Beamforming strategies in wireless multi-antenna communication systems

Zhang, Zhiyu

2016

Zhang, Z. (2016). Beamforming strategies in wireless multi-antenna communication systems. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/65922>

<https://doi.org/10.32657/10356/65922>

Beamforming Strategies in Wireless Multi-Antenna Communication Systems

ZHANG ZHIYU

School of Electrical and Electronic Engineering

A thesis submitted to the Nanyang Technological University in
conformity with the requirements for the degree of
Doctor of Philosophy

2016

Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Teh Kah Chan, and my co-supervisor Prof. Li Kwok Hung for their professional guidance, advice, encouragement, and continued support throughout my PhD research. Their assistance on both research as well as on my career has been priceless.

I am appreciative of my current and past colleagues in the Information Center of Excellence for their constructive advice and extensive support in the past several years. My special gratitude goes to Shiyang Han, Celeste Li and Yan Wang for their company, encouragement and constant help. I am also grateful to my friends around the world who made this period of my life enjoyable and unforgettable.

The financial support of the Nanyang Technological University and Research Student Scholarships are also greatly appreciated.

Finally, I would like to express my deepest gratitude to my parents for their spiritual and emotional support throughout my PhD study. Their love is the most precious possession in my life.

Contents

Acknowledgement	i
List of Figures	vii
List of Tables	xi
List of Abbreviations	xiii
Abstract	xvii
1 Introduction	1
1.1 Background Knowledge	3
1.1.1 Wireless Channel Model	4
1.1.2 Transmit Beamforming	6
1.2 Motivation	8
1.3 Main Contributions	9
1.4 Organization of Thesis	11
2 Literature Review	13
2.1 Capacity Region of MIMO Channels	13
2.2 Beamforming in Conventional MIMO Systems	15

Contents

2.2.1	Beamforming in Single Cell Systems	15
2.2.2	Beamforming in Multicell Systems	16
2.3	Beamforming in Massive MIMO Systems	21
2.4	Limited Feedback for Beamforming	25
2.5	Beamforming in Secure Communications	27
3	Beamforming Design in Cooperative Multicell Systems	30
3.1	System Description	31
3.2	Semidefinite Relaxation-Based Algorithm	36
3.3	Numerical Results and Discussions	44
3.4	Conclusions	47
4	Pilot Contamination and Beamforming in Massive MIMO Systems	48
4.1	Pilot Contamination Mitigation	50
4.1.1	System Description	50
4.1.2	Conditions for Pilot Contamination Mitigation	54
4.1.3	Beamforming of Large Antenna Arrays	58
4.1.4	Numerical Results and Discussions	59
4.1.5	Conclusions	64
4.2	3D Beamforming	64
4.2.1	System Description	64
4.2.2	Distribution of User's Position and Downtilt	67
4.2.3	Region-Partition Beamforming	71
4.2.4	Statistic-Based Beamforming	74
4.2.5	Numerical Results and Discussions	75
4.2.6	Conclusions	77

5	Limited Feedback for Beamforming	78
5.1	System Description	79
5.1.1	Downlink Signal Model	79
5.1.2	User Scheduling and Beamforming	82
5.2	Compressive Sensing Limited Feedback	83
5.2.1	Background Knowledge	83
5.2.2	Compressive Sensing for Limited Feedback	85
5.3	Sum-Rate Lower Bound	90
5.4	Numerical Results and Discussions	93
5.5	Conclusions	100
6	Distributed Beamforming for Secure Communications	102
6.1	System Description	103
6.2	A Decomposition Framework Based on SCA	107
6.3	Extension of the Framework	109
6.4	Distributed Optimization Algorithm	114
6.4.1	Analysis of Computational Complexity	119
6.5	Numerical Results and Discussions	120
6.6	Conclusions	125
7	Conclusions and Future Work	126
7.1	Conclusions	126
7.2	Future Work	127
	Appendix A	129
A.1	WMMSE Algorithm	129
A.2	DWSM Algorithm	131

Contents

A.3 RCI Precoding	132
Author's Publications	134
Bibliography	136

List of Figures

1.1	General classification of wireless channels.	5
1.2	System diagram for transmit beamforming.	7
2.1	Cooperative multicell systems.	18
2.2	Demonstration of pilot contamination.	23
2.3	System diagram of limited feedback.	26
2.4	A general wiretap channel.	28
3.1	Clustered cooperative multicell cellular system with cluster size $V = 3$. .	32
3.2	Sum rate per cluster as function of SNR with $V = 3$, $N_t = 4$, $K = 3$. . .	44
3.3	Sum rate per cluster of the proposed algorithm with different number of transmit antennas for $V = 3$ and $K = 3$	45
3.4	Sum rate per cluster with different cluster size for $N_t = K$ and SNR = 10dB.	46
4.1	Multi-cell massive multi-antenna system with 3D spatial channel and pilot contamination effect.	51
4.2	3D channel model with VMF distribution.	52
4.3	Estimation errors with different directional dispersion κ	60
4.4	Sum rates per cell with different directional dispersion and $K = 15$	61

