beamwidths of the BS and the off-boresight angle in the BS, respectively. Parameter α is non-zero and $\alpha \ll 1$ for narrow beams. In Fig. 5.2, the dashed line represents a realistic beamforming model and the solid line represents the modeled beamforming considered in this chapter. It is observed that in this model, the gain of the main-lobe is distributed uniformly in the entire beamwidth and the side-lobe gain is assumed to be constant. We use this model to keep the complexity of the presentation low, as iit is done in many other papers [109, 95, 127]. However, we should note that there is a better but more complicated model [45] that assumes the deviation from the boresight angle decreases the beamforming gain. In our system, due to the presence of far users, it makes sense to select narrow beamwidths. In such a case, the beamforming

gain degradation is negligible. Similarly, the model for the beamforming gain of $\mathrm{UE_i}$ is given by

$$G_{\mathrm{UE}}(\theta^{\mathrm{i}'},\mu^{\mathrm{i}}) = \begin{cases} \frac{2\pi-(2\pi-\mu^{\mathrm{i}})\alpha}{\mu^{\mathrm{i}}}, & \text{if } |\theta^{\mathrm{i}'}| \leq \frac{\mu^{\mathrm{i}}}{2}, \\ \alpha, & \text{otherwise}, \end{cases} \tag{5.7}$$

where μ^{i} and $\theta^{\mathrm{i}'}$ denote the beamwidths of the $\mathrm{UE_i}$ and the off-boresight angle in the $\mathrm{UE_i}$, respectively. Considering the signal after applying SIC, (5.6), and (5.7), the SNR of $\mathrm{UE_n}$ is expressed as

$$\gamma_{\mathrm{n}} = \frac{P_{\mathrm{n}} h^{\mathrm{n}^2} G_{\mathrm{BS}}(\phi',\eta) G_{\mathrm{UE}}(\theta^{\mathrm{n}'},\mu^{\mathrm{n}})}{\sigma^2}, \tag{5.8}$$

and the signal-to-interference-plus-noise ratio (SINR) of $\mathrm{UE_f}$ is obtained as

$$\gamma_{\mathrm{f}} = \frac{P_{\mathrm{f}} h^{\mathrm{f}^2} G_{\mathrm{BS}}(\phi',\eta) G_{\mathrm{UE}}(\theta^{\mathrm{f}'},\mu^{\mathrm{f}})}{P_{\mathrm{n}} h^{\mathrm{f}^2} G_{\mathrm{BS}}(\phi',\eta) G_{\mathrm{UE}}(\theta^{\mathrm{f}'},\mu^{\mathrm{f}}) + \sigma^2}. \tag{5.9}$$

5.3 Two-Step Beamwidth Design

5.3.1 The Existing Beamwidth Design

Most of the current methods determine the beamwidth in one step [109, 69, 34, 49]. That is, for mmWave single-user, multi-user, and NOMA scenarios, all the directions are examined in one step. In particular, for two UEs with comparable distances from the BS, the beam search algorithm should select a much narrower beamwidth to find the best beam for $\mathrm{UE_f}$ at the cost of considerable large beam alignment time. To better elaborate, we have the following example. We assume the beam resolutions of the BS, $\mathrm{UE_n}$, and $\mathrm{UE_f}$ to be respectively $\beta^{\mathrm{BS}} = \lceil \frac{\omega_{\mathrm{BS}}}{\eta} \rceil$, $\beta^{\mathrm{n}} = \lceil \frac{\omega_{\mathrm{n}}}{\mu^{\mathrm{n}}} \rceil$, and $\beta^{\mathrm{f}} = \lceil \frac{\omega_{\mathrm{f}}}{\mu^{\mathrm{f}}} \rceil$, and the required minimum SNR is satisfied. Recall that ω_{BS}, ω_{n}, and ω_{f} denote the sector-level beamwidths of the BS, the near user, and the far user, respectively. In this chapter, it

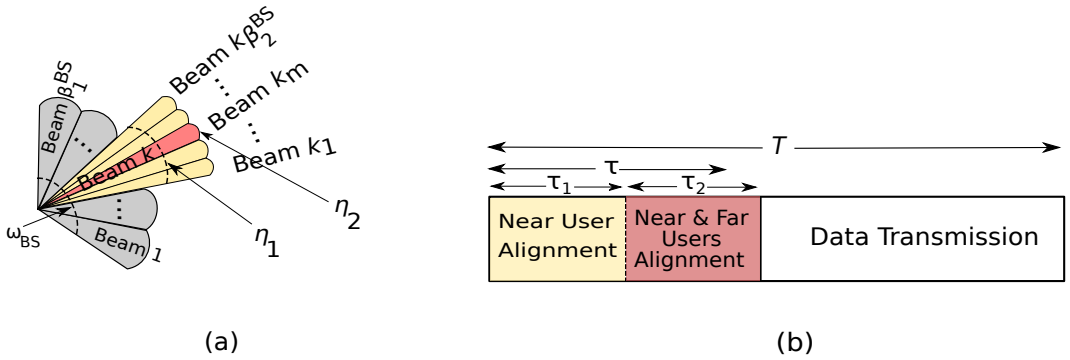

Figure 5.3: (a) The proposed two-step beam alignment algorithm. The steps are demonstrated only for the BS. At Step one, the BS has the beam resolution β_1^{BS} and the beamwidth η_1. It is assumed that $\mathrm{UE_n}$ is located in the direction cover by Beam k. At the second step, Beam k is divided into β_2^{BS} beams with the beamwidth η_2. $\mathrm{UE_n}$ and $\mathrm{UE_f}$ are found to be located in the direction of Beam k_m. To ease the representation, only main-lobes are illustrated and side-lobes are eliminated. (b) The beam alignment and data transmission frame. The channel coherence time is T seconds. The whole alignment time takes τ seconds. The first and second steps take τ_1 and τ_2 seconds, respectively.

is assumed that the sector-level alignment is performed ideally and we only focus on beam-level alignment [109]. We note that if the minimum SNR is satisfied for $\mathrm{UE_f}$, it will also be satisfied for $\mathrm{UE_n}$. Using the one-step beam search, the alignment time is $\tau_{\mathrm{conv}} = \max\{\beta^{\mathrm{BS}}\beta^{\mathrm{n}}T_p, \beta^{\mathrm{BS}}\beta^{\mathrm{f}}T_p\}$ where T_p denotes the pilot transmission time. Intuitively speaking, due to the higher path loss, β^{f} becomes much larger than β^{n}. Therefore, as the distance between the BS and $\mathrm{UE_f}$ increases, the alignment time increases too. In brief, the one-step beam alignment may not be the most efficient algorithm for the mmWave-NOMA systems. On the other hand, the hierarchical algorithm presented in [69] may not be suitable for the mmWave-NOMA systems since the large beamwidth in the first step probably cannot provide sufficient SNR to find the right direction between the BS and $\mathrm{UE_f}$.

5.3.2 The Proposed Beamwidth Design

The proposed algorithm consists of two steps as shown in Fig. 5.3. At the first step, the BS randomly selects one of $\mathrm{UE_n}$s. By solving Problem (4.18) defined in Section 5.4, the BS designs

the parameters $\beta_1^{\mathrm{BS}} = \lceil \frac{\omega_{\mathrm{BS}}}{\eta_1} \rceil$, $\beta_1^{\mathrm{n}} = \lceil \frac{\omega_{\mathrm{n}}}{\mu_1^{\mathrm{n}}} \rceil$, $\beta_2^{\mathrm{BS}} = \lceil \frac{\eta_1}{\eta_2} \rceil$, $\beta_2^{\mathrm{n}} = \lceil \frac{\mu_1^{\mathrm{n}}}{\mu_2^{\mathrm{n}}} \rceil$, and $\beta_2^{\mathrm{f}} = \lceil \frac{\mu_1^{\mathrm{n}}}{\mu_2^{\mathrm{f}}} \rceil$. The parameters η_1 and μ_1^{n} denote the beamwidth of the BS and the selected $\mathrm{UE_n}$ at the first step of the alignment. The parameters η_2, μ_2^{n}, and μ_2^{f} are respectively defined as the beamwidth of the BS, the selected $\mathrm{UE_n}$, and $\mathrm{UE_f}$s at the second step. At the first step, the beam alignment results in obtaining the AoD_1 and the AoA of the selected $\mathrm{UE_n}$, i.e., $\mathrm{AoA}_1^{\mathrm{n}}$. Let us assume that after the first step, the $\mathrm{UE_n}$'s location is found to be in Beam k as shown in Fig. 5.3(a).

To begin the second step, the BS determines those $\mathrm{UE_f}$s which are located in Beam k. Suppose there are M_{f}' ($1 \leq M_{\mathrm{f}}' \leq M_{\mathrm{f}}$) $\mathrm{UE_f}$s located in Beam k. This information suggests that the M_{f}' $\mathrm{UE_f}$s should look for the best direction within the range of Beam k. Therefore, the search time for determining the best direction can be remarkably reduced. In Fig. 5.3(a), k_m is supposed to be the best direction. To carry out the beam alignment at the second step, the BS forwards β_2^{n} to $\mathrm{UE_n}$ and β_1^{n}, β_2^{f}, and AoD_1 to the selected $\mathrm{UE_f}$s. Then, the beam alignment leads to finding the AoD_2, the AoA of the selected $\mathrm{UE_n}$, i.e., $\mathrm{AoA}_2^{\mathrm{n}}$, and the AoA of $\mathrm{UE_f}$s, i.e., $\mathrm{AoA}_2^{\mathrm{f}}$s. In summary, we first determine the AoA of the $\mathrm{UE_n}$. Then, using the obtained AoA, we determine the AoA of the M_{f}' $\mathrm{UE_f}$s. The proposed two-step algorithm is described in Alg. 8.

At the first step, the alignment time is $\tau_1 = \beta_1^{\mathrm{BS}} \beta_1^{\mathrm{n}} T_p$. The beam alignment time for the second step is $\tau_2 = \max\{\beta_2^{\mathrm{BS}} \beta_2^{\mathrm{n}} T_p, \beta_2^{\mathrm{BS}} \beta_2^{\mathrm{f}} T_p\}$. Therefore, the total beam alignment time is $\tau = \tau_1 + \tau_2 = (\beta_1^{\mathrm{BS}} \beta_1^{\mathrm{n}} + \max\{\beta_2^{\mathrm{BS}} \beta_2^{\mathrm{n}}, \beta_2^{\mathrm{BS}} \beta_2^{\mathrm{f}}\}) T_p$ as depicted in Fig. 5.3(b). It is clear that a large τ leaves a short time for data transmission. However, it can be shown that $\beta_1^{\mathrm{BS}} \beta_1^{\mathrm{n}} + \max\{\beta_2^{\mathrm{BS}} \beta_2^{\mathrm{n}}, \beta_2^{\mathrm{BS}} \beta_2^{\mathrm{f}}\} < \max\{\beta^{\mathrm{BS}} \beta^{\mathrm{n}}, \beta^{\mathrm{BS}} \beta^{\mathrm{f}}\}$ which provides $\tau < \tau_{\mathrm{conv}}$. It is worth mentioning that the paired $\mathrm{UE_f}$ takes the advantages of the mmWave-NOMA system. Especially, the $\mathrm{UE_f}$ is aligned by using the direction information of the $\mathrm{UE_n}$ which significantly reduces the beam alignment time.

As the $\mathrm{AoA}_2^{\mathrm{n}}$ and $\mathrm{AoA}_2^{\mathrm{f}}$s are found, the BS looks for the best $\mathrm{UE_f}$ that can be paired with the $\mathrm{UE_n}$. Specifically, the $\mathrm{UE_f}$ positioned in Beam k_m is paired with the $\mathrm{UE_n}$. To this end, the BS requests the $\mathrm{UE_f}$s to feed their $\mathrm{AoA}_2^{\mathrm{f}}$ back. If only one of the M_{f}' $\mathrm{AoA}_2^{\mathrm{f}}$s is the same as the $\mathrm{AoA}_2^{\mathrm{n}}$, the two UEs are paired. If more than one of the M_{f}' $\mathrm{AoA}_2^{\mathrm{f}}$s are the same as the $\mathrm{AoA}_2^{\mathrm{n}}$, the $\mathrm{UE_n}$ is paired

Algorithm 8: Two-Step Beam Alignment

Input: Long term channel informations, the probability ε, the minimum rate $R_{\min}$
Output: AoD_2, $\mathrm{AoA}_2^{\mathrm{n}}$, and $\mathrm{AoA}_2^{\mathrm{f}}$s
Step 1: AoD_1, and $\mathrm{AoA}_1^{\mathrm{n}}$ determination
    Using CQIs, the BS divides UEs into $\mathrm{UE_n}$s and $\mathrm{UE_f}$s.
    The BS selects a UE randomly from the $\mathrm{UE_n}$s set.
    Solving Problem (4.18), the parameters β_1^{BS}, β_1^{n}, β_2^{BS}, β_2^{n}, and β_2^{f} are designed.
    Using β_1^{BS} and β_1^{n}, the beam alignment procedure provides the AoD_1 and $\mathrm{AoA}_1^{\mathrm{n}}$.
Step 2: AoD_2, $\mathrm{AoA}_2^{\mathrm{n}}$, and $\mathrm{AoA}_2^{\mathrm{f}}$s determination
    Using β_1^{BS} and AoD_1, the BS determines the M_{f}' $\mathrm{UE_f}$s.
    The BS forwards β_2^{n} to $\mathrm{UE_n}$ and β_1^{n}, β_2^{f}, and AoD_1 to the M_{f}' $\mathrm{UE_f}$s.
    The beam alignment procedure leads to determination of AoD_2, $\mathrm{AoA}_2^{\mathrm{n}}$, and $\mathrm{AoA}_2^{\mathrm{f}}$s.

with the furthest $\mathrm{UE_f}$. The reason for the furthest UE selection is that NOMA achieves a better performance compared with the traditional multiple access techniques as the distance between the paired UEs increases [6].

5.4 Optimization Formulation

Here, we formulate an optimization problem with the goal of maximizing the sum-rate subject to the beamwidths and the allocated powers. To do this, we introduce the following variables. The rates of the $\mathrm{UE_n}$ and $\mathrm{UE_f}$ are denoted by $R_{\mathrm{n}} = (1 - \frac{\tau}{T})\log(1 + \gamma_{\mathrm{n}})$ and $R_{\mathrm{f}} = (1 - \frac{\tau}{T})\log(1 + \gamma_{\mathrm{f}})$, respectively, where τ is the total alignment time and T denotes the frame duration shown in

Fig. 5.3(b). Then, the optimization problem is formulated as

$$\underset{\eta_1,\mu_1^{\mathrm{n}},\eta_2,\mu_2^{\mathrm{n}},\mu_2^{\mathrm{f}},P_{\mathrm{n}},P_{\mathrm{f}}}{\text{maximize}} \quad R_{\mathrm{n}}+R_{\mathrm{f}} \tag{5.10a}$$

$$\text{subject to} \quad \eta_{\min}\leq\eta_2\leq\eta_1\leq\omega, \tag{5.10b}$$

$$\mu_{\min}\leq\mu_2^{\mathrm{n}}\leq\mu_1^{\mathrm{n}}\leq\omega, \tag{5.10c}$$

$$\mu_{\min}\leq\mu_2^{\mathrm{f}}\leq\mu_1^{\mathrm{n}}, \tag{5.10d}$$

$$R_{\mathrm{n}},R_{\mathrm{f}}\geq R_{\min}, \tag{5.10e}$$

$$P_{\mathrm{n}}+P_{\mathrm{f}}\leq P_{\mathrm{tot}}, \tag{5.10f}$$

where $\eta_{\min}$ and $\mu_{\min}$ denote the highest resolution of the arrays at the BS and the UEs, respectively. P_{tot} represents the total power.

In Problem (5.10), seven parameters should be optimized. To reduce the number of design parameters, we assume: (i) $\eta_1=\mu_1^{\mathrm{n}}$, (ii) $\eta_2=\mu_2^{\mathrm{n}}=\mu_2^{\mathrm{f}}$, and (iii) $\eta_{\min}=\mu_{\min}$. Thus, the problem is simplified to

$$\underset{\eta_1,\eta_2,P_{\mathrm{n}},P_{\mathrm{f}}}{\text{maximize}} \quad R_{\mathrm{n}}+R_{\mathrm{f}} \tag{5.11a}$$

$$\text{subject to} \quad \eta_{\min}\leq\eta_2\leq\eta_1\leq 2\pi, \tag{5.11b}$$

$$R_{\mathrm{n}},R_{\mathrm{f}}\geq R_{\min}, \tag{5.11c}$$

$$P_{\mathrm{n}}+P_{\mathrm{f}}\leq P_{\mathrm{tot}}. \tag{5.11d}$$

Now, the number of the parameters is reduced to four resulting in a more straightforward optimization problem.

