Low Complexity Precoding Schemes for Massive MIMO Systems

Yulong Chen

School of Engineering

Newcastle University

A thesis submitted for the degree of

Doctor of Philosophy

June, 2019

Abstract

In order to deal with the challenges of the exponentially growing communication traffic and spectrum bands with wider bandwidth, massive MIMO technology was been proposed, which employs an unprecedented number of base station antennas simultaneously to serve a smaller number of user terminals in the same channel. Although the very large antenna arrays for massive multiple-input multiple-output (MIMO) systems lead to unprecedented data throughputs and beamforming gains to meet these data traffic demands, they also lead to prohibitively high energy consumption and hardware complexity. In terms of precoding schemes, the conventional linear precoding entirely processes the complex signals in the digital domain and then upconverts to the carrier frequency after passing through radio frequency (RF) chains, which can achieve near-optimal performance with the large antenna arrays. However, it is infeasible because with fully digital precoding, every antenna element needs to be coupled with one RF chain, including the digital-to-analog convertors, mixers and filters, which is accountable for excessively high hardware cost and power consumption. This thesis focuses on the design and analysis of low complexity precoding schemes.

The novel contributions in this thesis are presented in three sections.

First, a low complexity hybrid precoding scheme is proposed for the downlink transmission of massive multi-user MIMO systems with a finite dimensional channel model. By analysing the structure of the channel model, the beamsteering codebooks are combined with extracting the

phase of the conjugate transpose of the fast fading matrix to design the RF precoder, which thereby harvests the large array gain achieved by an unprecedented number of base station antennas. Then a baseband precoder is designed based on the equivalent channel with zero forcing (ZF) precoding. In addition, a tight upper bound on the spectral efficiency is derived and the performance of hybrid precoding is investigated.

Second, based on successive refinement, a new iterative hybrid precoding scheme is proposed with a sub-connected architecture for mmWave MIMO systems.In each iteration, the first step is to design the RF precoder and the second step is to design the baseband precoder. The RF precoder is regarded as an input to update the baseband precoder until the stopping criterion is triggered. Phase extraction is used to obtain the RF precoder and then the baseband precoder is optimized by the orthogonal property. This algorithm effectively optimizes the hybrid precoders and reduces the hardware complexity with sub-connected architecture. A closed-form expression of upper bound for the spectral efficiency is derived and the energy efficiency and the complexity of the proposed hybrid precoding scheme are analyzed.

Finally, the use of low-resolution digital-to-analog converters (DACs) for each antenna and RF chain is considered. Moreover, in a more practical scenario, the hardware mismatch between the uplink and the downlink for the channel matrix is a focus, where the downlink is not the transpose of the uplink in time-division duplex mode. The impact of one-bit DACs on linear precoding is studied for the massive MIMO systems with hardware mismatch. Using the Bussgang theorem and random matrix theorem, a closed-form expression for the signal to quantization, interference and noise ratio with consideration of hardware mismatch and one-bit ZF precoding is derived, which can be used to derive the achiev-

able rate. Then a performance approximation is also derived in the high signal-to-noise ratio (SNR) region, which is related to the ratio of the number of base station antennas and the number of mobile users , and the statistics of the circuit gains at the base station.

In conclusion, analytical and numerical results show that the proposed techniques are able to achieve close-to-optimal performances with low hardware complexity, thus the low complexity precoding schemes can be valid candidates for practical implementations of modern communication systems.

Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my supervisor Prof. Said Boussakta, who has provided me with valuable guidance and unconditional support throughout my years in Newcastle University. I truly appreciate his patience and countless hours he spent with me, discussing my research. I am also thankful to my second supervisor Dr. Charalampos Tsimenidis for his time and great advice during my PhD period. Special thanks are also extended to Prof. Jonathon Chambers for his constructive comments and professional suggestions to my research progression.

I feel very happy to be part of Communications, Sensors, Signal and Information Processing Research Group, the colleagues here are so nice and kind. I would like to thank Zhen Mei, Pengming Feng, Jamal Ahmed Hussein, Wael Abd Alaziz, Achonu Adejo, Jiachen Yin, Yang Sun, Zeyu Fu, Safaa Nash'at Awny, Haicang Li and Yang Xian. The friendship between us is the most unforgettable memory during the past few years in Newcastle.

Additionally, my deepest gratitude is also expressed to my parents for their selfless support and encouragement. Even though we were far apart, their caring warms my heart all the way.

Finally, I would like to express my love to my wife Yu Hong, and my daughter Grace Chen. During my PhD study, my wife has been there for me and gave me all the love. Little Grace is the most precious gift

in my life. Words cannot express how happy I was, when I first held her in my arm. Their love and support is what keeps me going.

Contents

List of Figures

List of Tables

List of Acronyms & Symbols

Symbols

$(\cdot)^*$	Complex Conjugate Operation
$(\cdot)^H$	Hermitian transpose of a vector or matrix
$(\cdot)^T$	Transpose of a vector or matrix
$(\cdot)^{-1}$	Inverse of a matrix
$\forall$	For all
$\Im$	Imaginary component of a complex number
$\in$	$\mathcal{X} \in \mathcal{Y}$ indicates that $\mathcal{X}$ takes values from the set $\mathcal{Y}$
$\lceil \cdot \rceil$	Ceiling Operation
$\mathbb{C}$	Set of complex numbers
$\mathbb{E}[\cdot]$	Expectation Operation
$\mathbb{Q}(\cdot)$	Non-linear quantizer-mapping function
$\mathbb{R}$	Set of real numbers
$\mathcal{CN}(\cdot, \cdot)$	Complex normal statistical distribution
$\mathcal{O}(\cdot)$	Complexity order
$\triangleq$	Equality by definition

$\Re$	Real component of a complex number
σ_n^2	The variance of the Gaussian noise
$\mathbf{I}_N$	N by N identity matrix
angle$\{\cdot\}$	Phase Extraction
det$(\cdot)$	Determinant of a matrix
diag$(\cdot)$	Diagonal of a matrix
sign$(\cdot)$	Sign of arguments
Tr$(\cdot)$	Trace of a matrix
$\lvert\cdot\rvert$	Absolute value of a complex number
$\lVert\cdot\rVert_2$	2 norm of a vector or matrix
$\lVert\cdot\rVert_F$	Frobenius norm of a vector or matrix
$rank(\cdot)$	Rank of a matrix

Acronyms/Abbreviations

4G	Fourth Generation
5G	Fifth Generation
ADE	Asymptotic Deterministic Equivalent
BD	Block Diagonalization
BER	Bit Error Rate
CAGR	Compound Annual Growth Rate
CSI	Channel State Information
DAC	Digital-to-Analog Converter

DPC	Dirty Paper Coding
FDD	Frequency-Division Duplexing
IP	Internet Protocol
LTE	Long Term Evolution
MF	Matched Filter
MIMO	Multiple-Input Multiple-Output
MMSE	Minimum Mean Square Error
mmWave	Millimetre Wave
OFDM	Orthogonal Frequency Division Multiplexing
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
SER	Symbol Error Rate
SIC	Successive Interference Cancellation
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SQINR	Signal-to-Quantization-Interference-Noise Ratio
SVD	Singular Value Decomposition
TDD	Time-Division Duplexing
TH	Tomlinson-Harashima
ULA	Uniform Linear Array
VP	Vector Perturbation

WiMAX	Worldwide Interoperability for Microwave Access
ZF	Zero Forcing

Chapter 1

Introduction

1.1 Introduction

In the past few years, 4G wireless systems, which were standardized in 2012, have provided improved service quality in terms of throughput, spectral efficiency and latency [1]. There are two 4G candidate systems commercially deployed, the Mobile WiMAX 2.0 and the LTE Advanced [1]. The common feature of both candidate systems is that they will provide All-IP connectivity with flexible bit rates and quality of service guarantees for multiple classes of services including voice, mainly using voice over IP, data and video services. This offers theoretical speeds of up to 1.5 Gbps, but the current crop of LTE Advanced networks have a maximum potential speed of 300 Mbps with real world speeds falling a lot lower. With the benefits of 4G networks, wireless internet connectivity will be faster and more affordable. However, with the development of the information and communication technology, recent studies predict that the number of mobile-connected devices is expected to reach 11.6 billion by 2021, including machine-to-machine modules. According to a white paper from Cisco [2], mobile data traffic will grow at a compound annual growth rate (CAGR) of 47 percent from 2016 to 2021, reaching 49 exabytes per month by 2021. Consequently, 5G wireless systems are proposed to deal with the challenges of the exponentially growing communication traffic and spectrum bands with wider

bandwidth [3–5]. In order to achieve high array gain and high spatial multiplexing gain, massive MIMO [6–8] employs an unprecedented number of base station antennas simultaneously to serve a small number of single-antenna user terminals in the same channel. Based on the massive MIMO systems, research on hardware complexity and energy consumption has captured the attention of researchers all over the world.

Massive MIMO technology brings huge improvements in spectral efficiency and energy efficiency with the employment of very large antenna arrays at the base stations. Compared with conventional MIMO, massive MIMO can increase the spectral efficiency 10 times or more and simultaneously improve the energy efficiency on the order of 100 times [4]. Additionally, massive MIMO can reduce the latency, simplify the multiple access layer and increase the robustness against interference and intentional jamming [4]. Although massive MIMO systems offer huge advantages, there are still challenges ahead for practical implementation, such as hardware requirements and signal processing [7]. Conventional MIMO systems process the complex signals digitally in the digital domain and then upconvert to the carrier frequency, thus every antenna element needs to be coupled with one RF chain, which includes the digital-to-analog convertors, mixers and power amplifiers. When the number of antennas at the base station is very large, a large number of RF chains will result in excessively high hardware cost and power consumption [9]. Because of these considerations, this thesis focuses on the design of novel hybrid precoding schemes to overcome the constraints of limited number of RF chains. Moreover, in terms of the spectral efficiency, the performance of low complexity precoding schemes is analyzed.

1.2 Motivation and Challenges

With an increase in the number of antennas at the base station, traditional linear precoding schemes such as ZF and minimum mean square error (MMSE) are able to achieve near optimal performance achieved by the dirty paper coding in the downlink communication [6] [8]. However, when the antenna size scales large, traditional schemes require a large number of RF chains. Due to the tremendous number of RF chains, massive MIMO systems will suffer from huge fabrication cost and energy consumption [10]. In order to deal with this problem, cost-effective variable phase shifters are employed to handle the mismatch between the number of RF chains and of antennas. Variable phase shifters with high-dimensional phase-only RF processing are exploited to control the phases of the upconverted RF signal [11–13], which are digitally controlled and changed in a reasonably low time scale for variable channels. Therefore, hybrid precoding schemes are proposed [10,14–16], which exploit a phase-only RF precoder in the analog domain and a baseband precoder in the digital domain. Although the hybrid precoding contributes to reduce power consumption and increase energy efficiency, it leads to the performance degradations. The efficient hybrid precoding schemes should be developed to achieve the best tradeoff between performance and hardware complexity.

Moreover, the millimetre wave (mmWave) frequencies have been put forward as prime candidates for future generation cellular systems, with the potential bandwidth reaching 10 GHz [17] [18]. Thanks to the decrease in the wavelength of mmWave MIMO systems, the large scale antenna arrays at the transmitters can provide significant beamforming gains to overcome path loss. The hybrid baseband and RF processing is particularly appropriate for mmWave MIMO systems, because the mmWave systems rely heavily on RF processing and the hybrid processing can effectively reduce the excessive cost of RF chains. In order to further reduce the hardware complexity, the number of phase shifters in use can also be reduced. Therefore, the hybrid precoding architectures can be categorized into the fully-connected

and sub-connected architectures [19]. In the fully-connected architecture, each RF chain is connected to all transmitting antennas via phase shifters, while in the sub-connected architecture, each RF chain is connected to only a subset of transmitting antennas [20]. Although the sub-connected architecture sacrifices some beamforming gain, it can significantly reduce the hardware implementation complexity without obvious performance loss.

Likewise, another approach to reduce the power consumption is the use of low-resolution DACs for each antenna and RF chain [21] [22]. The power consumption of the DACs grows linearly with increases in bandwidth and exponentially with the number of quantization bits [23]. With a large number of required DACs in massive MIMO systems, the systems will suffer from prohibitively high power consumption. Therefore, the resolution of DACs must be limited to make the power consumption acceptable. It is worth studying the impact of quantized precoding for the downlink massive MIMO systems.

1.3 Aims and Objectives

The aim of this thesis is to provide a framework to achieve efficient massive MIMO systems with low computational and hardware complexity through the introduction of proposed hybrid precoding schemes. Moreover, the performance of different precoding schemes will be explored and analyzed. The objectives of this thesis are:

- To design the hybrid precoding scheme with fully-connected architecture for the finite dimensional channel in massive MIMO systems.
- To design the hybrid precoding scheme with sub-connected architecture for mmWave massive MIMO systems.
- To analyze the performance of the low complexity precoding schemes in terms of spectral efficiency and bit error rate (BER).

1.4 Summary of Contributions

The contributions of this thesis are focused on the design and performance analysis of novel low complexity precoding schemes for massive MIMO systems. The following list highlights and summarizes the main contributions of this thesis:

In Chapter 3, a low complexity hybrid precoding scheme is proposed for the downlink transmission of massive multi-user MIMO systems with a finite dimensional channel model. A tight upper bound on the spectral efficiency is derived and the performance of hybrid precoding is investigated. Simulation results show that the proposed hybrid precoding achieves spectral efficiency close to that achieved by the optimal ZF precoding and performs better than existing hybrid precoding schemes from the literature.

In Chapter 4, based on successive refinement, a new iterative hybrid precoding scheme is proposed with a sub-connected architecture for mmWave massive MIMO systems. Then an upper bound on the spectral efficiency is derived with a closed-form expression. The energy efficiency and the complexity of different hybrid precoding schemes are analyzed. Numerical results demonstrate that the proposed hybrid precoding scheme approaches the performance of the optimal unconstrained singular value decomposition (SVD) and has higher energy efficiency and better BER performance than the fully-connected architecture.

In Chapter 5, the impact of one-bit ZF precoding is studied for massive MIMO systems with the uplink and downlink hardware mismatch. The Bussgang theorem and random matrix theorem are used to derive the closed-form expressions for the achievable rate, through the evaluation of expectations and asymptotic deterministic equivalents (ADEs) of a series of random variables. The numerical simulations indicate the validation of the accuracy of the approximation expressions.

1.5 Organization of the Thesis

This thesis is organized as follows:

Chapter 2 provides the background theory and current research related to this work. The fundamentals of massive MIMO are introduced and millimetre wave communications are presented in detail. In addition, a comprehensive survey of the precoding schemes is described, such as, linear precoding, non-linear precoding and hybrid precoding, where the theoretical and methodological contributions to the MIMO precoding are summarized.

Chapter 3 introduces a hybrid precoding design for an M-dimensional channel model. By analyzing the structure of the channel model, the beamsteering codebooks are combined with extracting the phase of the conjugate transpose of the fast fading matrix to design the RF precoder. Then the baseband precoder is designed with ZF precoding based on the equivalent channel obtained from the product of the RF precoder and the channel matrix. With perfect channel state information, the spectral efficiency is analyzed and the performance of the proposed hybrid precoding scheme is evaluated in simulation.

Chapter 4 presents an iterative hybrid precoding scheme with a sub-connected architecture. Based on successive refinement, in each iteration, the first step is to design the RF precoder and the second step is to design the baseband precoder. The RF precoder is regarded as an input to update the baseband precoder until the stopping criterion is triggered. Phase extraction is used to obtain the RF precoder and then the baseband precoder is optimized by the orthogonal property. A comprehensive performance analysis is then carried out in terms of spectral efficiency, energy efficiency and computational complexity.

Chapter 5 studies the impact of one-bit ZF precoding for massive MIMO systems with the uplink and downlink hardware mismatch. With the use of low-resolution DACs for each antenna and RF chain, the hardware complexity can be reduced effectively. Moreover, in more practical scenario, the hardware mismatch is considered

between the uplink and the downlink for the channel matrix, where the downlink is not the transpose of the uplink in TDD mode. Based on these, using the Bussgang theorem and random matrix theorem, the closed-form analytical expressions are derived for the achievable rate and the performance approximation.

Chapter 6 concludes the thesis and future work in this field is also presented.

1.6 Publications Related to the Thesis

1. **Y. Chen**, S. Boussakta, C. Tsimenidis, J. Chambers and S. Jin, "Low complexity hybrid precoding in finite dimensional channel for massive MIMO systems", *in Proc. 25rd European Signal Processing Conference (EUSIPCO)*, Kos, Greece, 2017.

Chapter 2

Literature Review

2.1 MIMO Communications

MIMO communication systems were first investigated to focus on point-to-point scenarios. At the transmitter, the serial data is transformed into several parallel data sub-stream and sent by multiple transmit antennas. The multiplex signals received by multiple receive antennas are processed in space and time domain and then switched back to the serial original data at the receiver. The benefits offered by multiple antennas are mainly to harvest the spatial diversity and spatial multiplexing gains. The spatial diversity gain aims to enhance the communication reliability by reducing BER [24] [25] and the spatial multiplexing gain contributes to substantially increase the communication capacity of the system by multiplexing more data streams [26].

With the development of MIMO communication technology, multi-user MIMO is proposed [27], where a base station equipped with multiple antennas simultaneously communicates with a set of single antenna mobile terminals. In multi-user MIMO scenarios, the base station requires expensive equipment and the mobile terminals can be relatively cheap using single antenna devices. Therefore, multi-user MIMO is more practical to provide the high capacity, increased diversity and interference suppression and being deployed throughout the world.

Point-to-point MIMO channel is shown in Fig. 2.1, where the numbers of the transmitter antennas and the receiver antennas are denoted as N_t and N_r, respectively. It is assumed that $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ denotes the channel matrix and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{N_r})$ represents an additive white Gaussian noise vector with elements having zero mean and σ_n^2 variance. The processed received signal can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{y}$ is the $\mathbb{C}^{N_r \times 1}$ received signal vector and $\mathbf{s}$ is the $\mathbb{C}^{N_t \times 1}$ transmitted signal vector.

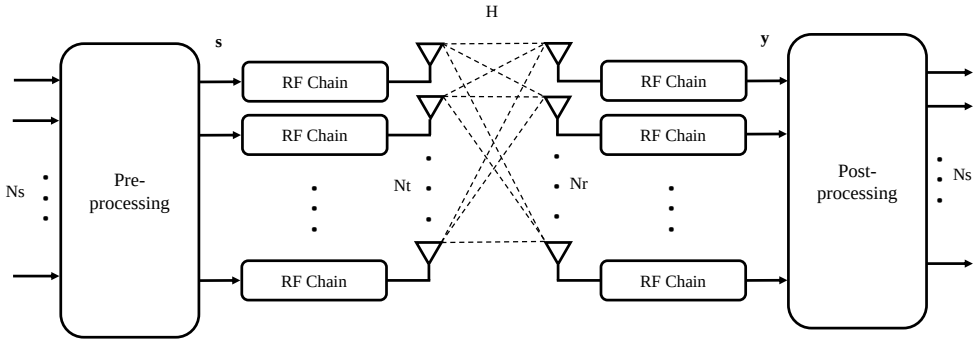


Figure 2.1: Point-to-point MIMO system.

It is assumed that perfect channel state information is known at both transmitter and receiver sides, the capacity of point-to-point MIMO systems can be expressed as

$$C = \log_2 \det \left(\mathbf{I} + \frac{P_t}{N_s \sigma_n^2} \mathbf{H}\mathbf{W}\mathbf{H}^H \right), \tag{2.2}$$

where P_t is the total average transmit power, N_s is the number of transmitted data streams and $\mathbf{W}$ is the power allocation matrix. $\mathbf{W} = \text{diag}\{w_1, w_2, \ldots, w_{N_t}\}$ with w_i being different transmit power. The SVD of the channel matrix is $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H$ [28], where $\mathbf{U} \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{V} \in \mathbb{C}^{N_t \times N_t}$ are unitary matrices and Σ is an $N_r \times N_t$ diagonal matrix of singular values λ_i in descendant order. Therefore, by decomposing the MIMO channel into N_{min} single-input single-output (SISO)

channels, (2.2) can be written as

$$C = \sum_{i=1}^{N_{min}} \log_2 \left(1 + \frac{P_t}{N_s \sigma_n^2} \hat{w_i} \lambda_i^2\right), \tag{2.3}$$

where $N_{min} = min(N_r, N_t)$ is the rank of channel matrix and $\hat{w_i}$ is the optimal power allocation value using the waterfilling algorithm to maximize the system capacity. $\hat{w_i}$ can be given as

$$\hat{w_i} = \max\left(0, \mu - \frac{N_s \sigma_n^2}{P_t \lambda_i^2}\right), \tag{2.4}$$

where μ is the waterfilling level, which is chosen to respect the total power constraint.

2.2 Massive MIMO

In recent years, massive MIMO, upgraded from the conventional MIMO technology, has been widely studied to achieve substantial improvements in spectral efficiency and energy efficiency [6] [7]. Massive MIMO was proposed by Thomas L.Marzetta in 2010 [8], which employs an unprecedented number of base station antennas simultaneously to serve a smaller number of mobile terminals in the same channel. When the number of base station antennas grows asymptotically to infinity, the effects of uncorrelated noise, small-scale fading and intra-cell interference will be eliminated, which makes the channel between the base station and each mobile terminal near-orthogonal. Moreover, large antenna arrays can also achieve large multiplexing and array gains.

Massive MIMO depends on spatial multiplexing, which further relies on the base station to have perfect channel state information, both on the uplink and downlink. Time-division duplex (TDD) is preferred in massive MIMO system, where channel reciprocity can be exploited to get channel state information (CSI) from channel estimation for the uplink. That means the mobile terminals could obtain the same CSI directly for the downlink as it is estimated by using uplink received

pilots. Of course in practical, compared with the downlink CSI, the uplink CSI could be inaccurate or outdated as the channel is fast time-varying, which would affect the performance of the system. Unlike TDD, frequency-division duplexing (FDD), where channel reciprocity cannot be exploitable, is less used. The reason is that the overhead scales linearly with the number of antennas in FDD [29], which makes it difficult to be deployed in massive MIMO system. Nonetheless, FDD mode is still a promising research direction and many researchers have been working on it.

2.2.1 Multi-user Channel

Consider a multi-user massive MIMO system in a single cell scenario shown in Fig. 2.2, where the base station is equipped with N antennas and serves K single-antenna mobile terminals. The channel coefficient is denoted from the n-th antenna of the base station to the k-th mobile terminal as $h_{k,n}$:

$$h_{k,n} = \sqrt{d_k} t_{k,n}, \tag{2.5}$$

where d_k represents the real large-scale fading coefficients and $t_{k,n}$ represents the

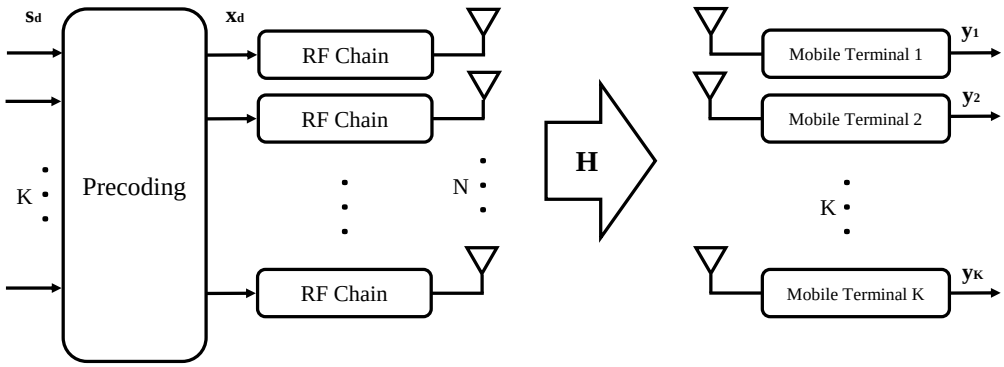


Figure 2.2: Multiuser massive MIMO system.

complex small-scale fading coefficients. The large-scale fading accounts for path loss and shadow fading, thus the fading coefficients $\mathbf{D}^{1/2}$ are assumed to be the same for different base station antennas. The small-scale fading coefficients $\mathbf{T}$ are assumed to be different for each mobile terminal and each antenna at the base station. Then,

the downlink channel matrix $\mathbf{H}$ from the base station to all K mobile terminals can be represented as the product of a $K \times K$ diagonal matrix $\mathbf{D}^{1/2}$ and a $K \times N$ matrix $\mathbf{T}$:

$$\mathbf{H} = \mathbf{D}^{1/2}\mathbf{T}, \tag{2.6}$$

where

$$\mathbf{D} = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_K \end{pmatrix}, \tag{2.7}$$

$$\mathbf{T} = \begin{pmatrix} t_{1,1} & \cdots & t_{1,N} \\ \vdots & \ddots & \vdots \\ t_{K,1} & \cdots & t_{K,N} \end{pmatrix}. \tag{2.8}$$

In the downlink transmission, the received signal $\mathbf{y}_d$ can be written as

$$\mathbf{y}_d = \sqrt{\rho_d}\mathbf{H}\mathbf{x}_d + \mathbf{n}_d, \tag{2.9}$$

where $\mathbf{x}_d = [x_1^d, x_2^d, \cdots, x_N^d]^T$ is the vector of the transmitted signal with $\mathbb{E}[\mathbf{x}_d\mathbf{x}_d^H] = \mathbf{I}_N$, $\mathbf{n}_d \in \mathbb{C}^{K\times 1}$ represents an additive white Gaussian noise vector with elements having zero mean and unit variance and ρ_d is the downlink transmit power.

The channel state information is assumed to be known at both base station and mobile terminals. Therefore, in order to maximize the sum transmission rate, the sum capacity of the downlink multi-user massive MIMO systems with power allocations is defined as [4] [30]

$$C = \max_{\mathbf{W}} \log_2 \det(\mathbf{I}_N + \rho_d \mathbf{H}^H \mathbf{W}\mathbf{H}), \tag{2.10}$$

where $\mathbf{W}$ is a positive diagonal matrix whose diagonal elements $(w_1, w_2, \cdots, w_K)$ represent the power allocations for each mobile terminal. The power constraint is denoted as $\sum_{k=1}^{K} w_k = 1$. In multi-user massive MIMO systems, the number of base station antennas N tends to infinity [8], which exceeds the number of mobile terminals. Because the small-scale fading coefficients of $\mathbf{T}$ are independent for different mobile terminals, the row-vectors of the channel matrix for different mobile terminals are asymptotically orthogonal, hence [31]

$$\begin{aligned}\mathbf{H}\mathbf{H}^H &= \mathbf{D}^{1/2}\mathbf{T}\mathbf{T}^H\mathbf{D}^{1/2} \\ &\approx N\mathbf{D}^{1/2}\mathbf{I}_K\mathbf{D}^{1/2} \\ &= N\mathbf{D}. \end{aligned} \tag{2.11}$$

Then the asymptotic sum capacity can be expressed as

$$C \approx \max_{\mathbf{W}} \log_2 \det(\mathbf{I}_K + \rho_d N\mathbf{W}\mathbf{D}). \tag{2.12}$$

For simplicity, matched filter (MF) precoder is used to process the vector of signal for all mobile terminals and the transmitted signal is given by

$$\begin{aligned}\mathbf{x}_d &= \mathbf{P}\mathbf{D}^{-1/2}\mathbf{W}^{1/2}\mathbf{s}_d \\ &= \mathbf{H}^H\mathbf{D}^{-1/2}\mathbf{W}^{1/2}\mathbf{s}_d, \end{aligned} \tag{2.13}$$

where $\mathbf{s}_d \in \mathbb{C}^{K\times 1}$ is the vector of source signal and $\mathbf{P}$ is the precoding matrix. The processed received signal in (2.9) can be rewritten as

$$\begin{aligned}\mathbf{y}_d &= \sqrt{\rho_d}\mathbf{H}\mathbf{H}^H\mathbf{D}^{-1/2}\mathbf{W}^{1/2}\mathbf{s}_d + \mathbf{n}_d \\ &\approx \sqrt{\rho_d}N\mathbf{D}^{1/2}\mathbf{W}^{1/2}\mathbf{s}_d + \mathbf{n}_d. \end{aligned} \tag{2.14}$$

From (2.12), due to the diagonal matrices $\mathbf{W}$ and $\mathbf{D}$, the MIMO transmission from

the base station to all mobile terminals can be decomposed to multiple SISO transmission. When the power allocation is optimized, the sum capacity can be maximized to achieve capacity optimization.

In the uplink transmission, TDD operation is assumed, thus the uplink channel matrix is the transpose of the downlink channel matrix. The received signal vector y_u can be described as

$$y_u = \sqrt{\rho_u}\mathbf{H}^T\mathbf{x}_u + \mathbf{n}_u, \tag{2.15}$$

where $\mathbf{x}_u = [x_1^u, x_2^u, \cdots, x_K^u]^T$ is the vector of the transmitted signal from all the mobile terminals to the base station with $\mathbb{E}[|x_k^u|^2] = 1$, $\mathbf{n}_u \in \mathbb{C}^{N\times 1}$ represents an additive white Gaussian noise vector with elements having zero mean and unit variance and ρ_u is the uplink transmit power. Assuming that the base station knows the CSI, the capacity for the uplink is

$$C = \log_2\det(\mathbf{I}_K + \rho_u\mathbf{H}^*\mathbf{H}^T). \tag{2.16}$$

Based on the result in (2.11), the asymptotic capacity is given by

$$\begin{aligned} C &\approx \log_2\det(\mathbf{I}_K + N\rho_u\mathbf{D}) \\ &= \sum_{k=1}^{K}\log_2(1 + N\rho_u d_k). \end{aligned} \tag{2.17}$$

Due to the asymptotic orthogonality of the channel vectors, MF processing at the base station becomes asymptotically optimal. Therefore, the processed received signal is the product of the conjugate-transpose of the uplink channel matrix and the received signal vector, as

$$\begin{aligned} \mathbf{H}^*\mathbf{y}_u &= \mathbf{H}^*(\sqrt{\rho_u}\mathbf{H}^T\mathbf{x}_u + \mathbf{n}_u) \\ &\approx N\sqrt{\rho_u}\mathbf{D}\mathbf{x}_u + \mathbf{H}^*\mathbf{n}_u. \end{aligned} \tag{2.18}$$

From (2.18), the processed received signals from different mobile terminals are efficiently and perfectly separated into different streams and the inter-user interference is asymptotically neglected. The signal transmission from each mobile terminal can be considered as a SISO channel with SNR = $N\rho_u d_k$.

2.2.2 Millimetre Wave Channel

The exponentially growth in communication traffic has exacerbated spectrum congestion in current frequency bands, thus new spectrum bands are exploited for future communications. MmWave wireless communications have been demonstrated as a promising approach to solve the spectrum congestion problem, which make the use of spectrum from 30 GHz to 300 GHz [17, 18, 32, 33]. Recent studies show that mmWave frequencies can be used to augment the current microwave frequencies in the range between 700 MHz - 2.6 GHz and the comparison of available bandwidth at microwave and mmWave frequencies is shown in Table 2.1 [34].

On one hand, the use of mmWave frequencies in wireless communications still needs to overcome some technical difficulties. Because of the atmospheric absorption [35], rain attenuation [36] and low penetration, the ten-fold increase in carrier frequency will suffer from less favourable propagation loss [34] [37], which is much higher than that of conventional frequency bands. On the other hand, the decreased wavelength of mmWave enables a large antenna array to be packed in small physical dimension [18]. The large antenna array can provide sufficient antenna gain to overcome the severe path loss of mmWave channel. Further, by the use of precoding techniques, the large antenna array may support the transmission of multiple data streams to improve spectral efficiency and allow systems to approach capacity [10] [38].

In the mmWave systems, due to the practical consideration of the high frequency and bandwidth, there are new constraints on the hardware, such as power consumption and circuit technology, which stimulate intensive interest in overcoming these

Table 2.1: The Comparison of Available Bandwidth at Microwave and MmWave Frequencies

Band	**Generation**	**Wavelength (m)**	**Frequency Range (GHz)**	**Available Bandwidth (GHz)**
microwave	2G	0.3750	0.791 - 0.862	2.5
		0.3333	0.880 - 0.959	
	3G	0.1667	1.710 - 1.880	
	4G	0.1429	1.920 - 2.169	
		0.1154	2.500 - 2.690	
	5G	0.5000	0.470 - 0.694	
		0.4286	0.694 - 0.790	
		0.2000	1.427 - 1.518	
		0.0857	3.300 - 3.800	
		0.0638	4.500 - 4.990	
		0.0536	5.500 - 5.700	
mmWave	5G	0.0130	22.55 - 23.55	23
		0.0107	27.50 - 31.23	
		0.0079	38.6 - 40.0	
		0.0075	40.5 - 42.5	
		0.0065	45.5 - 46.9	
		0.0064	47.2 - 48.2	
		0.0061	48.2 - 50.2	
		0.0041	71 - 76	
		0.0036	81 - 86	
		0.0032	92 - 95	

constraints. In order to reduce the number of digital-to-analog converters and their resolutions, the signal processing operations can be partitioned into digital and analog domains, which lead to the research on hybrid precoding schemes, beamspace signal processing [39] and low resolution DAC methods. Moreover, since the phase sifters may suffer from quantized phase and insertion loss [40], the performance analysis of the impairments of the analog components and the novel algorithms which can achieve good performance even in presence of impairments are promising research areas.