List of Figures

4.5	Beam patterns with non-overlapping AoAs and perfect channel estimation.	62
4.6	Beam patterns with imperfect channel estimation, overlapping AoAs between the desired and interfering UEs, $\kappa = 200$.	62
4.7	Beam patterns with imperfect channel estimation, overlapping AoAs between the desired and interfering UEs, $\kappa = 50$.	63
4.8	Two-dimensional massive multi-antenna network with multipath com- ponents of clustered scatterers.	65
4.9	Demonstration of two-tier user distribution.	68
4.10	CDF of antenna downtilt for the two-tier clustered distribution, $h_{BS} =$ $40m$ and $h_{UT} = 1.5m$.	73
4.11	Spatial capacity comparison among different beamforming strategies.	75
4.12	BER for different beamforming strategies.	76
5.1	Massive MISO downlink model with limited feedback, user scheduling and beamforming.	81
5.2	Channel decomposition.	91
5.3	Sum rates versus receive power with $N = 30$, $B = 8$ and compression ratio $\rho = 0.2$.	94
5.4	Sum rates for different feedback rates with $N = 30$, $\rho = 0.2$ and the receive power is set at 15 dBm.	97
5.5	Sum rate for different antenna numbers with $B = 6$ and the receive power is set at 15 dBm.	98
5.6	Sum rates with different compression ratio ρ for $N = 30$ and $B = 8$.	99
5.7	Sum rates versus different factor α with $N = 30$, $B = 8$ and the receive power is set at 15 dBm.	100
5.8	Average BER for different schemes with $N = 15$, $B = 8$ and $\rho = 0.33$.	101

6.1	Multicell multi-antenna interference wiretap channel model.	104
6.2	Average secrecy sum rates versus number of iterations with $M = 4$, $N = 2$, $N_e = 3$, and $K = 3$	121
6.3	Iteration error versus number of iterations with $M = 4$, $N = 2$, $N_e = 3$, and $K = 3$	122
6.4	Average secrecy sum rates versus P_t with $M = 4$, $N = 2$, $N_e = 3$, and $K = 3$	123
6.5	Average number of iterations versus number of users with $M = 4$, $N = 2$, $N_e = 4$, and $P_t = 5W$	125

List of Tables

6.1	Definition of symbols.	106
6.2	Average CPU time versus number of users.	124

List of Abbreviations

2D	Two-dimensional
3D	Three-dimensional
5G	Fifth generation
AoA	Angle of arrival
BC	Broadcast channel
BER	Bit error rate
BF	Beam-forming
BS	Base station
CCA	Concentric circular array
CDF	Cumulative distribution function
CDI	Channel direction information
CDMA	Code-division multiple access
CoMP	Coordinated multipoint

List of Abbreviations

CQI	Channel quality information
CS	Compressive sensing
CS-LF	Compressive sensing limited feedback
CSI	Channel state information
CSIT	Channel state information at transmitter
DC	Difference convex
DCT	Discrete cosine transform
DFT	Discrete Fourier transform
DoF	Degree of freedom
DPC	Dirty paper coding
DWSM	Distributed weighted sum-rate maximization
FDD	Frequency-division duplex
FDMA	Frequency-division multiple access
HPBW	Half-power beam width
i.i.d.	Independent and identically distributed
ICI	Inter-cell interference
ICIC	Inter-cell interference cancellation
IUI	Inter-user interference

List of Abbreviations

LTE	Long term evolution
MAC	Multiple-access channel
MIMO	Multiple-input multiple-output
MISO	Multiple-input single-output
MMSE	Minimum mean-square error
MRT	Maximum ratio transmission
MSE	Mean-square error
MU	Mobile user
NUQ	Non-uniform quantization
pdf	Probability density function
PSK	Phase-shift keying
QAM	Quadrature amplitude modulation
QCQP	Quadratically constrained quadratic program
RCI	Regularized channel inversion
RHS	Right-hand side
RVQ	Random vector quantization
SCA	Successive convex approximation
SDMA	Space-division multiple access

List of Abbreviations

SDP	Semidefinite programming
SDR	Semidefinite relaxation
SINR	Signal-to-interference-plus-noise ratio
SNR	Signal-to-noise ratio
SUS	Semi-orthogonal user selection
SVD	Singular value decomposition
TDD	Time-division duplex
TDMA	Time-division multiple access
UCA	Uniform circular array
ULA	Uniform linear array
URA	Uniform rectangular array
UT	User terminal
VMF	Von Mises Fisher
VP	Virtual point
VQ	Vector quantization
WMMSE	Weighted minimum mean-square error
WSR	Weighted sum rate
ZF	Zero-forcing

Abstract

In this thesis, we study beamforming strategies for downlink transmission in different wireless communication systems. The major objective is to properly design the beamformer at the transmitter to mitigate the interference at the receivers such that the overall system performance can be enhanced, e.g., the network throughput is maximized. Firstly, the sum-rate optimization problem in a multicell system is considered. The original problem is transformed to a convex semidefinite programming problem by applying the uplink-downlink duality followed by semidefinite relaxation, which can be solved by using convex optimization techniques. Moreover, beamforming strategies in the massive multiple-input multiple-output (MIMO) system of the fifth-generation (5G) network are investigated. Assuming that the base station (BS) is equipped with two-dimensional (2D) massive antenna arrays, we analyze the condition in which the pilot contamination can be mitigated by applying the Von Mises Fisher (VMF) channel model and exploiting the property of the channel covariance matrix. It is shown that the pilot contamination can be mitigated if the angle of arrivals (AoAs) of the interfering users have a nonoverlapping region with the AoA of the desired user. We propose two realistic three-dimensional (3D) beamforming strategies, namely, region-partition beamforming when the BS has perfect channel state information (CSI) and statistic-based beamforming when the BS only knows the statistical knowledge of a two-tier clustered user distribution. The probability