5.5 Joint Beamwidth and Power Optimization Algorithm

Problem (5.11) is still an intractable non-convex optimization problem with complex constraints. Thus, we propose an algorithm which iterates between the power allocation and the beamwidth optimization.

5.5.1 Power Allocation

When we fix the beamwidth, the alignment time and beamforming gain are accordingly fixed. We denote the beamforming gain as G.

To further simplify the problem, we introduce slack variables x_1, x_2, d_2, to transform objective in Problem (5.11) into $\log_2 2^{x_1} + \log_2 \frac{2^{x_2}}{2^{d_2}} = x_1 + x_2 - d_2$. Then, Problem (5.11) can be simplified to

$$\underset{P_{\mathrm{n}}, P_{\mathrm{f}}}{\text{maximize}} \quad x_1 + x_2 - d_2 \tag{5.12a}$$

$$\text{subject to} \quad 2^{x_1} \leq \frac{P_{\mathrm{n}} G h^{\mathrm{n}^2} + \sigma^2}{\sigma^2}, \tag{5.12b}$$

$$2^{x_2} \leq P_{\mathrm{n}} G h^{\mathrm{f}^2} + P_{\mathrm{f}} G h^{\mathrm{f}^2} + \sigma^2, \tag{5.12c}$$

$$2^{d_2} \geq P_{\mathrm{n}} G h^{\mathrm{f}^2} + \sigma^2, \tag{5.12d}$$

$$(1 - \frac{\tau}{T}) x_1 \geq R_{\min}, \tag{5.12e}$$

$$(1 - \frac{\tau}{T})(x_2 - d_2) \geq R_{\min}, \tag{5.12f}$$

$$P_{\mathrm{n}} + P_{\mathrm{f}} \leq P_{\mathrm{tot}}. \tag{5.12g}$$

In (5.12), the optimum is achieved when the constraints (5.12b)-(5.12d) satisfy with equality. However, (5.12d) is still non-convex. To relax it, we use an SPCA [14]. In this method, the non-convex feasible set is sequentially approximated by an inner convex approximation. For example, at Iter-

ation k, we have a lower bound for 2^{d_2} as:

$$2^{d_2^*[k-1]} \log 2(d_2 - d_2^*[k-1]) + 2^{d_2^*[k-1]} \leq 2^{d_2}, \tag{5.13}$$

where $d_2^*[k-1]$ is the optimal solution at Iteration $k-1$. Based on (5.13), we can relax (5.12d) into a convex constraint as:

$$2^{d_2^*[k-1]} \log 2(d_2 - d_2^*[k-1]) + 2^{d_2^*[k-1]} \geq P_{\mathrm{n}} G h^{\mathrm{f}^2} + \sigma^2. \tag{5.14}$$

Then, at each iteration, we can relax (5.12) to

$$\underset{P_{\mathrm{n}}, P_{\mathrm{f}}}{\text{maximize}} \quad x_1 + x_2 - d_2 \tag{5.15a}$$

$$\text{subject to} \quad (5.12b) - (5.12c),\ (5.12e) - (5.12g), \tag{5.15b}$$

$$(5.14). \tag{5.15c}$$

Problem (5.15) is a convex problem, which can be efficiently solved by off-the-shelf solutions, such as CVX [41]. Then, we can use an iterative algorithm to provide an approximation solution for Problem (5.12). Detailed steps are presented in Alg. 9. According to [14], Alg. 9 converges.

5.5.2 Beamwidth Optimization

To optimize the beamwidths, we make the following approximations:

$$G_1 = \frac{(2\pi - (2\pi - \eta_1)\alpha)^2}{\eta_1^2} \approx \frac{(2\pi - 2\pi\alpha)^2}{\eta_1^2}, \tag{5.16}$$

$$G_2 = \frac{(2\pi - (2\pi - \eta_2)\alpha)^2}{\eta_2^2} \approx \frac{(2\pi - 2\pi\alpha)^2}{\eta_2^2}, \tag{5.17}$$

Algorithm 9: Power Allocation for Two-Step Alignment

Set the sum-rate $R_{\text{sum}}[-1] \leftarrow 0$, the maximal iteration number $k_{max} \leftarrow 1000$ and the convergence threshold $\varepsilon \leftarrow 10^{-3}$;
repeat
  Choose a feasible start point $\mathbf{P}^*[0]$;
  $x_1^*[0] \leftarrow \log_2(\frac{P_{\text{n}} G h^{\text{n}^2}}{\sigma^2})$ and $x_2^*[0] \leftarrow \log_2(P_{\text{n}} G h^{\text{f}^2} + P_{\text{f}} G h^{\text{f}^2} + \sigma^2)$
  $d_2^*[0] \leftarrow \log_2(P_{\text{n}} G h^{\text{f}^2} + \sigma^2)$;
until (5.12e) and (5.12f) are satisfied
$k \leftarrow 0$;
while $R_{\text{sum}}[k] - R_{\text{sum}}[k-1] \geq \varepsilon R_{\text{sum}}[k-1]$ and $k \leq k_{max}$ **do**
  $k \leftarrow k+1$;
  Solve (5.15) to obtain $x_1^*[k]$, $x_2^*[k]$, $d_2^*[k]$, $P_{\text{n}}^*[k]$ and $P_{\text{f}}^*[k]$;
end while
Return P_{n}^* and P_{f}^*.

where G_1 and G_2 are the beamforming gains of the alignment stages 1 and 2, respectively. Denoting $(2\pi - 2\pi\alpha)$ as A, then, we have $\eta_1^2 = A^2/G_1$ and $\eta_2^2 = A^2/G_2$. As discussed in Section 5.3, the alignment time of our algorithm can be expressed as

$$\tau = \frac{(2\pi)^2}{\eta_1^2} + \frac{\eta_1^2}{\eta_2^2} = \frac{(2\pi)^2 G_1}{A^2} + \frac{G_2}{G_1}. \tag{5.18}$$

Instead of finding the optimal beamwidth, we try to find the optimal G_1 and G_2 by solving:

$$\underset{G_1, G_2}{\text{maximize}} \quad (1 - \frac{\tau}{T})(R_{\text{n}} + R_{\text{f}}) \tag{5.19a}$$

$$\text{subject to} \quad \frac{A^2}{4\pi^2} \leq G_1 \leq G_2, \tag{5.19b}$$

$$G_2 \leq \frac{A^2}{\eta_{\min}^2}, \tag{5.19c}$$

$$R_{\text{n}}, R_{\text{f}} \geq R_{\min}, \tag{5.19d}$$

where $R_{\text{n}} = \log_2(1 + \frac{P_n G_2 h^{\text{n}^2}}{\sigma^2})$ and $R_{\text{f}} = \log_2(1 + \frac{P_f G_2 h^{\text{f}^2}}{P_n G_2 h^{\text{f}^2} + \sigma^2})$. Note that R_{n} and R_{f} are functions of G_2 but not G_1.

Problem (5.19) is an intractable problem. However, we can observe that, when G_2 is fixed, the

Algorithm 10: Beamwidth Optimization for Two-Step Alignment

1: Set the sum-rate $R_{\text{sum}}[-1] \leftarrow 0$, the maximal iteration number $k_{max} \leftarrow 1000$ and the convergence threshold $\varepsilon \leftarrow 10^{-3}$;
2: Choose a feasible start point $G_1^{\text{opt}}[0]$ and $G_2^{\text{opt}}[0]$;
3: **while** $R_{\text{sum}}[k] - R_{\text{sum}}[k-1] \geq \varepsilon R_{\text{sum}}[k-1]$ and $k \leq k_{max}$ **do**
4: $k \leftarrow k+1$;
5: Calculate $G_1^{\text{opt}}[k]$ based on $G_2^{\text{opt}}[k-1]$;
6: Do a one-dimensional search for $G_2^{\text{opt}}[k]$ based on $G_1^{\text{opt}}[k]$;
7: Calculate $R_{\text{sum}}[k]$;
8: **end while**
9: Return G_1^{opt} and G_2^{opt}.

optimal G_1 is the one that maximizes $1-\tau/T$, which is the same as minimizing $\tau = \frac{(2\pi)^2 G_1}{A^2} + \frac{G_2}{G_1}$. For $\tau = \frac{(2\pi)^2 G_1}{A^2} + \frac{G_2}{G_1}$, the minimum point is $G_1^* = \frac{\sqrt{G_2}A}{2\pi}$. When $G_1 < G_1^*$, τ is a decreasing function of G_1 and when $G_1 > G_1^*$, τ is an increasing function in terms of G_1. Considering the constraints for G_1, we can pick the optimal G_1 as: (i) if $\frac{A^2}{4\pi^2} < G_1^*$, then $G_1^{\text{opt}} = \min\{G_1^*, G_2\}$; (ii) if $\frac{A^2}{4\pi^2} > G_1^*$, then $G_1^{\text{opt}} = \frac{A^2}{4\pi^2}$.

Then, if we fix G_1 to optimize G_2, we can perform a one-dimensional search to find the optimal G_2^{opt} for the following problem:

$$\underset{G_2}{\text{maximize}} \quad (1-\frac{\tau}{T})(R_{\text{n}}+R_{\text{f}}) \tag{5.20a}$$

$$\text{subject to} \quad G_1 \leq G_2 \leq \frac{A^2}{\eta_{\min}^2}, \tag{5.20b}$$

$$R_{\text{n}}, R_{\text{f}} \geq R_{\min}. \tag{5.20c}$$

Since we can find the optimal G_1 when G_2 is fixed and vice versa, we can use an iterative algorithm to find the optimal G_1 and G_2. The algorithm is described in Alg. 10.

Algorithm 11: Joint Optimization for Two-Step Alignment

Set the sum-rate $R_{\text{sum}}[-1] \leftarrow 0$, the maximal iteration number $k_{max} \leftarrow 1000$ and the convergence threshold $\varepsilon \leftarrow 10^{-3}$;
Choose a feasible start point $G_1^{\text{opt}}[0]$, $G_2^{\text{opt}}[0]$, $P_{\text{n}}^{\text{opt}}[0]$ and $P_{\text{f}}^{\text{opt}}[0]$;
while $R_{\text{sum}}[k] - R_{\text{sum}}[k-1] \geq \varepsilon R_{\text{sum}}[k-1]$ and $k \leq k_{max}$ **do**
    $k \leftarrow k+1$;
    Solve Problem (5.12) using the SPCA method;
    Solve Problem (5.19) by iterating between G_1 and G_2;
    Calculate $R_{\text{sum}}[k]$;
end while
Calculate η_1^{opt} and η_2^{opt} based on G_1^{opt} and G_2^{opt}
Return η_1^{opt}, η_2^{opt}, $P_{\text{n}}^{\text{opt}}$ and $P_{\text{f}}^{\text{opt}}$.

5.5.3 Joint Optimization

We propose a joint optimization algorithm, which iterates between the power allocation and the beamwidth optimization. The details of the algorithm are described in Alg. 11.

5.5.4 Convergence Analysis

The convergence of Alg. 9 has been proved in [14].

In Alg. 10, to maximize the sum-rate, we optimize the beamwidth for one stage while fixing the beamwidth of the other stage. Such a step cannot decrease the sum-rate and generates a non-decreasing sequence of sum-rate values. Therefore, the convergence of Alg. 10 is guaranteed because the algorithm generates a sequence of non-decreasing sum-rates with an upper bound (the maximum sum-rate).

In Alg. 11, when allocating the power, we increase the sum-rate while keeping the beams in the feasible region. When optimizing the beamwidth, we search the feasible region for the beams to find the maximum sum-rate while guaranteeing the minimum rate constraint and keeping the powers in the feasible region. By doing so, we generate a monotonically increasing sequence with

an upper bound (the maximum sum-rate), which proves the convergence.

5.5.5 Complexity Analysis

In this section, we analyze the complexity of the optimization algorithm. Since the optimization is performed before the beam alignment, its results can be used for the beam alignment as long as the users are static or slightly moving. Therefore, the optimization time can be neglected compared with the beam alignment time.

For the power allocation, we use an off-the-shelf solution, i.e., CVX to solve Problem (5.15), which uses the interior-point method. The computational complexity of CVX is $\mathcal{O}(N^{3.5})$, where N is the number of variables to be optimized in the interior-point method. In our algorithm, $N = 5$.

For the beamwidth optimization, we iteratively optimize the beamwidth for two alignment stages. When we optimize G_1, the complexity is $\mathcal{O}(1)$, since we provide a closed-form solution for G_1. When we optimize G_2, the main complexity lies in the one-dimensional search algorithm, with the complexity $\mathcal{O}(\frac{G_2^{\max}-G_2^{\min}}{\delta})$, where δ is the step size of the one-dimensional search algorithm.

5.6 Beamwidth Optimization with Beam Misalignment

In practice, the directional communication in outdoor mmWave systems is susceptible to beam misalignment during the alignment period due to longer distance and outdoor propagation [127, 147, 3]. Even if beam alignment is performed without error, the other environmental factors such as precipitation and wind cause movement and vibration at the BS which leads to frequent beam misalignment [58]. In this section, we only focus on beam misalignment imposed by the mmWave propagation characteristics and longer distance. In particular, in the mmWave-NOMA systems, far users are mostly located in longer distances from the BS and prone to beam misalignment. This

is our main motivation to evaluate the impact of beam misalignment on the system described in Section 5.2.

5.6.1 Beam Misalignment Probability

The probability of beam misalignment of the $\mathrm{UE_i}$ is defined as

$$\begin{aligned}\mathrm{P}^{\mathrm{i}}_{\mathrm{mis}} &= P\left(\cup_{k^{\mathrm{i}}=1}^{\beta}(\cup_{k^{\mathrm{BS}}\neq k^{\mathrm{i}}=1}^{\beta}\{|y^{\mathrm{i}}_{1,1}|^2 < |y^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}}|^2\})\right) \\ &\leq \sum_{k^{\mathrm{i}}=1}^{\beta}\sum_{k^{\mathrm{BS}}\neq k^{\mathrm{i}}=1}^{\beta} P\left(|y^{\mathrm{i}}_{1,1}|^2 < |y^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}}|^2\right), \end{aligned} \tag{5.21}$$

where we denote $y_{k^{\mathrm{BS}},k^{\mathrm{i}}}$ as the received signal at $\mathrm{UE_i}$ for the beam pair $\{k^{\mathrm{BS}},k^{\mathrm{i}}\}$. We assume that the pair $\{1,1\}$ is the correct pair, while other pairs are all misaligned pairs. The main-lob gain for the beam pair $\{1,1\}$ is denoted by $G^{\mathrm{i}}_{1,1} = \frac{(2\pi-(2\pi-\eta_{\mathrm{i}})\alpha)^2}{\eta_{\mathrm{i}}^2}$ and the misaligned beam pair gain is denoted by $E^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}} = \frac{2\pi-(2\pi-\eta_{\mathrm{i}})\alpha}{\eta_{\mathrm{i}}}\alpha$, where η_{i} is the beamwidth of $\mathrm{UE_i}$ and we assume $\eta_{\mathrm{BS}} = \eta_{\mathrm{i}}$. Note that we neglect the misaligned pairs whose $E^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}} = \alpha^2$, since $\alpha^2 \ll G^{\mathrm{i}}_{1,1}$.

We notice that $y^{\mathrm{i}}_{1,1}$ has a complex normal distribution with mean $\sqrt{P_{\mathrm{tot}}G^{\mathrm{i}}_{1,1}}h^{\mathrm{i}}$ and variance of 1, i.e., $\mathscr{CN}(\sqrt{P_{\mathrm{tot}}G^{\mathrm{i}}_{1,1}}h^{\mathrm{i}},1)$, and $y^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}} \sim \mathscr{CN}(\sqrt{P_{\mathrm{tot}}E^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}}}h^{\mathrm{i}},1)$. Using [110], for two Gaussian random variables $y^{\mathrm{i}}_{1,1}$ and $y^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}}$ we have

$$P\left(|y^{\mathrm{i}}_{1,1}|^2 < |y^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}}|^2\right) = Q_1(a,b) - \frac{1}{2}\exp\left(-\frac{a^2+b^2}{2}\right)I_0(ab), \tag{5.22}$$

where $a = \sqrt{P_{\mathrm{tot}}E^{\mathrm{i}}_{k^{\mathrm{BS}},k^{\mathrm{i}}}}h^{\mathrm{i}}$ and $b = \sqrt{P_{\mathrm{tot}}G^{\mathrm{i}}_{1,1}}h^{\mathrm{i}}$. Further, $Q_1(a,b)$ denotes the first-order Marcum Q-function and $I_0(ab)$ denotes the 0th order of the modified Bessel function of the first kind.