2.2.2.1 Propagation Characteristics

For mmWave massive MIMO systems with free space propagation, based on the Friis' Law [41], the ratio between received power P_r and transmitted power P_t is given by

$$\frac{P_r}{P_t} = G_r G_t \left(\frac{\lambda}{4\pi d}\right)^2, \tag{2.19}$$

where G_r and G_t are the receive and transmit antenna gains, respectively, λ is the wavelength and d is the distance between transmitter and receiver. The path loss in (2.19) is given by $16\pi^2(d/\lambda)^2$, which implies that if the system is operated at two different frequencies with the same transmitter and receiver gains, mmWave propagation will experience a higher path loss than conventional lower frequencies.

MmWave signals are very susceptible to blockages, because the reduced diffraction and specular propagation are exhibited in mmWave frequencies, which will lead to a nearly bimodal channel based on the presence or absence of line-of-sight. Due to the sensitivity to blockages, the transmission link can be easily influenced.

However, mmWave frequencies with shorter wavelengths lead to a tremendous increase in the number of antenna elements within the same physical area, which can provide higher array gains to compensate large path loss values. For an N-element uniform linear array (ULA) with constant length L, we define the inter-element

spacing is $\lambda/2$, thus the number of elements N can be expressed as

$$N = \frac{2L}{\lambda}. \tag{2.20}$$

The array gains are proportional to the number of elements N. Therefore, the higher array gains at high frequencies can compensate for the increased path loss in free space propagation.

The research on mmWave propagation characteristics has captured the attention of researchers over the past years. In small cell scenarios, for distance of up to 200m, the path loss with mmWave frequencies achieves as good performance as conventional cellular frequencies [17, 42, 43], which shows the feasibility of mmWave system in outdoor urban environments. For the urban cellular deployments, the absorption of atmospheric and rain can be beneficial because it further attenuates interference from more distant base stations. In addition, the limitation of mmWave, which can not penetrate brick and concrete, may make the mmWave cells located strictly indoor or strictly outdoor.

2.2.2.2 Channel Model

Due to the high free space path loss, mmWave propagation suffers from the limited spatial selectivity or scattering. The large antenna array is implemented to combat the high path loss, which leads to high levels of antenna correlation. In sparse scattering environments, the large antenna array makes many of the statistical fading distributions used in traditional MIMO analysis inaccurate for mmWave channel. Therefore, a clustered mmWave channel model is introduced to characterize its key features.

When describing the channel model for mmWave systems with N_t transmit and N_r receive antennas, we generally refer to the array response vectors as a function of the antenna array structure only. For an N-element ULA, the array response vector

can be given by [14]

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{N}}[1, e^{j\frac{2\pi}{\lambda}dsin(\theta)}, e^{j\frac{4\pi}{\lambda}dsin(\theta)}, \cdots, e^{j(N-1)\frac{2\pi}{\lambda}dsin(\theta)}]^T, \tag{2.21}$$

where λ is the wavelength of the signal and d denotes the distance between any two adjacent antenna elements and θ is the physical angle of arrival or departure. Based on the form of (2.21), the channel matrix is assumed to be the sum of all N_c scattering clusters, where each cluster contributes N_p propagation paths. Under these circumstances, the mmWave channel model can be defined as [10, 44, 45]

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{N_c N_p}} \sum_{i=1}^{N_c} \sum_{j=1}^{N_p} \beta_{ij} \Lambda_r(\theta_{ij}) \Lambda_t(\phi_{ij}) \mathbf{a}_r(\theta_{ij}) \mathbf{a}_t(\phi_{ij})^H, \tag{2.22}$$

where β_{ij} is the complex gain of the j-th path in the i-th scattering cluster. θ_{ij} and ϕ_{ij} represent the azimuth angles of arrival and departure, respectively. The functions $\Lambda_r(\theta_{ij})$ and $\Lambda_t(\phi_{ij})$ are the receive and transmit antenna element gains at the azimuth angles of θ_{ij} and ϕ_{ij}, respectively. $\mathbf{a}_r(\theta_{ij})$ and $\mathbf{a}_t(\phi_{ij})$ are the receive and transmit array response vectors at the azimuth angles of arrival and departure.

In terms of the N_p azimuth angles of arrival and departure in the i-th cluster, θ_{ij} and ϕ_{ij} are assumed to be randomly distributed with a uniformly random mean values of θ_i and ϕ_i, respectively. The ranges of θ_{ij} and ϕ_{ij} are defined as $[\theta_{\min}, \theta_{\max}]$ and $[\phi_{\min}, \phi_{\max}]$. In addition, the angular spreads of θ_{ij} and ϕ_{ij} in all clusters are assumed to be constant with σ_θ and σ_ϕ. The distribution for the angles of arrival and departure is found that the Laplacian distribution is a good choice to generate all the θ_{ij}'s and ϕ_{ij}'s [46]. Based on these definitions, it is assumed that the transmit and receive antenna elements are modelled as being ideal sectored elements. Thus, $\Lambda_t(\phi_{ij})$ can be given by [10]

$$\Lambda_t(\phi_{ij}) = \begin{cases} 1 & \forall \phi_{ij} \in [\phi_{\min}, \phi_{\max}], \\ 0 & \text{otherwise}, \end{cases} \tag{2.23}$$

and $\Lambda_r(\theta_{ij})$ can be given by

$$\Lambda_r(\theta_{ij}) = \begin{cases} 1 & \forall \theta_{ij} \in [\theta_{\min}, \theta_{\max}], \\ 0 & \text{otherwise}, \end{cases} \tag{2.24}$$

where the transmit and receive antenna element gains are unit over the azimuth sectors.

2.2.3 Potentials and Challenges

2.2.3.1 Potentials

1) Massive MIMO has the capability that it can improve the energy efficiency by 100 times and increase the capacity by 10 times or more [47]. In massive MIMO systems, the spatial multiplexing technique leads to the increase in capacity. The improvement of energy efficiency is because the energy can be concentrated in small regions in the space, with the large number of antennas. Based on the coherent superposition of wavefronts, after sending out the shaped signals from the antennas, the base station can confirm that all the wavefronts emitted from the antennas will add up constructively at the intended terminals' locations and destructively elsewhere. In order to suppress the interference between mobile terminals, ZF algorithm is used at the cost of increased transmitted power.

   Besides ZF, MF is also a good choice for massive MIMO systems, because it reduces the computational complexity with multiplication of the received signals by the conjugate channel responses and is performed in a distributed mode, independently at every antenna element. Although MF performs worse than ZF for the conventional MIMO systems, it works well for massive MIMO systems, because, with large number of base station antennas, the channel responses with different mobile terminals tend to be almost orthogonal. With

MF algorithm, the power can be scaled down as much as possible without seriously affecting the overall spectral efficiency and multi-user interference. Compared to conventional MIMO systems, massive MIMO achieves the overall 10 times higher spectral efficiency, because the systems serve more terminals simultaneously in the same time frequency resource [48].

2) Massive MIMO systems can be built with low-cost and low-power components. In conventional MIMO systems, the base station equips with few antennas, which are fed from high power amplifiers. However, in massive MIMO systems, a large number of antennas lead to hundreds of amplifiers, thus it is infeasible to use high power amplifiers. In order to overcome the problem, the low-cost amplifiers with output power in the milliwatt range are used in massive MIMO systems. Using a large number of antennas, the limits on accuracy and linearity of every amplifier and RF chain are reduced and their combination action becomes more significant. Moreover, the noise, fading and hardware imperfections are averaged with the signals from a large number of antennas combined together in the free space, which increase the robustness of massive MIMO systems.

   In addition, in order to reduce the huge energy consumed by the cellular base stations, the renewable resources such as solar or wind can be used to consume less power. If the base stations are deployed to the places where electricity is not available, the renewable resources can be a good choice to address this problem. Along with this, the electromagnetic interference generated by the base stations can also be substantially reduced.

3) Massive MIMO permits a significant decrease in latency on the air interface. When the signal is transmitted from the base station to the terminal, it will travel through multiple paths, which results from the scattering, reflection and diffraction. The fading makes the wireless communication systems suffer

from latency, because the strength of the signal through different paths can be reduced to a considerable low point. If the terminal is trapped in a fading dip, it has to wait until the transmission channel changes until any data can be received. However, in massive MIMO systems, based on the large antenna arrays, beamforming can effectively avoid fading dip and further decrease the latency [4].

Massive MIMO can also simplify the multiple access layer. Thanks to the use of large antenna arrays, the channel strengthens and the frequency domain scheduling is not suitable. In orthogonal frequency division multiplexing (OFDM) systems, massive MIMO provides each subcarrier with the same channel gain, because each terminal can be provided with the whole bandwidth. Therefore, most of the physical layer control signalling can be redundant [4].

4) Massive MIMO increases the strength against the unintended man-made interference and intentional jamming. For the cyber security, intentional jamming is a growing concern. Massive MIMO can provide the methods to improve robustness of wireless communications by the multiple antennas. Massive MIMO also offers an excess of degrees of freedom to cancel the signals from intended jammers. Using joint channel estimation and decoding, massive MIMO systems can reduce the harmful interference by smart jammers, instead of the conventional uplink pilots of channel estimation [47].

2.2.3.2 Challenges

1) Channel State Information Acquisition:

Due to channel estimation and feedback issues, the TDD transmission mode relying on channel reciprocity is regarded almost as a requirement for realistic implementations of massive MIMO systems [49] [50]. Channel reciprocity can be exploited to get CSI from channel estimation for the uplink. That means

the mobile terminals could obtain the same CSI directly for the downlink as it is estimated by using uplink received pilots. Of course in practical, compared with the downlink CSI, the uplink CSI could be inaccurate or outdated as the channel is fast time-varying, which would affect the performance of the system.

Unlike TDD, in the FDD transmission mode, channel reciprocity cannot be exploitable, where the downlink and uplink transmissions operate at different frequencies [51]. Using FDD, the uplink and downlink channels are characterized by two separate channel matrices. For the uplink channel estimation, all terminals send different pilot sequences to the base station and the time required for uplink pilot transmission is independent of the number of base station antennas. For the downlink channel estimation, the time required for downlink pilot transmission is proportional to the number of base station antennas, because the pilot signals are transmitted from the base station to all terminals first and then all terminals feed back estimated CSI for the downlink channels to the base station. The whole coherence time may be used for the downlink channel estimation, leaving no time for data transmission. Therefore, in massive MIMO systems, the large antenna arrays make the FDD mode infeasible.

However, the research on the FDD transmission has captured the attention of researchers as a very interesting approach in massive MIMO systems. There are several possible methods to enable FDD mode. One way is to design efficient precoding schemes based on partial CSI or even no CSI. Another way is to use the idea of compressed sensing to reduce the feedback overhead. Therefore, more investigations of the challenges and feasibility for FDD operation in massive MIMO are needed.

2) Pilot Contamination:

In typical multi-cell massive MIMO systems, the pilot sequences employed by the terminals in adjacent cells may no longer be orthogonal to those within the cell, because the number of orthogonal pilots is smaller than the number of terminals. This inevitably causes interference among pilots in different cells and incurs an ultimate limitation to obtain optimal system performance and sufficiently accurate channel estimation for the uplink, leading to the pilot contamination problem [7]. The channel parameters are estimated from not only the desired link in the target cell but also the interference links in neighbouring cells. The interference rejection performance on massive MIMO systems is challenging in practice due to huge CSI signalling overhead and backhaul signalling latency.

In the multi-cell scenario, it is assumed that the pilot sequence of the k-th terminal in the l-th cell is $\boldsymbol{\psi}_{k,l} = [\psi_{k,l}^1, \psi_{k,l}^2, \cdots, \psi_{k,l}^\tau]^T$, where τ denotes the length of the pilot sequence. In terms of non-interference between terminals, the pilot sequences employed by the terminals within the same cell and between neighbouring cells are orthogonal, thus

$$\boldsymbol{\psi}_{k,l}^H \boldsymbol{\psi}_{i,j} = \delta[k-i]\delta[l-j], \tag{2.25}$$

where $\delta[\cdot]$ is defined as

$$\delta[x] = \begin{cases} 1 & x = 0, \\ 0 & x \neq 0. \end{cases} \tag{2.26}$$

Therefore, the base station can obtain the estimation of the channel matrix without pilot contamination. However, if the period and bandwidth are limited, the number of orthogonal pilot sequences limits the number of terminals that can be served by the multi-cell multi-user systems. In massive MIMO systems, the base stations are expected to serve more terminals, thus non-

orthogonal pilot sequences are utilized in adjacent cells. In this case, suffering from pilot contamination,

$$\boldsymbol{\psi}_{k,l}^{H}\boldsymbol{\psi}_{i,j} \neq 0. \tag{2.27}$$

Thus, the estimation of channel matrix will be affected by the non-orthogonal pilot sequences.

Recently, various approaches are proposed to reduce and even eliminate the pilot contamination phenomenon. For example, the efficient channel estimation algorithms or blind transmission techniques that circumvent the use of pilots can be used to mitigate the influence of pilot contamination. The pilot contamination precoding is also a promising method to reduce the interference with cooperative transmission.

3) Hardware and Computational Complexity

With hundreds of antennas at the base station, massive MIMO systems require hundreds of RF chains, which include a large number of digital-to-analog convertors, mixers and power amplifiers. This leads to huge power consumption, hardware cost and complexity. Therefore, the need for cheaper and low power hardware becomes a significant factor for massive MIMO systems. To reduce implementation cost and complexity without obvious performance loss, an electromagnetic lens antenna is proposed in [52] [53] to provide spatial multipath separation and energy focusing functions.

In addition, the computational complexity of precoding schemes increases together with the number of antennas at the base station. Therefore, the design of low complexity precoding algorithms is needed in massive MIMO systems, which reduces the time for data transmission. Moreover, the hybrid precoding schemes can also contribute to reduce the hardware complexity by reducing the number of RF chains.

2.3 MIMO Precoding

Precoding is a generalization of beamforming to support multi-stream transmission in multi-antenna wireless communications. In multi-user MIMO systems, precoding techniques are generally preferred for the downlink communications. The precoding techniques are applied at the source signals before they are transmitted which can overcome the interference between mobile terminals. Moreover, because the base station has high computing ability and power supply, precoding is a reliable technique that exploits the channel state information available at the transmitting side in order to be convenient for signal detection at the receiving side, which can reduce the burdens of signal processing and simplify the structure at the mobile terminals.

For regular MIMO systems, precoding techniques can be divided into two categories, linear precoding [54–56] and non-linear precoding [57] and there are essential differences between them. Linear precoding has poor performance with low implementation complexity, while non-linear precoding has good performance with high implementation complexity. Theoretical analysis illustrates that the non-linear precoding methods such as Dirty Paper Coding (DPC) could achieve the capacity of MIMO Gaussian broadcast channel [58]. In massive MIMO systems, linear precoding schemes are shown to be near-optimal with the large antenna arrays. Thus, it is more practical to use low complexity linear precoding techniques in massive MIMO systems. However, the use of linear precoding brings us new problems in massive MIMO systems, which makes the systems suffer from excessively high hardware cost and power consumption. Then we introduce some efficient hybrid precoding schemes to overcome the constraints of hardware complexity.

2.3.1 Linear Precoding

With the perfect CSI known at the transmitter, we define the transmitted signal $\mathbf{x}$ is derived as a linear combination of the source signal $\mathbf{s}$. Therefore, the transmitted

signal can be calculated as

$$\mathbf{x} = \mathbf{P}\mathbf{s}, \tag{2.28}$$

where $\mathbf{P}$ is the precoding matrix. Since the transmitted power is limited by P_t, the precoding matrix has to be designed to satisfy the transmit power constraint, that is

$$\mathbb{E}[\|\mathbf{P}\mathbf{s}\|_2^2] = P_t. \tag{2.29}$$

Some fundamental linear precoding techniques are introduced for MIMO communications as follows.

MF is the simplest linear precoding scheme to maximize the received SNR, ignoring multi-user interference. The MF precoding matrix $\mathbf{P}_{\mathrm{MF}}$ can be computed as the Hermitian of the channel matrix $\mathbf{H}$,

$$\mathbf{P}_{\mathrm{MF}} = \frac{1}{\beta_{\mathrm{MF}}}\mathbf{H}^H, \tag{2.30}$$

where $\beta_{\mathrm{MF}} = \frac{1}{\sqrt{P_t}}\sqrt{\mathrm{tr}(\mathbf{H}\mathbf{H}^H)}$ is a power normalization factor.

In multi-user scenarios, if each terminal is equipped with a single antenna, interferences from other signals cannot be cancelled. In order to deal with the multi-user interference, ZF precoding is widely used. The ZF precoding matrix $\mathbf{P}_{\mathrm{ZF}}$ can be computed as the pseudoinverse of the channel matrix,

$$\mathbf{P}_{\mathrm{ZF}} = \frac{1}{\beta_{\mathrm{ZF}}}\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}, \tag{2.31}$$

where $\beta_{\mathrm{ZF}} = \frac{1}{\sqrt{P_t}}\sqrt{\mathrm{tr}((\mathbf{H}\mathbf{H}^H)^{-1})}$ is a power normalization factor. ZF precoding technique is attractive due to their simplicity, however, when the terminals are equipped with multiple antennas, block diagonalization (BD) method is more suitable. It is

defined that the base station with N_t antennas serves K terminals in the communication systems. The number of antennas for each terminal is N_r. For the i-th terminal, the received signal $\mathbf{y}_i \in \mathbb{C}^{N_r \times 1}$ can be expressed as

$$\mathbf{y}_i = \mathbf{H}_i \sum_{k=1}^{K} \mathbf{P}_k^{\mathrm{BD}} \mathbf{x}_k + \mathbf{n}_i, \tag{2.32}$$

where $\mathbf{H}_i \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix between the base station and the i-th terminal, $\mathbf{P}_k^{\mathrm{BD}} \in \mathbb{C}^{N_t \times N_r}$ is the BD precoding matrix for the k-th terminal, $\mathbf{x}_k \in \mathbb{C}^{N_r \times 1}$ is the source signal vector and $\mathbf{n}_i \in \mathbb{C}^{N_r \times 1}$ is the noise vector. The equivalent combined channel matrix of all terminals is given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_1 & \cdots & \mathbf{H}_1 \\ \mathbf{H}_2 & \mathbf{H}_2 & \cdots & \mathbf{H}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_K & \mathbf{H}_K & \cdots & \mathbf{H}_K \end{bmatrix}. \tag{2.33}$$

Therefore, the received signals for all terminals are written as

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_K \end{bmatrix} &= \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_1 & \cdots & \mathbf{H}_1 \\ \mathbf{H}_2 & \mathbf{H}_2 & \cdots & \mathbf{H}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_K & \mathbf{H}_K & \cdots & \mathbf{H}_K \end{bmatrix} \begin{bmatrix} \mathbf{P}_1^{\mathrm{BD}} \mathbf{x}_1 \\ \mathbf{P}_2^{\mathrm{BD}} \mathbf{x}_2 \\ \vdots \\ \mathbf{P}_K^{\mathrm{BD}} \mathbf{x}_K \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_K \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_1 \mathbf{P}_1^{\mathrm{BD}} & \mathbf{H}_1 \mathbf{P}_2^{\mathrm{BD}} & \cdots & \mathbf{H}_1 \mathbf{P}_K^{\mathrm{BD}} \\ \mathbf{H}_2 \mathbf{P}_1^{\mathrm{BD}} & \mathbf{H}_2 \mathbf{P}_2^{\mathrm{BD}} & \cdots & \mathbf{H}_2 \mathbf{P}_K^{\mathrm{BD}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_K \mathbf{P}_1^{\mathrm{BD}} & \mathbf{H}_K \mathbf{P}_2^{\mathrm{BD}} & \cdots & \mathbf{H}_K \mathbf{P}_K^{\mathrm{BD}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_K \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_K \end{bmatrix}. \end{aligned} \tag{2.34}$$

Based on (2.34), a matrix $\mathbf{H}^{-i}$ is constructed excluding the channel matrix of i-th

terminal as

$$\mathbf{H}^{-i} = [\mathbf{H}_1^H \cdots \mathbf{H}_{i-1}^H \mathbf{H}_{i+1}^H \cdots \mathbf{H}_K^H]^H, \tag{2.35}$$

where $\mathbf{H}^{-i} \in \mathbb{C}^{(N_t - N_r) \times N_t}$ and $N_t = KN_r$. In order to eliminate the multi-user interference, BD method is used to make the off-diagonal term $\mathbf{H}_i \mathbf{P}_j^{\mathrm{BD}}$ in (2.34) equal to $\mathbf{0}_{N_r \times N_r}$. Thus, (2.35) should be satisfied with the constraint

$$\mathbf{H}^{-i} \mathbf{P}_i^{\mathrm{BD}} = \mathbf{0}_{(N_t - N_r) \times N_r}, \quad i = 1, 2, \cdots, K. \tag{2.36}$$

This implies that the precoding matrix $\mathbf{P}_i^{\mathrm{BD}}$ must be designed to lie in the null space of $\mathbf{H}^{-i}$. With the use of BD, the received signals for all terminals can be rewritten as

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_K \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \mathbf{P}_1^{\mathrm{BD}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \mathbf{P}_2^{\mathrm{BD}} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H}_K \mathbf{P}_K^{\mathrm{BD}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_K \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_K \end{bmatrix}. \tag{2.37}$$

In terms of the design of BD precoding matrix, it is assumed the rank of $\mathbf{H}^{-i}$ is $N_{\mathrm{rank}} = N_t - N_r$. The SVD of $\mathbf{H}^{-i}$ can be expressed as

$$\begin{aligned} \mathbf{H}^{-i} &= \mathbf{U}_i \Sigma_i \mathbf{V}_i^H \\ &= \mathbf{U}_i \Sigma_i \begin{bmatrix} \mathbf{V}_i^{(1)} & \mathbf{V}_i^{(2)} \end{bmatrix}^H, \end{aligned} \tag{2.38}$$

where $\mathbf{U}_i$ is an $N_{\mathrm{rank}} \times N_{\mathrm{rank}}$ unitary matrix, $\mathbf{V}_i$ is an $N_t \times N_t$ unitary matrix and Σ_i is an $N_{\mathrm{rank}} \times N_t$ diagonal matrix of singular values in descendant order. $\mathbf{V}_i$ is divided into two parts, $\mathbf{V}_i^{(1)} \in \mathbb{C}^{N_t \times N_{\mathrm{rank}}}$ consists of the N_{rank} non-zero singular vectors and

$\mathbf{V}_i^{(2)} \in \mathbb{C}^{N_t \times N_r}$ consists of the N_r zero singular vectors. Multiplying $\mathbf{H}^{-i}$ with $\mathbf{V}_i^{(2)}$,

$$\begin{aligned}\mathbf{H}^{-i}\mathbf{V}_i^{(2)} &= \mathbf{U}_i \begin{bmatrix} \Sigma_i^{(1)} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_i^{(1)^H} \\ \mathbf{V}_i^{(2)^H} \end{bmatrix} \mathbf{V}_i^{(2)} \\ &= \mathbf{U}_i \Sigma_i^{(1)} \mathbf{V}_i^{(1)^H} \mathbf{V}_i^{(2)} \\ &= \mathbf{U}_i \Sigma_i^{(1)} \mathbf{0} \\ &= \mathbf{0}, \end{aligned} \tag{2.39}$$

where $\Sigma_i^{(1)}$ is the first partition of dimension $N_{\text{rank}} \times N_{\text{rank}}$. Thus, $\mathbf{V}_i^{(2)}$ forms an orthogonal for the null space of $\mathbf{H}^{-i}$ and $\mathbf{P}_i^{\text{BD}} = \mathbf{V}_i^{(2)}$ can be used for precoding the signal of the i-th terminal with BD constraint. However, the main computational complexity of BD precoding comes from the SVD operation, which makes the computational complexity increase with the number of terminals and the system dimensions.

2.3.2 Non-linear Precoding

Compared with linear precoding methods, non-linear methods, such as DPC [29], Tomlinson-Harashima (TH) precoding and vector perturbation (VP) offer significant sum-rates benefits, however, with the use of non-linear operations, the systems will suffer from higher implementation complexity.

DPC method was proposed by Costa in 1983, which showed that if the interference is known to the transmitter, the theoretical channel capacity can be achieved and the effect of the interference can be cancelled. In multi-user systems, when designing precoding matrix for the k-th terminal, the interferences caused by the first up to $(k-1)$-th terminals are considered to be cancelled. In addition, such remarkable performance can be achieved without the need of additional power in transmission nor of shared CSI with the receiver. However, DPC is infeasible in practical, because it requires infinite length codewords and complicated methods for

signal processing.

TH precoding is an equalization technique originally proposed to cancel the inter-symbol interference. In MIMO systems, TH precoding can be used to eliminate the interference between different sub-channels. Although TH precoding suffers a performance loss compared to DPC, it is feasible to implement in practice instead of DPC method.

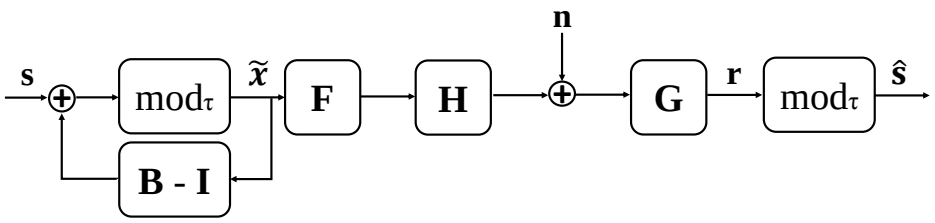


Figure 2.3: The block diagram of TH precoding based on LQ decomposition.

Fig. 2.3 shows the structure of TH precoding and we assume the numbers of transmitter and receiver antennas are equal to N for simplicity. TH precoding can be implemented on the basis of LQ decomposition, thus, the channel matrix $\mathbf{H}$ is decomposed into the multiplication of a lower triangular matrix $\mathbf{L}$ and a unitary matrix $\mathbf{Q}$,

$$\mathbf{H} = \mathbf{LQ}. \tag{2.40}$$

A scaling matrix $\mathbf{G}$ is defined to contain the corresponding weighted coefficient for each stream. Thus it should have a diagonal structure and the diagonal elements is corresponding to the inverse of the diagonal elements of $\mathbf{L}$, which is given by

$$\mathbf{G} = \begin{bmatrix} l_{11}^{-1} & 0 & \cdots & 0 \\ 0 & l_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{NN}^{-1} \end{bmatrix}, \tag{2.41}$$

where l_{ii} is the i-th diagonal element of the matrix $\mathbf{L}$. The feedback matrix $\mathbf{B}$ is

the product of **G** and **L**, which is utilized to cancel the interference caused by the previous streams from the current stream. The feedback matrix **B** is computed as

$$\mathbf{B} = \mathbf{GL}. \tag{2.42}$$

Therefore, the matrix **B** is a lower triangular matrix with ones on the main diagonal. It is defined the feedforward matrix **F** as the conjugate transpose of the unitary matrix **Q**, because the power of the transmitted signal keeps constant using the unitary matrix. The feedforward matrix **F** is implemented at the transmitter to enforce the spatial causality.

It is assumed that $\mathbf{s} = [s_1, s_2, \cdots, s_N]^T$ is the source signal and $\tilde{\mathbf{x}} = [\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_N]^T$ is the pre-signal. Based on the idea of serial interference cancellation, the interference of the transmit signals can be cancelled and the amplitude of pre-signals can be also limited by modulus operation. The pre-signal $\tilde{\mathbf{x}}$ is computed as

$$\begin{aligned}
&\tilde{x}_1 = s_1 \\
&\tilde{x}_2 = \mathrm{mod}_\tau \left[s_2 - b_{2,1}\tilde{x}_1 \right] \\
&\vdots \\
&\tilde{x}_i = \mathrm{mod}_\tau \left[s_i - \sum_{j=1}^{i-1} b_{i,j}\tilde{x}_j \right] \\
&\vdots \\
&\tilde{x}_N = \mathrm{mod}_\tau \left[s_N - \sum_{j=1}^{N-1} b_{N,j}\tilde{x}_j \right],
\end{aligned} \tag{2.43}$$

where $b_{i,j}$ is the element of the matrix **B** in i-th row and j-th column and mod_τ is the function of modulus operation to adjust the transmission power. τ is a real number that depends on the chosen modulation. From (2.43), the i-th terminal needs to cancel the interferences caused by the first up to $(i-1)$-th terminals, which is similar to DPC method.

With the effect of the modulus operation, a modified source signal can be obtained by adding a perturbation vector $\mathbf{d}$ to the source signal $\mathbf{s}$, which is given by

$$\mathbf{v} = \mathbf{s} + \mathbf{d}. \tag{2.44}$$

Based on Fig. 2.3, the feedback processing is mathematically equivalent to an inversion operation $\mathbf{B}^{-1}$, thus the received signal can be expressed as

$$\begin{aligned} \mathbf{r} &= \mathbf{G}(\mathbf{HFB}^{-1}\mathbf{v} + \mathbf{n}) \\ &= \mathbf{GHFB}^{-1}\mathbf{v} + \mathbf{Gn}. \end{aligned} \tag{2.45}$$

Finally, the receiver needs to apply an additional modulo operation in order to estimate the received signal.

In addition, [59] proposed a new non-linear precoding method called VP approach. When using VP, a perturbation vector is added to the source signal vector. However, the perturbation vector can not be arbitrary, because this vector is not known to the receivers, thus the receivers cannot eliminate the effect of the perturbation effectively. Based on TH precoding, the elements of the source signal are perturbed by an integer, then the effect of the perturbation vector can be eliminated by the modulus operation at the receiver. Assume the channel matrix is $\mathbf{H}$ of dimension $N \times N$ and the source signal vector is $\mathbf{s}$ of dimension $N \times 1$, and the perturbed signal vector $\tilde{\mathbf{s}}$ is written by

$$\tilde{\mathbf{s}} = \mathbf{s} + \tau\mathbf{l}, \tag{2.46}$$

where τ is a positive real number and $\mathbf{l}$ is a N-dimensional complex integer vector.

Based on the ZF-VP, the transmitted signal $\mathbf{x}$ is given by

$$\mathbf{x} = \frac{1}{\sqrt{\gamma}}\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}\tilde{\mathbf{s}}, \tag{2.47}$$

where $\gamma = \|\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}\tilde{\mathbf{s}}\|_F^2$ is the power of the transmitted signal vector. Then the received signal vector $\mathbf{y}$ is expressed as

$$\begin{aligned}\mathbf{y} &= \sqrt{\gamma}(\mathbf{H}\mathbf{x} + \mathbf{n}) \\ &= \sqrt{\gamma}(\frac{1}{\sqrt{\gamma}}\mathbf{H}\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}\tilde{\mathbf{s}} + \mathbf{n}) \\ &= \tilde{\mathbf{s}} + \sqrt{\gamma}\mathbf{n},\end{aligned} \tag{2.48}$$

where $\mathbf{n}$ is the noise vector. After the received signal is processed by the modulus operation, ignoring the effect of $\mathbf{n}$,

$$\mathrm{mod}_\tau(\mathbf{y}) = \mathrm{mod}_\tau(\mathbf{s} + \tau\mathbf{l}) = \mathbf{s}, \tag{2.49}$$

where τ is a real number that depends on the chosen modulation. Therefore, we can recover the source signal with VP method. From (2.48), $\sqrt{\gamma}$ is the main factor to design the perturbation vector $\mathbf{l}$. If γ is very large, the system performance will degrades significantly. Thus the transmit power normalization γ is minimized to design the optimal perturbation vector $\mathbf{l}$,

$$\begin{aligned}\mathbf{l} &= \arg\ \min_{\mathbf{l}} \gamma \\ &= \arg\ \min_{\mathbf{l}} \|\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}\tilde{\mathbf{s}}\|_F^2 \\ &= \arg\ \min_{\mathbf{l}} \|\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}(\mathbf{s} + \tau\mathbf{l})\|_F^2.\end{aligned} \tag{2.50}$$

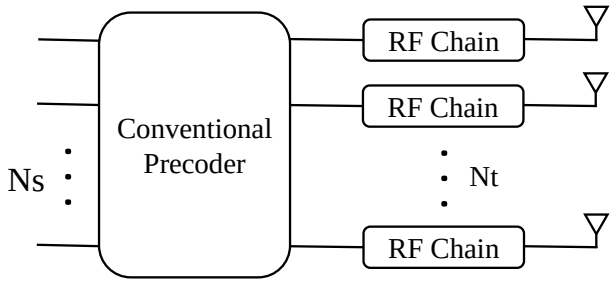
The choice of $\mathbf{l}$ is a N-dimensional integer-lattice least-squares problem, for which there exist a large number of approximate algorithms.