Abstract

density function (pdf) of the two-tier user distribution is derived accordingly. When the full CSI is not available at the BS, a novel limited feedback scheme based on the compressive sensing (CS) technique is proposed to offer accurate channel knowledge. The channel is compressed through linear measurement at the receiver and fed back to the BS. The CS measurement is modelled as a quantization procedure in which quantization noises are generated according to the number of feedback bits. The original channel can be recovered with high accuracy by solving an optimization problem at the BS. Lastly, the secrecy sum-rate maximization problem in Gaussian wiretap channels is considered. To guarantee the secure communication in an interference wiretap channel, we apply a novel decomposition framework to transform the original secrecy sum-rate function to decoupled convex approximation. The decomposition framework is based on Taylor-series expansion and is able to preserve the convexity of the secrecy sum-rate function and linearize the nonconvex part. Following that, we propose a distributed iterative optimization algorithm where the BSs iteratively solve a sequence of convex optimization subproblems in parallel until the optimal beamformers are obtained. Numerical results are presented to validate the theoretical analysis and effectiveness of the proposed approaches.

Chapter 1

Introduction

Wireless communication is one of the fastest growing field in information and communication industry. Cellular communication systems have experienced remarkable development and there are currently several billion users worldwide. Wireless local area networks (WLANs) are becoming the dominant way of communication in many situations. In addition, many other wireless applications such as wifi and wireless sensor networks have emerged as new technologies. People's lifestyle has also changed significantly during the past two decades due to the advances of communication technologies. Particularly, there are huge demands on data traffic as more and more multimedia services are required by various types of wireless devices. Therefore, a wireless communication system which is reliable, fast and resilient is highly desired.

An important system metric of the wireless communication is the network capacity (or throughput), which indicates the amount of information that can be transmitted to the receiver given limited spectrum resources. Several key technologies have been proposed to enhance the system capacity. Multiple-input multiple-output (MIMO) technique has been recognized as one of the most essential breakthroughs in modern wireless commu-

Introduction

nications [1]. Basically, MIMO is a radio antenna technology where multiple antennas are deployed at both the transmitter and receiver to enable a variety of signal paths to carry the data. One of the core ideas behind MIMO systems is the ability to effectively take advantage of multipath propagation fading [1]–[2]. As a result of the application of multiple antennas, MIMO wireless technology is able to considerably improve the network capacity. By increasing the number of transmit and receive antennas, it is possible to enhance the spectral efficiency of the system accordingly. MIMO technology has been standardized for WLANs, the third generation (3G) mobile communication networks, and the fourth generation (4G) mobile communication networks. Special cases of MIMO include multiple-input single-output (MISO) and single-input multiple-output (SIMO), which correspond to the deployment of single antenna at the transmitter and receiver, respectively.

Interference is a major constraint on wireless communication systems. System capacity improvement can be achieved by applying signal processing techniques at the transmitter (precoding) and receiver (decoding). In particular, researchers are more interested in finding effective linear precoding schemes to eliminate or mitigate the interference. Transmit beamforming with multiple antennas is a precoding technique that can realize the transmit diversity [3]–[4]. Each transmit antenna is weighted by a properly designed gain and phase shift before transmission. Physically, the transmitter will match the transmitted signal to the channel for the desired direction and suppress the signal power for the undesired direction. Basically, the beamforming technology requires the knowledge of the channel state information (CSI) as the beamformer is dependent on characteristics of the channel. In a frequency-division duplex (FDD) system, the CSI can be acquired through limited feedback [5], where the index of some codewords from a codebook is fed back to the transmitter with limited number of bits. In

a time-division duplex (TDD) system, the CSI can be obtained via the channel estimation since there exists a reciprocity between the downlink and uplink channel.

Upcoming wireless cellular systems are targeting universal frequency reuse to increase the data rates. This could, however, lead to high levels of co-channel interference (CCI) arising from simultaneous transmissions on the same frequency band by neighboring base stations (BSs). Multicell cooperation is one promising solution to deal with the CCI in future wireless communication systems [6]–[7]. It is also known as coordinated multipoint (CoMP) transmission in 3GPP long term evolution (LTE)-Advanced. BS cooperation entails sharing control signals, users' data, CSI and/or beamformers via high-capacity wired backhaul links to coordinate transmissions. In practice, however, backhaul will be bandwidth-limited due to the practical costs. This restricts the amount of information that can be exchanged among BSs, which in turn determines the level of cooperation. Basically, the BS cooperation can be performed at two levels. Partial cooperation can be realized if BSs only share the CSI of active users. The shared CSI can be used to design individual beamforming strategies to mitigate inter-user interference. This is also known as coordinated beamforming in 3GPP LTE-Advanced [8]. Furthermore, full BS cooperation is fulfilled by sharing the user data as well as the CSI among BSs, and the whole network forms a virtual “network MIMO” system [5]. Full cooperation leads to higher user sum rate at the cost of increased overhead because of the exchange of a greater amount of information among BSs.

1.1 Background Knowledge

Transmit beamforming is a signal processing technique at the physical layer of wireless communication systems. The design of beamformers is highly related to propagation

Introduction

channels. Hence, some basic knowledge of the wireless channel model will be reviewed in the following.