5.6.2 Optimization Reformulation

Here, we reformulate the optimization problem to take into account the beam misalignment. The problem is expressed as

$$\underset{\eta_1,\mu_1^{\mathrm{n}},\eta_2,\mu_2^{\mathrm{n}},\mu_2^{\mathrm{f}},P_{\mathrm{n}},P_{\mathrm{f}}}{\text{maximize}} \quad R_{\mathrm{n}}+R_{\mathrm{f}} \tag{5.23a}$$

$$\text{subject to} \quad (5.10\text{b})-(5.10\text{f}), \tag{5.23b}$$

$$\mathrm{P}_{\mathrm{mis}}^{\mathrm{n}},\mathrm{P}_{\mathrm{mis}}^{\mathrm{f}}\leq\varepsilon, \tag{5.23c}$$

where ε is the predefined beam misalignment probability. Further, $\mathrm{P}_{\mathrm{mis}}^{\mathrm{n}}$ and $\mathrm{P}_{\mathrm{mis}}^{\mathrm{f}}$ denote the beam misalignment probabilities of $\mathrm{UE_n}$ at the first step and $\mathrm{UE_f}$ at the second step, respectively.

In Problem (5.23), seven parameters should be optimized. As we did before, to reduce the number of design parameters, we assume: (i) $\eta_1=\mu_1^{\mathrm{n}}$, (ii) $\eta_2=\mu_2^{\mathrm{n}}=\mu_2^{\mathrm{f}}$, and (iii) $\eta_{\min}=\mu_{\min}$. Thus, the problem is simplified to

$$\underset{\eta_1,\eta_2,P_{\mathrm{n}},P_{\mathrm{f}}}{\text{maximize}} \quad R_{\mathrm{n}}+R_{\mathrm{f}} \tag{5.24a}$$

$$\text{subject to} \quad (5.11\text{b})-(5.11\text{d}), \tag{5.24b}$$

$$\mathrm{P}_{\mathrm{mis}}^{\mathrm{n}},\mathrm{P}_{\mathrm{mis}}^{\mathrm{f}}\leq\varepsilon. \tag{5.24c}$$

Further, since $G_{1,1}^{\mathrm{i}}$ and $E_{k^{\mathrm{BS}},k^{\mathrm{i}}}^{\mathrm{i}}$ are functions of the beamwidth η_{i}, the misalignment probability is also a function of the beamwidth. In fact, given P_{tot} and h^{i}, the appropriate range for η_{i} to satisfy the misalignment probability constraints can be found using an off-line search. For (5.24c), we denote the range for η_1 and η_2 from the off-line search as $[\eta_1^{\mathrm{l}},\eta_1^{\mathrm{u}}]$ and $[\eta_2^{\mathrm{l}},\eta_2^{\mathrm{u}}]$, respectively. Then,

we can simplify Problem (5.24) into

$$\underset{\eta_1,\eta_2,P_{\rm n},P_{\rm f}}{\text{maximize}} \quad R_{\rm n}+R_{\rm f} \tag{5.25a}$$

$$\text{subject to} \quad (5.11\text{b})-(5.11\text{d}), \tag{5.25b}$$

$$\eta_1 \in [\eta_1^{\rm l},\eta_1^{\rm u}], \tag{5.25c}$$

$$\eta_2 \in [\eta_2^{\rm l},\eta_2^{\rm u}]. \tag{5.25d}$$

We can use the same approach as the one we use for Problem (5.11), i.e., we iterate between the power allocation and the beamwidth optimization. When allocating the power, we can still use Alg. 9. However, when optimizing the beamwidth, we need to consider Constraints (5.25c) and (5.25d). Same as Section 5.5, we optimize G_1 and G_2 instead of the beamwidths. Then, Constraints (5.25c) and (5.25d) are transformed into $\frac{A^2}{\eta_1^{\rm u^2}} \leq G_1 \leq \frac{A^2}{\eta_1^{\rm l^2}}$ and $\frac{A^2}{\eta_2^{\rm u^2}} \leq G_2 \leq \frac{A^2}{\eta_2^{\rm l^2}}$. The optimization problem is expressed as follows:

$$\underset{G_1,G_2}{\text{maximize}} \quad (1-\frac{\tau}{T})(R_{\rm n}+R_{\rm f}) \tag{5.26a}$$

$$\text{subject to} \quad \max\{\frac{A^2}{4\pi^2},\frac{A^2}{\eta_1^{\rm u^2}}\} \leq G_1 \leq \min\{G_2,\frac{A^2}{\eta_1^{\rm l^2}}\}, \tag{5.26b}$$

$$\max\{G_1,\frac{A^2}{\eta_2^{\rm u^2}}\} \leq G_2 \leq \min\{\frac{A^2}{\eta_{\rm min}^2},\frac{A^2}{\eta_1^{\rm l^2}}\}, \tag{5.26c}$$

$$R_{\rm n},R_{\rm f} \geq R_{\rm min}. \tag{5.26d}$$

Same as Alg. 10, we iterate between G_1 and G_2 to solve Problem (5.26). When G_2 is fixed, we can pick the optimal G_1 as: (i) if $\max\{\frac{A^2}{4\pi^2},\frac{A^2}{\eta_1^{\rm u^2}}\} < G_1^*$, then $G_1^{\rm opt} = \min\{G_1^*,G_2,\frac{A^2}{\eta_1^{\rm l^2}}\}$; (ii) if $\max\{\frac{A^2}{4\pi^2},\frac{A^2}{\eta_1^{\rm u^2}}\} > G_1^*$, then $G_1^{\rm opt} = \max\{\frac{A^2}{4\pi^2},\frac{A^2}{\eta_1^{\rm u^2}}\}$. When G_1 is fixed, we can perform a one-dimension search for G_2 to find the optimal $G_2^{\rm opt}$.

The algorithm to solve Problem (5.24) has the same complexity as Alg. 11 since we only add lower bounds and upper bounds for η_1 and η_2. Also, the proof of convergence is the same as that of Alg.

11.

5.7 Simulation Results

In this section, we present simulation results for the proposed two-step algorithm in comparison with the one-step beam alignment algorithm [49]. In the simulations, the minimum rate is assumed to be $R_{\min} = 0.5$ bits/s/Hz. The SNR is defined as SNR=$\frac{P_{\text{tot}}}{\sigma^2}$ with $\sigma^2 = 1$. The side-lobe level is equal to $\alpha = 0.1$. Further, we define $\omega_{\text{BS}} = \omega_{\text{n}} = \omega_{\text{f}} = 120^\circ$ [109]. Unless otherwise specified, we assume that the near user experiences -15 dB channel gain, the far user has a channel gain of -25 dB, the predefined misalignment probability factor is given by $\varepsilon = 10^{-4}$, the channel coherence time is given by $T/T_p = 3 \times 10^3$, the minimum beamwidth at the transceiver is $\eta_{\min} = \mu_{\min} = 3^\circ$, and P_{tot} is normalized to 1. We incorporate the beam alignment procedure into our two-step beamwidth design and adopt the beamforming gain models (5.6) and (5.7) in our simulations. The advantage of the joint optimization has been shown in [10]. In this chapter, we have conducted all the simulations for three different scenarios: (i) the proposed two-step beam alignment algorithm without misalignment, (ii) the proposed two-step algorithm with misalignment, and (iii) the one-step beam alignment algorithm without misalignment. In these algorithms, the unknown parameters are the BS's beamwidth and the users' beamwidth and power coefficients. In the two-step alignment without misalignment, the optimized parameters are attained by solving Problem (4.18). In the two-step algorithm with misalignment, the parameters are optimized by solving Problem (5.23). Finally, the unknown parameters of the one-step algorithm are obtained by solving the optimization problem (9) in [49]. Reasonably, we presume that the one-step algorithm without misalignment outperforms that of with misalignment in terms of the sum-rate. Hence, the sum-rate evaluation of the one-step algorithm with misalignment is ignored.

Fig. 5.4 demonstrates the sum-rate performance versus SNR. By increasing the SNR, the sum-rate of all algorithms goes up. However, it is revealed that the proposed algorithm outperforms

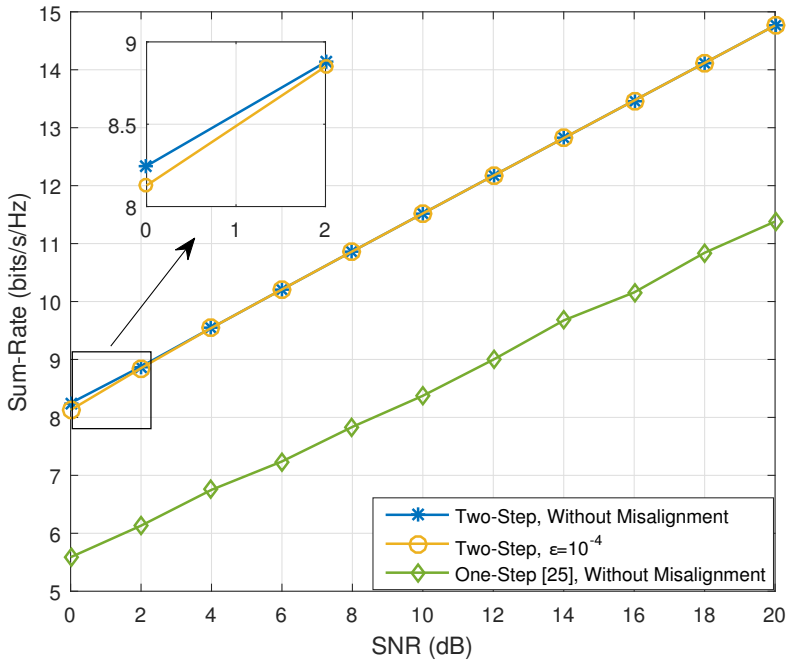

Figure 5.4: Performance of the sum-rate versus SNR with $h^{\mathrm{n}} = -15$ dB, $h^{\mathrm{f}} = -25$ dB, $T/T_p = 3 \times 10^3$, and $\eta_{\min} = \mu_{\min} = 3^\circ$.

the one-step algorithm by about 2.5 bits/s/Hz for low SNRs and about 3 bits/s/Hz for high SNRs. It is because, in our algorithm, the beam search is done for the $\mathrm{UE_n}$ first and then the $\mathrm{UE_f}$. In contrast, the one-step algorithm selects the beamwidth for $\mathrm{UE_n}$ and $\mathrm{UE_f}$, simultaneously [49]. Further, the two-step algorithm with misalignment achieves almost the same sum-rate as that of without misalignment except at SNRs between 0 dB to 2 dB. At very low SNR regions, say 0 dB, the algorithm experiences higher probability of misalignment at the second step due to the low channel gain of $\mathrm{UE_n}$. To compensate for the low channel gain, the algorithm generates a narrow beamwidth which increases the beamforming gain, although this requires spending more time on the beam alignment phase compared to the without beam misalignment. Thus, the achievable sum-rate decreases.

In Fig. 5.5, we conduct the simulation for the sum-rate versus the normalized channel coherence time, i.e., T/T_p. We assume that SNR is 0 dB. As the normalized channel coherence time increases, the sum-rate of all three algorithms increases as well. The main reason is that according to the time

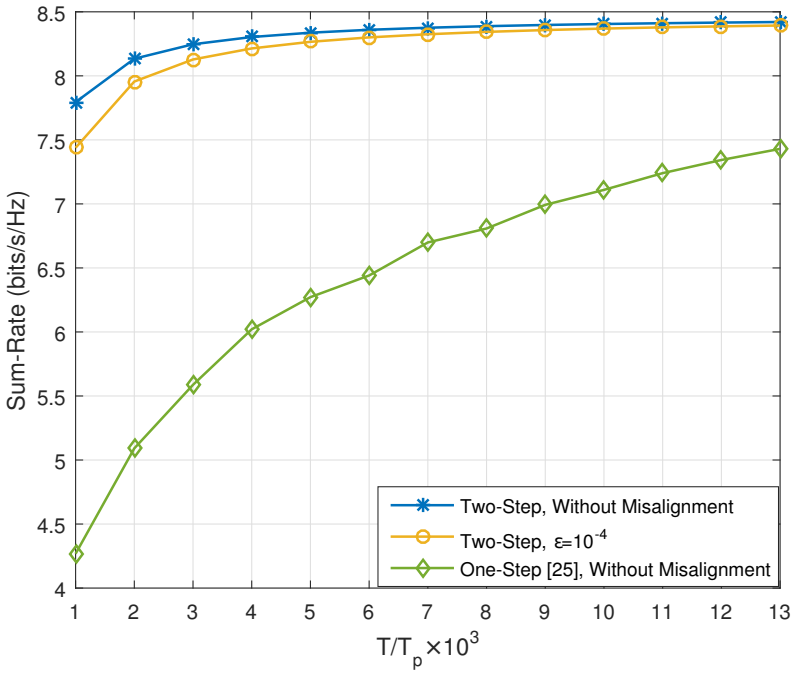

Figure 5.5: Performance of the sum-rate versus T/T_p for $h^{\mathrm{n}} = -15$ dB, $h^{\mathrm{f}} = -25$ dB, SNR$= 0$ dB, and $\eta_{\min} = \mu_{\min} = 3^{\circ}$.

frame in Fig. 5.3, when the channel coherence time is large, the algorithms have sufficient time to select a narrow beam and transmit their data for a longer period. Our two-step algorithm achieves a much higher sum-rate compared to the one-step algorithm although the sum-rate improvement depends on the coherence time. Especially, for a short channel coherence time, our algorithm is able to align the beam in a shorter time and transmit data for the remaining time which is far longer than that of the one-step algorithm. For very large normalized channel coherence times, the sum-rate of the one-step algorithm approaches to that of the two-step algorithm. Actually, by increasing the channel coherence time, the beam alignment time has less impact on the sum-rate. Therefore, it is expected that as the channel coherence time becomes larger and larger, the sum-rate of the one-step algorithm approaches to the sum-rate of our algorithm. The two-step algorithm without misalignment also outperforms the with misalignment algorithm. In particular, for a short normalized coherence channel time, the algorithm should ensure that all the constraints in Problem (5.23) are satisfied. In other words, the beam alignment during the two steps should

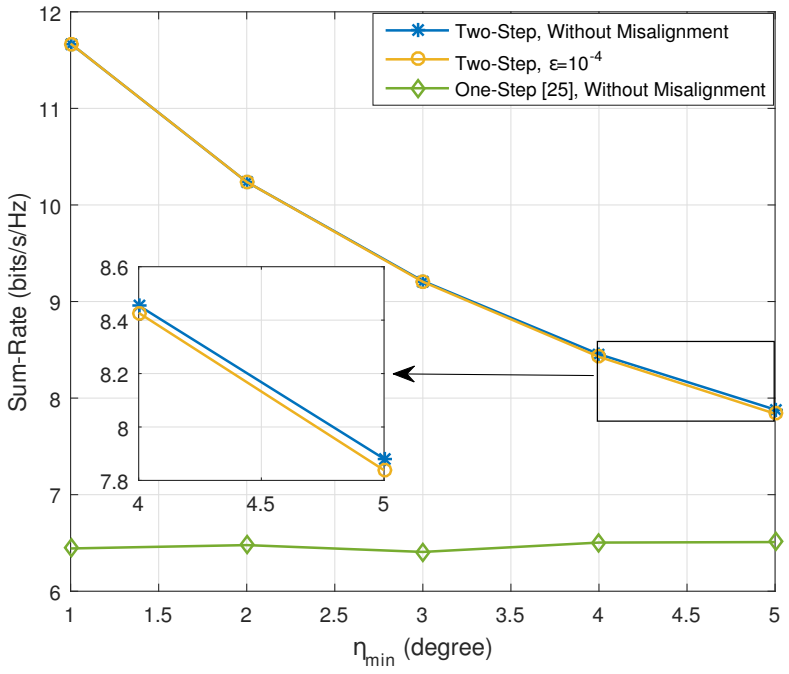

Figure 5.6: Performance of the sum-rate versus $\eta_{\min}$ for $h^{\mathrm{n}} = -15$ dB, $h^{\mathrm{f}} = -25$ dB, $T/T_p = 3 \times 10^3$, and SNR$=3$ dB.

result in a probability of misalignment less or equal to 10^{-4}. Thus, the beamwidths are selected sufficiently narrower than the without misalignment algorithm. This leads to assigning less time for the data transmission and achieving lower sum-rate. As the coherence time increases, the difference between the beam alignment times is negligible and the sum-rates are nearly the same.