2.3.3 Hybrid Precoding

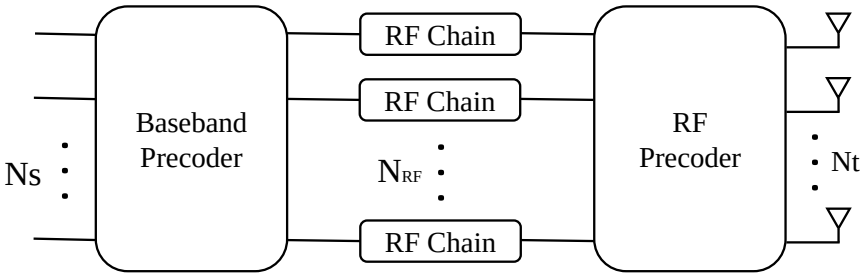
In mmWave MIMO system with a large number of antennas, the conventional linear and non-linear precoding schemes are infeasible, because the system will suffer from prohibitively high energy consumption and hardware complexity. To solve this problem, the hybrid precoding schemes are proposed to reduce the number of RF chains, with cost-effective variable phase shifters employed to handle the mismatch between the number of antennas and RF chains. The comparison between conventional precoding and hybrid precoding is shown in Fig. 2.4, where $N_{\mathrm{RF}} << N_t$. Hybrid precoding exploits a high-dimensional phase-only RF precoder in the analog domain and a low-dimensional baseband precoder in the digital domain. The transmitted signals are precoded by the baseband precoding first to guarantee the performance, and then precoded by the RF precoding to save the energy consumption and reduce the hardware complexity.

In mmWave massive MIMO systems, [60] and [10] use hybrid methods to obtain near optimal SVD performance, which decompose the optimal precoder and combiner through the concept of orthogonal matching pursuit. Due to the high complexity of searching columns of the overcomplete matrix in [10], a low-complexity hybrid sparse precoding method is proposed in [61] using a greedy method with the element-wise normalization of the first singular vector of the residual. Unlike the point-to-point MIMO systems, the multi-user MIMO systems not only suffer from the noise and inter-antenna interference but are also influenced by the multi-user interference. Thus, the methods above are not suitable for the multi-user scenario. Several hybrid precoding schemes are proposed to solve the problem. In [14], limited feedback hybrid precoding is proposed with short training and feedback overhead to achieve near optimal block diagonalization performance. A phased ZF precoding in [9] applies phase-only control at the RF domain and then performs a low-dimensional baseband ZF precoding based on the effective channel seen from baseband, which approaches the performance of the virtually optimal full complexity ZF precoding

in a massive multi-user MIMO scenario.



(a) Conventional precoding scheme



(b) Hybrid precoding scheme

Figure 2.4: The comparison between conventional precoding and hybrid precoding schemes.

The precoding schemes above are called fully-connected architecture, which means each RF chain is connected to all base station antennas. When the base station antenna number is fairly large, the fully-connected architecture requires a large number of phase shifters for the RF precoder, which leads to both high energy consumption and hardware complexity. In order to further reduce the energy consumption and hardware complexity, the hybrid precoding design with sub-connected architecture is considered in mmWave massive MIMO systems. In the sub-connected architecture, each RF chain is connected to only a subset of base station antennas, instead of all base station antennas. The advantage of the sub-connected architecture is to reduce the number of required phase shifters, which can be more energy-efficient and more practical for antenna deployment. In [19], some hybrid MIMO architectures are proposed to reduce the cost, complexity, and power consumption of mmWave MIMO systems, while incurring small loss in the system performance. The constraints on the original hybrid precoding problem with sub-connected architecture is different

from that with fully-connected architecture. Focusing on the optimization of achievable capacity, [62] proposes a successive interference cancellation (SIC)-based hybrid precoding with near-optimal performance and low complexity.

In this section, two effective hybrid precoding schemes are introduced in detail. In point-to-point communication scenario, formulating the achievable rate optimization problem as a sparse approximation problem, the spatially sparse precoding can obtain near optimal SVD performance through the concept of orthogonal matching pursuit. In multi-user communication scenario, limited feedback hybrid precoding is proposed to achieve near optimal block diagonalization performance, which involves an iterative searching procedure among the predefined codebook.

2.3.3.1 Spatially Sparse Precoding

In point-to-point massive MIMO scenario, the mmWave channel matrix $\mathbf{H}$ is defined as (2.22) and the spectral efficiency is given by

$$R = \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s} \mathbf{R}_n^{-1} \mathbf{W}_{\mathrm{BB}}^H \mathbf{W}_{\mathrm{RF}}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{H}^H \mathbf{W}_{\mathrm{RF}} \mathbf{W}_{\mathrm{BB}} \right), \quad (2.51)$$

where N_s is the number of data streams from the transmitter with N_t antennas to the receiver with N_r antennas. $\mathbf{W}_{\mathrm{RF}}$ is the $N_r \times N_r^{\mathrm{RF}}$ RF combining matrix and $\mathbf{W}_{\mathrm{BB}}$ is the $N_r^{\mathrm{RF}} \times N_s$ baseband combining matrix, where N_r^{RF} is the number of RF chains at the receiver. $\mathbf{F}_{\mathrm{RF}}$ is the $N_t \times N_t^{\mathrm{RF}}$ RF precoding matrix and $\mathbf{F}_{\mathrm{BB}}$ is the $N_t^{\mathrm{RF}} \times N_s$ baseband precoding matrix, where N_t^{RF} is the number of RF chains at the transmitter. $\mathbf{R}_n = \sigma_n^2 \mathbf{W}_{\mathrm{BB}}^H \mathbf{W}_{\mathrm{RF}}^H \mathbf{W}_{\mathrm{RF}} \mathbf{W}_{\mathrm{BB}}$ is the noise covariance matrix after combining.

The hybrid precoding scheme proposed in [10] is design to maximize the spectral efficiency expression in (2.51). To simplify the design, assuming that the optimal combiners are used at the receiver, the hybrid precoding is design to maximize the mutual information. The mutual information is equivalent to the spectral efficiency

over MIMO channel when Gaussian inputs are used, which is expressed as

$$I = \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{H}^H \right). \tag{2.52}$$

Therefore, the precoding optimization problem can be stated as

$$\begin{aligned} \max_{\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}} \quad & I, \\ s.t. \quad & \|\mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}}\|_F^2 = N_s, \\ & \mathbf{F}_{\mathrm{RF}} \in \mathcal{F}_{\mathrm{RF}}, \end{aligned} \tag{2.53}$$

where $\mathcal{F}_{\mathrm{RF}}$ is the set of matrices with all constant amplitude entries.

To obtain the optimal unconstrained precoder $\mathbf{F}_{\mathrm{opt}}$, the SVD of the channel matrix $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H$ is considered, where $\mathbf{U}$ is an $N_r \times N_r$ unitary matrix, Σ is an $N_r \times N_t$ diagonal matrix of singular values in descendant order and $\mathbf{V}$ is an $N_t \times N_t$ unitary matrix. Therefore, (2.52) can be rewritten as

$$I = \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma^2 \mathbf{V}^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V} \right). \tag{2.54}$$

The matrices Σ and $\mathbf{V}$ are defined as

$$\Sigma = \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{bmatrix}, \mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix}, \tag{2.55}$$

where Σ_1 is the first partition of dimension $N_\mathrm{s} \times N_\mathrm{s}$ and $\mathbf{V}_1$ is the first N_s columns of $\mathbf{V}$. The optimal precoder can be simply given by

$$\mathbf{F}_{\mathrm{opt}} = \mathbf{V}_1. \tag{2.56}$$

In order to find practical near-optimal precoders that can be implemented in Fig. 2.4(b), the hybrid precoder $\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}$ can be made sufficiently close to the optimal precoder

$\mathbf{V}_1$. Therefore, [10] makes the following system assumption,

$$\mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \approx \mathbf{I}_{N_s}, \mathbf{V}_2^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \approx \mathbf{0}. \tag{2.57}$$

Based on the partitions defined in (2.55), the following partition of the matrix $\mathbf{V}^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}$ can be defined as

$$\begin{aligned} \mathbf{V}^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V} &= \begin{bmatrix} \mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_1 & \mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_2 \\ \mathbf{V}_2^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_1 & \mathbf{V}_2^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_2 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix}. \end{aligned} \tag{2.58}$$

The equivalent approximations defined in (2.57) imply that $\mathbf{Q}_{12}$, $\mathbf{Q}_{21}$ and $\mathbf{Q}_{22}$ are approximately zero. Using (2.55) and (2.58), (2.54) can be approximated as

$$\begin{aligned} I &= \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma^2 \mathbf{V}^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V} \right) \\ &= \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \begin{bmatrix} \Sigma_1^2 & \mathbf{0} \\ \mathbf{0} & \Sigma_2^2 \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \right) \\ &\approx \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \mathbf{Q}_{11} \right) \\ &= \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_1 \right) \\ &\overset{(a)}{=} \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \right) \\ &\quad + \log_2 \det \left(\mathbf{I}_{N_s} - \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \right)^{-1} \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 (\mathbf{I}_{N_s} - \mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_1) \right) \\ &\overset{(b)}{\approx} \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \right) \\ &\quad - \mathrm{tr} \left(\left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \right)^{-1} \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 (\mathbf{I}_{N_s} - \mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_1) \right) \\ &\overset{(c)}{\approx} \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \right) - \mathrm{tr} (\mathbf{I}_{N_s} - \mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{V}_1) \\ &= \log_2 \det \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s \sigma_n^2} \Sigma_1^2 \right) - \left(N_s - \| \mathbf{V}_1^H \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}} \|_F^2 \right), \end{aligned} \tag{2.59}$$

where step (a) can be obtained by defining the matrices $\mathbf{B} = \frac{P_t}{N_s\sigma_n^2}\Sigma_1^2$ and $\mathbf{A} = \mathbf{V}_1^H\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\mathbf{F}_{\mathrm{BB}}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{V}_1$ and noting that $\mathbf{I}+\mathbf{BA} = (\mathbf{I}+\mathbf{B})(\mathbf{I}-(\mathbf{I}+\mathbf{B})^{-1}\mathbf{B}(\mathbf{I}-\mathbf{A}))$. Step (b) follows from (2.57), which implies that the eigenvalues of the matrix $\mathbf{X} = \left(\mathbf{I}_{N_s} + \frac{P_t}{N_s\sigma_n^2}\Sigma_1^2\right)^{-1}\frac{P_t}{N_s\sigma_n^2}\Sigma_1^2(\mathbf{I}_{N_s} - \mathbf{V}_1^H\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\mathbf{F}_{\mathrm{BB}}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{V}_1)$ are small and thus the approximation $\log_2\det(\mathbf{I}_{N_s} - \mathbf{X}) \approx \log_2(1-\mathrm{tr}(\mathbf{X})) \approx -\mathrm{tr}(\mathbf{X})$ is valid. Step (c) follows from adopting a high effective-SNR approximation, where $\left(\mathbf{I}_{N_s} + \frac{P_t}{N_s\sigma_n^2}\Sigma_1^2\right)^{-1}\frac{P_t}{N_s\sigma_n^2}\Sigma_1^2 \approx \mathbf{I}_{N_s}$.

Based on (2.59), the optimization problem in (2.53) is approximately equivalent to minimizing $\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F$, therefore, the hybrid precoding design problem can be rewritten as

$$\begin{aligned} \min_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{BB}}} \quad & \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F \\ s.t. \quad & \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = N_s, \\ & \mathbf{F}_{\mathrm{RF}} \in \mathcal{F}_{\mathrm{RF}}. \end{aligned} \tag{2.60}$$

By exploiting the structure of $\mathbf{H}$, the near-optimal RF precoder can be found by further restricting $\mathcal{F}_{\mathrm{RF}}$ to be the set of vectors of the form $\mathbf{a}_t(\phi_{i,j})$. Therefore, the design problem (2.60) can be expressed as

$$\begin{aligned} \min_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{BB}}} \quad & \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F \\ s.t. \quad & \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = N_s, \\ & \mathbf{F}_{\mathrm{RF}}^{(i)} \in \{\mathbf{a}_t(\phi_{i,j}), \forall i,j\}. \end{aligned} \tag{2.61}$$

In order to solve the design problem above, the solution consists of selecting the best N_t^{RF} array response vectors and finding their optimal baseband combination. An algorithmic solution is proposed with the concept of orthogonal matching pursuit. The algorithm finds the appropriate array response vectors $\mathbf{a}_t(\phi_{i,j})$ first, which has the maximum projection with the optimal unconstrained precoder $\mathbf{F}_{\mathrm{opt}}$ and the

selected column of $\mathbf{a}_t(\phi_{i,j})$ is appended to the RF precoder $\mathbf{F}_{\mathrm{RF}}$. After the dominant vector is found, the baseband precoder $\mathbf{F}_{\mathrm{BB}}$ is calculated by

$$\mathbf{F}_{\mathrm{BB}} = (\mathbf{F}_{\mathrm{RF}}^H \mathbf{F}_{\mathrm{RF}})^{-1} \mathbf{F}_{\mathrm{RF}}^H \mathbf{F}_{\mathrm{opt}}. \tag{2.62}$$

Then the residual matrix $\mathbf{F}_{\mathrm{res}}$ is obtained by removing the contribution of the selected vector, which is defined as

$$\mathbf{F}_{\mathrm{res}} = \frac{\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}}{\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F}. \tag{2.63}$$

With $\mathbf{F}_{\mathrm{res}}$, the iteration is utilized to find the remaining $N_t^{\mathrm{RF}} - 1$ vectors $\mathbf{a}_t(\phi_{i,j})$, which have the maximum projection with $\mathbf{F}_{\mathrm{res}}$. The process above repeats N_t^{RF} times to construct RF precoder $\mathbf{F}_{\mathrm{RF}}$ with dimension of $N_t \times N_t^{\mathrm{RF}}$ and baseband precoder $\mathbf{F}_{\mathrm{BB}}$ with dimension of $N_t^{\mathrm{RF}} \times N_s$, which can minimize the objective function $\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F$. Therefore, this algorithm succeeds in selecting the best N_t^{RF} steering directions and forming appropriate linear combinations of the obtained response vectors, which is close to the optimal precoder.

Fig. 2.5 shows the spectral efficiency achieved by optimal unconstrained precoding and sparse spatially precoding in mmWave massive MIMO systems, where the transmitter and the receiver are equipped with $N_t = 128$ and $N_r = 8$ antennas. Assuming that $N_t^{\mathrm{RF}} = N_s = 4$, sparse spatially precoding achieves spectral efficiency close to that achieved by optimal unconstrained precoding, which implies that this algorithm can accurately approximate the dominant singular vectors of channel matrix with a combination of four steering vectors.

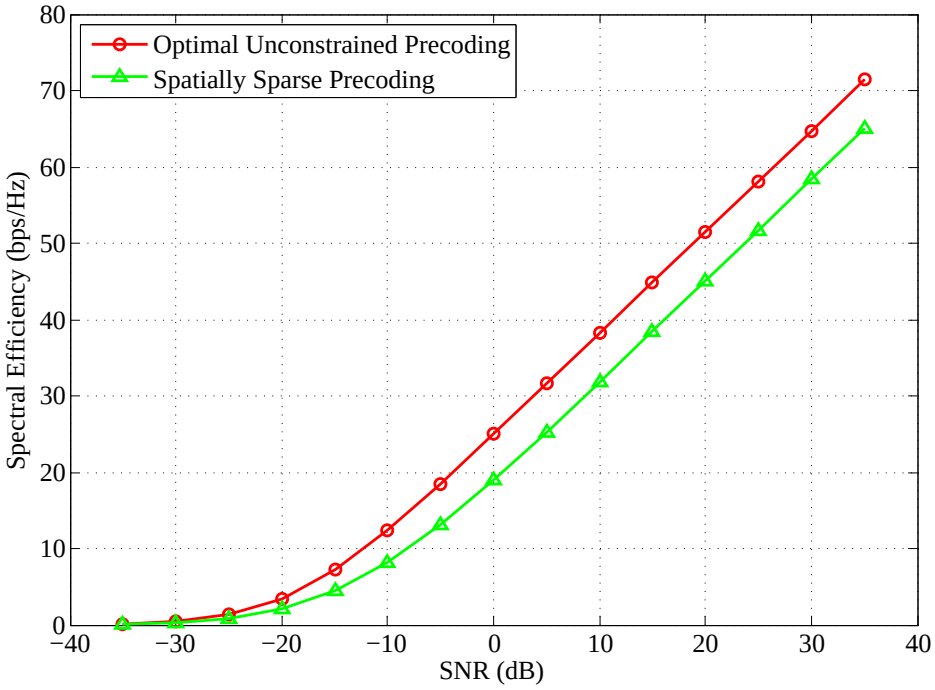


Figure 2.5: Spectral efficiency achieved by spatially sparse precoding for mmWave massive MIMO systems where $N_t = 128$, $N_r = 8$, $N_t^{\mathrm{RF}} = N_s = 4$.

2.3.3.2 Limited Feedback Hybrid Precoding

In multi-user massive MIMO scenario, the received signal at the k-th mobile terminal is defined as [14]

$$y_k = \mathbf{w}_k^{\mathrm{RF}^H} \mathbf{H}_k \sum_{i=1}^{K} \mathbf{F}_{\mathrm{RF}} \mathbf{f}_i^{\mathrm{BB}} s_i + \mathbf{w}_k^{\mathrm{RF}^H} \mathbf{n}_k, \tag{2.64}$$

where $\mathbf{H}_k$ is the $N_r \times N_t$ mmWave channel matrix between the base station and k-th mobile terminal. For simplicity, it is assumed the number of RF chains is equal to the number of mobile terminals K, thus $\mathbf{F}_{\mathrm{RF}} = [\mathbf{f}_1^{\mathrm{RF}}, \mathbf{f}_1^{\mathrm{RF}}, \cdots, \mathbf{f}_K^{\mathrm{RF}}]$ is the $N_t \times K$ RF precoder and $\mathbf{F}_{\mathrm{BB}} = [\mathbf{f}_1^{\mathrm{BB}}, \mathbf{f}_1^{\mathrm{BB}}, \cdots, \mathbf{f}_K^{\mathrm{BB}}]$ is the $K \times K$ baseband precoder. $\mathbf{w}_k^{\mathrm{RF}}$ is the RF combiner with the similar constraints as the RF precoder.

The achievable rate of the k-th mobile terminal can be expressed as

$$R_k = \log_2 \left(1 + \frac{\frac{P_t}{K} |\mathbf{w}_k^{\mathrm{RF}^H} \mathbf{H}_k \mathbf{F}_{\mathrm{RF}} \mathbf{f}_k^{\mathrm{BB}}|^2}{\frac{P_t}{K} \sum_{i \neq k} |\mathbf{w}_k^{\mathrm{RF}^H} \mathbf{H}_k \mathbf{F}_{\mathrm{RF}} \mathbf{f}_i^{\mathrm{BB}}|^2 + \sigma_n^2} \right). \tag{2.65}$$

The sum rate of the system is $R = \sum_{k=1}^{K} R_k$.

In [14], the beamsteering codebooks are adopted for the design of the analog beamforming vectors. The beamsteering codebooks have the same form of the array response vector and can be parameterized by a simple angle. The beamsteering codebook is defined as $\mathcal{F}$ and $\mathcal{F}$ consists of N_Q vectors $\mathbf{a}_t\left(\frac{2\pi n_Q}{N_Q}\right)$ where n_Q takes the values $0, 1, \cdots, N_Q - 1$. Similarly, the beamsteering codebook of the RF combining can be defined as $\mathcal{W}$.

In order to maximize the sum rate of the system, the precoding design problem is given by

$$\left\{\dot{\mathbf{F}}_{\mathrm{RF}}, \{\dot{\mathbf{f}}_k^{\mathrm{BB}}\}_{k=1}^K, \{\dot{\mathbf{w}}_k^{\mathrm{RF}}\}_{k=1}^K\right\} = \arg\max \sum_{k=1}^{K} \log_2\left(1 + \frac{\frac{P_t}{K}|\mathbf{w}_k^{\mathrm{RF}^H}\mathbf{H}_k\mathbf{F}_{\mathrm{RF}}\mathbf{f}_k^{\mathrm{BB}}|^2}{\frac{P_t}{K}\sum_{i\neq k}|\mathbf{w}_k^{\mathrm{RF}^H}\mathbf{H}_k\mathbf{F}_{\mathrm{RF}}\mathbf{f}_i^{\mathrm{BB}}|^2 + \sigma_n^2}\right) \tag{2.66}$$

$$\begin{aligned} s.t. \qquad & [\mathbf{F}_{\mathrm{RF}}]_{:,k} \in \mathcal{F}, \\ & \mathbf{w}_k^{\mathrm{RF}} \in \mathcal{W}, \\ & \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = K. \end{aligned}$$

In order to solve the problem above, [14] proposed to divide the calculation of the precoders into two stage with limited feedback. In the first stage, the RF precoding vectors $\mathbf{f}_k^{\mathrm{RF}}$ and combining vectors $\mathbf{w}_k^{\mathrm{RF}}$ are jointly designed to maximize the desired signal power for the k-th mobile terminal, neglecting the interference caused by other terminals. The design of RF precoder vectors is a typical single-user RF beamforming design problem, thus efficient beam training algorithms developed for single-user systems can be used [63] [64], which do not require explicit channel estimation and have a low training overhead. Then the selected RF precoding vectors are combined into the RF precoding matrx $\mathbf{F}_{\mathrm{RF}}$. In the second stage, the effective channel for each mobile terminal is computed as

$$\bar{\mathbf{h}}_k^H = \mathbf{w}_k^{\mathrm{RF}^H}\mathbf{H}_k\mathbf{F}_{\mathrm{RF}}, \quad k = 1, 2, \cdots K. \tag{2.67}$$

Based on (2.67), the effective channel vector $\bar{\mathbf{h}}_k$ for the system can be obtained with the dimension of $K \times 1$ and each mobile terminal quantizes its effective channel using a codebook. Therefore, the index of the quantized channel vector can be fed back to the base station and the baseband precoder is easily designed using ZF algorithm based on the quantized channel.

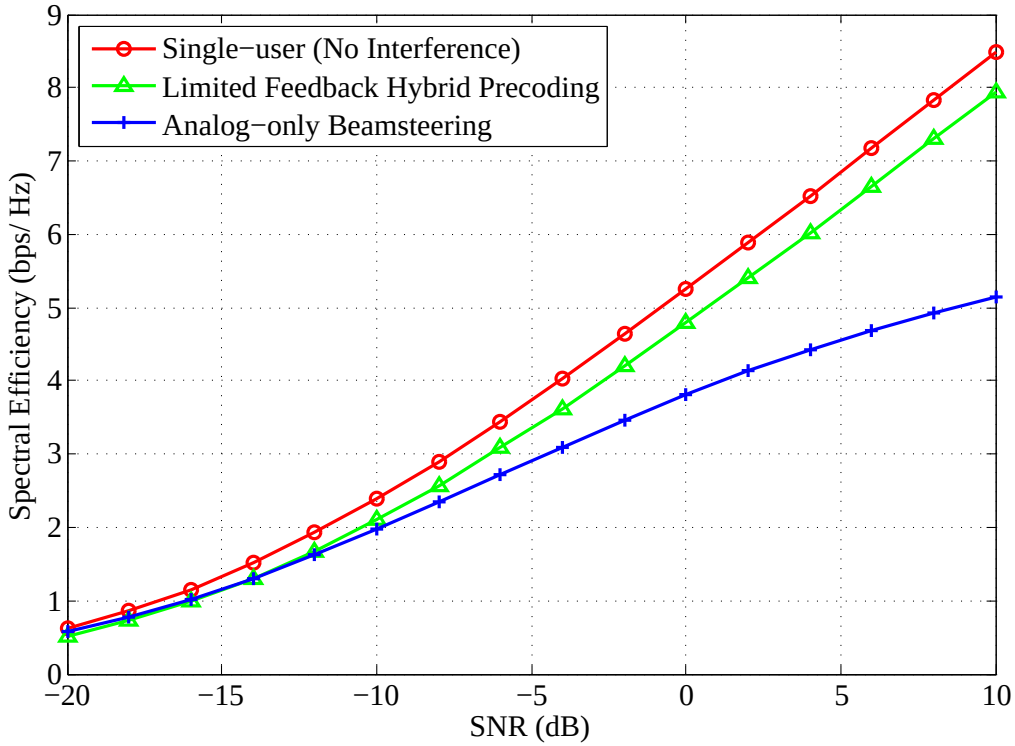


Figure 2.6: Spectral efficiency achieved by limited feedback hybrid precoding for mmWave massive MIMO systems where $N_t = 64$, $N_r = K = 4$.

As shown in Fig. 2.6, the spectral efficiency achieved by the limited feedback hybrid precoding is compared with the single user transmission and analog-only beamsteering. It is assumed that the base station equipped with 64 antennas serves 4 mobile terminals, each having 4 antennas. The figure indicates that the proposed hybrid precoding scheme achieves the spectral efficiency close to that achieved by single user, because the design of baseband precoder can cancel the inter-user interference in the second stage. Moreover, due to the interference limitation, analog-only beamsteering performs significantly worse than the proposed precoding in the high SNR region. Therefore, limited feedback hybrid precoding can achieve good perfor-

mance with low complexity.

2.4 Summary

In this chapter, based on the introduction of massive MIMO systems, the potentials and challenges have been presented, which gives me the inspiration for my research. In particular, the mmWave channel model has been explained in detail, because mmWave communication is a promising candidate for future cellular systems and widely used in massive MIMO systems. Moreover, this chapter has given the details of different precoding schemes, such as, linear precoding, non-linear precoding and hybrid precoding. With a good knowledge of different precoding schemes, we can further develop the novel precoding techniques and improve the performance of massive MIMO systems. To conclude, this chapter provides the essential background theory for the following novel chapters.

Chapter 3

Low Complexity Hybrid Precoding in Finite Dimensional Channel for Massive MIMO Systems

In this chapter, the downlink transmission of massive multi-user MIMO systems is considered with a finite dimensional channel model, in which a limited number of dimensions M is defined to model the channel matrix. The performance of MIMO systems depends critically on the complexity of the propagation environment and the properties of the antenna arrays being used. In massive MIMO systems, large antenna arrays operating at high frequency will imply channel vectors are correlated, hence many statistical fading models of traditional MIMO systems are not applicable. Many previous works focus on point-to-point mmWave channels using the clustered channel model [10,44,45], but in a realistic scenario, very little research has been reported on the multi-user channel model. Therefore, a finite dimensional channel model is designed and analyzed for the large multi-user MIMO systems. The reason is that the finite dimensional channel is able to reflect the property of the poor scattering channel environment caused by high pathloss at high frequency [65] and the angles of arrival at the mobile stations can be neglected. The finite dimensional channel vectors for different mobile stations can be more applicable to obtain

generally correlated or not asymptotically orthogonal channel [66]. In terms of the number of dimensions M, it is assumed that M is much smaller than the number of base station antennas, where M is the number of angular bins.

The hybrid precoding exploits a phase-only RF precoder in the analog domain and a baseband precoder in the digital domain. The RF precoder uses the phase shifters, which are digitally controlled and changed in reasonably low time scale for variable channels. In this chapter, a novel hybrid precoding method is introduced by analyzing the structure of the channel model. The RF precoder is designed by combining the beamsteering codebooks with extracting the phase of the conjugate transpose of the fast fading matrix, which thereby harvests the large array gain achieved by an unprecedented number of base station antennas. Then a baseband precoder is designed based on the equivalent channel with ZF precoding in order to cancel the multi-user interference.

3.1 System Model

The downlink of a massive multi-user MIMO system is considered in Fig. 3.1, where the base station employs N_t transmit antennas to communicate with a population of K single-antenna mobile stations, $N_t \gg K$. For simplicity, the base station is assumed to utilize K RF chains to serve the mobile stations and single data stream handled by the base station is transmitted to every mobile station.

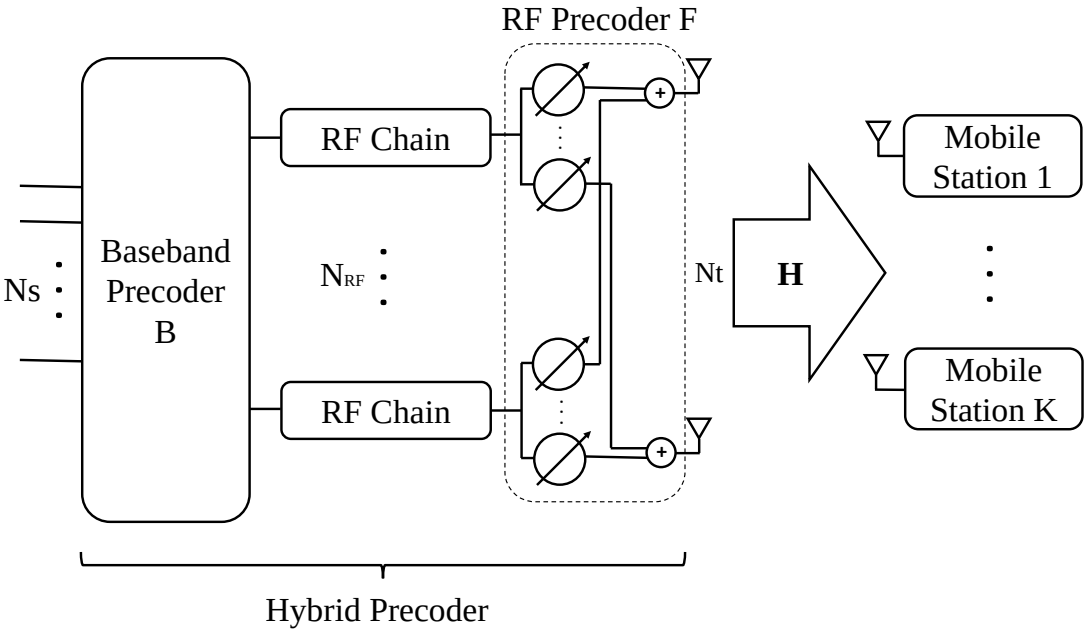


Figure 3.1: System diagram of massive multi-user MIMO systems with hybrid precoding for a finite dimensional channel.

On the downlink, the transmit signal is given by

$$\mathbf{x} = \mathbf{FBs}, \tag{3.1}$$

where $\mathbf{s} = [s_1, s_2, \cdots, s_K]^T$. $\mathbf{s} \in \mathbb{C}^{K\times 1}$ is the vector of signal for K mobile stations and $\mathbb{E}[\mathbf{ss}^H] = \frac{P_t}{K}\mathbf{I}_K$ where P_t is the total transmit power.

Furthermore, it is assumed that perfect channel state information is known at the base station. The signal received at the k-th mobile station is given by

$$y_k = \mathbf{H}_k \sum_{n=1}^{K} \mathbf{FB}_n s_n + n_k, \tag{3.2}$$

where $\mathbf{H}_k$ represents the $1 \times N_t$ downlink channel gains between the base station and k-th mobile station and n_k represents additive white Gaussian noise with i.i.d. entries of zero mean and unit variance, i.e. $\mathcal{CN}(0, \sigma^2)$. $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_K]$ is the RF precoder of dimension $N_t \times K$ and $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_K]$ is the baseband precoder of dimension $K \times K$. The baseband precoder modifies both amplitude and phase while only phase changes can be made by the RF precoder with a network of variable analog phase shifters. Thus, the entries $\mathbf{F}_{m,n}$ in the RF precoder are of constant modulus, which are normalized to satisfy $|\mathbf{F}_{m,n}|^2 = \frac{1}{N_t}$. Due to the total transmit power constraint, the entries of $\mathbf{B}$ are normalized to satisfy $\|\mathbf{FB}\|_F^2 = K$. In this system model, combiners are not used at the mobile stations, which reduces the hardware cost and power consumption.

3.2 Channel Model

The M-dimensional channel model divides the transmitted angular region into M directions, where $N_t \gg M$. In each direction, the angles of departure are denoted as $\theta_m \in [-\pi/2, \pi/2]$, where $m = 1, \cdots, M$. The ULA is employed in this chapter, thus the array response vector $\mathbf{a}_{BS}(\theta_m)$ associated with θ_m presents the transmit antenna array structure, which can be defined as

$$\mathbf{a}_{BS}(\theta_m) = \frac{1}{\sqrt{M}}[1, e^{-j\frac{2\pi}{\lambda}dsin\theta_m}, \cdots, e^{-j(N_t-1)\frac{2\pi}{\lambda}dsin\theta_m}]^T, \tag{3.3}$$

where λ is the wavelength of the signal and d is the distance between antenna elements.

The M-dimensional channel matrix between the base station and K mobile stations can be expressed as

$$\mathbf{H} = \mathbf{GA}^H, \tag{3.4}$$

where $\mathbf{A} = [\mathbf{a}_{BS}(\theta_1)\ \mathbf{a}_{BS}(\theta_2) \cdots \mathbf{a}_{BS}(\theta_M)]$ is an $N_t \times M$ matrix and $\mathbf{G}$ presents the propagation coefficient matrix of dimension $K \times M$ from the base station to mobile stations. The propagation matrix $\mathbf{G}$ is used to model independent fast fading, geometric attenuation and shadow fading [66], which can be given by

$$\mathbf{G} = \mathbf{DW}, \tag{3.5}$$

where $\mathbf{D}$, of dimension $K \times K$, is the diagonal matrix of the geometric attenuation and shadow fading coefficients for the multi-cell scenario. The diagonal elements of $\mathbf{D}$, which are given by β_{kk}, are assumed to be constant, because they vary slowly with time. β_{kk} represents the (k,k)-th entry of $\mathbf{D}$. The entries in the $K \times M$ matrix $\mathbf{W}$ represent the fast fading coefficients between the base station and mobile stations, which are assumed to have zero mean and unit variance [66]. Therefore, assuming $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, \cdots, \mathbf{W}_K]^T$, (3.2) can be written as

$$y_k = \beta_{kk}\mathbf{W}_k\mathbf{A}^H \sum_{n=1}^{K} \mathbf{FB}_n s_n + n_k, \tag{3.6}$$

3.3 Proposed Hybrid Precoding Scheme

The main objective is to achieve optimal spectral efficiency with low hardware complexity. The simplest precoding scheme is to invert the channel matrix using pseudoinverse, which is referred to as ZF precoding. However, in massive MIMO systems, although ZF precoding achieves near optimal capacity performance, a large number of RF chains are required to perform analog-to-digital conversion and frequency translation between baseband and RF, which suffers from high hardware cost and restricts the array size at the base station from scaling to a large dimension. Thus, the proposed precoding scheme leverages the structure of the finite dimensional channel to address the hardware constraints.