1.1.1 Wireless Channel Model

In practice, the transmitted signals typically undergo wireless fading propagation channels before arriving at the intended receiver. The fading channel can be basically classified into two types: the large-scale and small-scale fading [9]. Path loss and shadowing effect are two phenomenon commonly encountered for the large-scale fading. Path loss is the power attenuation of an electromagnetic wave as it propagates through space. It is proportional to the distance between the transmitter and receiver. Shadowing effect refers to the variation of signal strength because of diffraction caused by obstacles with large dimension such as buildings and hills in a propagation environment.

Moreover, some physical phenomenon such as reflection and scattering result in a multipath effect. After undergoing various paths, the signal will be added either constructively or destructively at the receiver, which leads to small-scale fluctuation in the received signal strength. The fluctuation may come from the variations in the amplitude, the phase, or the angle of arrival (AoA) of the received signal. The small-scale multipath fading can be classified according to different criteria. When the multipath delay spread is smaller than the symbol time, one may think that all the multipath components experience the same magnitude fading, and this is called the flat fading [9]. Different frequency components of the signal experience different magnitude fading, which is called frequency selective fading. If the channel coherent time over which the channel's response remains invariant is larger than the signal's symbol time, it is called slow fading. In this situation, the channel response remains constant for at least one symbol. Otherwise, we call it fast fading if the channel coherent time is smaller than the

Fig. 1.1 General classification of wireless channels.

symbol time. Most of the modern communication systems have slow fading channels. An overview of channel classification is shown in Fig. 1.1.

Channel coefficients can be represented by a mathematical model. In a MIMO system where the transmitter is equipped with M antennas and the receiver has N antennas, the overall channel response can be written as

$$\mathbf{G} = \sqrt{\beta} \mathbf{H}, \quad (1.1)$$

where β is the large-scale fading factor that is given by $\beta = \tau/d^\gamma$, where τ is a log-normal random variable with standard deviation of σ_{shad} , d is the distance between the transmitter and the receiver, and γ is the path-loss exponent. The small-scale fading $\mathbf{H}$ is the superposition of multipath components. In a rich scattering environment where the number of multipath components $P \gg 1$, the channel coefficient can be well represented by the independent and identically distributed (i.i.d.) Rayleigh fading model [9]. The entries of the channel matrix $\mathbf{H}$ are i.i.d. circularly symmetric complex Gaussian

variables whose magnitude and phase follow the Rayleigh distribution and the uniform distribution, respectively. The Rayleigh fading model is most frequently applied to evaluate the performance of MIMO systems due to its analytical tractability.

1.1.2 Transmit Beamforming

A higher demand on multimedia services calls for higher systems capacities. Over the past decades, researchers and engineers have made considerable progress in enhancing the capacity of wireless communication systems. With digital techniques, spectrum efficiency can be improved by applying multiple access techniques. Multiple access means that numerous users simultaneously send information to the BS while receiving information from the BS. There are four types of multiple access techniques with respect to different domains: time-division multiple access (TDMA), frequency-division multiple access (FDMA), code-division multiple access (CDMA), and space-division multiple access (SDMA). TDMA allows users to share the same frequency band by dividing the signal into different time slots. FDMA is the process of dividing one channel into multiple individual bands. CDMA uses the spread spectrum technology, and it allows many users to occupy the same time and frequency allocations over a single channel. In SDMA technique, users can share the wireless channel if they are physically separated from each other.

The ultimate purpose of SDMA is to use independently steered beams at the same frequency band to transmit signals to individual users. This technique has been widely applied in wireless communication systems. Latest technologies like smart antennas or adaptive antenna arrays use digital beamforming to put signals into narrow beams that can be focused on specific users. Digital beamforming combines the antenna technology with the digital technology [10]. The procedure of implementing the digital

Fig. 1.2 System diagram for transmit beamforming.

beamforming is shown in Fig. 1.2. The baseband signal will first go through the digital modulator, where some modulation schemes are applied, e.g., phase-shift keying (PSK) or quadrature amplitude modulation (QAM). Then the digital beamforming is carried out by weighting these digital signals, which includes adjusting their amplitudes and phases such that the desired beam can be formed when adding them together.

The beamformer or the beamforming matrix $\mathbf{W}_k$ needs to be properly designed to satisfy the specific system requirement. Adaptive beamforming technology, where the antenna has the ability to adapt to the wireless environment, is also referred to as smart antenna technology [10]. In this case, optimization algorithms are required such that the

Introduction

beamformer can be updated adaptively. Since the beamforming technology is dependent on channel characteristics, full or partial knowledge of the instantaneous or statistical CSI is needed at the transmitter. Note that the term beamforming is from the area of signal processing technique. Therefore, we focus on the signal processing and algorithm design aspects of wireless communication systems. The hardware implementation and the matter of antenna arrays will not be discussed. Furthermore, the beamforming technique can be discussed from two distinct aspects in the literature. Interference alignment (IA) technique is at the level of signal structure, i.e., how to construct the signal at the transmitter such that the interference can be made to overlap with each other at the receiver. On the other hand, beamforming can be viewed as a spatial matched filtering technique. Proper weights are applied to the signal by using some mapping (beamforming) matrices. Thus, they are at different levels and are two different research directions. In this thesis, we focus on the design of beamformers and the comparison among various beamforming methods. IA technique will not be covered by the thesis although it may improve the throughput performance as well.