Figure 5.6 demonstrates the sum-rate performance versus the minimum beamwidth. We notice that $\eta_{\min} = \mu_{\min}$. Also, we define SNR$=3$ dB to ensure that the two-step algorithm with misalignment, for the given minimum beamwidth range, satisfies Constraint (5.23c). For all the minimum beamwidths, our algorithm is able to achieve higher sum-rate. We observe that when the minimum beamwidth increases from 1 to 5, the sum-rate of the proposed two-step algorithm gradually decreases, while the sum-rate of the one-step algorithm approximately remains constant. This is because the one-step algorithm is not able to take advantages of the narrow minimum beamwidth. That is, the one-step algorithm should allocate much more time for beam alignment with a narrow beamwidth which is not an efficient solution in terms of the sum-rate. In contrast, our proposed

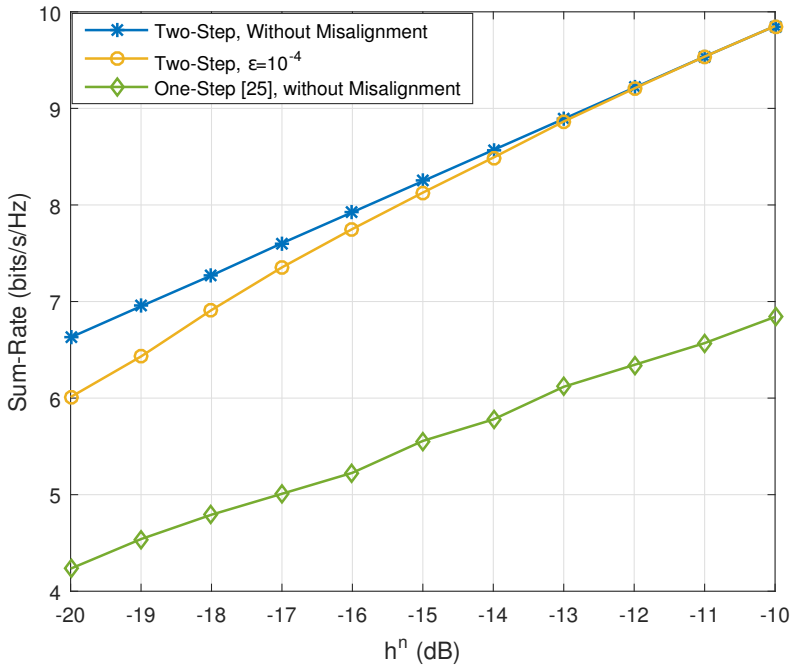

Figure 5.7: Performance of the sum-rate versus h^{n} for $h^{\mathrm{f}} = -25$ dB, $T/T_p = 3 \times 10^3$, SNR$=3$ dB, and $\eta_{\min} = 3^\circ$.

algorithm selects a fairly wide beamwidth at the first step of the beam alignment phase. Then, the algorithm chooses a narrower beamwidth in a short time within the range of the selected wide beamwidth. The narrow beamwidth remarkably increases the beamforming gain which leads to a higher sum-rate. When the minimum beamwidth increases, the algorithm has to choose a wider beamwidth which decreases the beamforming gain and consequently decreases the sum-rate. Further, since the proposed algorithm is able to choose a narrow beamwidth for $\eta_{\min} \in [1,3]$, the proposed algorithm with misalignment achieves exactly the same sum-rate as that of the without beamwidth algorithm. Due to the narrow beamwidth, the beamforming gain is sufficiently large and the beam misalignment probability is less than $\varepsilon = 10^{-4}$. For $\eta_{\min} \in (3,5]$, the sum-rate of the two-step algorithm with beam misalignment slightly degrades. The reason is that in order to keep the beam misalignment equal or less than $\varepsilon = 10^{-4}$, the proposed algorithm with beam misalignment allocates more time to the alignment phase to provide the necessary beamforming gain by selecting proper beamwidths.

In Fig. 5.7, we show the performance of the sum-rate versus the channel gain of $\mathrm{UE_n}$, i.e., h^{n}. As h^{n} increases and h^{f} is fixed, the sum-rate also increases for all algorithms. By making h^{n} larger, the sum-rate gap between our two-step algorithm and the one-step algorithm expands. When h^{n} is small, i.e., $\mathrm{UE_n}$ is located at the edge of the inner cell, the two-step algorithm spends more time on beam alignment at the first step. When $\mathrm{UE_n}$ approaches the BS, h^{n} becomes larger and the algorithm assigns less time for the beam alignment at the first step. Since $\mathrm{UE_f}$ is fixed, the time difference between the beam alignment at the second step when UE_n is close and far from the BS is negligible. Therefore, depending on the channel gain of $\mathrm{UE_n}$, the sum-rate gap between the proposed algorithm and the one-step algorithm is variable. Further, for a low channel gain, the sum-rate of the two-step with beam misalignment becomes lower than that of the without misalignment algorithm. To satisfy Constraint (5.23c), at the first step, the two-step with beam misalignment algorithm chooses narrower beamwidth compared to that of the without misalignment to provide the required beamforming gain. The same phenomenon may or may not occur for the second step of the beam alignment phase. Therefore, overall, the two-step with beam misalignment constraint assigns more time on beam alignment at low h^{n}. As the channel gain increases, Constraint (5.23c) can easily be met. In such a case, the same-rates are equal. It is worth mentioning that a similar simulation can be done for a variable h^{f} and constant h^{n}. Although, it is shown that the equivalent channel gain of two NOMA users with comparable channel gains is almost similar to the channel gain of the near user [6]. Thus, the sum-rate is highly affected by the channel gain of $\mathrm{UE_n}$.

Figure 5.8 evaluates the sum-rate performance versus the parameter ε. We perform simulation for $\eta_{\min} = 3^\circ$ and 4° and SNR$=$ 3 dB. For $\eta_{\min} = 3^\circ$, our solution for the optimization problem in (5.23) provides the optimized parameters while the misalignment probability can be as low as 10^{-8}. Note that the two-step algorithm without misalignment constraint may not satisfy the misalignment probability constraints on the x-axis. As it is mentioned before, for a small ε, the algorithm aims to choose a sufficiently narrow beamwidth which provides high beamforming gain. This causes the sum-rate degradation. As ε becomes smaller, the sum-rate gap becomes bigger. For $\eta_{\min} = 4^\circ$, our solution for the optimization problem in (5.23) exists for values up to $\varepsilon = 10^{-7}$.

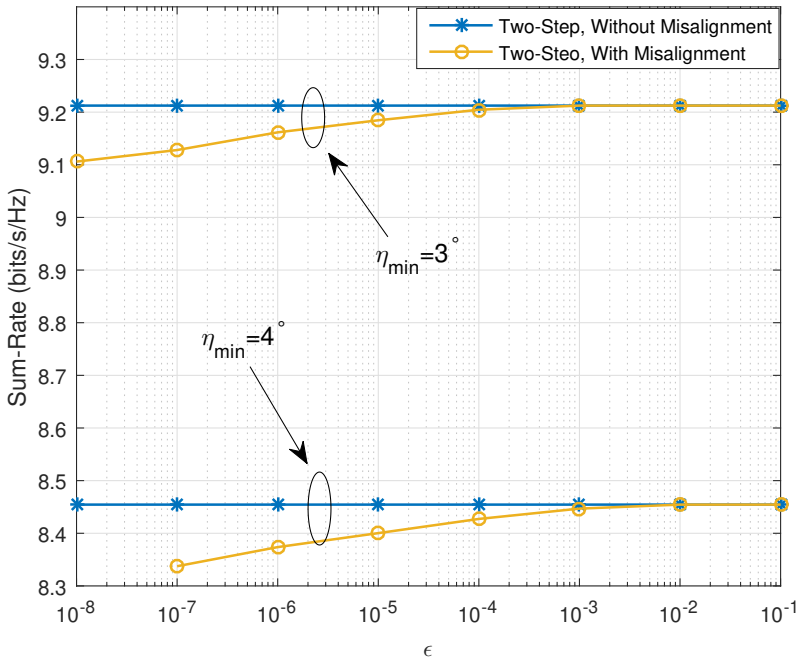

Figure 5.8: Performance of the sum-rate versus ε for $h^{\mathrm{n}} = -15$ dB, $h^{\mathrm{f}} = -25$ dB, $T/T_p = 3 \times 10^3$, SNR$=$ 3 dB, $\eta_{\min} = 3^\circ$ and 4°.

For smaller values, the unknown parameters cannot be determined. Therefore, as the minimum beamwidth increases, achieving lower misalignment probability becomes impractical.

5.8 Conclusion

In this chapter, we proposed a new two-step beam alignment algorithm for the analog beamforming mmWave-NOMA systems with constrained channel coherence time. The algorithm takes the advantages of NOMA, as such, first the beamwidth is selected for the near user and then for the far user. We defined an optimization problem to maximize the sum-rate. We then jointly optimized the beamwidths and the allocated powers while keeping the minimum required rate equal or higher than a predefined threshold. To evaluate the effect of beam misalignment due to the mmWave propagation characteristics and longer distance, we formulated another optimization problem, where we imposed extra constraints on the beamwidths to satisfy the misalignment probability requirements.

An off-line search was adopted to find the appropriate ranges for the beamwidths. Then, with the new ranges of the beamwidths, we used the similar joint optimization algorithm structure as we proposed for the problem without misalignment probability requirements to find the optimal power allocation and beamwidths. Our numerical results showed that the proposed algorithm with and without beam misalignment achieves higher sum-rate compared to the one-step beam alignment algorithm. Further, due to aligning beams in two steps, the proposed algorithm can take advantages of narrow beams. This leads to a higher beamforming gain and better sum-rate. Further, in the case of lower SNR, channel coherence time, and the lower channel of the near user, the proposed algorithm with beam misalignment marginally reduces the sum-rate.

Chapter 6

Reconfigurable Intelligent Surface Assisted mmWave UAV Wireless Cellular Networks

6.1 Introduction

Unmanned aerial vehicles (UAVs) have received increasing attention in the past decade because of their flexible, mobility, and fast deployment [132, 142]. There are several typical applications of UAV-assisted wireless networks including wireless sensor networks (WSNs) [45, 71, 44], caching aided wireless networks, cloud radio access networks (CRANs) [19], etc. Among these scenarios, UAV-assisted wireless cellular network is a promising technology to enable significantly enhanced UAV-ground communications [141]. In UAV-assisted wireless cellular networks, a UAV can serve as a flying base-station (BS), an aerial radio access point, and an aerial relay to expand wireless coverage and provide data transmission towards physical objects.

Millimeter wave (mmWave) communications are considered in the UAV wireless cellular networks to further enhance the available bandwidth and increase the data rate [132, 131]. The deployment and placement optimization of UAV operating in the mmWave band has been studied in the lit-

erature. In [146], a spatial interference channel model is established for UAV groups, and the expression of signal to interference plus noise ratio (SINR), which depends on codebook design and direction of arrival (DOA), is obtained. In [131], a joint optimization of the UAV-BS deployment and beamforming to maximize the achievable sum-rate in a multi-user mmWave-UAV system is proposed.

Although UAVs bring a lot of flexibility in deploying the networks, their high mobility and instability severely impair the quality of communication. One method to improve the reliability and quality of the UAV networks is to change the wireless scattering environment. Reconfigurable antennas have been proposed to change the transmission states and improve the performance [17, 42, 35]. Reconfigurable antennas for mmWave systems have been designed and provide similar benefits [52, 11]. Similar to reconfigurable antennas, reconfigurable intelligent surfaces (RIS) can intelligently configure the wireless environment to improve the transmission quality between the transmitter and receiver [129, 55]. However, different from reconfigurable antennas, which use active units to change the transmission state, RIS use passive units which only incur phase shift to the incident signal without power consumption. Moreover, RIS can help improve the channel quality when the line-of-sight (LoS) path is affected by physical obstacles or under harsh environments such as rains. In UAV networks, RIS can be implemented on building walls and remotely configured by central controllers to coherently direct the reflected radio waves towards specific users [84]. In [54], the UAV-BS link is assisted by the RIS. In [76], trajectory and beamforming are jointly designed for the scenario where one UAV serves one user.

In this chapter, we consider a RIS-assisted mmWave UAV wireless cellular network, where one UAV is serving several users with the help of multiple RIS. We propose a joint optimization problem, which considers the deployment, user scheduling, beamforming vector, and RIS phases to maximize the sum-rate. We consider the constraints of the minimum rate, the movement of the UAV, the analog beamforming, and the RIS phases. To solve this complex problem, we use an iterative method. In our method, we optimize one variable while fixing the other three variables.

When optimizing the deployment, we find the optimal position for the UAV by a sphere search. Then, we formulate an integer linear programming to find the best scheduling. We also design the analog beamforming vector by compensating the phases of the channel which combines the direct path and the RIS paths. When optimizing the RIS phases, we formulate a semi-definite programming to find the best phases. The proposed joint optimization outperforms the system without RIS assistance and the system without deployment optimization.

Our contributions are summarized as follows:

- We propose a new scenario where a UAV serves users with the help of multiple RIS.
- We further propose a joint optimization of the UAV deployment, the scheduling, the RIS phases and the beamforming to maximize the system throughput.

The reminder of the chapter is organized as follows. In Section 6.2, the system model is explained. In Section 6.3 the optimization problem is formulated. In Section 6.4, the joint algorithm is proposed. Numerical results are presented in Section 6.5. Section 6.6 concludes this chapter.

6.2 System model

We consider a single UAV, multi-RIS and multi-user scenario. The UAV functions as a flying BS to serve the ground users. The RIS are deployed on the ground and are controlled remotely by a central controller to help improve the communication quality between the UAV and ground users. The UAV is equipped with N_t antennas and each user is equipped with a single antenna. The total number of users is K. Each RIS is equipped with N_{RIS} reflecting elements and the total number of RIS is R.

In our system, we assume a quasi-static mobility model. That is, within a timeblock, the UAV

is static and then it can move one step. Each timeblock contains M' timeslots. We collect users' locations every M' timeslots.

At each timeblock, we fine tune the position of the UAV to fit in the locations of the users. In each timeblock, we aim to serve all the K users in M timeslots ($M < M'$). Note that the redundant $M' - M$ timelots are used for the data collection and UAV movement. To guarantee that all users can be served, we assume $K \leq M$.

The goal of our system is to maximize the system throughput in each timeblock. To do this, we need to jointly optimize the position of the UAV, design the scheduling, optimize the UAV beams, and adjust the RIS phases. The system model is shown in Fig. 6.1, where we illustrate a 2-RIS and 5-user system. In Timeslot 3 of the timeblock, the UAV serves User 2 with the help of RIS 1 and RIS 2. In different timeblock, the UAV will fine tune its location to optimize the performance within its users.

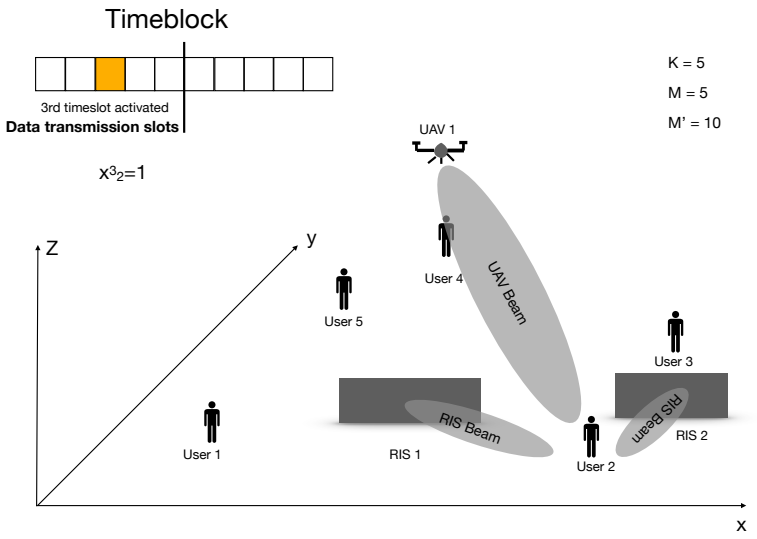

Figure 6.1: System model

6.3 Problem formulation

6.3.1 Channel model

UAV channel

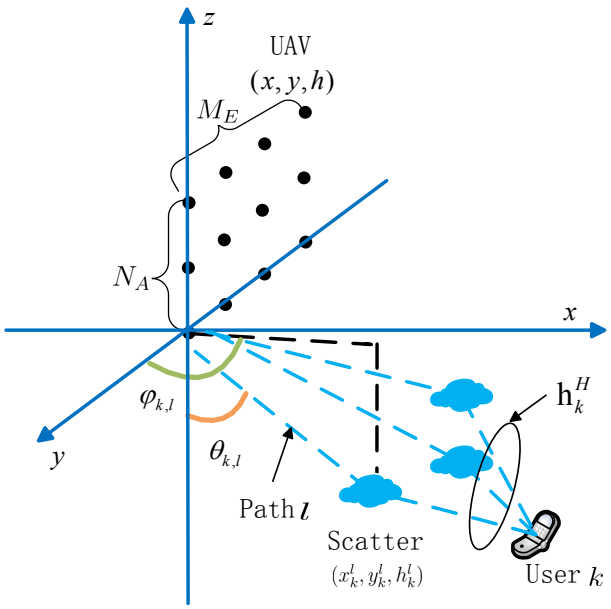

Figure 6.2: channel model with three clusters, where the dashed blue lines are the paths formed by reflecting scatters.