3.3.1 RF Precoder Design

The RF precoder $\mathbf{F}$ utilizes the phase shifters to couple K RF chains with N_t base station antennas based on phase-only control of the upconverted RF signal.

Beamsteering codebooks are used and have similar form to the array response vector. $\mathcal{A}$ is assumed to be the beamsteering codebook and because of the phase-only processing, it is designed by the normalized array response vector. The matrix $\widetilde{\mathbf{A}}$ is selected from $\mathcal{A}$ to maximize the desired signal power of mobile station, neglecting the other users' interference.

Then, perfect channel state information is assumed known at the base station. The phases of the conjugate transpose of the fast fading coefficient matrix $\mathbf{W}$ is extracted as $\widetilde{\mathbf{W}}_{m,n} = \frac{1}{\sqrt{M}} e^{j\varphi_{m,n}}$, where $\varphi_{m,n}$ is the phase of the (m,n)-th entry of the conjugate transpose of $\mathbf{W}$. The the diagonal term of $\mathbf{W}\widetilde{\mathbf{W}}$ can be expressed as

$$\mathbf{W}_j \widetilde{\mathbf{W}}_j = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} |\mathbf{W}_{i,j}|, \tag{3.7}$$

where $\widetilde{\mathbf{W}}_j$ is the the j-th column of $\widetilde{\mathbf{W}}$ and $\mathbf{W}_{i,j}$ is the (i,j)-th entry of $\mathbf{W}$. Because each entry of $\mathbf{W}$ is i.i.d. complex Gaussian random variable with unit variance and zero mean, based on the central limit theorem,

$$\mathbf{W}_j \widetilde{\mathbf{W}}_j \sim \mathcal{N}(\frac{\sqrt{\pi M}}{2}, 1 - \frac{\pi}{4}), \tag{3.8}$$

where M has a large value. In addition, in terms of the off-diagonal term of $\mathbf{W}\widetilde{\mathbf{W}}$, we have

$$\mathbf{W}_j \widetilde{\mathbf{W}}_k = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} \mathbf{W}_{i,j} e^{j\varphi_{i,k}}, \tag{3.9}$$

where $j \neq k$. Because $\mathbf{W}_j \widetilde{\mathbf{W}}_k$ follows the distribution with unit variance and zero mean, i.e., $\mathbf{W}_j \widetilde{\mathbf{W}}_k \sim \mathcal{CN}(0,1)$, $|\mathbf{W}_j \widetilde{\mathbf{W}}_k|$ has the mean of $\frac{\pi}{2}$. Compared with the diagonal terms, the off-diagonal terms can be negligible if M is large enough. $\widetilde{\mathbf{W}}$ is

used to further maximize the desired signal power of mobile stations.

Therefore, the design of the RF precoder $\mathbf{F}$ can be given by

$$\mathbf{F} = \widetilde{\mathbf{A}}\widetilde{\mathbf{W}}, \tag{3.10}$$

In realistic application scenarios, the practical constraints of phase shifters makes the phase of each entry of $\mathbf{F}$ suffer from the impact of coarse quantization, which influences the performance of hybrid precoding. Therefore, instead of the continuous values, the situation where the phases of the entries of $\mathbf{F}$ are quantized up to B bits of precision is also considered, i.e., $B = 4$ or 5. The quantized phase is computed by the closest Euclidean distance of its nearest neighbor phase.

3.3.2 Baseband Precoder Design

The baseband precoder $\mathbf{B}$ performs low-dimensional multiple stream processing to change the amplitude and phase of the transmitted complex signal based on ZF precoding.

After obtaining the RF precoder $\mathbf{F}$, an equivalent channel is defined as

$$\overline{\mathbf{H}} = \mathbf{HF}, \tag{3.11}$$

where $\overline{\mathbf{H}} = [\overline{\mathbf{H}}_1, \overline{\mathbf{H}}_2, \cdots, \overline{\mathbf{H}}_K]$.The equivalent channel $\overline{\mathbf{H}}$ is applied to design the baseband precoder $\mathbf{B}$. The dimension of $\overline{\mathbf{H}}$ is $K \times K$, which is much lower than the $N_t \times K$ dimension of the original channel. Based on ZF precoding, the direct baseband precoding matrix $\overline{\mathbf{B}}$ is then defined as

$$\overline{\mathbf{B}} = \overline{\mathbf{H}}^H(\overline{\mathbf{H}}\overline{\mathbf{H}}^H)^{-1}, \tag{3.12}$$

The normalized baseband precoder $\mathbf{B}$ is

$$\mathbf{B} = \frac{\overline{\mathbf{B}}}{\sqrt{\|\mathbf{F}\overline{\mathbf{B}}\|_F^2}}, \tag{3.13}$$

where $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_K]$ to satisfy the precoding power constraints $\|\mathbf{F}\mathbf{B}_k\|_2^2 = 1$. From (3.10) and (3.13), $\mathbf{F}$ and $\mathbf{B}$ of the hybrid precoding are both obtained. The transmitter supports simultaneous transmission of K data streams and the hardware complexity is obviously reduced because there are only K RF chains using the proposed precoding instead of N_t RF chains required by the traditional ZF precoding and $K \ll N_t$, giving a significant reduction of hardware complexity.

3.4 Analysis of Spectral Efficiency

In this section, the achievable spectral efficiency of the transmission is analyzed with perfect channel state information, which is given by

$$R_k = \log_2(1 + \frac{\frac{P_t}{K}|\mathbf{H}_k\mathbf{F}\mathbf{B}_k|^2}{\frac{P_t}{K}\sum_{n=1,n\neq k}^{K}|\mathbf{H}_k\mathbf{F}\mathbf{B}_n|^2 + \sigma^2}). \tag{3.14}$$

The sum spectral efficiency of the system is $R = \sum_{k=1}^{K} R_k$.

Based on (3.14), a tight upper bound on spectral efficiency achieved by hybrid precoding is derived for the finite dimensional channel model, which can be characterized by the following theorem.

Theorem 4.1 The upper bound of the spectral efficiency for the k-th mobile station can be expressed as

$$R_k \leq \log_2(1 + \frac{\mathrm{SNR}}{K}\lambda_{app}(\mathbf{A}\mathbf{A}^H)(\mathbf{W}_k\mathbf{W}_k^H)), \tag{3.15}$$

where $\mathrm{SNR} = \frac{P_t}{\sigma^2}$ and $\lambda_{app}(\mathbf{A}\mathbf{A}^H)$ is the appropriate eigenvalue of matrix $\mathbf{A}\mathbf{A}^H$.

Proof 4.1 The spectral efficiency for the k-th mobile station is

$$\begin{aligned} R_k &= \log_2(1+\frac{\mathrm{SNR}}{K}|\overline{\mathbf{H}}_k\mathbf{B}_k|^2) \\ &= \log_2(1+\frac{\mathrm{SNR}}{K}|\overline{\mathbf{H}}_k\frac{\overline{\mathbf{B}}_k}{\sqrt{\|\mathbf{F}\overline{\mathbf{B}}_k\|_2^2}}|^2) \\ &= \log_2(1+\frac{\mathrm{SNR}}{K}\frac{1}{\|\mathbf{F}\overline{\mathbf{B}}_k\|_2^2}). \end{aligned} \tag{3.16}$$

Based on (3.10) (3.11) and (3.12),

$$\mathbf{F}\overline{\mathbf{B}}_k = \widetilde{\mathbf{A}}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\widetilde{\mathbf{A}}^H\mathbf{A}\mathbf{W}_k^H(\mathbf{W}_k\mathbf{A}^H\widetilde{\mathbf{A}}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\widetilde{\mathbf{A}}^H\mathbf{A}\mathbf{W}_k^H)^{-1}. \tag{3.17}$$

Using the property of the beamsteering codebooks, $\widetilde{\mathbf{A}} = \sqrt{\frac{M}{N_t}}\mathbf{A}$ is obtained to maximize the desired signal power of mobile stations. Therefore, (3.17) can be rewritten by

$$\mathbf{F}\overline{\mathbf{B}}_k = \mathbf{A}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\mathbf{A}^H\mathbf{A}\mathbf{W}_k^H(\mathbf{W}_k\mathbf{A}^H\mathbf{A}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\mathbf{A}^H\mathbf{A}\mathbf{W}_k^H)^{-1}, \tag{3.18}$$

Next,

$$\frac{1}{\|\mathbf{F}\overline{\mathbf{B}}_k\|_2^2} = \frac{|\mathbf{W}_k\mathbf{A}^H\mathbf{A}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\mathbf{A}^H\mathbf{A}\mathbf{W}_k^H|^2}{(\mathbf{A}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\mathbf{A}^H\mathbf{A}\mathbf{W}_k^H)^H(\mathbf{A}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\mathbf{A}^H\mathbf{A}\mathbf{W}_k^H)}. \tag{3.19}$$

Briefly, by defining $\mathbf{X} \triangleq \mathbf{W}_k\mathbf{A}^H\mathbf{A}\widetilde{\mathbf{W}}^H$ and $\mathbf{Y} = \widetilde{\mathbf{W}}\mathbf{A}^H\mathbf{A}$, (3.19) can given by

$$\begin{aligned} \frac{1}{\|\mathbf{F}\overline{\mathbf{B}}_k\|_2^2} &= \frac{|\mathbf{X}\mathbf{Y}\mathbf{W}_k^H|^2}{\mathbf{X}\mathbf{Y}\widetilde{\mathbf{W}}^H\mathbf{X}^H} \\ &\overset{(a)}{\leqslant} \frac{\mathbf{X}\mathbf{Y}(\mathbf{X}\mathbf{Y})^H\mathbf{W}_k\mathbf{W}_k^H}{\mathbf{X}\mathbf{Y}\widetilde{\mathbf{W}}^H\mathbf{X}^H}, \end{aligned} \tag{3.20}$$

where (a) is valid by applying the Cauchy-Schwarz inequality [67]. Substituting $\mathbf{X}$

and $\mathbf{Y}$ into (3.20) and defining $\mathbf{Z} = \mathbf{W}_k\mathbf{A}^H\mathbf{A}\widetilde{\mathbf{W}}^H\widetilde{\mathbf{W}}\mathbf{A}^H$, (3.20) can be rewritten by

$$\begin{aligned}\frac{1}{\|\mathbf{F}\overline{\mathbf{B}}_k\|_2^2} &\leqslant \frac{\mathbf{Z}\mathbf{A}\mathbf{A}^H\mathbf{Z}^H}{\mathbf{Z}\mathbf{Z}^H}\mathbf{W}_k\mathbf{W}_k^H \\ &\overset{(b)}{\leqslant} \lambda_{app}(\mathbf{A}\mathbf{A}^H)\mathbf{W}_k\mathbf{W}_k^H, \end{aligned} \tag{3.21}$$

where the Rayleigh-Ritz theorem [68] is used in (b) because $\mathbf{A}\mathbf{A}^H$ is a Hermitian matrix. ■

Proof of (a) Because the vectors $\mathbf{X}\mathbf{Y} \in \mathbb{C}^{1\times M}$ and $\mathbf{W}_k \in \mathbb{C}^{1\times M}$, a vector $\mathbf{T}$ is defined as

$$\mathbf{T} = \mathbf{W}_k^H - \frac{(\mathbf{X}\mathbf{Y})\mathbf{W}_k^H(\mathbf{X}\mathbf{Y})^H}{\|(\mathbf{X}\mathbf{Y})^H\|_2^2} \tag{3.22}$$

Then due to $\mathbf{X}\mathbf{Y}\mathbf{T} = 0$,

$$\begin{aligned}\|\mathbf{T}\|_2^2 &= \mathbf{T}^H\mathbf{W}_k^H - \frac{(\mathbf{X}\mathbf{Y})\mathbf{W}_k^H\mathbf{T}^H(\mathbf{X}\mathbf{Y})^H}{\|(\mathbf{X}\mathbf{Y})^H\|_2^2} \\ &= \mathbf{W}_k\mathbf{W}_k^H - \frac{(\mathbf{X}\mathbf{Y})\mathbf{W}_k^H\mathbf{W}_k(\mathbf{X}\mathbf{Y})^H}{\|(\mathbf{X}\mathbf{Y})^H\|_2^2} \\ &= \mathbf{W}_k\mathbf{W}_k^H - \frac{|\mathbf{X}\mathbf{Y}\mathbf{W}_k^H|^2}{(\mathbf{X}\mathbf{Y})(\mathbf{X}\mathbf{Y})^H} \\ &\geqslant 0 \end{aligned} \tag{3.23}$$

Thus, the results in (3.20) are obtained. ■

The upper bound on the spectral efficiency is achieved by (3.15) with the proposed hybrid precoding scheme. The upper bound separates the dependence on the array response vectors and the fast fading coefficients matrix and shows the optimality of the hybrid precoding.

3.5 Simulation Results

In this section, the performance of the proposed hybrid precoding scheme is evaluated. In our simulation, the relative element spacing of the ULA is $\frac{d}{\lambda} = 0.3$ and the angles of departure are uniformly distributed as $\theta_m = -\pi/2 + (m-1)\pi/M$, where $m = 1, 2, \cdots, M$.

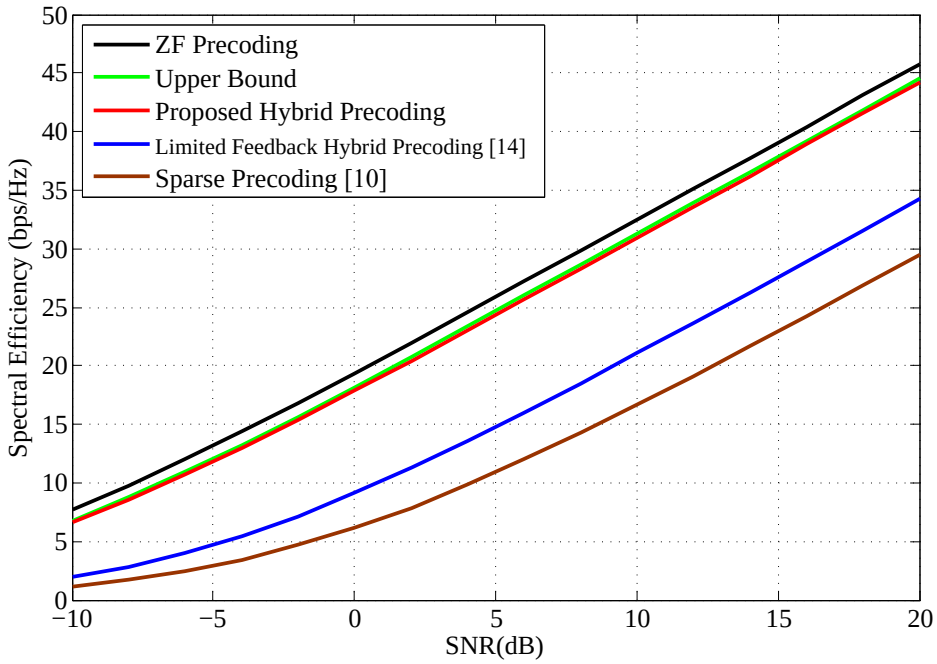


Figure 3.2: Spectral efficiency achieved by different precoding schemes with infinite resolution in downlink massive MU-MIMO systems where $N_t = 128, K = N_{RF} = 4$ and the finite dimension M is 64.

First, the spectral efficiency achieved by different precoding schemes with infinite resolution is compared, together with the tight upper bound in Fig. 3.2. The base station is assumed to employ the ULA with $N_t = 128$ antennas, serving $K = 4$ mobile stations each with a single antenna. There are $N_{RF} = 4$ RF chains and the finite dimension M is 64. Although ZF precoding is optimal in the simulation, it is infeasible because of the requirement of costly RF chains in the large antenna array. It is observed that the proposed hybrid precoding achieves spectral efficiency close to that achieved by the optimal ZF precoding with less than 1 dB loss. The tight upper bound derived in Section 3.4 is also plotted. Obviously, the proposed hybrid precoding method performs better than the limited feedback hybrid precoding and

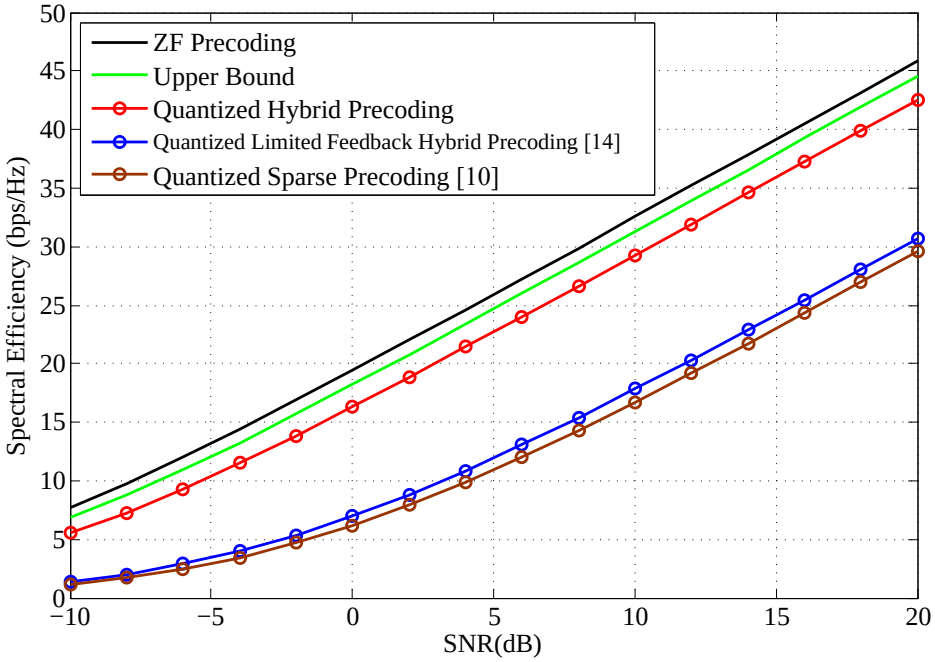


Figure 3.3: Spectral efficiency achieved by different quantized precoding schemes with 4 bits of precision in downlink massive MU-MIMO systems where $N_t = 128, K = N_{RF} = 4$ and the finite dimension M is 64.

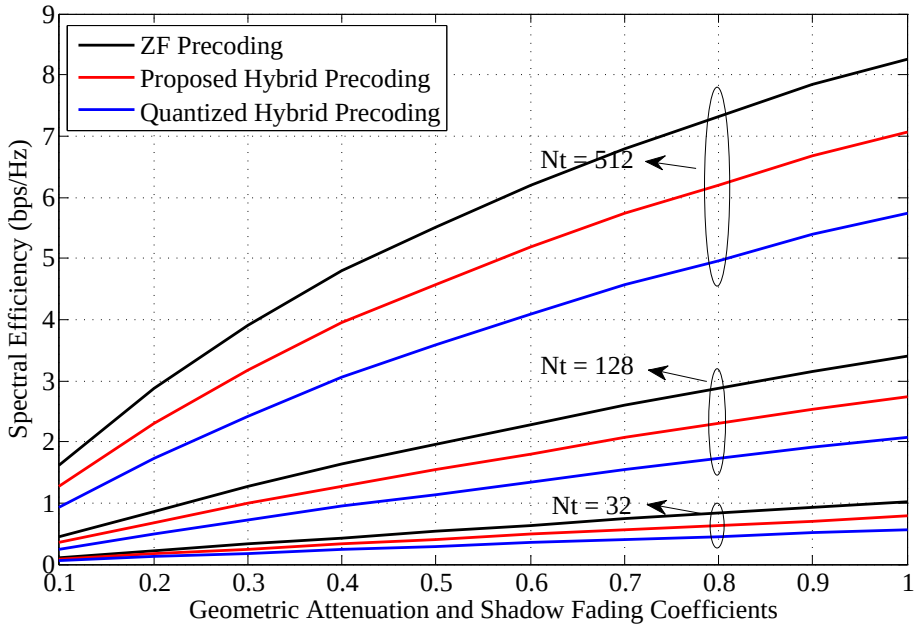


Figure 3.4: Spectral efficiency of precoding schemes versus the geometric attenuation and shadow fading coefficients.

the sparse precoding.

Then Fig. 3.3 illustrates the spectral efficiency achieved by different quantized precoding schemes, considering the same setup as Fig. 3.2. For the quantized phase, we simulate with $B = 4$ bits of precision and the phase control candidates are $(2\pi n)/(2^B)$ where $n \in \{0, \cdots, 2^B - 1\}$. It is observed that the quantized hybrid precoding schemes all suffer degradation, while the proposed hybrid precoding still has better performance than the limited feedback hybrid precoding and the sparse precoding.

In Fig. 3.4, the effort of the multi-cell transmission is considered at SNR $= -10$dB, but with different numbers of base station antennas, $N_t = 32$, 128 and 512. The spectral efficiency for different precoding schemes versus the geometric attenuation and shadow fading coefficients is simulated with $K = 4$ mobile stations. The number of RF chains is equal to the number of mobile stations and the finite dimension M is 16. It is assumed that β_{kk} are equal to 1 when the mobile stations are in the cell of the transmitting base station, but β_{kk} are equal to a when the mobile stations are in other cells, where $0 < a < 1$. When the value of the coefficients is low, the effect of the antenna number at the base station tends to be small, while when the coefficients becomes higher, the spectral efficiency gap between different numbers of base station antennas is increasingly obvious.

As shown in Figs. 3.5 - 3.7, the different beam patterns are generated by optimal ZF precoding, proposed precoding and quantized proposed precoding, respectively. It is assumed that the base station equipped with $N_t = 64$ antennas serves $K = 4$ mobile terminals at SNR = 5dB. The directions of 4 different mobile terminals are defined as $20°, 30°, 60°, 140°$. In Fig. 3.6, the beam patterns generated by the proposed hybrid precoding scheme closely resemble the patterns generated by optimal ZF precoding without obvious gain loss. However, due to the influence of the quantization, the beam patterns generated by the quantized proposed precoding suffer from degradation, which is shown in Fig. 3.7. Therefore, the proposed scheme is

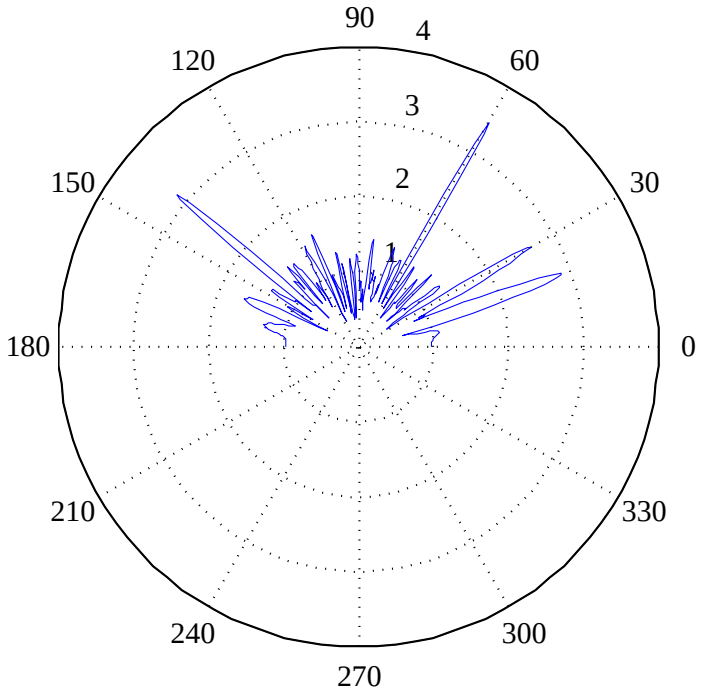


Figure 3.5: Beam pattern with optimal ZF precoding where $N_t = 64, K = N_{RF} = 4$ and SNR is 5dB.

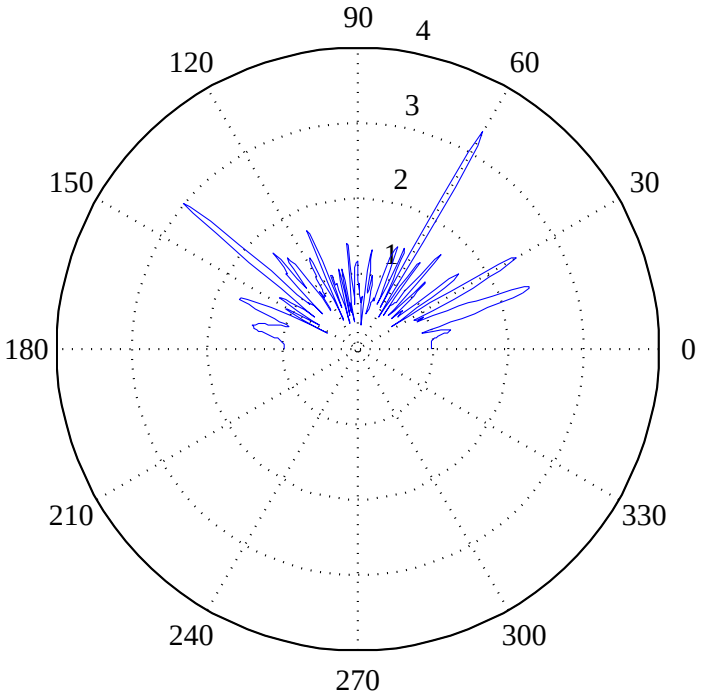


Figure 3.6: Beam pattern with proposed hybrid precoding where $N_t = 64, K = N_{RF} = 4$ and SNR is 5dB.

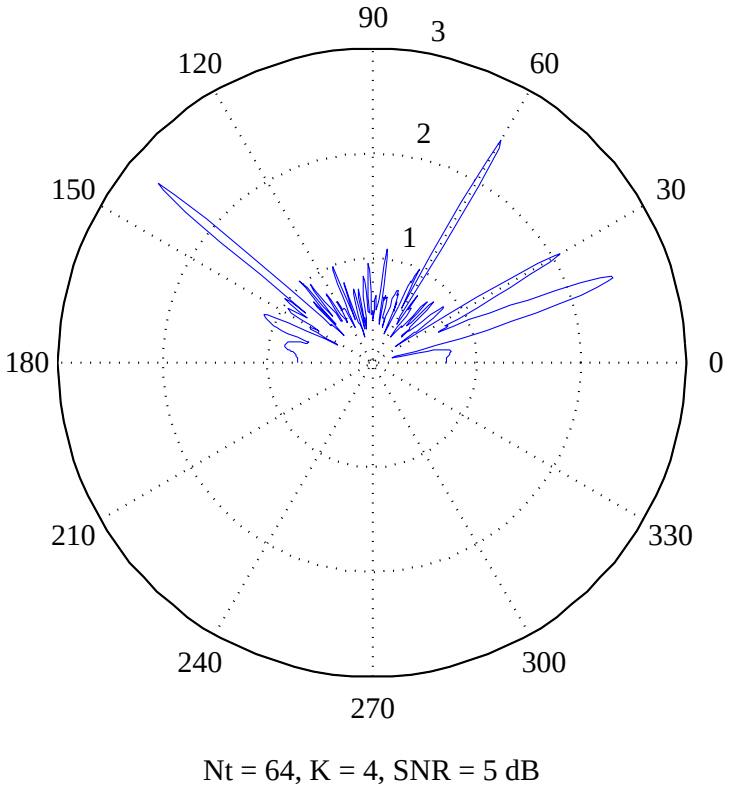


Figure 3.7: Beam pattern with quantized proposed hybrid precoding with 4 bits of precision where $N_t = 64, K = N_{RF} = 4$ and SNR is 5dB.

proved to perform good in beamforming.

In Fig. 3.8, the BER between different precoding schemes is compared, including optimal ZF precoding, the proposed hybrid precoding, spatially sparse precoding and limited feedback hybrid precoding. The massive MIMO system is considered with $N_t = 128$ base station antennas and $K = 4$ mobile terminals. Based on the simulation results, the gap between the performance of optimal ZF precoding and the performance of the proposed hybrid precoding is remarkably small, about 1dB for a target BER of 10^{-4}. Moreover, the BER of the proposed hybrid precoding outperforms the spatially sparse precoding and the BER for limited feedback hybrid precoding saturates at 10^{-2}. Thus, the proposed method can yield satisfactory BER performance for a range of SNR and offer a better trade-off between performance and complexity.

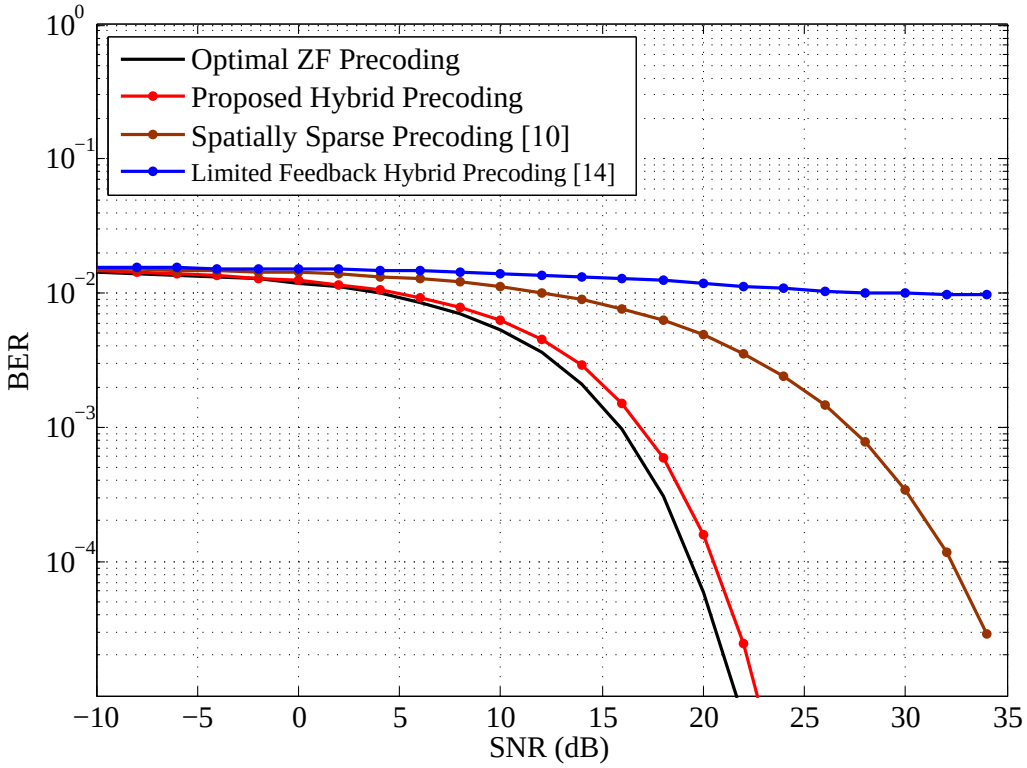


Figure 3.8: BER comparison of the proposed hybrid precoding with optimal ZF precoding, spatially sparse precoding and limited feedback hybrid precoding where $N_t = 128, K = N_{\mathrm{RF}} = 4$ and the finite dimension $M = 16$.

3.6 Summary

In this chapter, a low complexity hybrid precoding scheme is proposed for multiple mobile stations for the finite dimensional channel. The geometric channel model with finite dimension is applicable to obtain generally correlated or not asymptotically orthogonal channel. The proposed precoding with non-iterative design gives a significant reduction of RF chains, which reduces the hardware complexity in massive MIMO systems. In terms of the performance of spectral efficiency, the proposed method approaches the optimal ZF performance and performs better than other precoding schemes, which is showed in the simulation results. In addition, the numerical results also show the BER performance of the proposed precoding scheme is still better than other precoding schemes.

Chapter 4

Energy Efficient Iterative Hybrid Precoding Scheme with Sub-Connected Architecture for Massive MIMO Systems

4.1 Introduction

In massive MIMO systems, the very large arrays of transmit antennas lead to a prohibitive hardware complexity and the RF chains can be accountable for 50-80% of the total power consumption [69]. Therefore, a hybrid precoding scheme with fully-connected architecture is proposed to address the RF hardware constraints in Chapter 3. The variable phase shifters with high-dimensional phase-only RF processing are introduced, which are exploited to control the phases of the upconverted RF signal.