1.2 Motivation

The focus of this thesis is on the beamforming/precoding design for the downlink transmission in various wireless multi-antenna systems, i.e., cooperative multicell system, massive MIMO system, and Gaussian MIMO wiretap channel. CoMP transmission technique, massive MIMO technique and physical layer security are all the important issues in the current and future wireless communications. The design problems in these systems can usually be formulated as some optimization problems. Although some engineering optimization techniques have been applied to obtain (sub)optimal

solutions in the literature of CoMP system, the current methods are either centralized optimization requiring joint processing across the network or distributed approaches with relatively high complexity. More efficient algorithms with low complexity that are easier to implement need to be further explored. Particularly, distributed algorithm that can achieve almost the same performance with that of centralized schemes but requires local channel knowledge and limited BS coordination is preferred. This motivates us to investigate the distributed convex optimization technique for the CoMP system. Furthermore, the research on beamforming in massive MIMO (full-dimension MIMO) system for the next generation communication system is in the early stage in academia. The conventional MIMO system is considered in a two-dimensional (2D) plane where the elevation beamforming gain can not be achieved. More research efforts are required to explore the three-dimensional (3D) beamforming technique to fully utilize the multiplexing gain in the space. In the literature of physical layer security, many existing works consider a simple scenario with single transmitter and single eavesdropper. The problem of secure communication in the realistic multicell system is more complicated and difficult to handle. Thus, it motivates us to explore distributed solutions based on convex optimization to ensure the secure communication in the multicell scenario.

1.3 Main Contributions

In this thesis, we mainly consider the transmit beamforming design in different wireless communication systems, i.e., cooperative multicell systems, massive MIMO systems, and wiretap MIMO channels. The major contributions of the thesis are summarized as follows:

Introduction

- 1) We investigate the sum-rate maximization in a clustered cooperative multicell system. As the objective function needs to be jointly optimized, we adopt sub-optimal uplink-downlink duality to decouple the original problem into independent problems that could be solved locally. Then we use the semidefinite relaxation (SDR) to relax and transform the original non-convex optimization problem into a convex semidefinite programming (SDP), which can be efficiently solved in polynomial time.
- 2) We analyze the pilot contamination problem in a massive MIMO system where the antennas of BS are in the 2D uniform rectangular array (URA) topology. The 3D Von Mises Fisher (VMF) distribution is applied to model the AoA of the multipath signals. Moreover, we show that the pilot contamination can be mitigated if the AoA pairs of the desired user do not overlap with those of interfering users.
- 3) We derive the statistical distribution function of users' geographic position in a two-tier clustered distribution model, which will be used in the design of 3D beamforming. Two practical 3D beamforming strategies are proposed. In a region-partition 3D beamforming strategy, users in a certain partition are assigned a particular downtilt. We also propose a passive statistic-based 3D beamforming strategy when the BS only has users' statistical location knowledge of the two-tier distribution.
- 4) We propose a novel efficient limited feedback scheme based on the compressive sensing (CS) technique for multi-antenna systems. The CSI of downlink channels is compressed through linear measurement and fed back to the BS. The CS measurement is modelled as a quantization procedure that generates quantization errors. At the BS, the channel can be recovered by solving an l_1 -norm minimization problem.
- 5) We consider the secrecy sum-rate maximization in an interference wiretap channel model. We show that a novel decomposition framework based on the Taylor-series

expansion can be extended to a real-valued function of complex matrices variables. Based on the framework, the original sum-rate function is decomposed into a convex approximation by preserving its inherent convexity and linearizing the non-convex part. We propose an iterative distributed algorithm to find the optimal transmit covariance matrices. All the BSs solve in parallel a sequence of decoupled convex optimization problems to obtain the optimal transmit covariance matrices in the current iteration. Then the covariance matrices are updated in a gradient-based manner. The computational complexity of the proposed algorithm is also analyzed.

1.4 Organization of Thesis

This thesis is divided into seven chapters. In Chapter 1, we give a brief introduction on wireless channel model and basic knowledge of beamforming. We also discuss the objective and motivation of our research work. Lastly, the major contributions of the thesis are summarized.

In Chapter 2, we conduct detailed literature review on the beamforming technique in various wireless communication systems.

In Chapter 3, we consider the suboptimal beamforming to maximize the sum rate in cooperative multicell systems. We decouple the original problem into independent subproblems through the uplink-downlink duality. SDR technique is applied to transform the non-convex subproblem into an SDP which can be solved efficiently.

In Chapter 4, we consider the pilot contamination problem and 3D beamforming strategies in massive MIMO systems. We first analyze the condition for mitigating the pilot contamination by applying the VMF distribution to model the AoA of multipath components. Then we propose two practical 3D beamforming approaches.

Introduction

In Chapter 5, we consider the limited feedback problem in multi-antenna systems. We propose a novel limited feedback method based on the theory of compressive sensing technique. The proposed method is compared with some existing limited feedback methods.

In Chapter 6, we study the problem of finding the optimal transmit beamformers at the BSs to maximize the secrecy sum rate in an interference wiretap channel. We propose a parallel and distributed optimization algorithm based on a novel decomposition framework that fully exploits the convexity of the original problem.

In Chapter 7, we summarize our research works and provide some ideas and research directions for the future work.