In our scenario, we assume the UAV is carrying a uniform planar array (UPA) with one RF chain which operates on mmWave band. A multi-path channel (MPC) model is adopted. As Fig. 6.2 illustrates, we denote $\mathbf{h}_k$ as the channel vector, then it can be expressed as

$$\mathbf{h}_k^H = \sqrt{\frac{N_t}{L_k}} \sum_{l=1}^{L_k} a_{k,l} \alpha(\theta_{k,l}, \varphi_{k,l})^H, \tag{6.1}$$

where $a_{k,l}$ is the channel gain of the l^{th} path from User k to the UAV, $\theta_{k,l}$ and $\varphi_{k,l}$ are the elevation steering angle and azimuth angle, respectively, of the l^{th} path from User k to the UAV, and L_k is the total number of paths for User k to the UAV. Each path is formed through a scatter. Parameter $\alpha(\theta_{k,l})$ is the steering vector function for the UPA. For an $M_E \times N_A$ ($N_t = M_E N_A$) UPA, the steering

vector is defined as

$$\begin{aligned}\alpha(\theta_{k,l},\varphi_{k,l}) &= \tfrac{1}{\sqrt{N_t}}[1,...,e^{j\pi\sin(\theta_{k,l})f_{m,n}(\varphi_{k,l})},\\ &\quad ...,e^{j\pi\sin(\theta_{k,l})f_{M_E,N_A}(\varphi_{k,l})}]^T,\end{aligned} \tag{6.2}$$

where $f_{n,m}(\varphi_{k,l}) \triangleq (m-1)\cos(\varphi_{k,l})+(n-1)\sin(\varphi_{k,l})$, $0 \le m \le M_E-1$ and $0 \le n \le N_A-1$. The steering angles $\theta_{k,l}$ and $\varphi_{k,l}$ depend on the location of the UAV and the location of the scatter which forms the path l. We denote (x,y,h) as the location of the UAV and (x_k^l, y_k^l) as the location of the scatter of the l^{th} path for User k. Then, $\theta_{k,l}$ and $\varphi_{k,l}$ can be calculated as

$$\begin{cases} \theta_{k,l} = \arctan(\frac{\sqrt{(x_k^l-x)^2+(y_k^l-y)^2}}{h}) \\ \varphi_{k,l} = \arctan(\frac{y_k^l-y}{x_k^l-x}) \end{cases} \tag{6.3}$$

Model (6.1) describes the Non-LoS scenario between User k and the UAV. In most cases, an LoS path exists as well. When there is an LoS path, the channel becomes

$$\mathbf{h}_k^H = \sqrt{N_t}a_k\alpha(\theta_k,\varphi_k)^H, \tag{6.4}$$

where a_k is the channel gain of the LoS path from User k to the UAV, θ_k and φ_k are the elevation steering angle and azimuth angle of the LoS path from User k to the UAV, respectively.

The LoS path between the UAV and the serving user can be blocked if the propagation environment contains physical obstacles. The probability of existing an LoS component can be described as a function of the angle ξ_k as follows

$$P_{\text{LOS}}(\xi_k) = \frac{1}{1+a\exp(-b(\xi_k-a))}, \tag{6.5}$$

where a and b are the positive modeling parameters depending on the propagation environment, e.g., rural, urban, or dense urban. ξ_k is calculated by $\xi_k = \arctan(h/D_k)$ with the horizontal distance

from the UAV to User k denoted as $D_k = \sqrt{(x-x_k)^2+(y-y_k)^2}$. The probability of existing an LoS component increases as the elevation angle increases, and it approaches 1 when h is large enough.

RIS channel

We denote the channel between RIS r and the UAV as $\mathbf{G}_r$. The channel between RIS r and User k is denoted by $\mathbf{h}_k^{r,H}$. We use the widely adopted MPC model to model $\mathbf{G}_r$ and $\mathbf{h}_k^{r,H}$. For $\mathbf{h}_k^{r,H}$, it can be expressed as

$$\mathbf{h}_k^{r,H} = \sqrt{\frac{N_{\text{RIS}}}{L_k^r}} \sum_{l=1}^{L_k^r} a_{k,l}^r \alpha(\theta_{k,l}^r, \varphi_{k,l}^r)^H, \tag{6.6}$$

where $a_{k,l}^r$ is the channel gain of the l^{th} path from RIS r to User k. $\alpha(\theta_{k,l}^r, \varphi_{k,l}^r)$ is the steering vector using the same model as (6.2).

For $\mathbf{G}_r$, it can be expressed as

$$\mathbf{G}_r = \sqrt{\frac{N_t N_{\text{RIS}}}{L_r}} \sum_{l=1}^{L_r} a_{r,l} \alpha^{\mathbf{r}}(\theta_{\mathbf{r},\mathbf{l}}^{\mathbf{r}}, \varphi_{\mathbf{r},\mathbf{l}}^{\mathbf{r}}) \alpha^t(\theta_{r,l}^t, \varphi_{r,l}^t)^H, \tag{6.7}$$

where $a_{r,l}$ is the channel gain of the l^{th} path from the UAV to RIS r. $\alpha^t(\theta_{r,l}^t, \varphi_{r,l}^t)$ and $\alpha^r(\theta_{r,l}^r, \varphi_{r,l}^r)$ are the transmitting steering vector and receiving steering vector, respectively.

The overall channel between User k and the UAV via RIS r can be expressed as

$$\mathbf{h}_{k,r}^H = \mathbf{h}_k^{r,H} \Theta_r \mathbf{G}_r, \tag{6.8}$$

where $\Theta_r = \text{diag}(e^{j\theta_{r,1}}, ..., e^{j\theta_{r,N_{\text{RIS}}}})$ is the phase-shift matrix of the r^{th} RIS. $\theta_{r,m} \in [0, 2\pi]$ denotes the phase shift associated with the m^{th} passive element of the r^{th} RIS.

6.3.2 Scheduling

In each timeblock, we have 2 rules for scheduling: (i) in one timeslot, the UAV can only serve at most one user; and (ii) across all timeslots, all K users should be scheduled at least once.

To describe the process of scheduling, we denote the binary variable $x_k^m \in \{0,1\}$ to indicate whether User k is scheduled by the UAV in Timeslot m, i.e.,

$$x_k^m = \begin{cases} 1, \text{if User } k \text{ is scheduled in Timeslot } m \\ 0, \text{otherwise} \end{cases} \tag{6.9}$$

For Rule (i), we have

$$0 \leq \sum_{k=1}^{K} x_k^m \leq 1 \tag{6.10}$$

For Rule (ii), we have

$$\sum_{m=1}^{M} x_k^m \geq 1 \tag{6.11}$$

6.3.3 UAV Beamforming and RIS Reflecting

We denote $\mathbf{w}_k$ as the beamforming vector from the UAV to User k with the constant-modulus constraint $|[\mathbf{w}_k]_m| = \frac{1}{\sqrt{N_t}}, m = 1, ..., N_t$. Then, at Timeslot m, the received signal from the UAV to User k is

$$y_k^m = x_k^m \sqrt{P}(\mathbf{h}_k^H + \sum_{r=1}^{R} \mathbf{h}_k^{r,H} \Theta_r \mathbf{G}_r) \mathbf{w}_k \mathbf{s}_k + n_k. \tag{6.12}$$

The achievable data rate from User k to the UAV at Timeslots m can be expressed as

$$R_k^m = \log_2(1 + \frac{x_k^m P|(\mathbf{h}_k^H + \sum_{r=1}^{R} \mathbf{h}_k^{r,H} \Theta_r \mathbf{G}_r)\mathbf{w}_k|^2}{\sigma^2}), \tag{6.13}$$

where σ^2 is the power of Gaussian white noise at User k, and P is the total transmission power at the UAV.

6.3.4 Joint optimization

In each timeblock, we want to maximize the thoughput of the system. The optimal way is to find the best position of the UAV in the entire space. However, this is impractical since the UAV cannot rapidly move from one position to another position far away within a short time. To this end, we assume that the UAV can only move one small step in each timeblock or stay static. This means $|\mathbf{p} - \mathbf{p}_{\text{pre}}| = d$ or 0, where $\mathbf{p} \triangleq (x, y, h)$ is the position of the UAV and $\mathbf{p}_{\text{pre}}$ is the position of the UAV in the previous timeblock. The parameter d is decided by the UAV's energy constraint. Then, the optimization problem is described as follows:

$$\underset{\substack{\{\mathbf{p}\},\{x_k^m\},\\ \{\mathbf{w}_k\},\{\Theta_r\}}}{\text{maximize}} \quad \sum_{m=1}^{M}\sum_{k=1}^{K} R_k^m \tag{6.14a}$$

$$\text{subject to} \quad (6.10)-(6.11) \quad , \tag{6.14b}$$

$$|\mathbf{p}-\mathbf{p}_{\text{pre}}| = d \text{ or } 0 \quad , \tag{6.14c}$$

$$\sum_{m=1}^{M} R_k^m \geq \gamma_k \quad , \tag{6.14d}$$

$$|[\mathbf{w}_k]_m| = \frac{1}{\sqrt{N_t}}, m = 1,...,N_t \quad , \tag{6.14e}$$

$$\Theta_r = \text{diag}(e^{j\theta_{r,1}},...,e^{j\theta_{r,N_{\text{RIS}}}}). \tag{6.14f}$$

Constraint (6.14d) is the constraint for the minimum data rate for each user. Constraint (6.14e) is for analog beamforming. Constraint (6.14f) is for the RIS phases.

6.4 Solution

To solve Problem (6.14), we will iterate among the deployment, scheduling, beamforming, and RIS phases. When we optimize one variable, we fix the other three variables.

6.4.1 Deployment

When optimizing the deployment, we fix the scheduling, the beamforming vector, and the RIS phases. We denote the index of the scheduled user at Timeslot m by i_m, then the sub-problem for

deployment can be expressed as

$$\underset{\mathbf{p}}{\text{maximize}} \quad \sum_{m=1}^{M} R_{i_m} \tag{6.15a}$$

$$\text{subject to} \quad |\mathbf{p}-\mathbf{p}_{\text{pre}}| = d \text{ or } 0, \tag{6.15b}$$

$$R_{i_m} \geq \gamma_m \quad , \tag{6.15c}$$

where $R_{i_m} = \log_2(1+\frac{P|\tilde{\mathbf{h}}_{i_m}^H \mathbf{w}_{i_m}|^2}{\sigma^2})$, $\tilde{\mathbf{h}}_{i_m}^H = \mathbf{h}_{i_m}^H + \sum_{r=1}^{R} \mathbf{h}_{i_m}^{r,H} \Theta_r \mathbf{G}_r$, and γ_m is calculated by $\gamma_{i_m}/(\sum_{n=1}^{M} x_{i_m}^n)$.

The position of the moved UAV can be expressed as

$$\mathbf{p} = \mathbf{p}_{\text{pre}} + d[\sin\theta_{\text{mv}}\cos\varphi_{\text{mv}}, \sin\theta_{\text{mv}}\sin\varphi_{\text{mv}}, \cos\theta_{\text{mv}}]^T, \tag{6.16}$$

where θ_{mv} and φ_{mv} are the movement elevation angle and azimuth from $\mathbf{p}$ to $\mathbf{p}_{\text{pre}}$, respectively. To find the best position for the UAV, we perform a sphere search for $\mathbf{p}$ based on $\mathbf{p}_{\text{pre}}$, i.e., we find the optimal moving direction from $\mathbf{p}_{\text{pre}}$ to $\mathbf{p}$. The detailed algorithm is described in Alg. 12.

Algorithm 12: Best deployment

Input:
Searching step size Δ, $\mathbf{p}_{\text{pre}}$ and the sum-rate of the previous timeblock $R_{\text{sum}}^{\text{pre}}$;
$\mathbf{p}^{\text{opt}} \leftarrow \mathbf{p}_{\text{pre}}$, $R_{\text{sum}}^{\text{max}} \leftarrow R_{\text{sum}}^{\text{pre}}$;
for $\theta_{\text{mv}} = 0 : \Delta : 2\pi$ **do**
  for $\varphi_{\text{mv}} = 0 : \Delta : 2\pi$ **do**
    Update $\mathbf{p}$ and R_{sum};
    if $R_{\text{sum}} \geq R_{\text{sum}}^{\text{max}}$ and Constraint (6.15c) is satisfied **then**
      $\mathbf{p}^{\textbf{opt}} \leftarrow \mathbf{p}$, $R_{\textbf{sum}}^{\textbf{max}} \leftarrow R_{\textbf{sum}}$;
    end if
  end for
end for
Return the optimal $\mathbf{p}^{\text{opt}}$.

6.4.2 Scheduling

When the deployment, beamforming, and RIS phases are fixed, the throughput maximization with respect to the scheduling is formulated as

$$\underset{\mathbf{x}}{\text{maximize}} \quad \sum_{k=1}^{K}[R_k \sum_{m=1}^{M} x_k^m] \tag{6.17a}$$

$$\text{subject to} \quad (6.10)-(6.11), \tag{6.17b}$$

$$\sum_{m=1}^{M} x_k^m R_k \geq \gamma_k \quad , \tag{6.17c}$$

where $\mathbf{x} = \{x_k^n[m]\}$ is the set of scheduling indicators and $R_k = \log_2(1 + \frac{P|\tilde{\mathbf{h}}_k^H \mathbf{w}_k|^2}{\sigma^2})$. Problem (6.17) is an integer linear programming (ILP) since all variables are binary and all constraints are linear. Optimizers such as the Gurobi [91] can be employed to obtain the optimal solution to Problem (6.17).

6.4.3 Beamforming vector optimization

Given the scheduling order, the optimal deployment and the RIS phases, we can simplify the beamforming vector design problem into

$$\underset{\{\mathbf{w}_k\}}{\text{maximize}} \quad \sum_{m=1}^{M} R_{i_m} \tag{6.18a}$$

$$\text{subject to} \quad R_{i_m} \geq \gamma_m \quad , \tag{6.18b}$$

$$|[\mathbf{w}_k]_m| = \frac{1}{\sqrt{N_t}}, m = 1, ..., N_t. \tag{6.18c}$$

To further simplify the problem, we decouple Problem (6.18) by timeslots. At Timeslot m, we need

to design the beamforming vector for the scheduled user according to

$$\underset{\mathbf{w}_{i_m}}{\text{maximize}} \quad R_{i_m} \tag{6.19a}$$

$$\text{subject to} \quad R_{i_m^n} \geq \gamma_{i_m} \quad , \tag{6.19b}$$

$$|[\mathbf{w}_{i_m}]_t| = \frac{1}{\sqrt{N_t}}, t = 1, ..., N_t. \tag{6.19c}$$

Since we are maximizing the rate R_{i_m} in (6.19), we can drop the lower-bound constraint (6.19b) to simplify the problem. Note that maximizing R_{i_m} is equivalent to maximizing the power of the receiving signal, since there is no interference. Then, we can re-formulate the optimization problem as

$$\underset{\mathbf{w}_{i_m}}{\text{maximize}} \quad |\tilde{\mathbf{h}}_{i_m}^H \mathbf{w}_{i_m}|^2 \tag{6.20a}$$

$$\text{subject to} \quad |[\mathbf{w}_{i_m}]_t| = \frac{1}{\sqrt{N_t}}, t = 1, ..., N_t, n = 1, ..., N. \tag{6.20b}$$

According to the Cauchy-Schwartz inequality, the optimal $\mathbf{w}_{i_m}^{\text{opt}}$ is the one which compensates the phases of the channel, i.e.,

$$\mathbf{w}_{i_m}^{\text{opt}} = \frac{1}{\sqrt{N_t}} e^{j \arg(\tilde{h}_{i_m})} \tag{6.21}$$

6.4.4 RIS phase design

When designing the RIS phases, similar to the beamforming vector designing, we decompose the problem by timeslots. At Timeslot m, we reformulate the optimization problem as

$$\underset{\{\Theta_r\}}{\text{maximize}} \quad \sum_r^R \text{Tr}(\mathbf{G}_r^H \Theta_r^H \mathbf{h}_{i_m}^r \mathbf{h}_{i_m}^{r,H} \Theta_r \mathbf{G}_r \mathbf{W}_{i_m}) \tag{6.22a}$$

$$\text{subject to} \quad \Theta_r = \text{diag}(e^{j\theta_{r,1}}, ..., e^{j\theta_{r,N_{\text{RIS}}}}), \tag{6.22b}$$

where we ignore the term $\mathbf{w}_{i_m}^H \mathbf{h}_{i_m} \mathbf{w}_{i_m}$, since it is a constant term for fixed $\mathbf{w}_{i_m}$. To simplify the optimization problem, we define $\mathbf{H}_{i_m^n}^{\mathrm{RIS},H} = [\mathbf{h}_{i_m^n}^{1,H}, ..., \mathbf{h}_{i_m^n}^{R,H}] \in \mathscr{C}^{1 \times RN_{RIS}}$, $\blacksquare = \mathrm{diag}(\Theta_1, ..., \Theta_R) \in \mathscr{C}^{RN_{RIS} \times RN_{RIS}}$ and $\mathbf{G}_{\mathrm{RIS}} = [\mathbf{G}_1^H, ..., \mathbf{G}_R^H]^H \in \mathscr{C}^{RN_{RIS} \times N_t}$. Then, the objective in (6.22) can be reformulated as $\mathrm{Tr}(\blacksquare^H \mathbf{H}_{\mathrm{RIS}} \mathbf{H}_{i_m^n}^{\mathrm{RIS},H} \blacksquare \mathbf{G}_{\mathrm{RIS}} \mathbf{W}_{i_m} \mathbf{G}_{\mathrm{RIS}}^H)$. Further, by defining $\mathbf{v}_\Theta$ as the vector collecting the diagonal elements of $\blacksquare$, according to [144], we can transform the objective into $\mathbf{v}_\Theta^H \mathbf{E} \mathbf{v}_\Theta$, where $\mathbf{E} = (\mathbf{H}_{\mathrm{RIS}} \mathbf{H}_{i_m^n}^{\mathrm{RIS},H}) \odot (\mathbf{G}_{\mathrm{RIS}} \mathbf{W}_{i_m} \mathbf{G}_{\mathrm{RIS}}^H)$. Operator $\odot$ represents the Hadamard product.