In this chapter, in order to further reduce the energy consumption and hardware complexity, the hybrid precoding design with sub-connected architecture is considered in mmWave massive MIMO systems. The total achievable rate optimization problem can be formulated to minimize the Frobenius norm of the matrix of differ-

ence between the unconstrained precoding matrix and products of the hybrid RF and baseband precoding matrices. Therefore, the closed-form expression of baseband precoder is obtained by mathematical analysis and the phase extraction is used to obtain the RF precoder. Based on successive refinement [70], the RF precoder becomes the new input for the next iteration to update the baseband precoder until the stopping criterion is triggered. Therefore, the proposed hybrid scheme can be close to the optimal precoder with less performance loss.

4.2 System Model

A communication scenario consisting of a single-user mmWave massive MIMO system is considered as shown in Fig. 4.1, where the numbers of the transmitter antennas and the receiver antennas are denoted as N_t and N_r respectively. It is assumed there are only N_{RF} RF chains at the transmitter and the number of transmitted data streams is equal to N_{RF} for simplicity. In the sub-connected architecture, each RF chain is only connected with N_a antennas where N_a is assumed to be integer and $N_a = N_t/N_{\mathrm{RF}}$, thus the total number of phase shifters is N_t. However, the fully-connected architecture requires $N_t N_{\mathrm{RF}}$ phase shifters, because each RF chain is connected with N_t antennas.

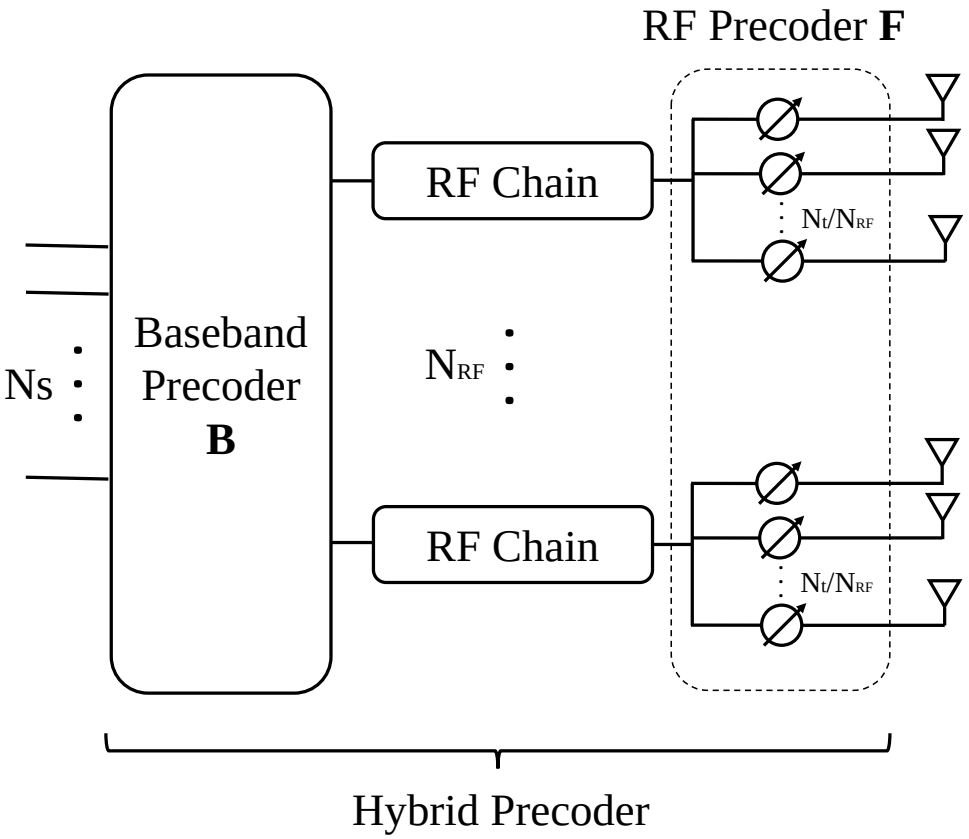


Figure 4.1: Hybrid precoding scheme with sub-connected architecture for mmWave massive MIMO systems.

A hybrid precoding scheme with a sub-connected architecture is proposed. In the digital domain, data streams are precoded by the baseband precoder **B** of dimension $N_{\mathrm{RF}} \times N_{\mathrm{RF}}$. Then in the analog domain, let **F** be the RF precoder of dimension

$N_t \times N_{\mathrm{RF}}$ comprising N_{RF} vectors $\mathbf{f}_i$ of dimension N_a as

$$\mathbf{F} = \begin{bmatrix} \mathbf{f}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2 & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{f}_{N_{\mathrm{RF}}} \end{bmatrix}. \tag{4.1}$$

The elements of $\mathbf{f}_i$ have constant amplitude and satisfy $|\mathbf{f}_i(j)|^2 = N_{\mathrm{RF}}/N_t$, where $\mathbf{f}_i(j)$ is the j-th element of $\mathbf{f}_i$. Therefore, the precoded signal can be calculated as

$$\mathbf{x} = \mathbf{FBs}, \tag{4.2}$$

where $\mathbf{s} = [s_1, s_2, \cdots, s_{N_{\mathrm{RF}}}]^T$ is the vector of the transmitted signal with $\mathbb{E}[\mathbf{ss}^H] = \frac{1}{N_{\mathrm{RF}}}\mathbf{I}_{N_{\mathrm{RF}}}$. Due to the total transmit power constraint, the entries of $\mathbf{B}$ are normalized to satisfy $\|\mathbf{FB}\|_F^2 = N_{\mathrm{RF}}$. The processed received signal can be written as

$$\mathbf{y} = \sqrt{\rho}\mathbf{WHFBs} + \mathbf{n}, \tag{4.3}$$

where $\mathbf{y} = [y_1, y_2, \cdots, y_{N_r}]^T$ is the received signal vector and ρ is the average received power of the transmission. The matrix $\mathbf{W}$ of dimension $N_r \times N_r$ is the combiner at the receiver. $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ denotes the channel matrix and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I}_{N_r})$ represents an additive white Gaussian noise vector with elements having zero mean and unit variance. The channel state information is known at both the transmitter and the receiver.

4.3 Channel Model

In this section, the narrowband mmWave channel matrix follows a geometric channel model [9] given by

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{N_p}} \sum_{p=1}^{N_p} \beta_p \mathbf{a}_r(\theta_p) \mathbf{a}_t(\phi_p)^H, \tag{4.4}$$

where N_p donates the number of propagation paths for mmWave communication systems. β_p is the gain of the p-th path. θ_p and ϕ_p are the angles of arrival and departure, respectively. $\mathbf{a}_r(\theta_p)$ and $\mathbf{a}_t(\phi_p)$ are the receive and transmit array response vectors determined by the array structures. For an U-element ULA, the array response vector can be given by

$$\mathbf{a}_{ULA}(\varphi) = \frac{1}{\sqrt{U}}[1, e^{j\frac{2\pi}{\lambda}dsin(\varphi)}, \cdots, e^{j(U-1)\frac{2\pi}{\lambda}dsin(\varphi)}]^T, \tag{4.5}$$

where λ is the wavelength of the signal and d denotes the distance between antenna elements.

4.4 Problem Formulation

Assuming perfect channel state information known at the receiver, the design of the hybrid precoders is explored to maximize the spectral efficiency of the system

$$R = \log_2 \det \left(\mathbf{I}_{N_r} + \frac{\rho}{N_{\mathrm{RF}}} \mathbf{R}_n^{-1} \mathbf{W}^H \mathbf{H} \mathbf{F} \mathbf{B} \mathbf{B}^H \mathbf{F}^H \mathbf{H}^H \mathbf{W} \right), \tag{4.6}$$

where $\mathbf{R}_n = \sigma_n^2 \mathbf{W}^H \mathbf{W}$ is the noise covariance matrix after combining. We consider the SVD of the channel matrix $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H$ to obtain the optimal precoder $\mathbf{P}_{\mathrm{opt}}$. Assuming that the right unitary matrix $\mathbf{U}$ of size $N_r \times N_r$ is used for the optimal combiner at the receiver, we only focus on the design of the precoding matrices at

the transmitter. Therefore, (4.6) can be rewritten as

$$R = \log_2 \det \left(\mathbf{I}_{N_r} + \frac{\rho}{N_{\mathrm{RF}} \sigma_n^2} \mathbf{H}\mathbf{F}\mathbf{B}\mathbf{B}^H \mathbf{F}^H \mathbf{H}^H \right). \tag{4.7}$$

Because $\mathbf{V}$ is an $N_t \times N_t$ unitary matrix and Σ is an $N_r \times N_t$ diagonal matrix of singular values in descendant order, the matrices are defined as

$$\Sigma = \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{bmatrix}, \mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix}, \tag{4.8}$$

where Σ_1 is the first partition of dimension $N_{\mathrm{RF}} \times N_{\mathrm{RF}}$ and $\mathbf{V}_1$ is the first N_{RF} columns of $\mathbf{V}$. The optimal precoder can be given by

$$\mathbf{P}_{\mathrm{opt}} = \mathbf{V}_1. \tag{4.9}$$

The optimal hybrid precoding scheme ensures the product of the RF precoder and the baseband precoder is close to the optimal precoder [10]. The hybrid precoding design problem can be written as

$$\begin{aligned} \underset{\mathbf{F},\mathbf{B}}{\text{minimize}} \quad & \|\mathbf{P}_{\mathrm{opt}} - \mathbf{F}\mathbf{B}\|_F \\ \textit{subject to} \quad & \|\mathbf{F}\mathbf{B}\|_F^2 = N_{\mathrm{RF}}, \\ & |\mathbf{f}_i(j)| = \sqrt{\frac{N_{\mathrm{RF}}}{N_t}}. \end{aligned} \tag{4.10}$$

To deal with the nonconvex constraint of $\mathbf{F}$, assuming $\mathbf{B}$ is a unitary matrix, the objective function can be reformulated. When $\mathbf{B}^H\mathbf{B} = \mathbf{I}$, $\|\mathbf{F}\mathbf{B}\|_F^2 = \mathrm{Tr}(\mathbf{B}^H\mathbf{F}^H\mathbf{F}\mathbf{B}) =$

$\mathrm{Tr}(\mathbf{I}_{N_{\mathrm{RF}}}) = N_{\mathrm{RF}}$. Then (4.10) has an upper bound expressed as

$$\begin{aligned}\|\mathbf{P}_{\mathrm{opt}} - \mathbf{FB}\|_F &= \|\mathbf{P}_{\mathrm{opt}}\mathbf{B}^H\mathbf{B} - \mathbf{FB}\|_F \\ &\leq \|\mathbf{P}_{\mathrm{opt}}\mathbf{B}^H - \mathbf{F}\|_F\|\mathbf{B}\|_F \\ &= \sqrt{N_{\mathrm{RF}}}\|\mathbf{P}_{\mathrm{opt}}\mathbf{B}^H - \mathbf{F}\|_F. \end{aligned} \tag{4.11}$$

By adopting (4.11) as the objective function, the hybrid precoder design problem is reformulated as

$$\begin{aligned} \underset{\mathbf{F},\mathbf{B}}{\text{minimize}} \quad & \|\mathbf{P}_{\mathrm{opt}}\mathbf{B}^H - \mathbf{F}\|_F^2 \\ \textit{subject to} \quad & \mathbf{B}^H\mathbf{B} = \mathbf{I}, \\ & |\mathbf{f}_i(j)| = \sqrt{\frac{N_{\mathrm{RF}}}{N_t}}. \end{aligned} \tag{4.12}$$

In the following section, a hybrid precoding algorithm is proposed with a convergent method to solve the problem in (4.12).

4.5 The design of the hybrid precoding with sub-connected architecture

In this section, an iterative design of hybrid precoding algorithm with sub-connected architecture is proposed in order to achieve the near-optimal performance and reduce the hardware complexity.

4.5.1 RF Precoder Design

By adopting (4.12) as the objective function, the design of $\mathbf{F}$ is considered assuming that the unitary matrix $\mathbf{B}$ is fixed. The RF precoder design problem is formulated

as

$$\begin{aligned} \underset{\mathbf{F}}{\text{minimize}} \quad & \|\mathbf{P}_{\text{opt}}\mathbf{B}^H - \mathbf{F}\|_F^2 \\ \textit{subject to} \quad & |\mathbf{f}_i(j)| = \sqrt{\frac{N_{\text{RF}}}{N_t}}. \end{aligned} \tag{4.13}$$

Because of the special characteristic of $\mathbf{F}$, (4.13) can be expressed as

$$\underset{\theta_m}{\text{minimize}} \quad \|(\mathbf{P}_{\text{opt}})_{m,:}(\mathbf{B})_{n,:}{}^H - \sqrt{\frac{N_{\text{RF}}}{N_t}} e^{j\theta_m}\|_2^2, \tag{4.14}$$

where $1 \leq m \leq N_t$ and $n = \lceil m\frac{N_{\text{RF}}}{N_t} \rceil$. Therefore, $\mathbf{F}$ can be presented by a closed-form expression based on phase rotation. Because there are zero elements in $\mathbf{F}$, the nonzero elements $\mathbf{F}_{m,n}$ in $\mathbf{F}$ can be written as

$$\mathbf{F}_{m,n} = \frac{\sqrt{N_{\text{RF}}}}{\sqrt{N_t}} e^{j\text{angle}\{(\mathbf{P}_{\text{opt}})_{m,:}(\mathbf{B})_{n,:}{}^H\}} \tag{4.15}$$

The constant amplitude of the elements in $\mathbf{F}$ is $\sqrt{N_{\text{RF}}}/\sqrt{N_t}$ so that the constraint in (4.12) is satisfied.

4.5.2 Baseband Precoder Design

In this section, a closed-form solution for $\mathbf{B}$ is proposed assuming that $\mathbf{F}$ is fixed. To deal with the constrained problem

$$\begin{aligned} \underset{\mathbf{B}}{\text{minimize}} \quad & \|\mathbf{P}_{\text{opt}}\mathbf{B}^H - \mathbf{F}\|_F^2 \\ \textit{subject to} \quad & \mathbf{B}^H\mathbf{B} = \mathbf{I}, \end{aligned} \tag{4.16}$$

(4.16) can be further expressed as

$$\begin{aligned} & \|\mathbf{P}_{\text{opt}}\mathbf{B}^H - \mathbf{F}\|_F^2 \\ =& \text{Tr}(\mathbf{P}_{\text{opt}}^H\mathbf{P}_{\text{opt}}) - 2\Re\{\text{Tr}(\mathbf{B}\mathbf{P}_{\text{opt}}^H\mathbf{F})\} + \text{Tr}(\mathbf{F}^H\mathbf{F}). \end{aligned} \tag{4.17}$$

Therefore, in terms of optimising the variable matrix $\mathbf{B}$, (4.16) can be reformulated as an optimization problem as

$$\begin{aligned} \underset{\mathbf{B}}{\text{maximize}} \quad & \Re\{\text{Tr}(\mathbf{B}\mathbf{P}_{\text{opt}}^{H}\mathbf{F})\} \\ \textit{subject to} \quad & \mathbf{B}^{H}\mathbf{B} = \mathbf{I}. \end{aligned} \tag{4.18}$$

*Theorem*1 : Denote the hybrid precoders at the t-th iteration by $(\mathbf{F}^{(t)}\mathbf{B}^{(t)})$ and assume $\mathbf{F}^{(t)}$ is fixed. An optimal closed-form solution for $\mathbf{B}$ is obtained to solve the problem (4.18) by

$$\mathbf{B}^{(t)} = (\hat{\mathbf{U}}^{(t)}\hat{\mathbf{V}}^{(t)^{H}})^{-1}, \tag{4.19}$$

where $\mathbf{P}_{\text{opt}}^{H}\mathbf{F}^{(t)} = \hat{\mathbf{U}}^{(t)}\hat{\Sigma}^{(t)}\hat{\mathbf{V}}^{(t)^{H}}$.

Proof : For notational convenience, the iteration number is omitted in the proof. The upper bound of the problem $\underset{\mathbf{B}}{\text{minimize}}\, \|\mathbf{P}_{\text{opt}} - \mathbf{F}\mathbf{B}\|_F^2$ is derived as

$$\begin{aligned} \Re\{\text{Tr}(\mathbf{B}\mathbf{P}_{\text{opt}}^{H}\mathbf{F})\} &= \Re\{\text{Tr}(\mathbf{B}\hat{\mathbf{U}}\hat{\Sigma}\hat{\mathbf{V}}^{H})\} \\ &= \Re\{\text{Tr}(\hat{\mathbf{V}}^{H}\mathbf{B}\hat{\mathbf{U}}\hat{\Sigma})\} \\ &= \Re\{\text{Tr}(\mathbf{T}\hat{\Sigma})\} \\ &\overset{(a)}{\leq} \sum_{i=1}^{N_{\text{RF}}} \delta_i\sigma_i, \end{aligned} \tag{4.20}$$

where $\mathbf{T}$ is defined as $\mathbf{T} = \hat{\mathbf{V}}^{H}\mathbf{B}\hat{\mathbf{U}}$. The parameters $\delta_1, ..., \delta_{N_{\text{RF}}}$ and $\sigma_1, ..., \sigma_{N_{\text{RF}}}$ are the non-zero singular values of $\mathbf{T}$ and $\hat{\Sigma}$, respectively. $\Re\{\text{Tr}(\mathbf{B}\mathbf{P}_{\text{opt}}^{H}\mathbf{F})\}$ reaches the maximum, when the equation in (a) is satisfied. Assume $rank(\mathbf{T}) = N_{\text{RF}}$ and

$\delta_1 \geq \cdots \geq \delta_{N_{\mathrm{RF}}} > 0$. Define $\tilde{\mathbf{T}} = \begin{pmatrix} \mathbf{0} & \mathbf{T} \\ \mathbf{T}^H & \mathbf{0} \end{pmatrix}$, then

$$\begin{aligned} \det(\lambda \mathbf{I} - \tilde{\mathbf{T}}) &= \det \begin{pmatrix} \lambda \mathbf{I}_{N_{\mathrm{RF}}} & -\mathbf{T} \\ -\mathbf{T}^H & \lambda \mathbf{I}_{N_{\mathrm{RF}}} \end{pmatrix} \\ &= \det(\lambda \mathbf{I}_{N_{\mathrm{RF}}}) \det(\lambda \mathbf{I}_{N_{\mathrm{RF}}} - \lambda^{-1} \mathbf{T}^H \mathbf{T}) \\ &= \det(\lambda^2 \mathbf{I} - \mathbf{T}^H \mathbf{T}) = 0, \end{aligned} \tag{4.21}$$

where λ denotes the eigenvalues of $\tilde{\mathbf{T}}$ given by $\delta_1, \cdots, \delta_{N_{\mathrm{RF}}}, -\delta_{N_{\mathrm{RF}}}, \cdots, -\delta_1$. Similarly, defining $\tilde{\Sigma} = \begin{pmatrix} \mathbf{0} & \hat{\Sigma} \\ \hat{\Sigma}^H & \mathbf{0} \end{pmatrix}$, the eigenvalues of $\tilde{\Sigma}$ can be presented as $\sigma_1, \cdots, \sigma_{N_{\mathrm{RF}}}, -\sigma_{N_{\mathrm{RF}}}, \cdots, -\sigma_1$, so

$$\begin{aligned} \mathrm{Tr}(\tilde{\mathbf{T}}\tilde{\Sigma}) &= \mathrm{Tr}(\mathbf{T}\hat{\Sigma}^H) + \mathrm{Tr}(\mathbf{T}^H \hat{\Sigma}) \\ &= 2\Re\{\mathrm{Tr}(\mathbf{T}\hat{\Sigma}^H)\} \\ &\overset{(b)}{\leq} \sum_{i=1}^{N_{\mathrm{RF}}} \delta_i \sigma_i + \sum_{i=1}^{N_{\mathrm{RF}}} (-\delta_i)(-\sigma_i) \\ &= 2 \sum_{i=1}^{N_{\mathrm{RF}}} \delta_i \sigma_i. \end{aligned} \tag{4.22}$$

Therefore, (4.20) is proved. When the equality in (b) is achieved, then $\tilde{\mathbf{T}} = \sum_{i=1}^{2N_{\mathrm{RF}}} \lambda_i(\tilde{\mathbf{T}}) \mathbf{e}_i \mathbf{e}_i^H$ using Von Neumann's trace inequality [71], where $\mathbf{e}_i$ denote the eigenvector of $\tilde{\Sigma}$. Because $\hat{\Sigma}$ is a diagonal matrix, the eigenvectors of $\tilde{\Sigma}$ are obtained though simple computations as

$$\begin{aligned} &\mathbf{e}_1 = \begin{pmatrix} \tilde{\mathbf{e}}_1 \\ \tilde{\mathbf{e}}_1 \end{pmatrix}, \cdots, \mathbf{e}_{N_{\mathrm{RF}}} = \begin{pmatrix} \tilde{\mathbf{e}}_{N_{\mathrm{RF}}} \\ \tilde{\mathbf{e}}_{N_{\mathrm{RF}}} \end{pmatrix}, \mathbf{e}_{N_{\mathrm{RF}}+1} = \begin{pmatrix} -\tilde{\mathbf{e}}_{N_{\mathrm{RF}}} \\ \tilde{\mathbf{e}}_{N_{\mathrm{RF}}} \end{pmatrix}, \\ &\cdots, \mathbf{e}_{2N_{\mathrm{RF}}} = \begin{pmatrix} -\tilde{\mathbf{e}}_1 \\ \tilde{\mathbf{e}}_1 \end{pmatrix}, \end{aligned} \tag{4.23}$$

where $\tilde{\mathbf{e}}_i \in \mathbb{R}^{N_{\mathrm{RF}} \times 1}$ has only one non-zero element. The i-th non-zero element in $\tilde{\mathbf{e}}_i$ is defined as w_i where $|w_i| = \frac{1}{\sqrt{2}}$. Therefore,

$$\begin{aligned}\tilde{\mathbf{T}} &= \sum_{i=1}^{2N_{\mathrm{RF}}} \lambda_i(\tilde{\mathbf{T}}) \mathbf{e}_i \mathbf{e}_i^H \\ &= \sum_{i=1}^{N_{\mathrm{RF}}} \delta_i \begin{pmatrix} \tilde{\mathbf{e}}_i \\ \tilde{\mathbf{e}}_i \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{e}}_i^H & \tilde{\mathbf{e}}_i^H \end{pmatrix} + \sum_{i=1}^{N_{\mathrm{RF}}} -\delta_i \begin{pmatrix} -\tilde{\mathbf{e}}_i \\ \tilde{\mathbf{e}}_i \end{pmatrix} \begin{pmatrix} -\tilde{\mathbf{e}}_i^H & \tilde{\mathbf{e}}_i^H \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0} & 2\sum_{i=1}^{N_{\mathrm{RF}}} \delta_i \tilde{\mathbf{e}}_i \tilde{\mathbf{e}}_i^H \\ 2\sum_{i=1}^{N_{\mathrm{RF}}} \delta_i \tilde{\mathbf{e}}_i \tilde{\mathbf{e}}_i^H & \mathbf{0} \end{pmatrix}. \end{aligned} \tag{4.24}$$

Then the diagonal matrix

$$\mathbf{T} = 2\sum_{i=1}^{N_{\mathrm{RF}}} \delta_i \tilde{\mathbf{e}}_i \tilde{\mathbf{e}}_i^H = \begin{pmatrix} \delta_1 & & & \\ & \delta_2 & & \\ & & \ddots & \\ & & & \delta_{N_{\mathrm{RF}}} \end{pmatrix}. \tag{4.25}$$

Moreover, because $\hat{\mathbf{V}}$, $\mathbf{B}$ and $\hat{\mathbf{U}}$ are unitary matrices, hence

$$\mathbf{T}\mathbf{T}^H = \mathbf{I}. \tag{4.26}$$

Based on (4.25) and (4.26), the matrix $\mathbf{T}$ can be written by

$$\mathbf{T} = \hat{\mathbf{V}}^H \mathbf{B} \hat{\mathbf{U}} = \mathbf{I}, \tag{4.27}$$

therefore the closed-form expression of $\mathbf{B}$ can be given as

$$\mathbf{B} = (\hat{\mathbf{U}}\hat{\mathbf{V}}^H)^{-1}. \tag{4.28}$$

Due to the property of the unitary matrix, $\mathbf{B}$ satisfies the constraint $\mathbf{B}^H\mathbf{B} = \mathbf{I}$.

■

4.5.3 Successive Refinement

The hybrid precoders are designed by minimizing the objective function in (4.12) instead of the original objective function, which will result in some performance loss. In order to solve this problem, the process of successive refinement is used to update the baseband precoder and the RF precoder until a stopping criterion is triggered. The stopping criterion is determined by the difference between the chordal distances of two successive iterations. The chordal distance is defined as the difference between the optimal precoder and the product of the baseband and RF precoders, which can be expressed as

$$\eta^{(t)} = \frac{1}{\sqrt{2}} \|\mathbf{P}_{\text{opt}}\mathbf{P}_{\text{opt}}^{H} - (\mathbf{F}^{(t)}\mathbf{B}^{(t)})(\mathbf{F}^{(t)}\mathbf{B}^{(t)})^{H}\|_{F}. \tag{4.29}$$

The difference between the chordal distances of two consecutive successive iterations is denoted as $\varepsilon^{(t)} = \eta^{(t-1)} - \eta^{(t)}$. The iteration continues until $\varepsilon^{(t)}$ falls below a given convergence threshold ε_0, where ε_0 is sufficiently small. $\mathbf{B}^{(t-1)}$ is used as a new input for the next iteration to design $\mathbf{F}^{(t)}$ and $\mathbf{B}^{(t)}$. A step-by-step summary of the proposed hybrid precoder design is given in Algorithm 1.

Algorithm 1 Hybrid precoding design with sub-connected architecture.

Input: The optimal precoder, $\mathbf{P}_{\text{opt}}$; The number of RF chains, N_{RF};
Output: The RF precoder, $\mathbf{F}$; The baseband precoder, $\mathbf{B}$;

1: Initialize $t = 0$ and $\mathbf{F}^{(0)}$ with random phases;
2: **repeat**
3: Compute the SVD of $\mathbf{P}_{\text{opt}}^{H}\mathbf{F}^{(t)}$, $\mathbf{P}_{\text{opt}}^{H}\mathbf{F}^{(t)} = \hat{\mathbf{U}}^{(t)}\hat{\Sigma}^{(t)}\hat{\mathbf{V}}^{(t)^{H}}$;
4: Obtain $\mathbf{B}^{(t)} = (\hat{\mathbf{U}}^{(t)}\hat{\mathbf{V}}^{(t)^{H}})^{-1}$;
5: $\eta^{(t)} = \frac{1}{\sqrt{2}}\|\mathbf{P}_{\text{opt}}\mathbf{P}_{\text{opt}}^{H} - (\mathbf{F}^{(t)}\mathbf{B}^{(t)})(\mathbf{F}^{(t)}\mathbf{B}^{(t)})^{H}\|_{F}$;
6: **if** $(\eta^{(t-1)} - \eta^{(t)}) > \varepsilon_0$ **then**
7: $t \leftarrow t + 1$;
8: Obtain $\mathbf{F}^{(t+1)}$ by (4.15);
9: **end if**
10: **until** $(\eta^{(t-1)} - \eta^{(t)}) \leq \varepsilon_0$;
11: **return** $\mathbf{F} = \mathbf{F}^{(t)}$, $\mathbf{B} = \mathbf{B}^{(t)}$;

4.6 Analysis of Spectral Efficiency

In this section, with perfect channel state information, an upper bound on the spectral efficiency achieved by the proposed hybrid precoding is derived. The upper bound can be expressed as

$$R \leq \sum_{i=1}^{N_{\mathrm{RF}}} \log_2 \left(1 + \frac{\rho}{N_{\mathrm{RF}}\sigma_n^2}\lambda_1(\mathbf{H}_i^H\mathbf{H}_i)\right), \tag{4.30}$$

where $\lambda_1(\mathbf{H}_i^H\mathbf{H}_i)$ is the largest eigenvalue of $\mathbf{H}_i^H\mathbf{H}_i$.

Proof : The spectral efficiency of the system is given by

$$R = \log_2 \det \left(\mathbf{I}_{N_r} + \frac{\rho}{N_{\mathrm{RF}}}\mathbf{R}_n^{-1}\mathbf{W}^H\mathbf{H}\mathbf{F}\mathbf{B}\mathbf{B}^H\mathbf{F}^H\mathbf{H}^H\mathbf{W}\right), \tag{4.31}$$

where $\mathbf{R}_n = \sigma_n^2\mathbf{W}^H\mathbf{W}$. Because the optimal combiner $\mathbf{W}$ and the baseband precoder $\mathbf{B}$ are both unitary matrices, (4.31) can be rewritten as

$$R = \log_2 \det \left(\mathbf{I}_{N_r} + \gamma\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H\right), \tag{4.32}$$

where $\gamma = \frac{\rho}{N_{\mathrm{RF}}\sigma_n^2}$. The RF precoder is set as $\mathbf{F} = [\widetilde{\mathbf{f}}_1, \widetilde{\mathbf{f}}_2, \cdots, \widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}]$ and it is defined as

$$\mathbf{F} \triangleq \left(\mathbf{F}_{N_{\mathrm{RF}}-1} \quad \widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}\right), \tag{4.33}$$

where $\mathbf{F}_{N_{\mathrm{RF}}-1}$ denotes the sub-matrix containing the first $(N_{\mathrm{RF}}-1)$ columns of $\mathbf{F}$ and $\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}$ is the N_{RF}-th column of $\mathbf{F}$. Thus, from (4.32),

$$\begin{aligned} R &= \log_2 \det \left(\mathbf{I}_{N_r} + \gamma\mathbf{H}\left(\mathbf{F}_{N_{\mathrm{RF}}-1} \quad \widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}\right)\left(\mathbf{F}_{N_{\mathrm{RF}}-1} \quad \widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}\right)^H\mathbf{H}^H\right) \\ &= \log_2 \det \left(\mathbf{I}_{N_r} + \gamma\mathbf{H}\mathbf{F}_{N_{\mathrm{RF}}-1}\mathbf{F}_{N_{\mathrm{RF}}-1}^H\mathbf{H}^H + \gamma\mathbf{H}\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}^H\mathbf{H}^H\right). \end{aligned} \tag{4.34}$$

Defining

$$\mathbf{Q}_{N_{\mathrm{RF}}-1} \triangleq \mathbf{I}_{N_r} + \gamma \mathbf{H}\mathbf{F}_{N_{\mathrm{RF}}-1}\mathbf{F}_{N_{\mathrm{RF}}-1}^{H}\mathbf{H}^{H}, \tag{4.35}$$

(4.34) can be rewritten as

$$\begin{aligned} R &= \log_2 \det\left(\mathbf{Q}_{N_{\mathrm{RF}}-1}\right) \\ &\quad + \log_2 \det\left(\mathbf{I}_{N_r} + \gamma \mathbf{Q}_{N_{\mathrm{RF}}-1}^{-1}\mathbf{H}\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}^{H}\mathbf{H}^{H}\right) \\ &\overset{(c)}{=} \log_2 \det\left(\mathbf{Q}_{N_{\mathrm{RF}}-1}\right) \\ &\quad + \log_2\left(1 + \gamma\, \widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}^{H}\mathbf{H}^{H}\mathbf{Q}_{N_{\mathrm{RF}}-1}^{-1}\mathbf{H}\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}\right), \end{aligned} \tag{4.36}$$

where (c) is valid by applying $\det(\mathbf{I}+\mathbf{X}\mathbf{Y}) = \det(\mathbf{I}+\mathbf{Y}\mathbf{X})$ by defining $\mathbf{Q}_{N_{\mathrm{RF}}-1}^{-1}\mathbf{H}\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}$ and $\widetilde{\mathbf{f}}_{N_{\mathrm{RF}}}^{H}\mathbf{H}^{H}$. It is easy to know that (4.32) and (4.35) have similar expressions, thus we can obtain the expression for $\mathbf{Q}_{N_{\mathrm{RF}}-1}$ with $\mathbf{Q}_{N_{\mathrm{RF}}-2}$ using the same method. After N_{RF} such transformations, (4.36) can be further stated as

$$R = \sum_{i=1}^{N_{\mathrm{RF}}} \log_2\left(1 + \gamma\, \widetilde{\mathbf{f}}_i^{H}\mathbf{H}^{H}\mathbf{Q}_{i-1}^{-1}\mathbf{H}\widetilde{\mathbf{f}}_i\right), \tag{4.37}$$

where $\mathbf{Q}_i$ can be presented as

$$\mathbf{Q}_i = \begin{cases} \mathbf{I}_{N_r}, & i = 0, \\ \mathbf{I}_{N_r} + \gamma \mathbf{H}\mathbf{F}_{i-1}\mathbf{F}_{i-1}^{H}\mathbf{H}^{H}, & i \geq 1. \end{cases} \tag{4.38}$$

From (4.37), the total spectral efficiency is the sum of N_{RF} sub-rates. Next, focusing on $\widetilde{\mathbf{f}}_i^{H}\mathbf{H}^{H}\mathbf{Q}_{i-1}^{-1}\mathbf{H}\widetilde{\mathbf{f}}_i$, the upper bound can be derived. Due to the sub-connected architecture, we define $\widetilde{\mathbf{f}}_i \triangleq [\mathbf{0}_{1\times N_a(i-1)}, \mathbf{f}_i, \mathbf{0}_{1\times N_a(N_{\mathrm{RF}}-i)}]^T$, where $\mathbf{f}_i$ is the $N_a \times 1$ non-zero vector of the i-th column $\widetilde{\mathbf{f}}_i$ of $\mathbf{F}$. The channel matrix is divided into N_{RF}

sub-matrices $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_{N_{\mathrm{RF}}}]$, where $\mathbf{H}_i \in \mathbb{C}^{N_r \times N_a}$. Then

$$\begin{aligned}
\widetilde{\mathbf{f}}_i^H \mathbf{H}^H \mathbf{Q}_{i-1}^{-1} \mathbf{H} \widetilde{\mathbf{f}}_i &= \mathbf{f}_i^H \mathbf{H}_i^H \mathbf{Q}_{i-1}^{-1} \mathbf{H}_i \mathbf{f}_i \\
&= \mathrm{Tr}(\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{Q}_{i-1}^{-1} \mathbf{H}_i \mathbf{f}_i) \\
&= \mathrm{Tr}(\mathbf{H}_i \mathbf{f}_i \mathbf{f}_i^H \mathbf{H}_i^H \mathbf{Q}_{i-1}^{-1}) \\
&\overset{(d)}{\leq} \mathrm{Tr}(\mathbf{H}_i \mathbf{f}_i \mathbf{f}_i^H \mathbf{H}_i^H) \lambda_1(\mathbf{Q}_{i-1}^{-1}) \\
&\overset{(e)}{\leq} \mathrm{Tr}(\mathbf{H}_i \mathbf{f}_i \mathbf{f}_i^H \mathbf{H}_i^H) \\
&= \mathrm{Tr}(\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_i) \\
&= \lambda_1(\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_i) \\
&\overset{(f)}{\leq} \lambda_1(\mathbf{H}_i^H \mathbf{H}_i). \qquad (4.39)
\end{aligned}$$

Defining $\mathbf{X} \triangleq \mathbf{H}_i \mathbf{f}_i \mathbf{f}_i^H \mathbf{H}_i^H$ and $\mathbf{Y} \triangleq \mathbf{Q}_{i-1}^{-1}$, since $\mathbf{X}$ and $\mathbf{Y}$ are both positive semi-definite Hermitian matrices, step (d) is obtained due to the fact that $\mathrm{Tr}(\mathbf{XY}) \leq \mathrm{Tr}(\mathbf{X})\lambda_1(\mathbf{Y})$. As can be observed from (4.38), each eigenvalue of $\mathbf{Q}_{i-1}$ is equal to or larger than 1 because $\mathbf{Q}_{i-1}$ is the sum of a unit matrix and a positive semi-definite Hermitian matrix when $i \geq 2$. Therefore, step (e) is valid since the largest eigenvalue of $\mathbf{Q}_{i-1}$ is smaller than 1, i.e. $\lambda_1(\mathbf{Q}_{i-1}^{-1}) < 1$. Step (f) follows from adopting Poincaré separation theorem, where $\mathbf{H}_i^H \mathbf{H}_i$ is a Hermitian matrix and $\mathbf{f}_i^H \mathbf{f}_i = 1$.