Chapter 2

Literature Review

2.1 Capacity Region of MIMO Channels

In wireless communication systems, a key performance metric is the capacity or throughput that indicates how much information can be transmitted given limited spectral resources. Thus, it is important to analyze the capacity region. The well-known Shannon-Hartley theorem provides a fundamental formula for the channel capacity. The channel capacity of an additive white Gaussian noise (AWGN) channel with bandwidth B Hz, average received power $\bar{P}$ and white noise power spectral density N_0 is given by [9]

$$C = B \times \log_2 \left(1 + \frac{\bar{P}}{BN_0} \right) \text{ bits/s.} \quad (2.1)$$

The term $\bar{P}/BN_0$ is referred to as signal-to-noise ratio (SNR). Information theoretical results of the capacity region of MIMO broadcast channels (BC) have been well characterized in [11]–[15]. In [11], the achievable throughput of a generally nondegraded Gaussian BC where the transmitter has multiple antennas and the receivers have one antenna each is investigated. The channel knowledge is assumed to be perfectly known

Literature Review

to all terminals. In a two-user situation, when the transmitter applies the successive dirty-paper coding (DPC), the optimal throughput can be achieved. In the general case, a suboptimal precoding, i.e., zero-forcing (ZF) DPC strategy is shown to achieve asymptotically optimal throughput for high SNR if the channel matrix has a full row rank.

In [12], the authors consider a multiuser MIMO Gaussian BC in which the transmitter and receiver have multiple antennas. A duality between the dirty paper (DP) region of the MIMO BC and the capacity region of the MIMO multiple access channels (MAC) is established. It has been shown that the DP region is equivalent to the capacity region of the dual MIMO MAC, by assuming that the transmitter has the same sum power constraint as the MIMO BC. This duality is established by showing that rates achievable in the dual MIMO MAC with some power constraints are also achievable in the MIMO BC. Moreover, this duality can be easily shown to hold for fading MIMO Gaussian BCs and MACs. It is very difficult to find the capacity region of the MIMO BC due to its nondegraded nature. However, according to the duality, the DP achievable region is shown to achieve the sum rate capacity of the MIMO BC.

The capacity region of the Gaussian MIMO BC is well characterized in [13]. The authors have also shown that the obtained region coincides with the DPC rate region for different types of input constraints, such as the average total power constraint and the transmit covariance matrix constraint. More generally, the results hold for any input constraint where the input covariance matrix lies in a set of positive semidefinite matrices. The sum capacity of MIMO interference channel where both the transmitter and receiver have perfect CSI knowledge is studied in [14]. A procedure based on the simultaneous superposition code approach is developed to compute a lower bound on the sum capacity.

2.2 Beamforming in Conventional MIMO Systems

The achievable ergodic rates of the MIMO BC where ZF beamforming is applied and the imperfect CSI is available at the transmitter are studied in [15]. The CSI is acquired at the receivers via downlink training and it is fed back to the transmitter with limited bit rate. Both unquantized and quantized channel state feedback schemes are analyzed and compared under different situations. It is shown that if the feedback link can be properly designed, significant downlink throughput can be achieved by applying simple and efficient channel state feedback strategy.

2.2 Beamforming in Conventional MIMO Systems

2.2.1 Beamforming in Single Cell Systems

A typical scenario of the downlink communication in one cell is the multiuser MIMO (MU-MIMO), which refers to the simultaneous communication with a number of users. The spatial degrees of freedom (DoF) provided by multiple antennas at the BS can be exploited to simultaneously serve multiple users equipped with multiple antennas, over the same time-frequency resource. It has been shown that the capacity region of the MIMO BC coincides with the DPC rate region. Hence, the DPC technique is considered as the theoretical optimal precoding technique [16]. However, DPC is difficult to be carried out in practical wireless systems due to its highly nonlinear nature and the computational complexity. A lot of research efforts have been made to find the suboptimal linear precoding schemes to deal with the inter-user interference and to get close to the DPC capacity region. In [17], a leakage-based beamforming approach is proposed to optimize the signal-to-leakage-and-noise ratio (SLNR) for designing multiuser transmit beamforming. The leakage-based criterion results in a decoupled optimization problem which is easy to solve. There is no restriction on the

Literature Review

number of transmit antennas of the proposed design scheme. The authors investigate the leakage-based beamforming in MIMO systems where Alamouti coding is applied. The beamforming scheme is also studied for the case of multiple streams per user. Lastly, channel estimation error is considered and the corresponding system performance is examined.

In [18], an iterative optimization algorithm based on the classic Arimoto-Blahut algorithm is proposed for the linear beamforming approach in MIMO BC. Each iteration involves updating the test conditional probabilities and beamforming matrices by solving some convex subproblems. By using the proposed algorithm together with uplink-downlink duality, weighted sum rate maximization, weighted rate balancing, and proportional fair rate allocation in MIMO BC are investigated. In [19], the authors have shown that achieving the maximum sum rate is equivalent to minimizing the product of mean-square error (MSE) matrix determinants. An iterative algorithm of low computational complexity is proposed to solve the corresponding minimization problem. In [20], linear transceiver design for weighted sum-rate maximization (WSRM) and per-user rate balancing are studied. Iterative algorithms based on the MSE duality and a heuristic alternating optimization framework are proposed. In [21], the concept, principles, and analysis of maximum ratio transmission (MRT) beamforming for MIMO systems are presented. Moreover, there have been some works on the investigation of ZF beamforming that can remove the multiuser interference in [22], [23].