We can transform the RIS phase design problem into a semi-definite program (SDP) as follows:

$$\underset{\mathbf{V}_\Theta}{\text{maximize}} \quad \mathrm{Tr}(\mathbf{E}\mathbf{V}_\Theta) \tag{6.23a}$$

$$\text{subject to} \quad [\mathbf{V}_\Theta]_{t,t} = 1, t = 1, ..., RN_{RIS}, \tag{6.23b}$$

$$\mathbf{V}_\Theta \succeq 0 \quad , \tag{6.23c}$$

$$\mathrm{rank}(\mathbf{V}_\Theta) = 1 \quad , \tag{6.23d}$$

where $\mathbf{V}_\Theta \triangleq \mathbf{v}_\Theta \mathbf{v}_\Theta^H$. To deal with the rank-one constraint in (6.23), we introduce the semi-definite programming relaxation (SDR) technique by dropping the rank-one constraint to solve the optimization problem below

$$\underset{\mathbf{V}_\Theta}{\text{maximize}} \quad \mathrm{Tr}(\mathbf{E}\mathbf{V}_\Theta) \tag{6.24a}$$

$$\text{subject to} \quad [\mathbf{V}_\Theta]_{t,t} = 1, t = 1, ..., RN_{RIS}, \tag{6.24b}$$

$$\mathbf{V}_\Theta \succeq 0 \tag{6.24c}$$

Problem (6.24) provides an upper bound for Problem (6.23) and its optimal solution can be found by standard tools of mathematical programming such as CVX [41]. Note that Problem (6.24) is the relaxed version of Problem (6.23), which means we cannot guarantee $\mathbf{V}_\Theta^{\mathrm{opt}}$ is rank-one. When the rank of $\mathbf{V}_\Theta^{\mathrm{opt}}$ is larger than one, we cannot recover $\mathbf{v}_\Theta^{\mathrm{opt}}$ from $\mathbf{V}_\Theta^{\mathrm{opt}}$ straightforwardly. In such cases, we use the same technique as [61], in which we generate a set of candidates which obey the

distribution of $\mathscr{CN}(\mathbf{0}, \mathbf{V}_{\Theta}^{\text{opt}})$. Then, we normalize the vector elements of each candidate. At last, we pick the normalized candidate $\mathbf{v}_{\Theta}^{\text{opt}}$ which maximizes $\text{Tr}(\mathbf{E}\mathbf{v}_{\Theta}^{\text{opt}}\mathbf{v}_{\Theta}^{\text{opt}^H})$.

6.4.5 Joint optimization

The details of the joint optimization algorithm are described in Alg. 13. Obviously, Alg. 13 converges, since we generate a monotonically increasing sequence with an upper bound (the maximum sum-rate).

Algorithm 13: Joint Optimization for RIS-assisted mmWave UAV networks

1: Set the sum-rate $R_{\text{sum}}[-1] \leftarrow 0$, the maximal iteration number $k_{max} \leftarrow 1000$ and the convergence threshold $\varepsilon \leftarrow 10^{-3}$;
2: Choose feasible start points $\mathbf{p}^{\text{opt}}[0]$, $\mathbf{x}^{\text{opt}}[0]$, $\{\mathbf{w}_k^{\text{opt}}[0]\}$, and $\blacksquare^{\text{opt}}[0]$;
3: **while** $R_{\text{sum}}[k] - R_{\text{sum}}[k-1] \geq \varepsilon R_{\text{sum}}[k-1]$ and $k \leq k_{max}$ **do**
4: $k \leftarrow k+1$;
5: Use Alg. 12 to find the optimal deployment;
6: Solve (6.17) to obtain the optimal scheduling;
7: Obtain the optimal beamforming vector by (6.21);
8: Solve (6.23) to get the optimal RIS phases;
9: Calculate $R_{\text{sum}}[k]$;
10: **end while**
11: Return $\mathbf{p}^{\text{opt}}$, $\mathbf{x}^{\text{opt}}$, $\{\mathbf{w}_k^{\text{opt}}\}$, and $\blacksquare^{\text{opt}}$.

6.5 Simulation results

In this section, we provide some simulation results for our proposed joint optimization algorithm. We consider a scenario where a UAV serves 4 users in 10 timeslots with the assistance of 2 RIS. The UAV serves the users using a mmWave carrier. We choose 28 GHz as the carrier's frequency, since 28 GHz is a typical frequency band in urban areas [100]. The parameters in Eq. (6.5) are set as $a = 11.95$ and $b = 0.14$. The channel gain coefficient a_k^n is generated according to a complex

Gaussian distribution $a_k^n \sim \mathscr{CN}(0, 10^{-0.1\kappa})$, where $\kappa = e + 10f\log_{10}(s) + \eta$. Parameter s is the distance between the UAV and the user. We calculate s according to the UAV's position in the previous timeblock. Parameters f and e are constants and $\eta \sim \mathcal{N}(0, \sigma_\eta)$. When the channel is an LoS channel, $f = 2$, $e = 61.4$ and $\sigma_\eta = 5.8$. When the channel is a Non-LoS channel, $f = 2.92$, $e = 72$ and $\sigma_\eta = 8.7$.

In our simulations, the positions of the RIS are $(10, 10, 0)$ and $(40, 40, 0)$. The UAV and the RIS are all equipped with a 64 (16×4) antenna array. We set the amplitude of the moving step for the UAV to be 5 meters. The initial position $\mathbf{p}[0] = (25, 25, 50)$. At each timblock, we randomly generate the positions of the users. The total number of timeblocks is 1000. We use the averaged sum-rate and minimum rate per timeblock as the measurements of our system.

In Figs. 6.3 and 6.4, we compare the sum-rate and the minimum rate among the system which uses our proposed joint optimization method, the system that optimizes the deployment without the assistance of RIS, the system which optimizes the beamforming vector and RIS phases but not the deployment, and the system without the best deployment and the optimal beamforming vector and RIS phases. The power of the Gaussian white noise is set to be -100 dBm and the minimum rate constraint is 1 bit/Hz. The results show that our joint optimization method brings great gains over the other three systems in both the sum-rate and the minimum rate.

6.6 Conclusion

In this chapter, we jointly optimized the deployment, user scheduling, beamforming vector, and RIS phases in a RIS assisted UAV wireless network. To solve the problem, we iterated among the 4 variables. While optimizing one variable, we fixed the other 3 variables. For the deployment, we found the optimal position by a sphere search. Then, we formulated an integer linear programming to find the best scheduling. We also designed the analog beamforming vector by compensating the

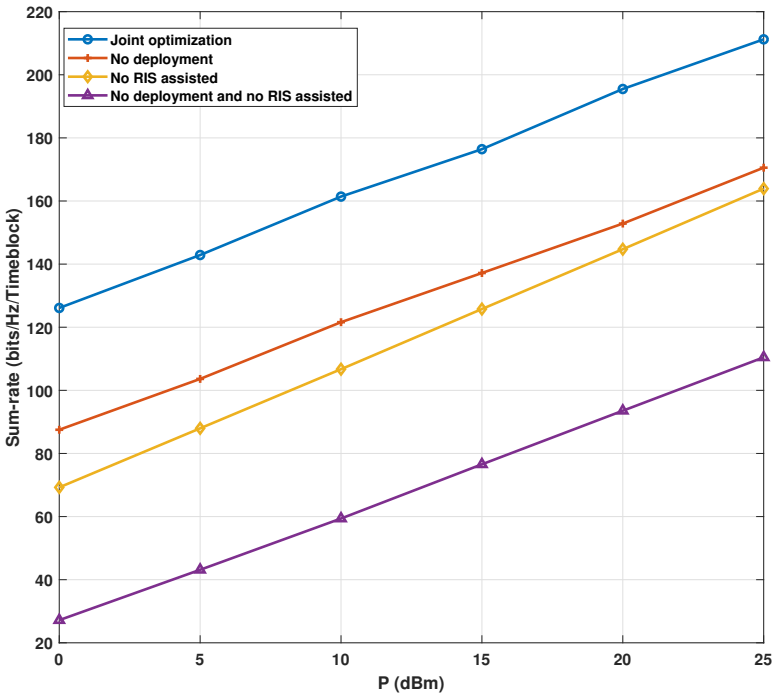

Figure 6.3: Sum-rate comparison

phases of the channel. When optimizing the RIS phases, we formulated a semi-definite programming to find the best phases. The proposed joint optimization outperforms the system without RIS assistance, and the system without deployment optimization.

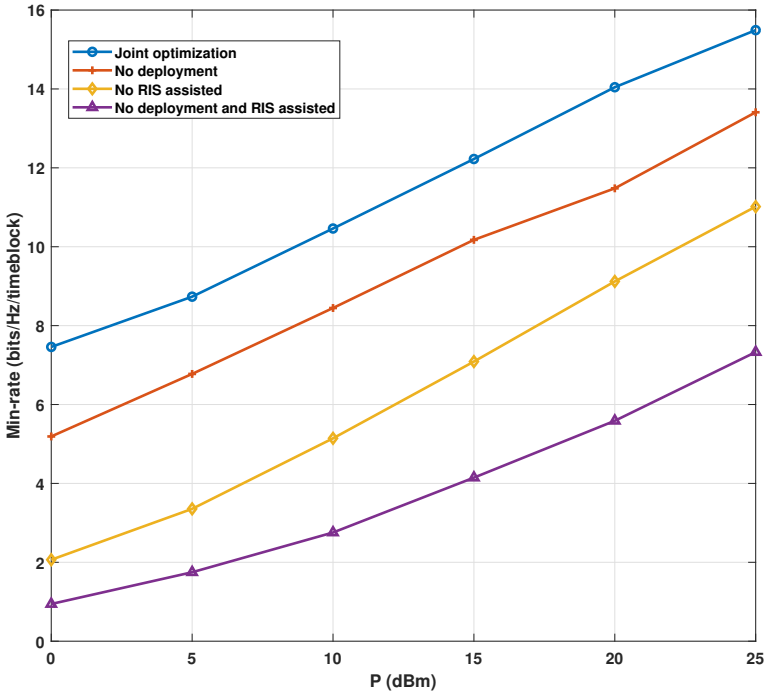

Figure 6.4: Minimum rate comparison

Chapter 7

Conclusions and future work

In this chapter, we conclude each chapter and we talk about the future work directions.

7.1 Conclusions and summaries

Table 7.1 summarize the main contents of this dissertation. In general, the dissertation can be divided into two categories: the beamforming algorithms design and the combination of mmWave beamforming and other technologies. For the beamforming algorithm design, we first designed an analog beamforming algorithm for multi-user system to suppress the interference and maximize the beamforming gain at the same time. Further, to support multi-stream transmission and extend the coverage range, we proposed a hybrid precoding algorithm for mmWave relay networks. Moreover, to achieve higher capacity and larger coverage, we combined mmWave with other technologies, such as NOMA and UAV networks. For mmWave-NOMA scenarios, we proposed a joint power and beamwidth algorithm to improve the system throughput. We also propsed a two-step beam alignment algorithm to reduce the training time. For UAV networks, we jointly optimize the deployment, the scheduling, the beamforming, and the RIS phases to greatly improve the system

throughput.

In particular, in chapter 2, we proposed an analog beamforming scheme which strikes a balance between the beamforming gain and the inter-user interference. We formulated an MOP that maximizes the beamforming gain and minimizes the interference at the same time. The weighted-sum method was used to transform the MOP into an SOP and the SDP was adopted to make the constant-magnitude constraints for the analog beamforming tractable. Furthermore, to alleviate the effects of the channel estimation and feedback quantization errors, we designed a robust beamforming scheme to overcome the channel uncertainty. A probabilistic constraint was used and an MOP similar with the non-robust beamforming scheme was formulated. For the non-robust case, simulation results showed that the proposed beamformer provides a better balance between the beamforming gain and the inter-user interference compared with other analog beamformers in the high SNR region. For the robust case, the simulation results demonstrated the highest robustness of our beamforming scheme against channel errors.

In chapter 3, we considered mmWave AF relay networks in the domain of massive MIMO. We designed the hybrid precoding/combining matrices for the source node, the relay node, and the destination node. We first performed the RF processing to decompose the channel into parallel sub-channels by compensating the phase of each eigenmode of the channel. Given the RF processing matrices, we designed the baseband matrices to maximize the mutual information. The baseband processing is divided into two parts. We first jointly designed the source node and the relay node by making use of the equivalence between maximizing the mutual information and the WMMSE. Given the optimal baseband source and relay filters, we implemented MMSE-SIC for baseband destination node to obtain the maximal mutual information. In addition, a robust hybrid precoding/combining design was proposed for the imperfect CSI. Simulation results show that our algorithm achieves better performance with lower complexity compared with other algorithms in the literature.

In chapter 4, NOMA is incorporated into mmWave hybrid beamforming systems. We also con-

sider the beam-training time because of the limited channel coherence time in mmWave directional communications. By combining the exhaustive search and tone-based beam-training algorithms, a new beam-training algorithm is employed. The formulated sum-rate expression consists of the channel coherence time and beam-training time. To maximize the sum-rate, a joint power allocation and beamwidth control optimization problem is solved by an algorithm which iterates between the power allocation and the beamwidth optimization. The non-convex power allocation is solved by the log-reformulation and SPCA. The beamwidth optimization is solved by iterating between the two clusters. A boundary-search algorithm is proposed to reduce the search complexity for the beamwidth in each cluster. The numerical results demonstrate that an efficient power allocation and beam-training time can lead to higher sum-rates compared to the conventional mmWave-NOMA without optimized parameters, NOMA-OMA, and OMA. The only exception is that for a short channel coherence time and high SNR, the optimized-NOMA and the fixed-NOMA have identical sum-rate performance. Also, at low SNRs, the size of the antenna array is a major obstacle in achieving higher sum-rates.

In chapter 5, we proposed a new two-step beam alignment algorithm for the analog beamforming mmWave-NOMA systems with constrained channel coherence time. The algorithm takes the advantages of NOMA, as such, first the beamwidth is selected for the near user and then for the far user. We defined an optimization problem to maximize the sum-rate. We then jointly optimized the beamwidths and the allocated powers while keeping the minimum required rate equal or higher than a predefined threshold. To evaluate the effect of beam misalignment due to the mmWave propagation characteristics and longer distance, we formulated another optimization problem, where we imposed extra constraints on the beamwidths to satisfy the misalignment probability requirements. An off-line search was adopted to find the appropriate ranges for the beamwidths. Then, with the new ranges of the beamwidths, we used the similar joint optimization algorithm structure as we proposed for the problem without misalignment probability requirements to find the optimal power allocation and beamwidths. Our numerical results showed that the proposed algorithm with and without beam misalignment achieves higher sum-rate compared to the one-step beam alignment al-

Table 7.1: Summary of the dissertation

Beamforming design	**Combination of the mmWave technology and other technologies**
Analog beamforming for multi-user scenario	mmWave-NOMA
Hybrid precoding for relay networks	mmWave-UAV

gorithm. Further, due to aligning beams in two steps, the proposed algorithm can take advantages of narrow beams. This leads to a higher beamforming gain and better sum-rate. Further, in the case of lower SNR, channel coherence time, and the lower channel of the near user, the proposed algorithm with beam misalignment marginally reduces the sum-rate.

In chapter 6, we jointly optimized the deployment, user scheduling, beamforming vector, and RIS phases in a RIS assisted UAV wireless network. To solve the problem, we iterated among the 4 variables. While optimizing one variable, we fixed the other 3 variables. For the deployment, we found the optimal position by a sphere search. Then, we formulated an integer linear programming to find the best scheduling. We also designed the analog beamforming vector by compensating the phases of the channel. When optimizing the RIS phases, we formulated a semi-definite programming to find the best phases. The proposed joint optimization outperforms the system without RIS assistance, and the system without deployment optimization.