■

Proof of (f) For the term $\lambda_1(\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_i)$ in (4.39), due to $\mathbf{f}_i^H \mathbf{f}_i = 1$, we first extend the vector $\mathbf{f}_i$ to an unitary matrix $\mathbf{W}$, which can be expressed as

$$\mathbf{W} = [\mathbf{f}_i \mid \mathbf{V}]. \qquad (4.40)$$

Then

$$\begin{aligned}\mathbf{A} &= \mathbf{W}^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{W} \\ &= \begin{bmatrix} \mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_i & \mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{V} \\ \mathbf{V}^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_i & \mathbf{V}^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{V} \end{bmatrix},\end{aligned} \tag{4.41}$$

and

$$\lambda_1(\mathbf{A}) = \lambda_1(\mathbf{H}_i^H \mathbf{H}_i). \tag{4.42}$$

Based on Sturm's theorem, since $\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_i$ is the principal sub-matrix of $\mathbf{A}$, we get

$$\begin{aligned}\lambda_1(\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_i) &\leq \lambda_1(\mathbf{A}) \\ &= \lambda_1(\mathbf{H}_i^H \mathbf{H}_i).\end{aligned} \tag{4.43}$$

■

Then the upper bound on the spectral efficiency in (4.30) is obtained using the proposed hybrid precoding scheme with sub-connected architecture, which is only based on the channel state information. The derived upper bound can be shown to be tight and verified in the simulation results in Fig. 4.2. Thus, this bound shows the asymptotic optimality of the spectral efficiency achieved by the proposed precoding.

4.7 Energy Efficiency

The energy efficiency is defined as the ratio of capacity versus transmitted power consumption, presenting the number of bits transmitted per Joule. The hybrid precoding with the sub-connected architecture is proposed in order to reduce the hardware complexity and power consumption. Therefore, the energy efficiency is a main factor to evaluate the performance of the hybrid precoding schemes.

Compared with the fully-connected architecture, the hybrid precoding with the

sub-connected architecture uses fewer phase shifters. Because of more design degrees of freedom in the RF domain, the spectral efficiency of the fully-connected architecture performs better than that of the sub-connected architecture. However, the energy efficiency of the sub-connected architecture outperforms the fully-connected architecture. Since a large number of antennas are involved, the power consumption is additionally yielded by RF chains, phase shifters and power amplifiers. The energy efficiency is defined as follow,

$$\xi \triangleq \frac{R}{P_a}, \tag{4.44}$$

where R is the spectral efficiency and P_a is the total power consumption. P_a can be calculated by $P_a = P_{\mathrm{S}} + N_{\mathrm{RF}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + N_{\mathrm{LNA}}P_{\mathrm{LNA}}$. P_{S} is the idle power consumption. P_{RF} is the RF chain power consumption. N_{PS} is the number of phase shifters and P_{PS} is the phase shifter power consumption, which depends on the type and the resolution of the quantized phases. N_{LNA} is the number of the low noise amplifiers and P_{LNA} is the low noise amplifier power consumption. The number of phase shifters is equal to $N_{\mathrm{RF}}N_t$ and the number of low noise amplifiers is equal to $N_t(N_{\mathrm{RF}}+1)$ in the fully-connected architecture, while it is reduced to N_t phase shifters and N_t low noise amplifiers in the sub-connected architecture, respectively. The numerical comparison will be presented in the simulation results in Section 4.9.

4.8 Analysis of Complexity

In this section, the complexity of the proposed hybrid precoding described in Algorithm 1 is analyzed. The complexity of the successive refinement is dominated by the calculation of the baseband and RF precoders.

In each iteration, the complexity of the baseband precoder comes from three parts. The first part comes from computing the product of the matrix $\mathbf{P}_{\mathrm{opt}}^{H}$ of size $N_{\mathrm{RF}} \times N_t$ with the matrix $\mathbf{F}^{(t)}$ of size $N_t \times N_{\mathrm{RF}}$, which is $\mathcal{O}(N_t N_{\mathrm{RF}}^2)$. The second part

originates from the computation of the SVD of $\mathbf{P}_{\text{opt}}^{H}\mathbf{F}^{(t)}$. Note that $\mathbf{P}_{\text{opt}}^{H}\mathbf{F}^{(t)}$ has size $N_{\text{RF}} \times N_{\text{RF}}$. Therefore, the order of complexity of this part is $\mathcal{O}(N_{\text{RF}}^3)$. The last part is from calculating $\mathbf{B}^{(t)}$ in (4.19), which can be expressed as $\mathbf{B}^{(t)} = (\hat{\mathbf{U}}^{(t)}\hat{\mathbf{V}}^{(t)^H})^{-1}$. Thus the complexity of the last part is equal to the complexity of computing the product of the matrix $\hat{\mathbf{U}}^{(t)}$ of size $N_{\text{RF}} \times N_{\text{RF}}$ with the matrix $\hat{\mathbf{V}}^{(t)^H}$ of size $N_{\text{RF}} \times N_{\text{RF}}$, which is $\mathcal{O}(N_{\text{RF}}^3)$, plus that of computing the inverse of a matrix of size $N_{\text{RF}} \times N_{\text{RF}}$, which is also $\mathcal{O}(N_{\text{RF}}^3)$.

In terms of the complexity of updating the RF precoder to $\mathbf{F}^{(t+1)}$, the RF precoder is computed using (4.15), which is an element-wise normalization stage that enforces the unit magnitude constraint. Therefore, the order of the complexity of the RF precoder is $\mathcal{O}(N_t N_{\text{RF}})$.

To sum up, the computational complexity of a single iteration in Algorithm 1 is $\mathcal{O}(N_t N_{\text{RF}}^2)$ and assuming the algorithm will run for K iterations, the complexity of Algorithm 1 is in the order of $\mathcal{O}(KN_t N_{\text{RF}}^2)$. Compared with the method in [20], the computational complexity of SIC-based precoding scheme is $\mathcal{O}(SN_t N_a)$, where S is the number of iterations. Since the numbers of iterations K and S have the same order of magnitude, the computational complexity of the proposed scheme is comparable with that of [20]. In contrast, the complexity of the spatially sparse precoding [10] is in the order of $\mathcal{O}(N_t^2 N_{\text{RF}}^2)$. Due to the large number of the transmitter antennas, the proposed precoding scheme has much lower computational complexity. In next section, the performance of the proposed precoding scheme will be numerically evaluated.

4.9 Results and Discussion

In this section, the simulation results are presented to demonstrate the performance of the proposed hybrid precoding algorithm with sub-connected architecture using the mmWave MIMO channel model in (4.4). In the simulation, the relative element

spacing of the ULA is $\frac{d}{\lambda} = 0.5$ and the number of propagation paths is set as 8. Using the directional antennas, the angles of departure are assumed to follow the uniform distribution within $[-\frac{\pi}{4}, \frac{\pi}{4}]$. The angles of arrival follow the uniform distribution within $[-\pi, \pi]$, because the mobile stations will locate random position. The convergence threshold ε_0 is defined as 1×10^{-6}. We evaluate the spectral effi-

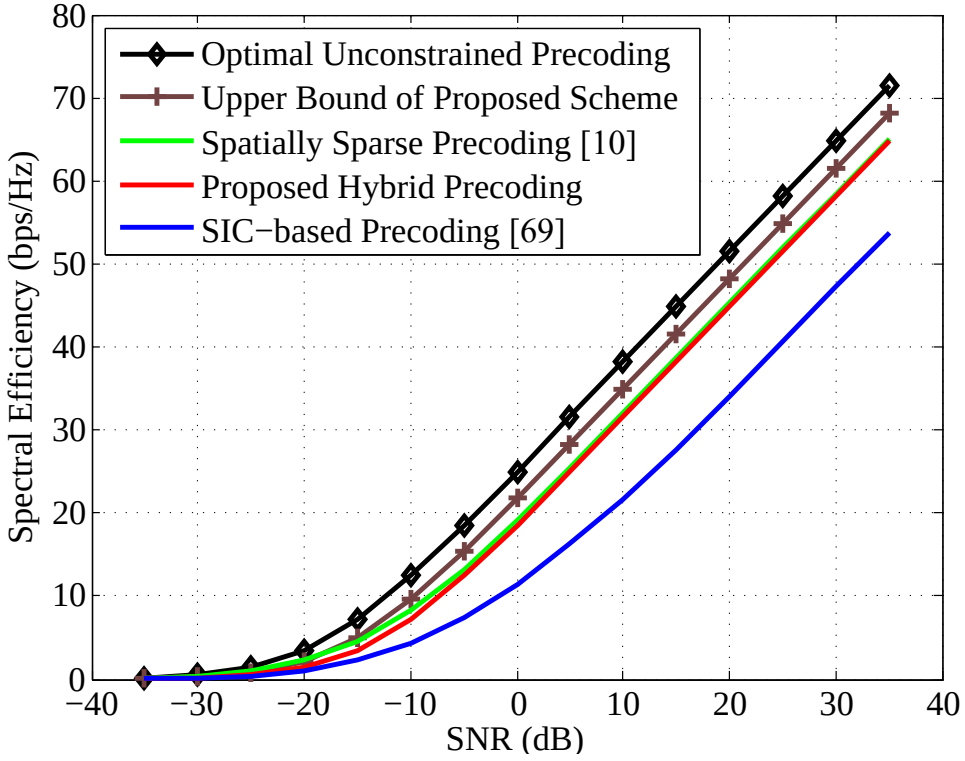


Figure 4.2: Spectral efficiency achieved by different precoding schemes with infinite resolution in mmWave massive MIMO systems where $N_t = 128, N_r = 8, N_{\mathrm{RF}} = 4$.

ciency achieved by different precoding schemes in mmWave MIMO systems, where the transmitter and the receiver are equipped with $N_t = 128$ and $N_r = 8$ antennas, respectively. It is assumed that the number of RF chains is $N_{\mathrm{RF}} = 4$. In Fig. 4.2, the proposed hybrid precoding performs slightly worse than the optimal unconstrained precoding, with less than 5 dB loss but substantially reduced complexity. However, it is observed that the proposed hybrid precoding with sub-connected architecture achieves the spectral efficiency close to that achieved by the spatially sparse precoding with fully-connected architecture and outperforms the SIC-based precoding with sub-connected architecture in the whole simulated SNR range. The proposed hybrid

precoding reduces the hardware complexity, achieving a better trade-off between the spectral efficiency and the hardware complexity.

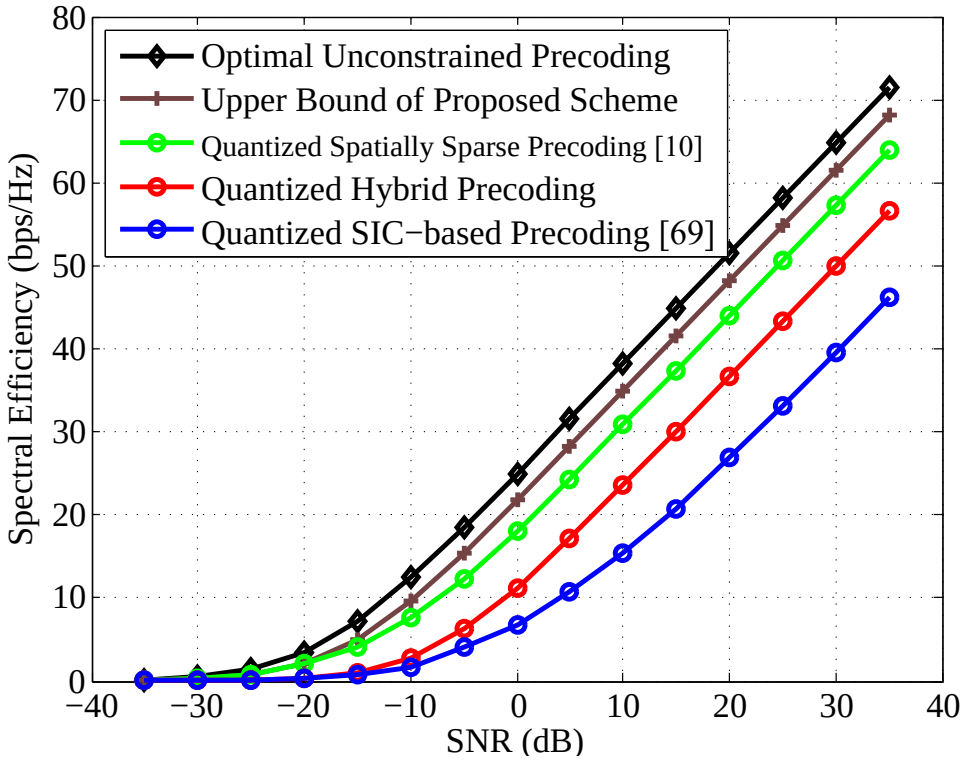


Figure 4.3: Spectral efficiency achieved by different quantized precoding schemes with 4 bits of precision in mmWave massive MIMO systems where $N_t = 128, N_r = 8, N_{\text{RF}} = 4$.

Due to the high cost of variable phase shifters with infinite resolution, the RF precoder suffers from the heavily quantized phase control. The phases of the RF precoder are quantized to a finite set with finite resolution phase shifters using the closet neighbor based on the nearest Euclidean distance, which can be expressed as $\widetilde{\varphi} = (2\pi\widetilde{n})/2^B$, where $\widetilde{n} = \arg\min_{n\in\{0,\cdots,2^B-1\}}|\varphi - \frac{2\pi n}{2^B}|$ and φ is the unquantized phase. We assume $B = 4$ bits of precision in the simulation. In terms of the heavily quantized phase control, Fig. 4.3 illustrates the spectral efficiency achieved by different quantized precoding schemes, considering the same setup as Fig. 4.2. It is observed that both hybrid precoding schemes with sub-connected architecture suffer degradation, thus there is a small gap between the fully-connected and sub-connected architectures. However, the performance of the proposed precoding still outperforms the SIC-based precoding.

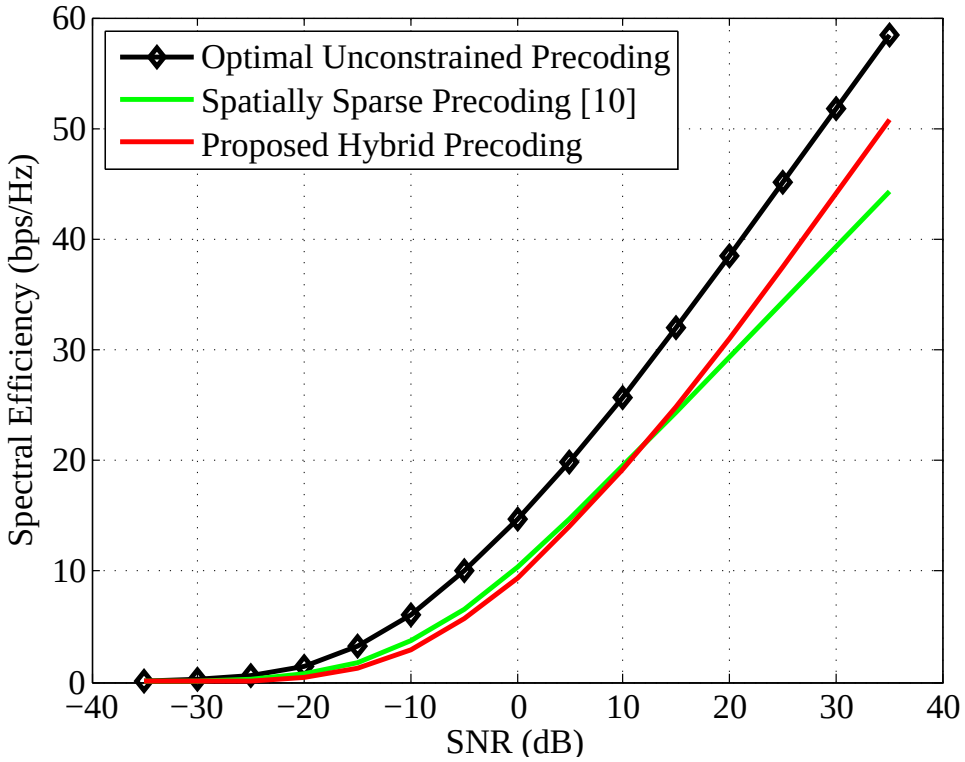


Figure 4.4: Spectral efficiency achieved by different precoding schemes with infinite resolution in mmWave massive MIMO systems where $N_t = 128, N_r = N_{\text{RF}} = 4$.

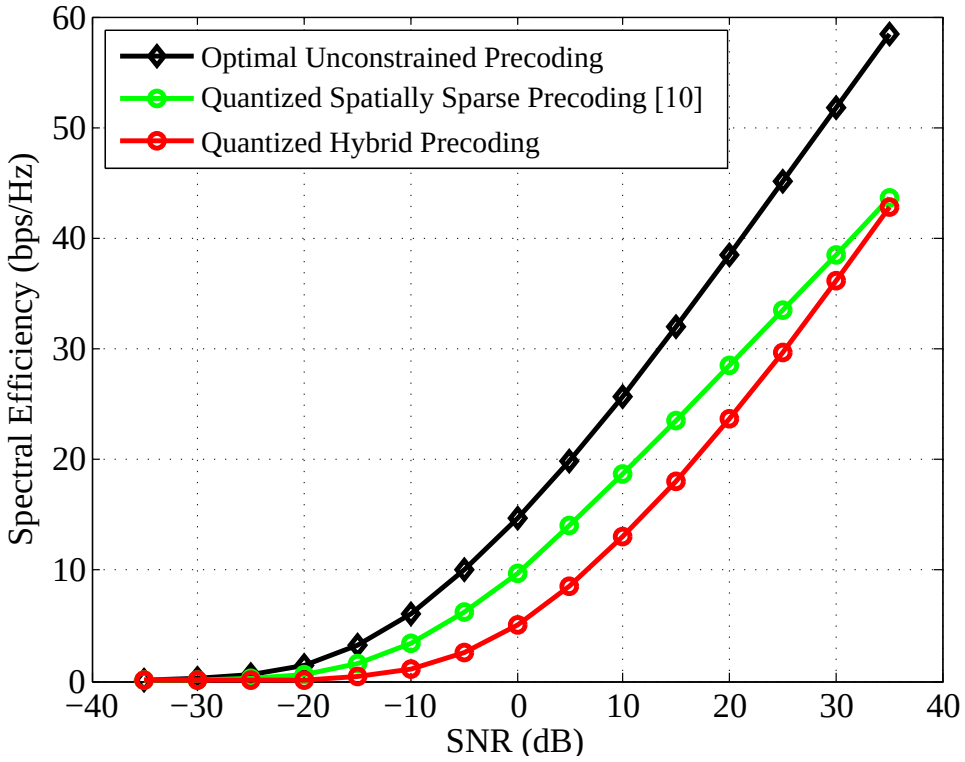


Figure 4.5: Spectral efficiency achieved by different precoding schemes with infinite resolution in mmWave massive MIMO systems where $N_t = 128, N_r = N_{\text{RF}} = 4$.

Under a special condition that the number of the receiver antennas is equal to the number of RF chains in the transmitter, Fig. 4.4 demonstrates the spectral efficiency of the above precoding schemes. Assuming $N_t = 128$ and $N_r = N_{\text{RF}} = 4$, note that in the high SNR region, the spectral efficiency of the proposed hybrid scheme outperforms that of the precoding scheme with fully-connected architecture. However, when the SNR is less than 11dB, the proposed scheme performs worse than the scheme with fully-connected architecture. Nevertheless, the resluts show the potential of our proposed algorithm, because the sub-connected architecture can even have better spectral efficiency performance than fully-connected architecture. In Fig. 4.5, the quantized versions results in 6dB loss in performance, which illustrates that the sub-connected architecture will be more affected by the quantization.

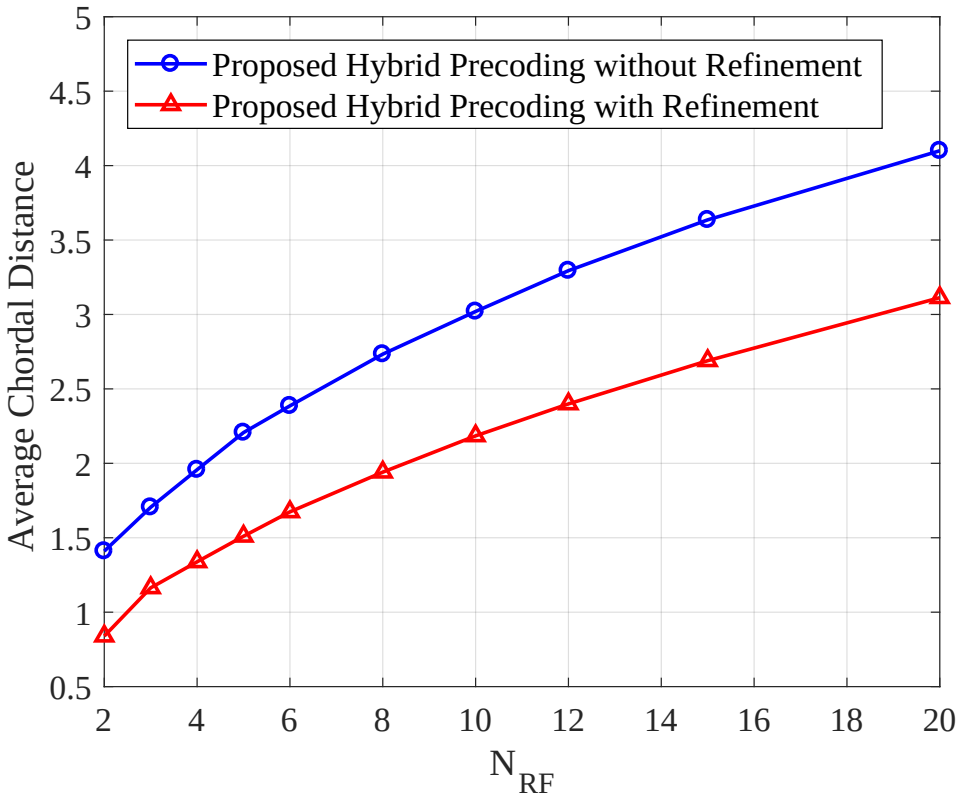


Figure 4.6: The average chordal distance as a function of the number of RF chains where $N_t = 120, N_r = 20$ and SNR $= 0$ dB.

Fig. 4.6 shows the chordal distance given by (4.29) to compare the performance of the proposed hybrid precoding with and without refinement process. In Fig. 4.6,

the numbers of antennas of the transmitters and the receivers are assumed to be $N_t = 120$ and $N_r = 20$, respectively. It is shown that the chordal distance with the refinement process is much smaller than that without refinement. Thus, the successive refinement improves the proposed algorithm to further approach the performance of the optimal precoding.

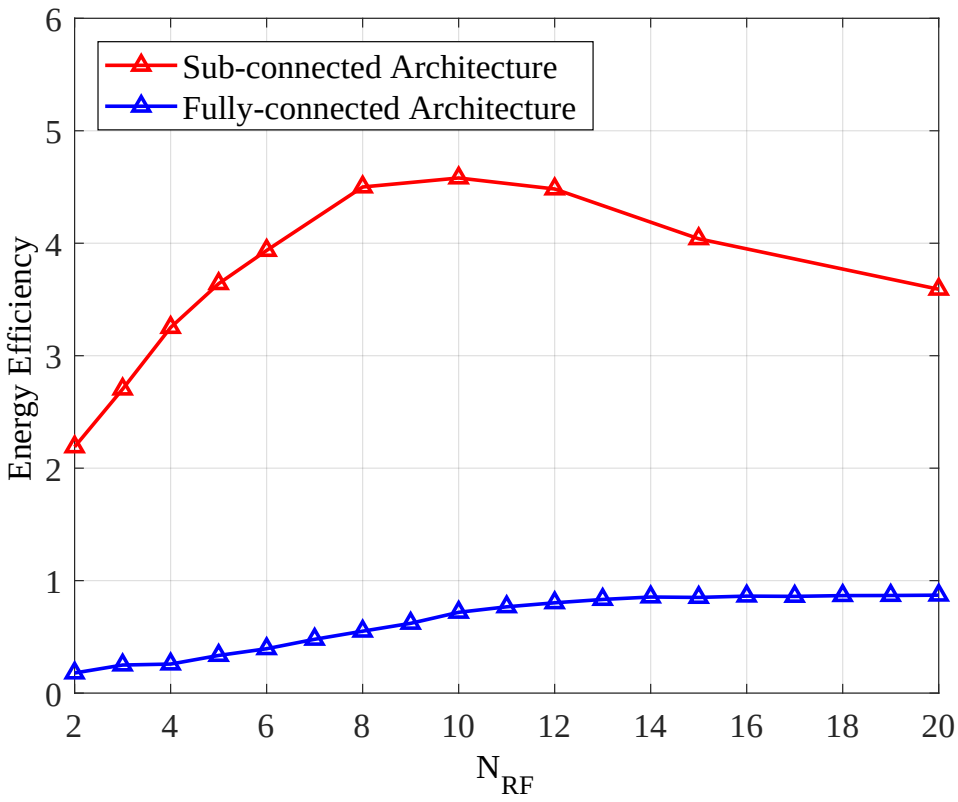


Figure 4.7: Energy efficiency of the fully-connected and sub-connected architectures against the number of RF chains where $N_t = 120, N_r = 20$ and SNR = 20 dB.

The energy efficiency defined in (4.44) of the sub-connected and fully-connected architectures is evaluated in Fig. 4.7. The simulation parameters are defined as $P_{\mathrm{S}} = 5\mathrm{W}$, $P_{\mathrm{RF}} = 40\mathrm{mW}$, $P_{\mathrm{PS}} = 30\mathrm{mW}$ and $P_{\mathrm{LNA}} = 20\mathrm{mW}$ [19]. When increasing N_{RF}, the numbers of phase shifters and low noise amplifiers increase in the fully-connected architecture, which makes the power consumption grow faster than that in the sub-connected architecture. The reason is that the numbers of phase shifters and low noise amplifiers are both independent of N_{RF} in the sub-connected architecture, thus the power consumption remains almost unchanged compared with the fully-connected architecture. Equation (4.44) illustrates substantially different behaviours

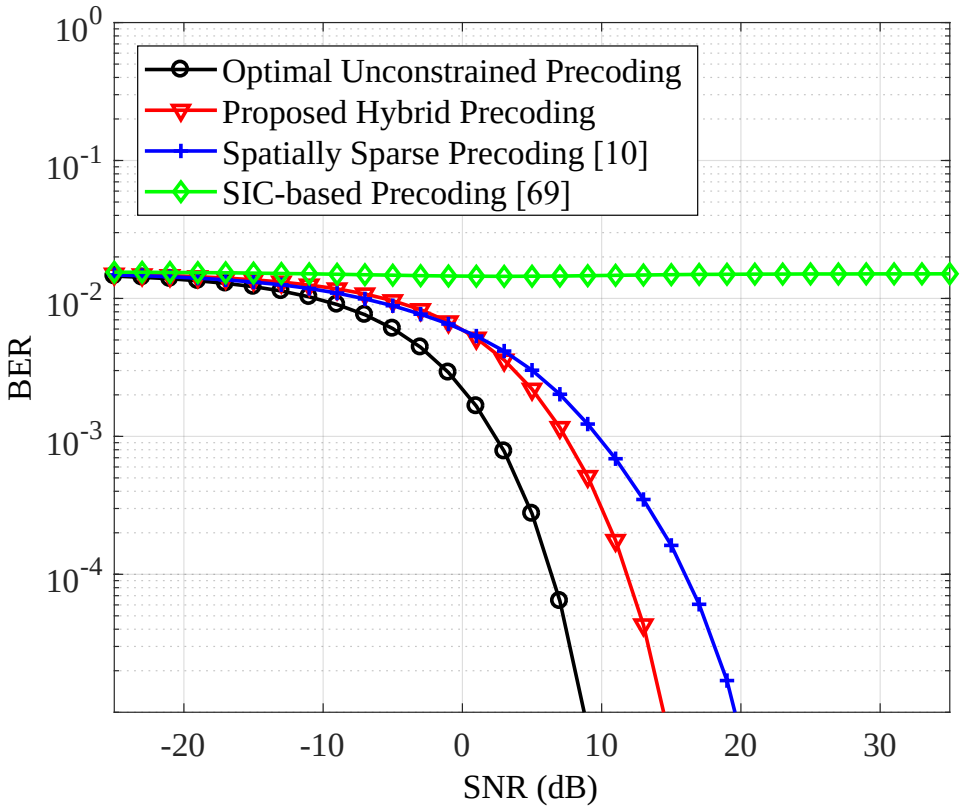


Figure 4.8: BER comparison of the proposed hybrid precoding with optimal unconstrained precoding, spatially sparse precoding and SIC-based preocidng where $N_t = 128, N_r = 8, N_{\mathrm{RF}} = 4$.

for the two architectures. Obviously, the sub-connected architecture achieves higher energy efficiency than the fully-connected architecture in the whole simulated N_{RF} range. Note that there is a peak at $N_{\mathrm{RF}} = 10$. When designing the hybrid schemes in the future, this phenomenon inspires us to consider the appropriate number of RF chains, which can maximize the energy efficiency.

Fig. 4.8 plots the BER results of each for the various hybrid precoding schemes. As shown in Fig. 4.8, there is a small gap between the optimal precoding and the proposed hybrid precoding. The BER of the proposed hybrid precoding outperforms the spatially sparse precodjng and the SIC-based precoding. Therefore, with lower hardware complexity, the proposed hybrid scheme performs better than the other schemes with sub-connected architecture, even that with fully-connected architecture.

4.10 Summary

In this chapter, a new iterative hybrid precoding scheme with a sub-connected architecture is proposed for mmWave MIMO systems. The main advantage of the proposed method is the reduction of the hardware complexity and the power consumption. On the basis of mathematical analysis, the closed-form expressions for the baseband and RF precoders are developed along with the successive refinement process. The simulation results of the performance of the proposed hybrid precoding scheme are presented and it is showed that it performs close to the spectral efficiency of the fully-connected architecture. Moreover, it performs much better than other existing hybrid precoding schemes with sub-connected architecture. Based on the evaluation of energy efficiency, the proposed hybrid precoding can have a good trade-off between the spectral efficiency and the power consumption.