2.2.2 Beamforming in Multicell Systems

The performance of wireless MIMO systems in a multicell situation is degraded by inter-cell interference (ICI), which is caused by simultaneous transmission when the universal frequency reuse is applied. The cell-edge user will experience more severe performance

2.2 Beamforming in Conventional MIMO Systems

degradation. One of the key technologies to combat the effect of ICI is to exploit the coordination among BSs. It is referred to as the CoMP transmission/reception in the 4G LTE [24]. The basic idea behind the CoMP is to make use of the global resources to coordinate the communication so that the overall system performance can be improved dramatically. There are two forms of coordination based on the level of signal exchange among BSs. In the joint processing (or network MIMO), both the user data and the CSI are shared among multiple geographically separated BSs. In this way, data is transmitted to the user simultaneously from a number of BSs. The received signal quality and strength can be improved accordingly. It may also efficiently cancel the interference from transmissions that are intended for other users. In coordinated beamforming, only the CSI and some control signaling are exchanged among BSs. The data to one user is transmitted from its serving BS, but the beamforming strategies of various BSs are coordinated to strike a good balance between maximizing the desired signal power and minimizing the unwanted interference that may be generated. Joint processing technique can achieve higher sum rate, however, at the cost of a higher demand onto the backhaul links since more information needs to be exchanged among BSs. The system diagram of a general cooperative multicell system is shown in Fig. 2.1.

A comprehensive study on the theory and key techniques of multi-cell MIMO cooperative networks is conducted in [6]. Unlike conventional MIMO systems where the cost of multi-antenna processing lies in the extra processing at individual devices, the cost of multi-cell cooperative MIMO lies in the additional exchange of information between BSs or between the BSs and the central controller. Various transmission and coding techniques to approach the theoretical capacity are reviewed, e.g., coordinated beamforming in this section. Moreover, some other practical issues, such as scalability, synchronization, and channel estimation are addressed. It has been pointed out that

Fig. 2.1 Cooperative multicell systems.

distributed optimization approach is preferred as most of the processing can be done locally, with limited signaling exchange between nearby BSs. Some challenges that need to be addressed are listed, such as coordinated power control, rate allocation across the spectrum, and trade-off between cooperation and information exchange.

Cooperative multicell precoding in a network MISO system is studied in [7]. The outer boundary of the achievable rate region is characterized when the BSs have either instantaneous or statistical CSI. It is shown that beamforming vectors that can attain the boundary depend only on local CSI, and can be interpreted as linear combinations of MRT and ZF beamformers. The authors propose distributed beamforming that only requires local CSI and processing based on a novel framework of distributed virtual

2.2 Beamforming in Conventional MIMO Systems

signal-to-interference-plus-noise ratio (SINR) beamforming that achieves the optimal multiplexing gain.

In [8], the uplink-downlink duality is applied by using the Lagrangian theory for two different design criteria: minimizing the total weighted transmitted power and minimizing the maximum antenna power subject to SINR constraints. An iterative algorithm is proposed to find the global optimal coordinated beamformers. In [25]–[26], the theory of uplink-downlink duality is applied to transform the downlink beamforming problem into an equivalent uplink problem which is easier to handle. The global optimal of the downlink beamforming problem can be equivalently obtained from solving a dual uplink problem.

In [27], high SINR approximation is applied to decouple the WSR maximization problem which involves joint optimization across all the BSs into a distributed weighted sum-rate (WSR) problem as a function of local CSI. The decoupled problems are solved iteratively by using a zero-gradient based algorithm which converges to a local optimal point. In [28], a distributed method for WSR maximization is developed through the BS cooperation. The authors propose an algorithm based on primal decomposition and subgradient methods, where the original nonconvex problem is split into a master problem and a number of subproblems. A novel sequential convex approximation strategy and a subgradient method are used to solve the nonconvex master problem. A descent algorithm based on second-order cone programming (SOCP) and a geometric programming are adopted for the subproblems. As a result, a good trade-off between the performance improvement and the implementation-level simplicity is achieved.

In [29], the authors propose both nonlinear and linear precoding methods to tackle the out-of-cell and intra-cell interference in a decentralized manner. A robust precoding scheme in the presence of CSI uncertainty is also proposed. In [30], the trans-

Literature Review

mit beamforming problem is recast as a virtual linear minimum mean-square error (LMMSE) problem. The authors adopt the Kalman smoothing framework to obtain a forward-backward and distributed beamforming algorithm for a simple linear array model multicell networks. The authors also present a distributed downlink beamforming algorithm for a hexagonal cellular array model based on the theory of factor graph and the sum-product algorithm. Distributed optimization schemes of multicell system are studied in [31], where there is only limited intercell coordination. A hierarchical iterative algorithm to jointly optimize downlink beamforming and intercell power allocation in a distributed manner is proposed.

In [32], the authors propose interference-aware coordinated beamforming algorithms for a two-cell scenario to obtain the optimal and suboptimal transmit/receive beamforming vectors. In [33], a novel iterative distributed algorithm is proposed to jointly optimize the beamforming vectors across a cluster of coordinated BSs and resource slots. In [34], game theory is applied to deal with the coordinated beamforming problem in multi-cell systems such that the transmit power of each BS can be minimized subject to some SINR constraints. Energy efficient joint power allocation and beamforming for coordinated multicell downlink systems is investigated in [35]. The original optimization problem is transformed into an equivalent subtractive form problem and an efficient optimization algorithm is developed to find the optimal solution.