7.2 Future work

MmWave communications have been considered as the key technology for the future wireless systems. This dissertation has explored how to implement beamforming in mmWave system to improve the system performance while reducing the implementation complexity. There are still some topic needs to be explored in the future.

- Limited feedback: how to quantize the mmWave channels is critical to improve the system performance. Although we have proposed some robust algorithm to combat the estimation

and quantization errors, the specific limited feedback scheme has not been explored. Future work can be focused on the design of effective limited feedback scheme for the mmWave communications.

- New scenarios: combining mmWave with other technologies is the mainstream trend for the wireless communication systems. Although we have already explored some scenarios such as mmWave-NOMA and UAV networks, a lot more scenarios remained unexplored, such as wireless cache, cloud radio access network, and so on. For the future work, we can implement the mmWave communications in more scenarios.

Bibliography

[1] New SID proposal: Study on full dimension MIMO for LTE. Technical report, 3GPP TSG RAN Meeting 58, Dec. 2012.

[2] 3GPP R1-154999. TP for classification of MUST schemes. *TSG-RAN WG1 #82, Beijing, China*, Aug. 24-28, 2015.

[3] M. A. Almasi, M. Vaezi, and H. Mehrpouyan. Impact of beam misalignment on hybrid beamforming NOMA for mmwave communications. *IEEE Transactions on Communications*, 67(6):4505–4518, June 2019.

[4] F. Afroz, R. Subramanian, R. Heidary, K. Sandrasegaran, and S. Ahmed. Sinr, rsrp, rssi and rsrq measurements in long term evolution networks. *International Journal of Wireless & Mobile Networks*, 2015.

[5] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip. Millimeter wave channel modeling and cellular capacity evaluation. *IEEE Journal on Selected Areas in Communications*, 32(6):1164–1179, June 2014.

[6] S. Ali, E. Hossain, and D. I. Kim. Non-orthogonal multiple access (NOMA) for downlink multiuser MIMO systems: User clustering, beamforming, and power allocation. *IEEE Access*, 5:565–577, Dec. 2017.

[7] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):831–846, Oct. 2014.

[8] A. Alkhateeb, G. Leus, and R. W. Heath. Limited feedback hybrid precoding for multi-user millimeter wave systems. *IEEE Transactions on Wireless Communications*, 14(11):6481–6494, Nov. 2015.

[9] M. A. Almasi, R. Amiri, M. Vaezi, and H. Mehrpouyan. Lens-based millimeter wave reconfigurable antenna NOMA. *in Proc. IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–5, May 2019.

[10] M. A. Almasi, L. Jiang, H. Jafarkhani, and H. Mehrpouyan. Joint beamwidth and power optimization in mmWave hybrid beamforming-NOMA systems. *IEEE Transactions on Wireless Communications*, 20(4):2442 – 2456, 2020.

[11] M. A. Almasi, H. Mehrpouyan, V. Vakilian, N. Behdad, and H. Jafarkhani. A new reconfigurable antenna MIMO architecture for mmwave communication. *in Proc. IEEE International Conference on Communications (ICC)*, May 2018.

[12] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 13(3):1499–1513, Mar. 2014.

[13] C. A. Balanis. *Antenna theory: analysis and design*. John wiley & sons, 2016.

[14] A. Beck et al. A sequential parametric convex approximation method with applications to nonconvex truss topology design problems. *Journal of Global Optimization*, 47(1):29–51, 2010.

[15] J. Brady, N. Behdad, and A. M. Sayeed. Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements. *IEEE Transactions on Antennas and Propagation*, 61(7):3814–3827, July 2013.

[16] J. Brady, N. Behdad, and A. M. Sayeed. Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements. *IEEE Transactions on Antennas and Propagation*, 61(7):3814–3827, July 2013.

[17] B. A. Cetiner, H. Jafarkhani, J.-Y. Qian, H. J. Yoo, A. Grau, and F. De Flaviis. Multifunctional reconfigurable MEMS integrated antennas for adaptive MIMO systems. *IEEE Communications Magazine*, 42(12):62–70, 2004.

[18] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gershman. Robust downlink beamforming based on outage probability specifications. *IEEE Transactions on Wireless Communications*, 6(10):3498–3503, Oct. 2007.

[19] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong. Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience. *IEEE Journal on Selected Areas in Communications*, 35(5):1046–1061, 2017.

[20] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi. Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design. *IEEE Transactions on Wireless Communications*, 7(12):4792–4799, 2008.

[21] P. Chung, H. Du, and J. Gondzio. A probabilistic constraint approach for robust transmit beamforming with imperfect channel information. *IEEE Transactions on Signal Processing*, 59(6):2773–2782, Jun. 2011.

[22] J. Cui, Z. Ding, P. Fan, and N. Al-Dhahir. Unsupervised machine learning-based user clustering in millimeter-Wave-NOMA systems. *IEEE Transactions on Wireless Communications*, 17(11):7425–7440, Nov. 2018.

[23] L. Dai, B. Wang, Y. Yuan, S. Han, C. I, and Z. Wang. Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Communications Magazine*, 53(9):74–81, Sept. 2015.

[24] D. De Donno, J. Palacios, and J. Widmer. Millimeter-wave beam training acceleration through low-complexity hybrid transceivers. *IEEE Transactions on Wireless Communications*, 16(6):3646–3660, June 2017.

[25] Z. Ding, L. Dai, R. Schober, and H. V. Poor. NOMA meets finite resolution analog beamforming in massive MIMO and millimeter-wave networks. *IEEE Communications Letters*, 21(8):1879–1882, Aug. 2017.

[26] Z. Ding, P. Fan, and H. V. Poor. Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions. *IEEE Transactions on Vehicle Technology*, 65(8):6010–6023, Aug 2016.

[27] Z. Ding, P. Fan, and H. V. Poor. Random beamforming in millimeter-wave NOMA networks. *IEEE Access*, 5:7667–7681, Feb. 2017.

[28] H. Du and P. Chung. A probabilistic approach for robust leakage-based MU-MIMO downlink beamforming with imperfect channel state information. *IEEE Transactions on Wireless Communications*, 11(3):1239–1247, Mar. 2012.

[29] S. Ekbatani, F. Etemadi, and H. Jafarkhani. Throughput maximization over slowly fading channels using quantized and erroneous feedback. *IEEE Transactions on Communications*, 57(9):2528–2533, Sep. 2009.

[30] S. Ekbatani, F. Etemadi, and H. Jafarkhani. Outage behavior of slow fading channels with power control using partial and erroneous csit. *IEEE Transactions on Information Theory*, 56(12):6097–6102, Dec. 2010.

[31] S. Ekbatani and H. Jafarkhani. Combining beamforming and space-time coding using noisy quantized feedback. *IEEE Transactions on Communications*, 57(5):1280–1286, May. 2009.

[32] O. El Ayach, R. W. Heath, S. AbuSurra, S. Rajagopal, and Z. Pi. The capacity optimality of beam steering in large millimeter wave MIMO systems. In *Proc. of IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Cesme, Turkey, Jun. 2012.

[33] O. El Ayach, S. Rajagopal, S. AbuSurra, Z. Pi, and R. W. Heath. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 13(3):1499–1513, Mar. 2014.

[34] N. Eshraghi, V. Shah-Mansouri, and B. Maham. Fair beamwidth selection and resource allocation for indoor millimeter-wave networks. *in Proc. IEEE International Conference on Communications (ICC)*, pages 1–6, May 2017.

[35] F. Fazel, A. Grau, H. Jafarkhani, and F. De Flaviis. Space-time-state block coded MIMO communication systems using reconfigurable antennas. *IEEE Transactions on Wireless Communications*, 8(12):6019–6029, 2009.

[36] M. Ganji and H. Jafarkhani. Improving NOMA multi-carrier systems with intentional frequency offsets. *IEEE Wireless Communications Letters*, 8(4):1060–1063, Aug. 2019.

[37] M. Ganji and H. Jafarkhani. Asynchronous transmission for multiple access channels: Rate-region analysis and system design for uplink NOMA. *IEEE Transactions on Wireless Communications*, 2021.

[38] X. Gao, L. Dai, and A. M. Sayeed. Low RF-complexity technologies to enable millimeter-wave MIMO with large antenna array for 5G wireless communications. *IEEE Communications Magazine*, 56(4):211–217, Apr. 2018.

[39] L. C. Godara. *Handbook of antennas in wireless communications*. CRC press, Aug. 2001.

[40] A. Goldsmith. *Wireless communications*. Cambridge university press, Aug. 2005.

[41] M. Grant, S. Boyd, and Y. Ye. CVX: Matlab software for disciplined convex programming, 2008.

[42] A. Grau, H. Jafarkhani, and F. De Flaviis. A reconfigurable multiple-input multiple-output communication system. *IEEE Transactions on Wireless Communications*, 7(5):1719–1733, 2008.

[43] R. M. Gray. *Entropy and information theory*. Springer Science & Business Media, 2011.

[44] J. Guo and H. Jafarkhani. Sensor deployment with limited communication range in homogeneous and heterogeneous wireless sensor networks. *IEEE Transactions on Wireless Communications*, 15(10):6771–6784, 2016.

[45] J. Guo, P. Walk, and H. Jafarkhani. Optimal deployments of UAVs with directional antennas for a power-efficient coverage. *IEEE Transactions on Communications*, 68(8):5159–5174, Aug. 2020.

[46] A. K. Gupta and D. K. Nagar. *Matrix variate distributions*. Chapman and Hall/CRC, 2018.

[47] S. Han, I. Chih-Lin, Z. Xu, and C. Rowell. Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G. *IEEE Communications Magazine*, 53(1):186–194, 2015.

[48] W. Hao, M. Zeng, Z. Chu, and S. Yang. Energy-efficient power allocation in millimeter wave massive MIMO with non-orthogonal multiple access. *IEEE Wireless Communications Letters*, 6(6):782–785, Dec. 2017.

[49] W. Hao, F. Zhou, Z. Chu, P. Xiao, R. Tafazolli, and N. Al-Dhahir. Beam alignment for MIMO-NOMA millimeter wave communication systems. *in Proc. IEEE International Conference on Communications (ICC)*, pages 1–6, May 2019.

[50] R. A. Hassan and N. Michelusi. Multi-user beam-alignment for millimeter-wave networks. *in Proc. Inf. Theory and Appl. Workshop*, pages 1–7, Feb. 2018.

[51] B. He and H. Jafarkhani. Low-complexity reconfigurable MIMO for millimeter wave communications. *IEEE Transactions on Communications*, 66(11):5278–5291, Nov. 2018.

[52] B. He and H. Jafarkhani. Low-complexity reconfigurable MIMO for millimeter wave communications. *IEEE Transactions on Communications*, 66(11):5278–5291, 2018.

[53] A. Hottinen, O. Tirkkonen, and R. Wichman. *Multi-antenna transceiver techniques for 3G and beyond*. John Wiley & Sons, Aug. 2004.

[54] M. Hua, L. Yang, Q. Wu, C. Pan, C. Li, and A. L. Swindlehurst. UAV-assisted intelligent reflecting surface symbiotic radio system. *arXiv preprint arXiv:2007.14029*, 2020.

[55] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Transactions on Wireless Communications*, 18(8):4157–4170, 2019.

[56] S. Huang, M. Xiao, and H. V. Poor. On the physical layer security of millimeter wave NOMA networks. *IEEE Transactions on Vehicle Technology*, 69(10):11697–11711, Oct. 2020.

[57] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh. Millimeter wave beamforming for wireless backhaul and access in small cell networks. *IEEE Transactions on Communications*, 61(10):4391–4403, Oct. 2013.

[58] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh. Millimeter wave beamforming for wireless backhaul and access in small cell networks. *IEEE Transactions on Communications*, 61(10):4391–4403, Oct. 2013.

[59] M. Hussain and N. Michelusi. Energy-efficient interactive beam alignment for millimeter-wave networks. *IEEE Transactions on Wireless Communications*, 18(2):838–851, Feb. 2019.

[60] H. Jafarkhani. *Space-time coding: theory and practice*. Cambridge university press, Sep. 2005.

[61] L. Jiang and H. Jafarkhani. Multi-User Analog Beamforming in Millimeter Wave MIMO Systems Based on Path Angle Information. *IEEE Transactions on Wireless Communications*, 18(1):608–619, 2018.

[62] L. Jiang and H. Jafarkhani. Multi-user analog beamforming in millimeter wave MIMO systems based on path angle information. *IEEE Transactions on Wireless Communications*, 18(1):608–619, Jan. 2019.

[63] L. Jiang and H. Jafarkhani. Multi-user analog beamforming in mmWave MIMO systems based on path angle information. *IEEE Transactions on Wireless Communications*, 18(1):608–619, Jan. 2019.

[64] L. Jiang and H. Jafarkhani. MmWave amplify-and-forward MIMO relay networks with hybrid precoding/combining design. *IEEE Transactions on Wireless Communications*, 19(2):1333–1346, Feb. 2020.

[65] L. Jiang, X. L. Liu, and H. Jafarkhani. Hybrid precoding/combining design in mmwave amplify-and-forward MIMO relay networks. In *Proc. of IEEE International Conference on Communications (ICC)*, Shanghai, China, May 2019.

[66] Y. Jing and H. Jafarkhani. Network beamforming with channel means and covariances at relays. In *Proc. of IEEE International Conference on Communications (ICC)*, Beijing, China, May 2008.

[67] Y. Jing and H. Jafarkhani. Network beamforming using relays with perfect channel information. *IEEE Transactions on Information Theory*, 55(6):2499–2517, 2009.

[68] J. Kazemitabar and H. Jafarkhani. Multiuser interference cancellation and detection for users with more than two transmit antennas. *IEEE Transactions on Communications*, 56(4):574–583, Apr. 2008.

[69] C. Kojima et al. Novel two-step beam search method for multi user millimeter-wave communication. *in Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–6, Oct. 2017.

[70] E. Koyuncu and H. Jafarkhani. Distributed beamforming in wireless multiuser relay-interference networks with quantized feedback. *IEEE Transactions on Information Theory*, 58(7):4538–4576, Jul. 2012.

[71] E. Koyuncu, M. Shabanighazikelayeh, and H. Seferoglu. Deployment and trajectory optimization of UAVs: A quantization theory approach. *IEEE Transactions on Wireless Communications*, 17(12):8531–8546, 2018.

[72] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Transactions on Information Technology*, 50(12):3062–3180, Dec. 2004.

[73] H. Lee and Y. Ko. Non-iterative symbol-wise beamforming for MIMO-OFDM systems. *IEEE Transactions on Wireless Communications*, 11(10):3788–3798, Oct. 2012.

[74] J. Lee and Y. H. Lee. AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing. In *IEEE International Conference on Communications (ICC)*, pages 5838–5842, June 2014.

[75] B. Li, Z. Zhou, W. Zou, X. Sun, and G. Du. On the efficient beam-forming training for 60GHz wireless personal area networks. *IEEE Transactions on Wireless Communications*, 12(2):504–515, Feb. 2013.

[76] S. Li, B. Duo, X. Yuan, Y.-C. Liang, and M. Di Renzo. Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming. *IEEE Wireless Communications Letters*, 9(5):716–720, 2020.

[77] W.-C. Li, T.-H. Chang, C. Lin, and C.-Y. Chi. Coordinated beamforming for multiuser MISO interference channel under rate outage constraints. *IEEE Transactions on Signal Processing*, 61(5):1087–1103, Mar. 2012.

[78] X. Li, Y. Zhu, and P. Xia. Enhanced analog beamforming for single carrier millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 16(7):4261–4274, Jul. 2017.

[79] C. Liu, M. Li, S. V. Hanly, I. B. Collings, and P. Whiting. Millimeter wave beam alignment: Large deviations analysis and design insights. *IEEE Journal on Selected Areas in Communications*, 35(7):1619–1631, July 2017.

[80] L. Liu and H. Jafarkhani. Space-time trellis codes based on channel-phase feedback. *IEEE Transactions on Communications*, 54(12):2186–2198, Dec. 2006.

[81] L. Liu, Y. Nam, and J. Zhang. Proportional fair scheduling for multi-cell multi-user mimo systems. In *Proc. of IEEE 44th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, New Jersey, USA, Mar. 2010.

[82] X. Liu, H. Jafarkhani, and E. Koyuncu. Amplify-and-forward relay networks with variable-length limited feedback. *IEEE Transactions on Wireless Communications*, 15(11):7725–7737, Nov. 2016.

[83] Z. Luo, C. Zhan, L. Zhang, and R. Zhang. Robust hybrid beamforming in millimeter wave relay networks with imperfect CSI. *IEEE Access*, 6:73093–73101, Nov. 2018.