Chapter 5

Performance Analysis of Linear Quantized Precoding for the Multiuser Massive MIMO Systems with Hardware Mismatch

5.1 Introduction

In massive multi-user MIMO systems, the data converters at the base station are main sources for power consumption. Although hybrid precoding schemes can efficiently reduce the number of required number of RF chains, they can not scale well for wideband systems. In this chapter, another direction of reducing the power consumption is used for massive MIMO systems, which employs low-resolution DACs for every RF chain connected to an antenna. In the downlink communications, the transmitted signal at each RF chain is generated by a pair of DACs. The power consumption of DACs grows exponentially with the number of quantization bits and linearly with the bandwidth and sampling rate [22] [23]. Obviously, if each RF chain is equipped with a pair of high-resolution DACs in massive MIMO systems, the systems will suffer from prohibitively high power consumption due to the large

number of required DACs. Therefore, the resolution of DACs should be limited to keep the power consumption within bearable levels.

In practice, the performance of massive MIMO systems is not only affected by the wireless channel, but also by the transceiver RF circuits [72]. Due to the difference between the hardware employed at the base station and the mobile terminals, the practical channel coefficients for the uplink and the downlink will be different with the analog circuit gains, which is called hardware mismatch. This means the downlink is not the transpose of the uplink in TDD mode. Therefore, the hardware mismatch leads to an significant constraint on the performance of the downlink communications, because the precoding matrix becomes inaccurate with the use of uplink channel estimation.

There is little theoretical analysis on the impact of hardware mismatch on system performance for massive MIMO systems. In [73], the analysis on the performance of massive MIMO systems with hardware mismatch is presented, using regularized ZF and MF precoding schemes. However, as mentioned above, the conventional ZF and MF precoding schemes are infeasible for massive MIMO systems, thus the quantized precoding is taken into consideration. There has only been little research that consider the influence of low-resolution DACs for the downlink massive multi-user MIMO systems. For the case of one-bit DACs, [22] studies the performance of one-bit quantized ZF precoding on Rayleigh fading channel. In [21], simple closed-form approximations are developed for the achievable rate with both one-bit and multi-bit DACs.

Therefore, in consideration of both hardware mismatch and low-resolution DACs, we study the impact of one-bit ZF precoding with hardware mismatch for the downlink massive multi-user MIMO systems. Hardware mismatch makes the communication channel be corrupted by two diagonal matrices of the transmit and receive circuit gains, which increases the difficulty of analysis. With the use of the Bussgang theorem and random matrix theorem, a analytical approximation expression

is derived for the achievable rate. In addition, for a special case, a performance approximation is also derived in the high SNR region, which is related to the ratio of the number of base station antennas and the number of mobile users , and the statistics of the circuit gains at the base station.

5.2 System Model and Quantized Precoding

5.2.1 System Model

The downlink of a single-cell massive multi-user MIMO system is considered operating in TDD mode. In this system, a base station with N antennas serves K single antenna mobile users simultaneously and usually $N \gg K$. In consideration of hardware mismatch, $\mathbf{h}_k = [h_{k1}, h_{k2}, \cdots, h_{kN}]$ and $\tilde{\mathbf{h}}_k = [\tilde{h}_{k1}, \tilde{h}_{k2}, \cdots, \tilde{h}_{kN}]$ are the uplink and downlink channel vectors for the k-th user, respectively. h_{kn} is the channel gain from the k-th user to the n-th base station antenna for the uplink and $\tilde{h}_{kn}$ is the channel gain from the n-th base station antenna to the k-th user for the downlink. Because the communication channel comprises not only the wireless propagation channel, but also the transmit and receive circuits, as shown in Fig. 5.1, we model the channel gains h_{kn} and $\tilde{h}_{kn}$ as

$$\begin{aligned} h_{kn} &= \tilde{t}_k v_{kn} r_n, \\ \tilde{h}_{kn} &= \tilde{r}_k \tilde{v}_{kn} t_n, \end{aligned} \tag{5.1}$$

where v_{kn} and $\tilde{v}_{kn}$ are the corresponding uplink and downlink wireless channel gains.

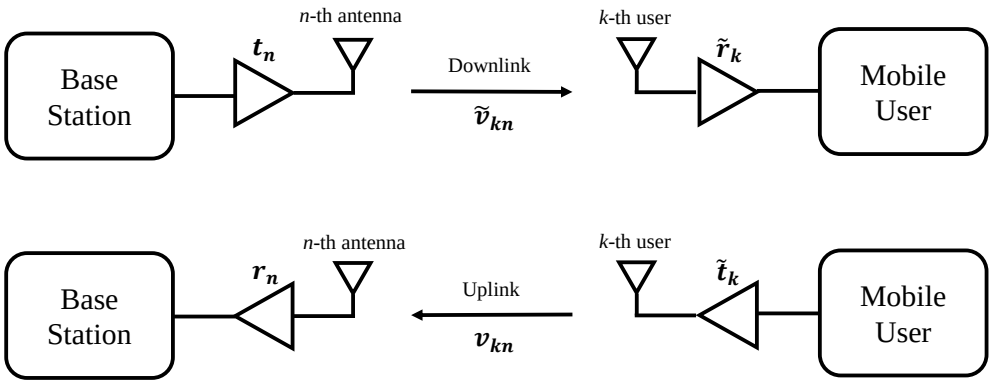


Figure 5.1: Hardware mismatch for the downlink and uplink transmissions.

It is assumed the wireless channel gains in the uplink and downlink transmission are

the same within the channel coherent time, thus

$$v_{kn} = \tilde{v}_{kn}. \tag{5.2}$$

$\tilde{t}_k$ and $\tilde{r}_k$ are the transmit and receive circuit gains for the k-th user and t_n and r_n are the transmit and receive circuit gains for the n-th base station antenna. Because of different hardware employed by the base station and the mobile users in the uplink and downlink, the circuit gains satisfy the condition,

$$\tilde{t}_i \neq \tilde{r}_i, t_i \neq r_i, \tilde{t}_k \neq t_i, \tilde{r}_i \neq r_i. \tag{5.3}$$

Although the circuit gains change with the working conditions, they vary slowly compared with the wireless channel gains, thus the circuit gains are regarded as constant in this chapter.

Under these assumption, the diagonal elements of $\mathbf{T}$ and $\mathbf{R}$ denote the coefficients of the transmit and receive circuit gains of the base station, respectively. Similarly, the diagonal elements of $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{R}}$ denote the coefficients of the transmit and receive circuit gains of all the mobile users, respectively. Thus, all these diagonal matrices of the circuit gains are defined as

$$\mathbf{T} = \mathrm{diag}(t_1, t_2, \cdots, t_N), \tag{5.4}$$

$$\mathbf{R} = \mathrm{diag}(r_1, r_2, \cdots, r_N), \tag{5.5}$$

$$\tilde{\mathbf{T}} = \mathrm{diag}(\tilde{t}_1, \tilde{t}_2, \cdots, \tilde{t}_K), \tag{5.6}$$

$$\tilde{\mathbf{R}} = \mathrm{diag}(\tilde{r}_1, \tilde{r}_2, \cdots, \tilde{r}_K). \tag{5.7}$$

Then the entries in the matrix $\mathbf{V} = [v_{kn}]_{K\times N}$ represent the wireless channel coefficients, which contains the i.i.d. complex Gaussian entries with zero mean and unit

variance. The uplink and downlink channels can be expressed as

$$\mathbf{H}_U = \tilde{\mathbf{T}}\mathbf{V}\mathbf{R}, \tag{5.8}$$

$$\mathbf{H}_D = \tilde{\mathbf{R}}\mathbf{V}\mathbf{T}. \tag{5.9}$$

Due to the hardware mismatch between the uplink and downlink communication, $\mathbf{H}_U$ and $\mathbf{H}_D$ are not exactly the same. The received signal at all the users can be expressed as

$$\mathbf{y} = \sqrt{\mathbf{P}}\mathbf{H}_D\mathbf{x} + \mathbf{n}, \tag{5.10}$$

where $\mathbf{P} = \text{diag}(p_1, p_2, \cdots, p_K)$ is the power loading matrix with p_k being the power loading factor for the k-th user, $\mathbf{x} = [x_1, x_2, \cdots, x_K]^T \in \mathbb{C}^{K\times 1}$ is the vector of transmitted signal for K mobile users and the vector $\mathbf{n} \in \mathbb{C}^{K\times 1}$ represents the additive white Gaussian noise with $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I}_K)$.

5.2.2 Linear Quantized Precoding

The transmitted signal vector $\mathbf{x}$ is the product of the precoding matrix $\mathbf{F} \in \mathbb{C}^{N\times K}$ and the symbol vector $\mathbf{s} \in \mathbb{C}^{K\times 1}$, which is given by

$$\mathbf{x} = \mathbf{F}\mathbf{s}, \tag{5.11}$$

where $\mathbf{s}$ is the vector of Quadrature Phase Shift Keying (QPSK) symbols with zero mean and unit variance. Because the linear precoding can achieve near optimal performance with relatively low computational complexity in massive MIMO systems, the linear precoding is used for the analysis. We first design the linear precoding matrix with infinite-resolution DACs and then quantize the resulting precoded vector. Quantizing the precoded vector implies no additional computational complexity. Thus, in Fig. 5.2, with the use of quantization, the transmitted signal vector

in (5.11) is given by

$$\mathbf{x} = \mathbb{Q}(\mathbf{Fs}), \tag{5.12}$$

where $\mathbb{Q}(\cdot)$ denotes the non-linear quantizer-mapping function. The one-bit quan-

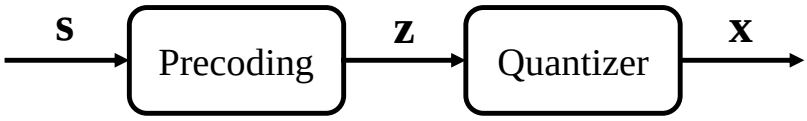


Figure 5.2: Diagram of linear quantized precoding.

tization operation is expressed as

$$\mathbb{Q}(\mathbf{z}) = \text{sign}(\Re(\mathbf{z})) + j\text{sign}(\Im(\mathbf{z})), \tag{5.13}$$

where $\Re(\cdot)$ is the real part, $\Im(\cdot)$ is the imaginary part and sign$(\cdot)$ is the sign of their arguments. Therefore, the received signal vector can be rewritten as

$$\mathbf{y} = \sqrt{\mathbf{P}}\mathbf{H}_D\mathbb{Q}(\mathbf{Fs}) + \mathbf{n}, \tag{5.14}$$

5.3 Analysis of One-bit Quantized Precoding Using Bussgang Theorem

The Bussgang theorem has recently been used to analyze the massive multi-user MIMO downlink with one-bit quantization. Based on the Bussgang theorem, the quantized signal vector $\mathbf{x}$ can be decomposed into a linear function of the input to the quantizers and a distortion that is uncorrelated with the input to the quantizer, thus

$$\mathbf{x} = \mathbb{Q}(\mathbf{z}) = \mathbf{Gz} + \mathbf{q}, \tag{5.15}$$

where $\mathbf{z} = \mathbf{Fs} \in \mathbb{C}^{N\times 1}$ and $\mathbf{d}$ is the quantization noise. The matrix $\mathbf{G}$ is a $N \times N$ diagonal matrix to satisfy the covariance matrix $\mathbf{R_{zq}} = \mathbb{E}[\mathbf{zq}^H] = \mathbf{0}$, hence

$$\begin{aligned}\mathbb{E}[\mathbf{qz}^H] &= \mathbb{E}[(\mathbf{x} - \mathbf{Gz})\mathbf{z}^H], \\ &= \mathbb{E}[\mathbf{xz}^H] - \mathbb{E}[\mathbf{Gzz}^H], \\ &= \mathbf{R_{xz}} - \mathbf{GR_{zz}}, \\ &= \mathbf{0}, \end{aligned} \tag{5.16}$$

where $\mathbf{R_{xz}}$ denotes the the covariance matrix of the output vector $\mathbf{x}$ of the quantizer with the input vector $\mathbf{z}$ of the quantizer and $\mathbf{R_{zz}}$ denotes the variance matrix of the precoded signal vector $\mathbf{z}$. Note that

$$\mathbf{R_{zz}} = \mathbb{E}[\mathbf{zz}^H] = \mathbf{F}\mathbb{E}[\mathbf{ss}^H]\mathbf{F}^H = \mathbf{FF}^H. \tag{5.17}$$

Assuming the elements of $\mathbf{z}$ is Gaussian distributed, the inter-correlation between the one-bit quantized x_i and unquantized z_j can be represented as the normalized inter-correlation of the unquantized signals with a factor $\sqrt{\frac{2}{\pi}}$ [74],

$$\mathbb{E}[x_i z_j^*] = \sqrt{\frac{2}{\pi}} \frac{\mathbb{E}[z_i z_j^*]}{\sqrt{\mathbb{E}[z_i z_i^*]}}. \tag{5.18}$$

Based on (5.18), $\mathbf{R_{xz}}$ in (5.17) can be defined as

$$\begin{aligned}\mathbf{R_{xz}} &= \sqrt{\frac{2}{\pi}}\{\mathrm{diag}(\mathbb{E}[\mathbf{zz}^H])\}^{-\frac{1}{2}}\mathbb{E}[\mathbf{zz}^H] \\ &= \sqrt{\frac{2}{\pi}}\{\mathrm{diag}(\mathbf{FF}^H)\}^{-\frac{1}{2}}\mathbf{FF}^H. \end{aligned} \tag{5.19}$$

Therefore, substituting (5.17) and (5.19) into (5.16),

$$\mathbf{G} = \sqrt{\frac{2}{\pi}}\{\mathrm{diag}(\mathbf{FF}^H)\}^{-\frac{1}{2}}. \tag{5.20}$$

Next, (5.15) is used to write the covariance matrix $\mathbf{R_{qq}}$ of the quantization noise vector $\mathbf{q}$ as

$$\begin{aligned}\mathbf{R_{qq}} &= \mathbb{E}[(\mathbf{x}-\mathbf{Gz})(\mathbf{x}-\mathbf{Gz})^H] \\ &= \mathbb{E}[\mathbf{xx}^H] - \mathbb{E}[\mathbf{xz}^H]\mathbf{G}^H - \mathbf{G}\mathbb{E}[\mathbf{zx}^H] + \mathbf{G}\mathbb{E}[\mathbf{zz}^H]\mathbf{G}^H \\ &= \mathbf{R_{xx}} - \mathbf{R}_{xz}\mathbf{G}^H - \mathbf{G}\mathbf{R_{zx}} + \mathbf{G}\mathbf{R}_{zz}\mathbf{G}^H \\ &= \mathbf{R_{xx}} - \mathbf{G}\mathbf{F}\mathbf{F}^H\mathbf{G}^H, \end{aligned} \tag{5.21}$$

where $\mathbf{R_{xx}}$ is the covariance matrix of the quantized signal vector $\mathbf{x}$. Based on the arcsin law, for a one-bit hard limiting quantizer, the inter-correlation of the quantized signals can be given by [74]

$$\mathbb{E}[x_i x_j^*] = \frac{2}{\pi}\arcsin\left(\Re\left(\frac{\mathbb{E}[z_i z_j^*]}{\sqrt{\mathbb{E}[z_i z_i^*]\mathbb{E}[z_j z_j^*]}}\right)\right) + j\frac{2}{\pi}\arcsin\left(\Im\left(\frac{\mathbb{E}[z_i z_j^*]}{\sqrt{\mathbb{E}[z_i z_i^*]\mathbb{E}[z_j z_j^*]}}\right)\right). \tag{5.22}$$

From (5.22), $\mathbf{R_{xx}}$ can be expressed as

$$\begin{aligned}\mathbf{R_{xx}} &= \frac{2}{\pi}\arcsin\{\{\mathrm{diag}(\mathbf{R_{zz}})\}^{-\frac{1}{2}}\Re(\mathbf{R_{zz}})\{\mathrm{diag}(\mathbf{R_{zz}})\}^{-\frac{1}{2}}\} \\ &\quad + j\frac{2}{\pi}\arcsin\{\{\mathrm{diag}(\mathbf{R_{zz}})\}^{-\frac{1}{2}}\Im(\mathbf{R_{zz}})\{\mathrm{diag}(\mathbf{R_{zz}})\}^{-\frac{1}{2}}\} \\ &= \frac{2}{\pi}\arcsin\{\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\Re(\mathbf{F}\mathbf{F}^H)\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\} \\ &\quad + j\frac{2}{\pi}\arcsin\{\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\Im(\mathbf{F}\mathbf{F}^H)\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\}. \end{aligned} \tag{5.23}$$

Thus,

$$\begin{aligned}\mathbf{R_{qq}} &= \frac{2}{\pi}\{\arcsin\{\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\Re(\mathbf{F}\mathbf{F}^H)\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\} \\ &\quad + j\arcsin\{\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\Im(\mathbf{F}\mathbf{F}^H)\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\} \\ &\quad - \{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\mathbf{F}\mathbf{F}^H\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}\}. \end{aligned} \tag{5.24}$$

5.4 Analysis of Achievable Rate

Based on the Bussgang theorem, the received signal vector $\mathbf{y}$ after quantization can be expressed as follows,

$$\begin{aligned}\mathbf{y} &= \sqrt{\mathbf{P}}\mathbf{H}_D\mathbf{x} + \mathbf{n} \\ &= \sqrt{\mathbf{P}}\mathbf{H}_D\mathbf{G}\mathbf{F}\mathbf{s} + \sqrt{\mathbf{P}}\mathbf{H}_D\mathbf{q} + \mathbf{n}.\end{aligned} \tag{5.25}$$

For the k-th user, the received signal y_k is given by

$$\begin{aligned}y_k &= \sqrt{p_k}\tilde{\mathbf{h}}_k^T\mathbf{G}\mathbf{F}\mathbf{s} + \sqrt{p_k}\tilde{\mathbf{h}}_k^T\mathbf{q} + n_k \\ &= \sqrt{p_k}\tilde{\mathbf{h}}_k^T\mathbf{G}\mathbf{f}_k s_k + \sum_{i=1,i\neq k}^{K}\sqrt{p_k}\tilde{\mathbf{h}}_k^T\mathbf{G}\mathbf{f}_i s_i + \sqrt{p_k}\tilde{\mathbf{h}}_k^T\mathbf{q} + n_k \\ &= \sqrt{p_k}g_{k,k}s_k + \sum_{i=1,i\neq k}^{K}\sqrt{p_k}g_{k,i}s_i + \sqrt{p_k}\tilde{\mathbf{h}}_k^T\mathbf{q} + n_k,\end{aligned} \tag{5.26}$$

where $g_{k,k}$ denotes the equivalent channel gain and $g_{k,i}$ denotes correlated channel gain. Therefore, the achievable rate R_k for the k-th user can be expressed as

$$R_k = \log_2(1+\gamma_k), \quad k = 1, 2, \cdots, K, \tag{5.27}$$

where

$$\gamma_k = \frac{p_k|\mathbb{E}[g_{k,k}]|^2}{\sum\limits_{i=1,i\neq k}^{K} p_k\mathbb{E}[|g_{k,i}|^2] + p_k\tilde{\mathbf{h}}_k^T\mathbf{R}_{\mathbf{qq}}\tilde{\mathbf{h}}_k^* + \sigma_n^2} \tag{5.28}$$

is the signal-to-quantization-interference-noise ratio (SQINR) for the k-th user. $\sum\limits_{i=1,i\neq k}^{K} p_k\mathbb{E}[|g_{k,i}|^2]$ accounts for the multi-user interference and $p_k\tilde{\mathbf{h}}_k^T\mathbf{R}_{\mathbf{qq}}\tilde{\mathbf{h}}_k^*$ accounts for the quantization noise.

Then with the effect of hardware mismatch, the analytical expressions of the achievable rate is derived for one-bit linear precoding. In this section, focusing on

the special case of ZF precoding, the precoding matrix is defined by

$$\mathbf{F} = \mathbf{H}_U^H (\mathbf{H}_U \mathbf{H}_U^H)^{-1}. \tag{5.29}$$

To deal with the mathematical expectation in (5.28), the asymptotic deterministic equivalents (ADEs) of the random variables are derived. Assuming N and K are large with the ratio $\frac{N}{K} > 1$, the expectation can be approximated by their corresponding ADEs. That means if a random variable $x \longrightarrow x_0$ and x_0 is deterministic, x_0 is the ADE of x.

In the asymptotic analysis, Lemma 1 is used, which is stated below,

Lemma 1 : Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ and $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N}\Phi_x)$, $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N}\Phi_y)$. Assume that $\mathbf{A}$ has uniformly bounded spectral norm (with respect to N) and that $\mathbf{x}$ and $\mathbf{y}$ are mutually independent and independent of $\mathbf{A}$. Then for all $p \geq 1$.

$$\begin{aligned}
&(1) \quad \mathbb{E}\left\{ \left| \mathbf{x}^H \mathbf{A} \mathbf{x} - \frac{1}{N} \mathrm{Tr}(\mathbf{A}\Phi_x) \right|^p \right\} = \mathcal{O}\left(N^{-\frac{p}{2}} \right), \\
&(2) \quad \mathbf{x}^H \mathbf{A} \mathbf{x} - \frac{1}{N} \mathrm{Tr}(\mathbf{A}\Phi_x) \to 0, \\
&(3) \quad \mathbf{x}^H \mathbf{A} \mathbf{y} \to 0, \\
&(4) \quad \mathbb{E}\left\{ \left| (\mathbf{x}^H \mathbf{A} \mathbf{x})^2 - \left(\frac{1}{N} \mathrm{Tr}(\mathbf{A}\Phi_x) \right)^2 \right| \right\} \to 0.
\end{aligned} \tag{5.30}$$

For analyzing the ZF precoding matrix, we consider the asymptotic behaviour of the matrix $\mathbf{H}_U \mathbf{H}_U^H$ first and have

$$\mathbf{H}_U \mathbf{H}_U^H = \tilde{\mathbf{T}} \mathbf{V} \mathbf{R} \mathbf{R}^H \mathbf{V}^H \tilde{\mathbf{T}}^H. \tag{5.31}$$

Since $\tilde{\mathbf{T}}$ and $\mathbf{R}$ are known constant matrices,

$$\begin{aligned}\mathbf{VRR}^H\mathbf{V}^H &= [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N] \begin{bmatrix} r_1 r_1^* & 0 & \cdots & 0 \\ 0 & r_2 r_2^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_N r_N^* \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^H \\ \mathbf{v}_2^H \\ \vdots \\ \mathbf{v}_N^H \end{bmatrix} \\ &= |r_1|^2 \mathbf{v}_1 \mathbf{v}_1^H + |r_2|^2 \mathbf{v}_2 \mathbf{v}_2^H + \cdots + |r_N|^2 \mathbf{v}_N \mathbf{v}_N^H, \end{aligned} \tag{5.32}$$

where $\mathbf{v}_i$ is the i-th column of the wireless channel matrix $\mathbf{V}$. The term $|r_i|^2 \mathbf{v}_i \mathbf{v}_i^H$ can be expressed as

$$\begin{aligned}|r_i|^2 \mathbf{v}_i \mathbf{v}_i^H &= |r_i|^2 \begin{bmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ v_{Ki} \end{bmatrix} \begin{bmatrix} v_{1i}^* & v_{2i}^* & \cdots & v_{Ki}^* \end{bmatrix} \\ &= |r_i|^2 \begin{bmatrix} v_{1i}v_{1i}^* & v_{1i}v_{2i}^* & \cdots & v_{1i}v_{Ki}^* \\ v_{2i}v_{1i}^* & v_{2i}v_{2i}^* & \cdots & v_{2i}v_{Ki}^* \\ \vdots & \vdots & \ddots & \vdots \\ v_{Ki}v_{1i}^* & v_{Ki}v_{2i}^* & \cdots & v_{Ki}v_{Ki}^* \end{bmatrix}. \end{aligned} \tag{5.33}$$

Therefore, the diagonal elements d_{jj} of the matrix $\mathbf{VRR}^H\mathbf{V}^H$ can be computed by

$$\begin{aligned} d_{jj} &= \sum_{i=1}^{N} |r_i|^2 v_{ji} v_{ji}^* \\ &= \begin{bmatrix} v_{j1} & v_{j2} & \cdots & v_{jN} \end{bmatrix} \begin{bmatrix} |r_1|^2 & 0 & \cdots & 0 \\ 0 & |r_2|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |r_N|^2 \end{bmatrix} \begin{bmatrix} v_{j1}^* \\ v_{j2}^* \\ \vdots \\ v_{jN}^* \end{bmatrix}. \end{aligned} \tag{5.34}$$

$\mathbf{v}'_j$ is defined as the j-th row of the matrix $\mathbf{V}$. Because $\mathbf{v}'_j \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, we have

$\sqrt{\frac{1}{N}}\mathbf{v}'_j \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N}\mathbf{I})$. Thus,

$$d_{jj} = N\left(\sqrt{\frac{1}{N}}\mathbf{v}'_j\right)\mathbf{R}\mathbf{R}^H\left(\sqrt{\frac{1}{N}}\mathbf{v}'^H_j\right). \tag{5.35}$$

According to Lemma 1,

$$d_{jj} \to \mathrm{Tr}(\mathbf{R}\mathbf{R}^H). \tag{5.36}$$

Similarly, the off-diagonal elements $d_{jl,j\neq l}$ of $\mathbf{V}\mathbf{R}\mathbf{R}^H\mathbf{V}^H$ will approximate to 0. Therefore,

$$\mathbf{V}\mathbf{R}\mathbf{R}^H\mathbf{V}^H \to \mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\mathbf{I}_K. \tag{5.37}$$

From (5.31),

$$\mathbf{H}_U\mathbf{H}_U^H \to \mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H. \tag{5.38}$$

Using the results of the previous section,

$$\begin{aligned}\mathbf{F}\mathbf{F}^H &= \mathbf{H}_U^H\left(\mathbf{H}_U\mathbf{H}_U^H\right)^{-2}\mathbf{H}_U \\ &\to \mathbf{R}^H\mathbf{V}^H\tilde{\mathbf{T}}^H\left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H\right)^{-2}\tilde{\mathbf{T}}\mathbf{V}\mathbf{R} \\ &= \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\right)^{-2}\mathbf{R}^H\mathbf{V}^H\left(\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H\right)^{-1}\mathbf{V}\mathbf{R}.\end{aligned} \tag{5.39}$$

Because the matrix $\mathbf{G} = \sqrt{\frac{2}{\pi}}\{\mathrm{diag}(\mathbf{F}\mathbf{F}^H)\}^{-\frac{1}{2}}$ only focuses on the diagonal elements of $\mathbf{F}\mathbf{F}^H$, when K is large, using Lemma 1 for the diagonal elements, the approximation of the matrix $\mathbf{G}$ can be expressed as

$$\mathbf{G} \to \sqrt{\frac{2}{\pi}}\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\mathrm{Tr}\left((\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H)^{-1}\right)^{-\frac{1}{2}}\left(\mathbf{R}^H\mathbf{R}\right)^{-\frac{1}{2}}. \tag{5.40}$$

In the same way, based on (5.38), the precoding matrix $\mathbf{F}$ has the asymptotic

approximation as

$$\mathbf{F} \to \mathbf{R}^H \mathbf{V}^H \tilde{\mathbf{T}}^H \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H) \tilde{\mathbf{T}} \tilde{\mathbf{T}}^H \right)^{-1}. \tag{5.41}$$

It is assumed that $\mathbf{f}_k$ denotes the k-th column of $\mathbf{F}$, then

$$\mathbf{f}_k \to \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H) |\tilde{t}_k|^2 \right)^{-1} \mathbf{R}^H \mathbf{v}_k^H \tilde{t}_k^*. \tag{5.42}$$

From (5.26), $g_{k,k}$ is given by

$$\begin{aligned} g_{k,k} &= \tilde{\mathbf{h}}_k^T \mathbf{G} \mathbf{f}_k \\ &= \tilde{r}_k \mathbf{v}_k \mathbf{T} \mathbf{G} \mathbf{f}_k \\ &\to \tilde{r}_k \tilde{t}_k^* \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H) |\tilde{t}_k|^2 \right)^{-1} \mathbf{v}_k \mathbf{T} \mathbf{G} \mathbf{R}^H \mathbf{v}_k^H. \end{aligned} \tag{5.43}$$

Using Lemma 1,

$$g_{k,k} \to \tilde{r}_k \tilde{t}_k^* \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H) |\tilde{t}_k|^2 \right)^{-1} \mathrm{Tr}\left(\mathbf{T}\mathbf{G}\mathbf{R}^H \right) \triangleq \mathbb{E}[g_{k,k}]. \tag{5.44}$$

Next, handling the ADE of $|g_{k,i}|^2$, $g_{k,i}$ can be expressed as

$$g_{k,i} = \tilde{r}_k \mathbf{v}_k \mathbf{T} \mathbf{G} \mathbf{f}_i, \tag{5.45}$$

then

$$\begin{aligned} |g_{k,i}|^2 &= |\tilde{r}_k|^2 \mathbf{f}_i^H \mathbf{G} \mathbf{T}^H \mathbf{v}_k^H \mathbf{v}_k \mathbf{T} \mathbf{G} \mathbf{f}_i \\ &\to |\tilde{r}_k|^2 \left(|\tilde{t}_i| \mathrm{Tr}(\mathbf{R}\mathbf{R}^H) \right)^{-2} \mathbf{v}_i \mathbf{R} \mathbf{G} \mathbf{T}^H \mathbf{v}_k^H \mathbf{v}_k \mathbf{T} \mathbf{G} \mathbf{R}^H \mathbf{v}_i^H. \end{aligned} \tag{5.46}$$

Using Lemma 1, (5.46) can be rewritten as

$$\begin{aligned}|g_{k,i}|^2 &\to |\tilde{r}_k|^2 \left(|\tilde{t}_i|\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\right)^{-2} \mathrm{Tr}\left(\mathbf{R}\mathbf{G}\mathbf{T}^H\mathbf{v}_k^H\mathbf{v}_k\mathbf{T}\mathbf{G}\mathbf{R}^H\right) \\ &= |\tilde{r}_k|^2 \left(|\tilde{t}_i|\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\right)^{-2} \mathrm{Tr}\left(\mathbf{v}_k\mathbf{T}\mathbf{G}\mathbf{R}^H\mathbf{R}\mathbf{G}\mathbf{T}^H\mathbf{v}_k^H\right) \\ &\to |\tilde{r}_k|^2 \left(|\tilde{t}_i|\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\right)^{-2} \mathrm{Tr}\left(\mathbf{T}\mathbf{G}\mathbf{R}^H\mathbf{R}\mathbf{G}\mathbf{T}^H\right) \triangleq \mathbb{E}[|g_{k,i}|^2]. \end{aligned} \tag{5.47}$$

In terms of the quantization noise $\tilde{\mathbf{h}}_k^T\mathbf{R}_{\mathbf{qq}}\tilde{\mathbf{h}}_k^*$, the expression of $\mathbf{F}\mathbf{F}^H$ should be first considered. Although $\mathbf{F}\mathbf{F}^H$ is rank deficient, if assuming that N and K tend to infinity, $\mathbf{F}\mathbf{F}^H$ can approximatively converge to a diagonal matrix. From (5.39),

$$\mathbf{F}\mathbf{F}^H \to \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\right)^{-2} \mathrm{Tr}\left((\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H)^{-1}\right)\mathbf{R}^H\mathbf{R}. \tag{5.48}$$

Therefore,

$$\Re(\mathbf{F}\mathbf{F}^H) \to \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H)\right)^{-2} \mathrm{Tr}\left((\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H)^{-1}\right)\mathbf{R}^H\mathbf{R}, \tag{5.49}$$

$$\Im(\mathbf{F}\mathbf{F}^H) \to \mathbf{0}. \tag{5.50}$$

Based on (5.24), the covariance matrix $\mathbf{R}_{\mathbf{qq}}$ can be approximated as

$$\mathbf{R}_{\mathbf{qq}} \to \frac{2}{\pi}\left(\frac{\pi}{2} - 1\right) = 1 - \frac{2}{\pi}, \tag{5.51}$$

and the quantization noise $\tilde{\mathbf{h}}_k^T\mathbf{R}_{\mathbf{qq}}\tilde{\mathbf{h}}_k^*$

$$\begin{aligned}\tilde{\mathbf{h}}_k^T\mathbf{R}_{\mathbf{qq}}\tilde{\mathbf{h}}_k^* &= \tilde{r}_k\mathbf{v}_k\mathbf{T}\mathbf{R}_{\mathbf{qq}}\mathbf{T}^H\mathbf{v}_k^H\tilde{r}_k^* \\ &\to \left(1 - \frac{2}{\pi}\right)|\tilde{r}_k|^2\mathrm{Tr}(\mathbf{T}\mathbf{T}^H). \end{aligned} \tag{5.52}$$

Therefore, substituting (5.44), (5.47) and (5.52) into (5.28), the SQINR can be

expressed in closed form as

$$\gamma_k \to \frac{p_k \left| \tilde{r}_k \tilde{t}_k^* \left(\mathrm{Tr}(\mathbf{R}\mathbf{R}^H) |\tilde{t}_k|^2 \right)^{-1} \mathrm{Tr}\left(\mathbf{T}\mathbf{G}\mathbf{R}^H \right) \right|^2}{\sum\limits_{i=1, i\neq k}^{K} p_k |\tilde{r}_k|^2 \left(|\tilde{t}_i| \mathrm{Tr}(\mathbf{R}\mathbf{R}^H) \right)^{-2} \mathrm{Tr}\left(\mathbf{T}\mathbf{G}\mathbf{R}^H\mathbf{R}\mathbf{G}\mathbf{T}^H \right) + p_k \left(1 - \frac{2}{\pi} \right) |\tilde{r}_k|^2 \mathrm{Tr}(\mathbf{T}\mathbf{T}^H) + \sigma_n^2}. \tag{5.53}$$

Then the achievable rate of one-bit quantized ZF precoding with the effect of hardware mismatch can be calculated according to (5.53).

For a simple scenario, in the high SNR region, the coefficients of the transmit and receive circuit gains of the base station are assumed to be random variables with independent and identical distribution, while those of the mobile users are constant and well compensated with $\tilde{t}_i = \tilde{r}_i = 1$. Then assuming equal power allocation is used, N and K tend to infinite with the ratio $\alpha = \frac{N}{K}$. With these assumption, in the high SNR region, the SQINR of the k-th user can be given as

$$\gamma_k \to \frac{\frac{2}{\pi K} \left| \mathrm{Tr}\left(\mathbf{R}^H \mathbf{T} (\mathbf{R}^H \mathbf{R})^{-\frac{1}{2}} \right) \right|^2}{\sum\limits_{i=1, i\neq k}^{K} \frac{2}{\pi K} \mathrm{Tr}(\mathbf{T}\mathbf{T}^H) + \left(1 - \frac{2}{\pi} \right) \mathrm{Tr}(\mathbf{T}\mathbf{T}^H)}. \tag{5.54}$$

Based on the Law of Large Number (LLN), as N and $K \to \infty$,

$$\begin{aligned} &\frac{1}{N} \mathrm{Tr}\left(\mathbf{R}^H \mathbf{T} (\mathbf{R}^H \mathbf{R})^{-\frac{1}{2}} \right) \to \mathbb{E}\left[r_i^* t_i \frac{1}{|r_i|} \right], \\ &\frac{1}{N} \mathrm{Tr}\left(\mathbf{T}\mathbf{T}^H \right) \to \mathbb{E}\left[|t_i|^2 \right]. \end{aligned} \tag{5.55}$$

Substituting (5.55) into (5.54),

$$\gamma_k \to \frac{\frac{2\alpha}{\pi} \left| \mathbb{E}\left[r_i^* t_i \frac{1}{|r_i|} \right] \right|^2}{\mathbb{E}\left[|t_i|^2 \right]}. \tag{5.56}$$

Therefore, (5.56) can be regarded as the performance approximation with constant value for the one-bit quantized ZF precoding in the high SNR region and

the performance are related to α and the statistics of the circuit gains at the base station.

5.5 Results and Discussion

In this section, the performances of achievable rate and SER are evaluated for one-bit ZF precoding with hardware mismatch. In the simulation, it is assumed that the circuit gains $t_i, r_i, \tilde{t}_i, \tilde{r}_i$ are identically and uniformly distributed as $U(1-0.5\delta, 1+0.5\delta)$ and δ is set as 0.4. Note that, if δ is large, the influence of hardware mismatch will be severe. All the simulation results are averaged over randomly generated coefficients of hardware mismatch and channel realizations.

In Fig. 5.3, the performance loss caused by the effect of the hardware mismatch is evaluated. It is assumed that the transmit power is equally allocated among all the mobile users and the number of base station antennas is $N = 100$. Compared with one-bit ZF precoding without hardware mismatch, achievable rate of one-bit ZF precoding with hardware mismatch suffers obvious degradation. This indicates that ZF precoding is sensitive to the influence of hardware mismatch, because the ZF preocoding matrix is computed based on the estimation of uplink channel matrix. In addition, as shown in Fig. 5.3, achievable rate of K=5 performs better than that of K=20. Thus, if the number of base station antennas is fixed, fewer number of mobile users can achieve better performance of achievable rate.

In Fig. 5.4, the performance loss is evaluated with different conditions. With the constant ratio $\alpha = \frac{N}{K}$, the results of achievable rate are simulated for different numbers of base station antennas and mobile users. The achievable rate with large values $N = 100,\ K = 5$ performs better than that with small values $N = 40,\ K = 2$, which shows the benefits of massive MIMO systems. Therefore, the results reinforces the observation that the performance is influenced by specific numbers of both base station antennas and mobile users.

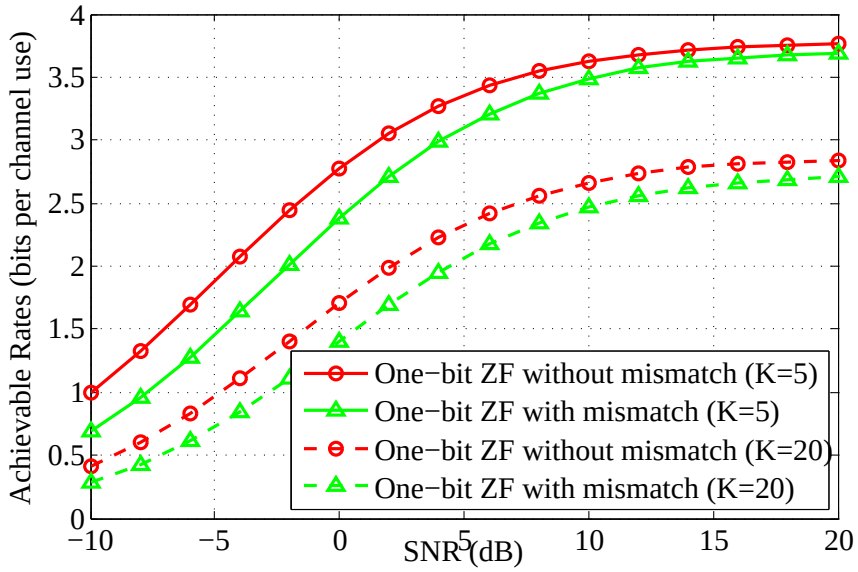


Figure 5.3: Achievable rates of one-bit ZF precoding with and without hardware mismatch, where $N = 100$.

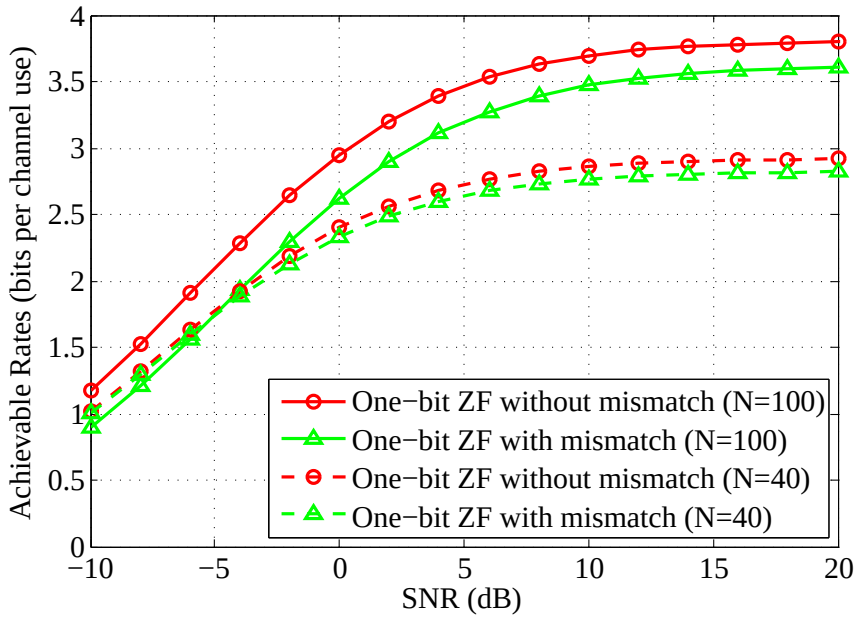


Figure 5.4: Achievable rates of one-bit ZF precoding with and without hardware mismatch, where $\alpha = 2$.

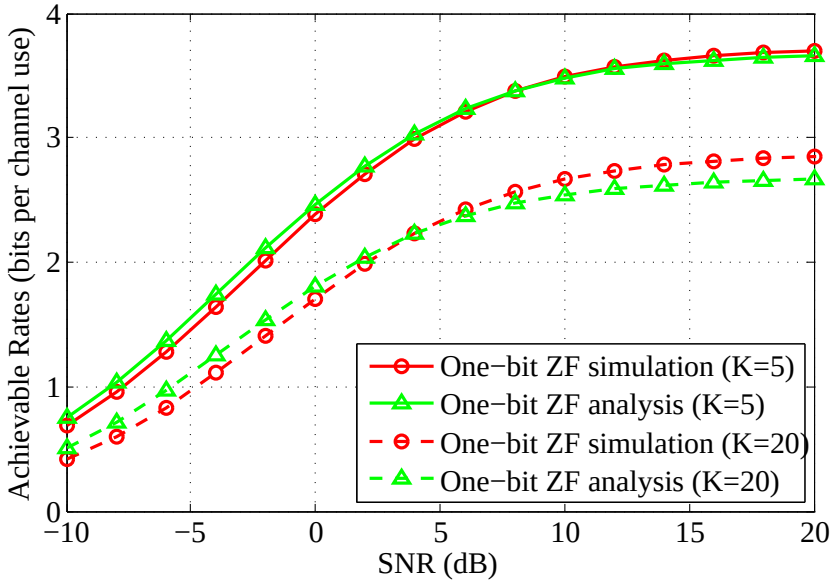


Figure 5.5: Achievable rates of one-bit ZF precoding with hardware mismatch, where $N = 100$.

The simulation results in Fig. 5.5 are conducted to validate the accuracy of the derived closed-form expression of the achievable rate, with $N = 100$. From Fig. 5.5, with the effect of hardware mismatch, our analytical approximation of one-bit ZF precoding in (5.53) can match well with the simulations for the cases of $K = 5$ and $K = 20$,which shows the accuracy of the analysis. The achievable rate for the case of $K = 5$ is better than that for the case of $K = 20$. Moreover, in the range of SNR, the analytical approximation with $K = 5$ is more accurate than that with $K = 20$. Therefore, larger value of the ratio $\alpha = N/K$ we use, higher achievable rate per user and more accurate approximation expression can be obtained for the one-bit ZF precoding with hardware mismatch.

Then Fig. 5.6 and Fig. 5.7 illustrate the performance approximation of achievable rate for one-bit ZF precoding in the high SNR region, considering the same setup as Fig. 5.5. Based on (5.56), the circuit gains at the mobile user is defined as $\tilde{\mathbf{T}} = \tilde{\mathbf{R}} = \mathbf{I}_K$. When the effect of hardware mismatch exists, the performance

approximation for one-bit ZF precoding is tight, which does not increase linearly with respect to SNR. When SNR is more than 20 dB, the achievable rates will saturate.

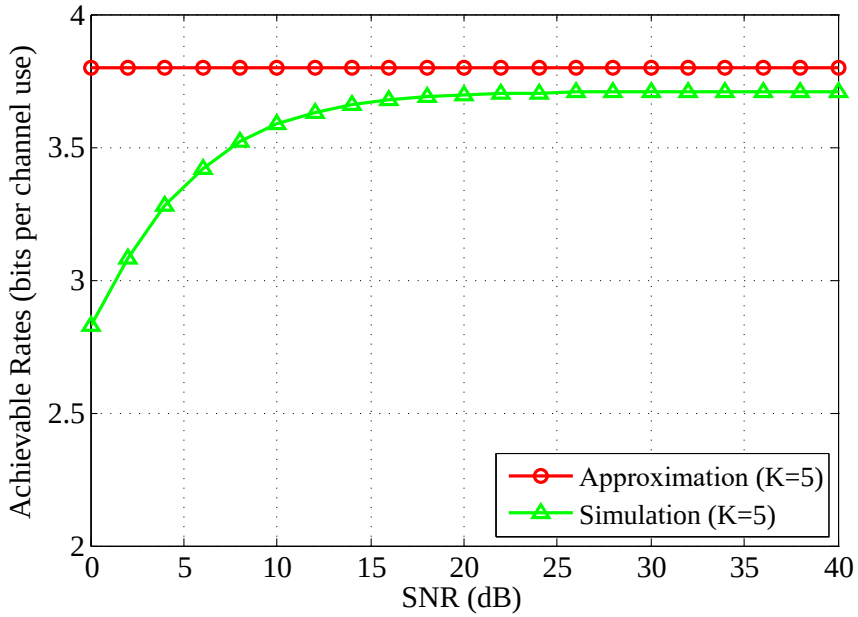


Figure 5.6: Performance approximation of one-bit ZF precoding with hardware mismatch, where $N = 100$ and $K = 5$.

In Fig. 5.8, the average SER for the unquantized ZF precoding is plotted to evaluate the influence of hardware mismatch with the number of base station antennas $N = 100$. For varied number of mobile users, the SER without hardware mismatch obviously outperforms that with hardware mismatch. For a target SER of 10^{-4}, the SER with hardware mismatch suffers from 3 dB performance loss. Similar to Fig. 5.8, Fig. 5.9 illustrates that the SER of one-bit ZF precoding is also influenced by hardware mismatch. In the high SNR region, the SER of one-bit ZF precoding with $K = 20$ approaches an error floor around 10^{-2} with and without hardware mismatch. The SER floor with $K = 5$ is of the order of 10^{-4} without hardware mismatch, while that with $K = 5$ is of the order of 10^{-3} with hardware mismatch.

Fig. 5.10 plots the SER results of unquantized and quantized ZF precoding when

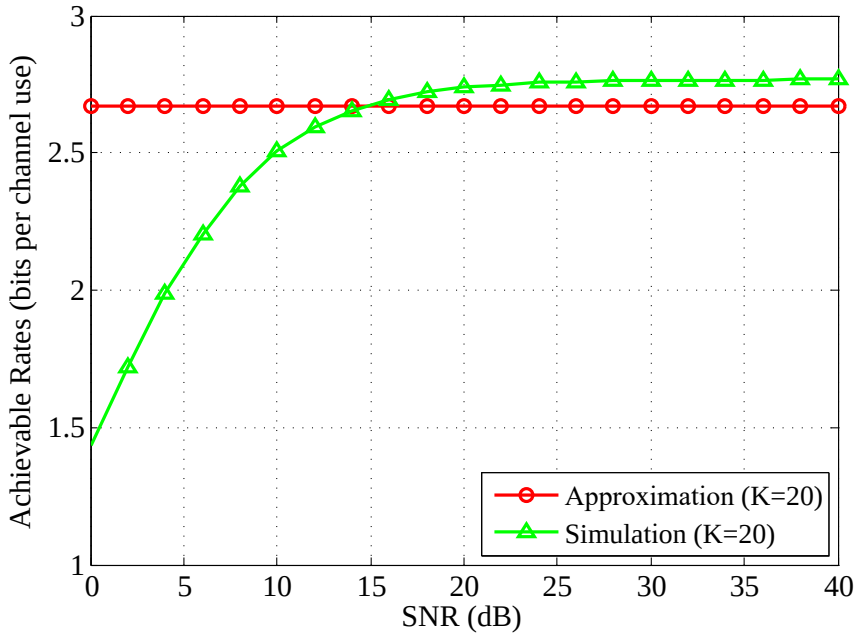


Figure 5.7: Performance approximation of one-bit ZF precoding with hardware mismatch, where $N = 100$ and $K = 20$.

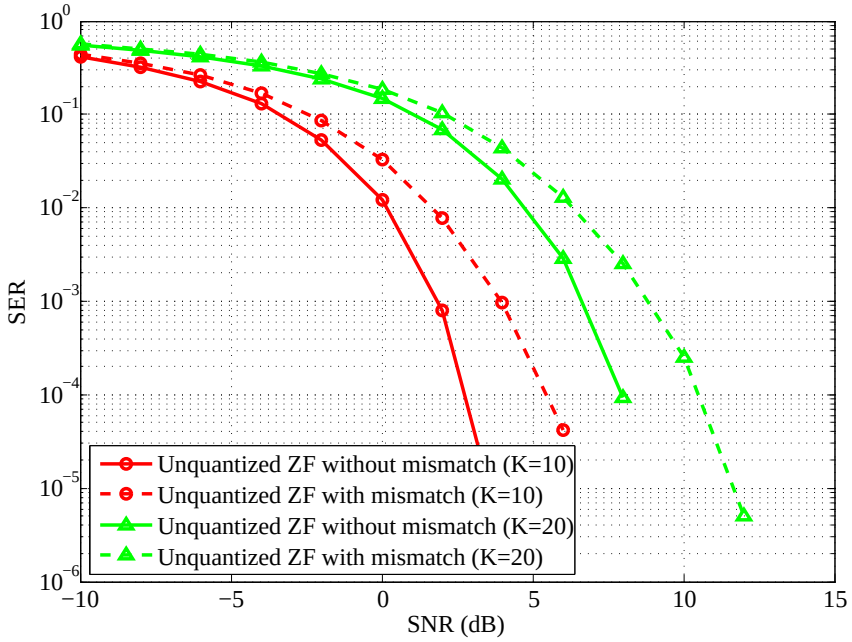


Figure 5.8: SER of unquantized ZF precoding with and without hardware mismatch, where $N = 100$.

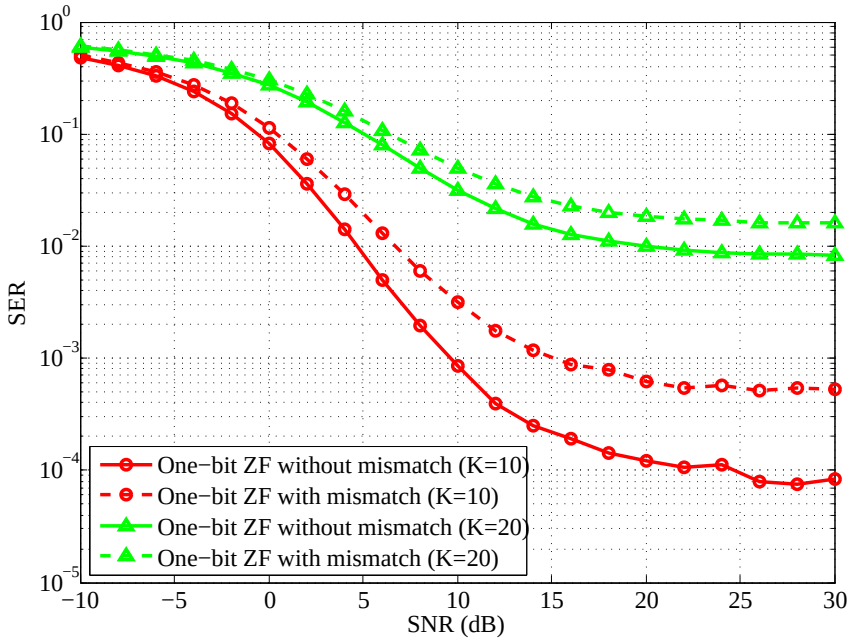


Figure 5.9: SER of one-bit ZF precoding with and without hardware mismatch, where $N = 100$.

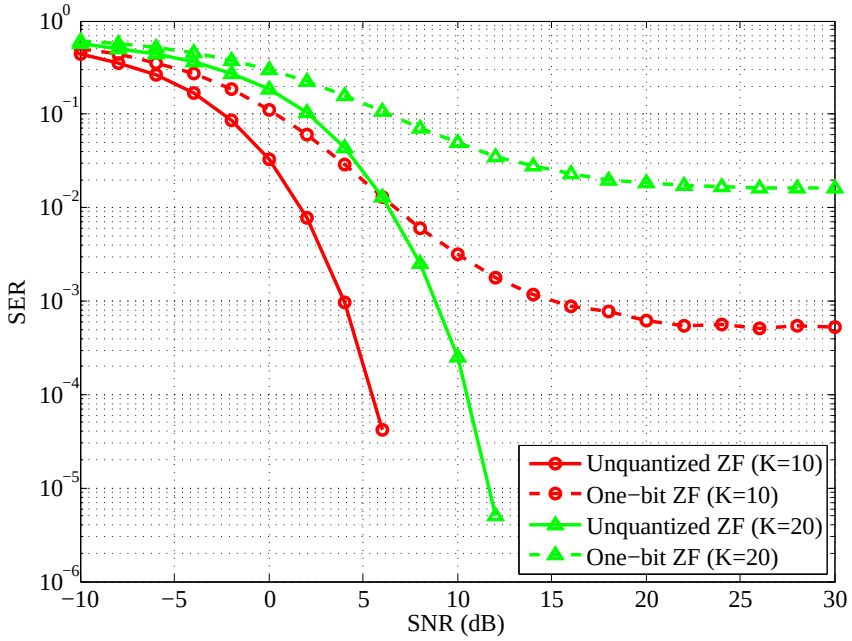


Figure 5.10: SER of unquantized and quantized ZF precoding with hardware mismatch, where $N = 100$.

hardware mismatch exists. It is observed that the SER of ZF precoding is adversely affected by the coarse one-bit quantization. When $N = 100$ and $K = 10$, the SER undergoes an error floor around 10^{-3}. However, when $N = 100$ and $K = 20$, the SER undergoes an error floor around 10^{-2}. Therefore, the smaller value of the ratio $\alpha = \frac{N}{K}$ will lead to a better SER performance.

5.6 Summary

In this chapter, the performance of one-bit ZF precoding is investigated with the hardware mismatch for massive multi-user MIMO systems. A closed-form expression of the achievable rate is derived using the Bussgang theorem and random matrix theorem. In addition, a performance approximation on the achievable rate is also derived. As discussed in the result section, the accuracy of the derived approximation expression is validated and the performance approximation is related to the ratio $\alpha = \frac{N}{K}$ and the statistics of hardware mismatch parameters. Moreover, the ratio α is proved to be a significant factor to obtain higher achievable rate and more accuracy approximation expression in consideration of both hardware mismatch and one-bit ZF precoding.

Chapter 6

Conclusions and Further Work

6.1 Conclusions

In downlink massive MIMO systems, due to the prohibitively high energy consumption and hardware complexity caused by the very large antenna arrays, the low complexity precoding schemes are often used to address this problem. Therefore, this thesis focuses on the design and analysis of the low complexity precoding schemes for massive MIMO systems. In this thesis, we have proposed novel hybrid precoding schemes with a comparable performance as the optimal methods, which can effectively reduce the energy consumption and hardware complexity without obvious performance loss. Moreover, the performance of the low complexity precoding schemes is analyzed .

In Chapter 3, a low complexity hybrid precoding scheme is proposed in massive multi-user MIMO systems with a finite dimensional channel model. The hybrid precoding exploits a high-dimensional RF precoder in the analog domain and a low-dimensional baseband precoder in the digital domain, which employs fewer RF chains to reduce the hardware complexity. In terms of the channel model, the finite dimensional channel is considered, because it can reflect the property of the poor scattering channel environment caused by high path loss at high frequency. Based on this channel model, the proposed method can achieve spectral efficiency

close to that achieved by the optimal ZF precoding. Then a tight upper bound on spectral efficiency achieved by the proposed precoding is derived, which separates the dependence on the array response vectors and the fast fading coefficients matrix and shows the optimality of the proposed hybrid precoding.

In Chapter 4, using successive refinement, an iterative hybrid precoding scheme is proposed with a sub-connected architecture for mmWave MIMO systems. Base on the hybrid precoding scheme with the fully-connected architecture in Chapter 3, the sub-connected architecture is utilized to reduce the number of phase shifters for the hybrid precoding design, which can further reduce the hardware complexity. In mmWave channel model, the iterative algorithm is used to effectively optimize the hybrid precoding matrices. In addition, an upper bound on the spectral efficiency achieved by the proposed hybrid precoding is derived with a closed-form expression and the comparable performance of the spectral efficiency as the hybrid precoding schemes with the fully-connected architecture is obtained by the proposed precoding scheme. Then the energy efficiency and the complexity of the hybrid precoding schemes are both analyzed, which show the sub-connected architecture achieves higher energy efficiency and lower complexity than the fully-connected architecture with the increase of the number of RF chains.

In Chapter 5, the impact of one-bit ZF precoding is studied for massive MIMO system with the uplink and downlink hardware mismatch. Due to the limitation of the hybrid precoding schemes, the low-resolution DACs are employed to reduce the hardware complexity and the power consumption. Because the hardware employed at the base station and the mobile terminals is different, the channel for massive MIMO system with hardware mismatch is considered. Based on these, a analytical approximation expression for the achievable rate is derived with the use of the Bussgang theorem and random matrix theorem. Then a performance approximation of the achievable rate is also derived in the high SNR region, which is related to the ratio of the number of base station antennas and the number of mobile users , and

the statistics of the circuit gains at the base station. The simulation results show the accuracy of the derived approximation expression. In addition, the performance of SER is also simulated to study the influence of one-bit ZF precoding and hardware mismatch.

6.2 Further Work

In order to reduce the hardware complexity deriving from the use of very large antenna arrays in massive MIMO system, antenna selection is an effective method to exploit the diversity gains and achieve complexity reduction. However, the antenna selection and the precoding schemes are two disjointed optimization problems. For future work, the research on a combination of low complexity precoding schemes with the antenna selection can be carried out to achieve the highest benefits from both techniques.

Due to the large number of required DACs in massive MIMO system, the low resolution DACs are used to reduce prohibitively high power consumption. The conventional linear precoding methods with quantization are far from optimal in terms of both BER and achievable rate, thus the design of low complexity precoding schemes with quantization poses an appealing challenge for future research, aiming to achieve near-optimal performance.

Furthermore, most prior work employs the low resolution DACs with a number of RF chains equal to the number of transmitting antennas, which will lead to a great performance loss. To deal with this problem, a mixed DAC architecture can be considered, in which some antennas are equipped with low-resolution DACs, while the rest of the antennas are composed of high-resolution DACs. Hence, a study of the trade-off between the achievable rate and power consumption for the different numbers of high and low resolution DACs is also a very challenging topic that will keep researchers busy in the future.

References

[1] C. Lim, T. Yoo, B. Clerckx, B. Lee, and B. Shim, "Recent trend of multiuser MIMO in LTE-advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 127–135, March 2013.

[2] "Cisco visual networking index: Global mobile data traffic forecast update," *White Paper*, 2017.

[3] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, February 2014.

[5] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. D. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, June 2017.

[6] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[7] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct 2014.

[8] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[9] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653–656, Dec 2014.

[10] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, March 2014.

[11] D. J. Love and R. W. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1102–1110, July 2003.

[12] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, Nov 2005.

[13] X. Zheng, Y. Xie, J. Li, and P. Stoica, "MIMO transmit beamforming under uniform elemental power constraint," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5395–5406, Nov 2007.

[14] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov 2015.

[15] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201–211, Jan 2016.

[16] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, April 2016.

[17] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[18] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, June 2011.

[19] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.

[20] X. Gao, L. Dai, S. Han, C. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, April 2016.

[21] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670–4684, Nov 2017.

[22] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4624–4634, Sept 2017.

[23] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, April 1999.

[24] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct 1998.

[25] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 451–460, March 1999.

[26] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals, Syst., Electron*, Oct 1998, pp. 295–300.

[27] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 60–67, Oct 2004.

[28] N. Chiurtu, B. Rimoldi, and E. Telatar, "On the capacity of multi-antenna Gaussian channels," in *Proc. 2001 IEEE Int. Symp. Inf. Theory*, June 2001, pp. 53–.

[29] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.

[30] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct 2003.

[31] M. Matthaiou, M. R. Mckay, P. J. Smith, and J. A. Nossek, "On the condition number distribution of complex wishart matrices," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1705–1717, June 2010.

[32] W. H. Chin, Z. Fan, and R. Haines, "Emerging technologies and research challenges for 5G wireless networks," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 106–112, April 2014.

[33] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: the next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, September 2014.

[34] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tutorials*, vol. 20, no. 2, pp. 836–869, Secondquarter 2018.

[35] Z. Qingling and J. Li, "Rain attenuation in millimeter wave ranges," in *Proc. 7th Int. Symp. Antennas, Propag. EM Theory*, Oct 2006, pp. 1–4.

[36] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proc. IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug 2011.

[37] F. Khan and Z. Pi, "Mmwave mobile broadband (MMB): Unleashing the 3–300GHz spectrum," in *Proc. 34th IEEE Sarnoff*, May 2011, pp. 1–6.

[38] E. Torkildson, C. Sheldon, U. Madhow, and M. Rodwell, "Millimeter-wave spatial multiplexing in an indoor environment," in *Proc. 2009 IEEE GLOBECOM Workshops*, Nov 2009, pp. 1–6.

[39] P. V. Amadori and C. Masouros, "Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212–2223, June 2015.

[40] J. Du, W. Xu, H. Shen, X. Dong, and C. Zhao, "Quantized hybrid precoding for massive multiuser MIMO with insertion loss," in *Proc. IEEE GLOBECOM*, Dec 2017, pp. 1–6.

[41] T. S. Rappaport, *Wireless Communications: Principles and Practice*. USA:Prentice Hall, 2002.

[42] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, 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 Trans. Antennas Propag.*, vol. 61, no. 4, pp. 1850–1859, April 2013.

[43] M. Samimi, K. Wang, Y. Azar, G. N. Wong, R. Mayzus, H. Zhao, J. K. Schulz, S. Sun, F. Gutierrez, and T. S. Rappaport, "28 GHz angle of arrival and angle of departure analysis for outdoor cellular communications using steerable beam antennas in New York city," in *Proc. IEEE Veh. Technol. Conf*, June 2013, pp. 1–6.

[44] P. F. M. Smulders and L. M. Correia, "Characterisation of propagation in 60 GHz radio channels," *Electron. Commun. Eng. J.*, vol. 9, no. 2, pp. 73–80, April 1997.

[45] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct 2002.

[46] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1924–1934, July 2007.

[47] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[48] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "On the optimality of single-carrier transmission in large-scale antenna systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 276–279, August 2012.

[49] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, February 2013.

[50] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. 40th Asilomar Conf. Signals, Syst. Comput. (ACSSC)*, Oct 2006, pp. 359–363.

[51] Y. Xu, G. Yue, and S. Mao, "User grouping for massive MIMO in FDD systems: New design methods and analysis," *IEEE Access*, vol. 2, pp. 947–959, 2014.

[52] Y. Zeng, R. Zhang, and Z. N. Chen, "Electromagnetic lens-focusing antenna enabled massive MIMO: Performance improvement and cost reduction," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1194–1206, June 2014.

[53] Y.Zeng, R.Zhang, and Z.N.Chen, "Electromagnetic lens-focusing antenna enabled massive MIMO," in *Proc. IEEE ICCC*, Aug 2013, pp. 454–459.

[54] A. H. Mehana and A. Nosratinia, "Diversity of MIMO linear precoding," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1019–1038, Feb 2014.

[55] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug 2005.

[56] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," in *IEEE Global Telecommunications Conference (GLOBECOM 2008)*, Nov 2008, pp. 1–4.

[57] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1305–1316, July 2004.

[58] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[59] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part II: perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, March 2005.

[60] O. E. Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *Proc. 2012 IEEE International Conf. Commun.*, June 2012, pp. 3724–3729.

[61] R. Méndez-Rial, C. Rusu, N. González-Prelcic, and R. W. Heath, "Dictionary-free hybrid precoders and combiners for mmwave MIMO systems," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, June 2015, pp. 151–155.

[62] L. Dai, X. Gao, J. Quan, S. Han, and C. I, "Near-optimal hybrid analog and digital precoding for downlink mmwave massive MIMO systems," in *Proc. 2015 IEEE International Conf. Commun.*, June 2015, pp. 1334–1339.

[63] J. Wang, Z. Lan, C. Pyo, T. Baykas, C. Sum, M. A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, October 2009.

[64] 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 Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, October 2013.

[65] V. V. Mai, J. Kim, S. Jeon, S. W. Choi, B. Seo, and W. Shin, "Degrees of freedom of millimeter wave full-duplex systems with partial CSIT," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1042–1045, May 2016.

[66] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, June 2013.

[67] J. M. Steele, *The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities.* Cambridge University Press, 2004.

[68] L. N. Trefethen and D. Bau, *Numerical linear algebra*. Siam, 1997,vol. 50.

[69] N. Li, Z. Wei, H. Yang, X. Zhang, and D. Yang, "Hybrid precoding for mmwave massive MIMO systems with partially connected structure," *IEEE Access*, vol. 5, pp. 15 142–15 151, 2017.

[70] J. Mirza, B. Ali, S. S. Naqvi, and S. Saleem, "Hybrid precoding via successive refinement for millimeter wave MIMO communication systems," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 991–994, May 2017.

[71] L. Mirsky, "A trace inequality of John von Neumann," *Monatshefte für Mathematik*, vol. 79, no. 4, pp. 303–306, 1973.

[72] X. Wang, Y. Wang, W. Ni, R. Sun, and S. Meng, "Sum rate analysis and power allocation for massive MIMO systems with mismatch channel," *IEEE Access*, vol. 6, pp. 16 997–17 009, 2018.

[73] W. Zhang, H. Ren, C. Pan, M. Chen, R. C. de Lamare, B. Du, and J. Dai, "Large-scale antenna systems with UL/DL hardware mismatch: Achievable rates analysis and calibration," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1216–1229, April 2015.

[74] A. Papoulis and S. Pillai, *Probability, Random Variables, and Stochastic Processes*. USA: McGraw-Hil, 2002.