[84] D. Ma, M. Ding, and M. Hassan. Enhancing cellular communications for UAVs via intelligent reflective surface. In *Proc. 2020 IEEE Wireless Communications and Networking Conference (WCNC)*, May. 2020.

[85] R. T. Marler and J. S. Arora. The weighted sum method for multi-objective optimization: new insights. *Structural and Multidisciplinary Optimization*, 41(6):853–862, Jun. 2010.

[86] J. N. Murdock, E. Ben-Dor, Y. Qiao, J. I. Tamir, and T. S. Rappaport. A 38 GHz cellular outage study for an urban outdoor campus environment. In *Proc. of 2012 IEEE Wireless Communications and Networking Conference (WCNC)*, Paris, France, Apr. 2012.

[87] S. Nadarajah. A generalized normal distribution. *Journal of Applied Statistics*, 32(7):685–694, Sep. 2005.

[88] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell. Efficient use of side information in multiple-antenna data transmission over fading channels. *IEEE Journal on Selected Areas in Communications*, 16(8):1423–1436, Oct. 1998.

[89] W. Ni and X. Dong. Hybrid block diagonalization for massive multiuser MIMO systems. *IEEE Transactions on Communications*, 64(1):201–211, Jan. 2016.

[90] S. Noh, M. D. Zoltowski, and D. J. Love. Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment. *IEEE Transactions on Wireless Communications*, 16(9):5689–5701, Sept. 2017.

[91] G. OPTIMIZATION. Inc. gurobi optimizer reference manual, 2015. *URL: http://www. gurobi. com*, page 29, 2014.

[92] J. Palacios, D. De Donno, D. Giustiniano, and J. Widmer. Speeding up mmwave beam training through low-complexity hybrid transceivers. *in Proc. IEEE 27th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–7, Sept. 2016.

[93] Z. Pi and F. Khan. An introduction to millimeter-wave mobile broadband systems. *IEEE Communications Magazine*, 49(6):101–107, June 2011.

[94] A. Pollok, W. G. Cowley, and N. Letzepis. Symbol-wise beamforming for MIMO-OFDM transceivers in the presence of co-channel interference and spatial correlation. *IEEE Transactions on Wireless Communications*, 8(12):5755–5760, Dec. 2009.

[95] C. Pradhan, H. Chen, Y. Li, and B. Vucetic. Joint beamwidth and energy optimization for multi-user millimeter wave communications. *in Proc. IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6, May 2018.

[96] R. Rajashekar and L. Hanzo. Iterative matrix decomposition aided block diagonalization for mm-wave multiuser MIMO systems. *IEEE Transactions on Wireless Communications*, 16(3):1372–1384, Mar. 2017.

[97] S. Rangan, T. S. Rappaport, and E. Erkip. Millimeter-wave cellular wireless networks: Potentials and challenges. *Proceedings of the IEEE*, 102(3):366–385, Mar. 2014.

[98] T. S. Rappaport, F. Gutierrez, E. BenDor, J. N. Murdock, Y. Qiao, and J. I. Tamir. Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications. *IEEE Transactions on Antennas and Propagation*, 61(4):1850–1859, Apr. 2013.

[99] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock. *Millimeter wave wireless communications*. Pearson Education, Sep. 2014.

[100] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun. Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE transactions on Communications*, 63(9):3029–3056, 2015.

[101] Y. Rong. Robust design for linear non-regenerative MIMO relays with imperfect channel state information. *IEEE Transactions on Signal Processing*, 59(5):2455–2460, May 2011.

[102] M. Sadek, A. Tarighat, and A. H. Sayed. Active antenna selection in multiuser MIMO communications. *IEEE Transactions on Signal Processing*, 55(4):1498–1510, Apr. 2007.

[103] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi. Non-orthogonal multiple access (NOMA) for cellular future radio access. *in Proc. IEEE 77th Vehicular Technology Conference (VTC) Spring*, pages 1–5, June 2013.

[104] A. A. M. Saleh and R. Valenzuela. A statistical model for indoor multipath propagation. *IEEE Journal on selected areas in communications*, 5(2):128–137, Feb. 1987.

[105] A. M. Sayeed. Deconstructing multiantenna fading channels. *IEEE Transactions on Signal Processing*, 50(10):2563–2579, Oct. 2002.

[106] M. Schubert and H. Boche. Solution of the multiuser downlink beamforming problem with individual SINR constraints. *IEEE Transactions on Vehicular Technology*, 53(1):18–28, Jan. 2004.

[107] J. Seo, Y. Sung, and H. Jafarkhani. A high-diversity transceiver design for MISO broadcast channels. *IEEE Transactions on Wireless Communications*, 18(5):2591–2606, May 2019.

[108] J.-S. Sheu. Hybrid digital and analogue beamforming design for millimeter wave relaying systems. *Journal of Communications and Networks*, 19(5):461–469, Nov. 2017.

[109] H. Shokri-Ghadikolaei, L. Gkatzikis, and C. Fischione. Beam-searching and transmission scheduling in millimeter wave communications. *in Proc. IEEE International Conference on Communications (ICC)*, pages 1292–1297, June 2015.

[110] M. K. Simon and M. . Alouini. A unified approach to the probability of error for non-coherent and differentially coherent modulations over generalized fading channels. *IEEE Transactions on Communications*, 46(12):1625–1638, Dec. 1998.

[111] P. F. Smulders and L. Correia. Characterisation of propagation in 60 GHz radio channels. *Electronics & Communication Engineering Journal*, 9(2):73–80, Apr. 1997.

[112] F. Sohrabi and W. Yu. Hybrid digital and analog beamforming design for large-scale antenna arrays. *IEEE Journal of Selected Topics in Signal Processing*, 10(3):501–513, Apr. 2016.

[113] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *IEEE Transactions on Signal Processing*, 52(2):461–471, Feb. 2004.

[114] A. Tarighat, M. Sadek, and A. H. Sayed. A multi user beamforming scheme for downlink MIMO channels based on maximizing signal-to-leakage ratios. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings (ICASSP'05)*, Philadelphia, Pennsylvania, USA, Mar. 2005.

[115] E. Torkildson, B. Ananthasubramaniam, U. Madhow, and M. Rodwell. Millimeter-wave MIMO: Wireless links at optical speeds. In *Proc. of 44th Allerton Conference on Communication, Control and Computing*, Monticello, Illinois, USA, Sep. 2006.

[116] Y. M. Tsang, A. S. Poon, and S. Addepalli. Coding the beams: Improving beamforming training in mmWave communication system. In *Proc. of 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Houston, Texas, USA, Dec. 2011.

[117] D. Tse and P. Viswanath. *Fundamentals of wireless communication*. Cambridge university press, May. 2005.

[118] C. G. Tsinos, S. Chatzinotas, and B. Ottersten. Hybrid analog-digital transceiver designs for mmwave amplify-and-forward relaying systems. In *Proc. of 41st IEEE International Conference on Telecommunications and Signal Processing (TSP)*, Athens, Greece, Jul. 2018.

[119] V. Va, J. Choi, and R. W. Heath. The impact of beamwidth on temporal channel variation in vehicular channels and its implications. *IEEE Transactions on Vehicular Technology*, 66(6):5014–5029, Nov. 2017.

[120] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath. Inverse multipath fingerprinting for millimeter wave V2I beam alignment. *IEEE Transactions on Vehicular Technology*, 67(5):4042–4058, Dec. 2018.

[121] M. Vaezi, Z. Ding, and H. V. Poor. *Multiple access techniques for 5G wireless networks and beyond*. Springer, 2019.

[122] I. Waldspurger, A. d'Aspremont, and S. Mallat. Phase recovery, maxcut and complex semidefinite programming. *Mathematical Programming*, 149(1-2):47–81, Feb. 2015.

[123] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou. Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array. *IEEE Journal on Selected Areas in Communications*, 35(10):2370–2382, Oct. 2017.

[124] J. Wang. Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems. *IEEE Journal on Selected Areas in Communications*, 27(8):1390–1399, Oct. 2009.

[125] Z. Wei, D. W. K. Ng, and J. Yuan. NOMA for hybrid mmwave communication systems with beamwidth control. *IEEE Journal of Selected Topics in Signal Processing*, 13(3):567–583, June 2019.

[126] A. Wiesel, Y. C. Eldar, and S. Shamai. Linear precoding via conic optimization for fixed MIMO receivers. *IEEE Transactions on Signal Processing*, 54(1):161–176, Jan. 2006.

[127] J. Wildman, P. H. J. Nardelli, M. Latva-aho, and S. Weber. On the joint impact of beamwidth and orientation error on throughput in directional wireless poisson networks. *IEEE Transactions on Wireless Communications*, 13(12):7072–7085, Dec. 2014.

[128] K. Wong, R. Cheng, K. B. Letaief, and R. D. Murch. Adaptive antennas at the mobile and base stations in an OFDM/TDMA system. *IEEE Transactions on Communications*, 49(1):195–206, Jan. 2001.

[129] Q. Wu and R. Zhang. Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design. In *Proc. 2018 IEEE GLOBECOM*, Dec. 2018.

[130] Y. Wu, Q. Yang, Q. He, and K. S. Kwak. Optimal energy harvesting-ratio and beamwidth selection in millimeter wave communications. *IEEE Signal Processing Letters*, 23(10):1364–1368, Oct. 2016.

[131] Z. Xiao, H. Dong, L. Bai, D. O. Wu, and X.-G. Xia. Unmanned Aerial Vehicle Base Station (UAV-BS) Deployment with Millimeter Wave Beamforming. *IEEE Internet of Things Journal*, 7(2):1336 – 1349, 2020.

[132] Z. Xiao, P. Xia, and X.-G. Xia. Enabling UAV cellular with millimeter-wave communication: Potentials and approaches. *IEEE Communications Magazine*, 54(5):66–73, 2016.

[133] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X. Xia. Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications. *IEEE Transactions on Wireless Communications*, 17(5):2961–2974, May 2018.

[134] C. Xing, S. Ma, and Y.-C. Wu. Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems. *IEEE Transactions on Signal Processing*, 58(4):2273–2283, Apr. 2010.

[135] C. Xing, S. Ma, Y.-C. Wu, and T.-S. Ng. Transceiver design for dual-hop nonregenerative MIMO-OFDM relay systems under channel uncertainties. *IEEE Transactions on Signal Processing*, 58(12):6325–6339, Dec. 2010.

[136] H. Xu, V. Kukshya, and T. S. Rappaport. Spatial and temporal characteristics of 60-ghz indoor channels. *IEEE Journal on selected areas in communications*, 20(3):620–630, 2002.

[137] X. Xue, Y. Wang, L. Dai, and C. Masouros. Relay hybrid precoding design in millimeter-wave massive mimo systems. *IEEE Transactions on Signal Processing*, 66(8):2011–2026, Apr. 2018.

[138] X. Xue, Y. Wang, X. Wang, and T. E. Bogale. Joint source and relay precoding in multiantenna millimeter-wave systems. *IEEE Transactions on Vehicular Technology*, 66(6):4924–4937, Oct. 2016.

[139] Y. Yapici, I. Guvenc, and H. Dai. Low-resolution limited-feedback NOMA for mmWave communications. *IEEE Transactions on Wireless Communications*, 19(8):5433–5446, Aug. 2020.

[140] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews. User association for load balancing in heterogeneous cellular networks. *IEEE Transactions on Wireless Communications*, 12(6):2706–2716, June 2013.

[141] Y. Zeng, J. Lyu, and R. Zhang. Cellular-connected UAV: Potential, challenges, and promising technologies. *IEEE Wireless Communications*, 26(1):120–127, 2018.

[142] L. Zhang, H. Zhao, S. Hou, Z. Zhao, H. Xu, X. Wu, Q. Wu, and R. Zhang. A survey on 5G millimeter wave communications for UAV-assisted wireless networks. *IEEE Access*, 7:117460–117504, 2019.

[143] X. Zhang, D. P. Palomar, and B. Ottersten. Statistically robust design of linear mimo transceivers. *IEEE Transactions on Signal Processing*, 56(8):3678–3689, Aug. 2008.

[144] X.-D. Zhang. *Matrix analysis and applications*. Cambridge University Press, 2017.

[145] L. Zhao, D. W. K. Ng, and J. Yuan. Multi-user precoding and channel estimation for hybrid millimeter wave systems. *IEEE Journal on Selected Areas in Communications*, 35(7):1576–1590, July 2017.

[146] W. Zhong, L. Xu, X. Lu, and L. Wang. Research on millimeter wave communication interference suppression of UAV based on beam optimization. In *Proc. International Conference on Machine Learning and Intelligent Communications*. Springer, Aug. 2017.

[147] Y. Zhou, V. W. S. Wong, and R. Schober. Coverage and rate analysis of millimeter wave NOMA networks with beam misalignment. *IEEE Transactions on Wireless Communications*, 17(12):8211–8227, Dec. 2018.

[148] X. Zou, M. Ganji, and H. Jafarkhani. Trellis-coded non-orthogonal multiple access. *IEEE Wireless Communications Letters*, 9(4):538–542, Apr. 2019.

Appendix A

Supplementary Proofs for Chapter 4

A.1 Proof of Proposition 2

Proof. Since $F(g_1, f_1)$ is a continuous function defined on a bounded closed set, it has a maximum point according to the extreme value theorem. Also, according to the critical point theorem, the maximum point should either be a stationary point or a boundary point. It is easy to observe that $F(g_1, f_1)$ is a monotonic increasing function of g_1 and a monotonic decreasing function of f_1. This means $\frac{\partial F}{\partial g_1} > 0$ and $\frac{\partial F}{\partial f_1} < 0$, i.e., there is no stationary point for $F(g_1, f_1)$ on the defined domain. Then, the maximum point should lie on the five boundaries: (i) $f_1 = lb_f,\ g_1 \in [lb_g, lb_f^2]$, (ii) $g_1 = lb_g,\ f_1 \in [lb_f, ub_f]$, (iii) $f_1 = ub_f,\ g_1 \in [lb_g, ub_g]$, (iv) $g_1 = ub_g,\ f_1 \in [\sqrt{ub_g}, ub_f]$, and (v) $g_1 = f_1^2,\ f_1 \in [lb_f, \sqrt{ub_g}]$.

For the boundary (i), since $F(g_1, f_1)$ is a monotonic increasing function of g_1, the maximum point can only lie on the point (lb_f, lb_f^2) which belongs to the boundary (v) as well. For the boundary (ii), since $F(g_1, f_1)$ is a monotonic increasing function of g_1, we can pick $g_1 > lb_g$ to increase the value of $F(g_1, f_1)$. This implies that the maximum point cannot lie on the boundary (ii). Similarly, the maximum point cannot lie on the boundary (iii) either. For the boundary (iv), since $F(g_1, f_1)$ is

a monotonic decreasing function of f_1, the maximum point can only lie on the point $(\sqrt{ub_g}, ub_g)$ which belongs to boundary (v) as well. Note that the possible maximum points on the boundaries (i) and (iv) also belong to the boundary (v). Therefore, the maximum point must lie on the boundary (v) and the proof is complete. □

A.2 Proof of Proposition 3

Proof. Since $G(g_1, f_1)$ is a continuous function defined on a bounded closed set, it has a maximum point according to the extreme value theorem. Also, according to the critical point theorem, the maximum point should either be a stationary point or a boundary point. It is easy to observe that $G(g_1, f_1)$ is a monotonic decreasing function of f_1. $\frac{\partial G}{\partial f_1} < 0$, i.e., there is no stationary point for $G(g_1, f_1)$ on the defined domain. Then, the maximum point should lie on the four boundaries: (i) $g_1 = lb_g,\ f_1 \in [\sqrt{lb_g}, ub_f]$, (ii) $f_1 = ub_f,\ g_1 \in [lb_g, ub_g]$, (iii) $g_1 = ub_g,\ f_1 \in [\sqrt{ub_g}, ub_f]$, and (iv) $g_1 = lb_f^2,\ f_1 \in [\sqrt{lb_g}, \sqrt{ub_g}]$.

For the boundary (i), since $G(g_1, f_1)$ is a monotonic decreasing function of f_1, the maximum point can only lie on the point $(\sqrt{lb_g}, lb_g)$ which belongs to the boundary (iv) as well. For the boundary (ii), since $G(g_1, f_1)$ is a monotonic decreasing function of f_1, we can pick $f_1 < ub_f$ to increase the value of $G(g_1, f_1)$, which implies that the maximum point cannot lie on this boundary. For the boundary (iii), since $G(g_1, f_1)$ is a monotonic decreasing function of f_1, the maximum point must lie on the point $(\sqrt{ub_g}, ub_g)$, which also belongs to the boundary (iv). Since the possible maximum points on the boundaries (i) and (iii) also belong to the boundary (iv), the maximum point must lie on the boundary (iv) and the proof is complete. □