Engineering optimization methods have been recognized as powerful techniques in helping with addressing the afore-mentioned challenges. The convex optimization framework has emerged as a useful tool for a wide variety of applications in wireless communications and signal processing [36]. Over the last few years, convex optimization has been recognized as the most useful technique for algorithm design and analysis of wireless communication systems. Semidefinite programming technique, which is a

2.3 Beamforming in Massive MIMO Systems

broad class of convex optimization problems, has been applied to solve the optimization problem in MIMO systems [37]–[39]. Based on the approximation, transformation or relaxation, the original optimization problem can be transformed to an equivalent problem in the form of SDP, which can be solved efficiently by using existing convex optimization algorithm. The optimal downlink beamforming to minimize the transmission power subject to some SINR constraints is considered in [40]. The optimization problem can be formulated as a separable homogeneous quadratically constrained quadratic program (QCQP). The authors studied rank-constrained solutions of general separable SDP and proposed rank reduction procedures to achieve a lower rank solution. In [41], distributed optimization method for obtaining the worst-case robust beamformer to minimize the weighted sum power of BSs is investigated. The original non-convex centralized problem is approximated through SDR. Then a distributed robust coordinated beamforming algorithm is proposed based on a distributed convex optimization technique called alternating direction method of multipliers (ADMM).

2.3 Beamforming in Massive MIMO Systems

Recently, the massive MIMO system, where the number of antennas at the BS far exceeds the number of users in one cell, has drawn great research attentions [42]–[45]. It is a key technology for the fifth generation (5G) wireless communication system. Since the number of antennas at the BS is typically assumed to be much larger than the number of users, a large number of degrees of freedom are available and can be used to shape the transmitted signals to deal with the interference. Massive MIMO has the potential of bringing orders-of-magnitude improvements in both spectral and energy efficiency. Extra antennas lead to huge improvements in network throughput

Literature Review

and radiated energy efficiency by focusing energy into smaller regions of space with very narrow beams. Exploiting the multi-dimensional properties of wireless channels with hundreds of antenna elements, massive MIMO systems can significantly enhance the system capacity and link reliability of wireless communication by applying simple transmit and receive signal processing strategies because of the quasi-orthogonal nature of the channels between the BS and the active users. The deployment of massive antennas results in a favorable propagation condition, which distinguishes the massive MIMO networks from the conventional small-scale multi-antenna network [43]. It has been shown that some simple precoding and detection schemes, such as the MRT and ZF beamformer and the matched filter (MF) detector, can achieve close-to-optimal performance for the massive MIMO system [46]. A multiuser massive MIMO BC under the constraint that each antenna emits signals having constant envelope is considered in [47]. Massive MIMO system has some other benefits. It can be built with inexpensive and low-power components. Massive MIMO technique enables a significant reduction of latency on the air interface. Furthermore, massive MIMO increases the robustness against unintended interference and the intentional jamming [43].

To keep the signaling overhead for CSI acquisition in massive MIMO systems manageable, TDD mode is preferred since the amount of CSI feedback grows with the number of BS antennas in FDD systems. Pilot-aided channel estimation method is used in TDD networks to obtain the channel knowledge as there exists a reciprocity between the downlink and uplink channels. A major impairment in massive MIMO systems is the so-called pilot contamination [45],[48]. Pilot contamination effect is caused by the reuse of the same pilot signal from cell to cell. As is shown in Fig. 2.2, the three users in different cells are assigned the same pilot signal. During the channel estimation procedure, the BS receives the same pilot signals from different users and cannot

Fig. 2.2 Demonstration of pilot contamination.

distinguish them. Hence the estimation of the channel of its own user is deteriorated. This reuse of pilot signal is unavoidable as, for a given pilot sequence length, the number of linear independent pilot signals is limited.

Recently, several efficient techniques have been proposed to overcome pilot contamination. In [49], the authors proposed an algorithm with polynomial complexity based on singular value decomposition (SVD) to mitigate pilot contamination without the need for coordination among cells. In [50], a novel approach based on coordination among cells during channel estimation is proposed. The coordination exploits the additional second-order statistical information about the user channels. The pilot contamination effect can be eliminated under certain conditions on the channel covariance. In [51], the low-rank property of users' channel covariance matrices that can be modeled by the an-

Literature Review

gular spread of multipath is characterized to realize the so-called pilot decontamination in both co-located and distributed massive MIMO systems.

Furthermore, as the antenna arrays in massive MIMO systems are to be deployed in the two-dimensional (2D) plane or in the three-dimensional (3D) space, the beamforming in the vertical dimension can, therefore, be fully explored. Full-dimensional MIMO or 3D beamforming is recognized as a promising technique in massive MIMO systems to enhance the cellular performance by deploying antenna arrays in both horizontal and vertical dimensions [52].

Recently, 3D beamforming has attracted some keen interests [53]–[59]. It combines the horizontal beam pattern adaptation, which refers to the conventional beamforming design, with the vertical beam pattern adaptation. An overview of 3D beamforming is provided in [53], in which the technological background, basic principles and several realization options for 3D beamforming are discussed. Concepts for 3D beamforming and the potential of its combination with interference coordination are also addressed. In [54], the probability density function (pdf) of the vertical angle distribution for 3D channel model is derived, based on which the authors provide the optimal vertical beamforming solutions for the active and passive antenna systems with MRT beamforming. In [55], a novel coordinated 3D beamforming strategy is proposed, where inter-cell interference cancellation (ICIC) technique is used in the horizontal plane and the downtilt of the BS antenna pattern is coordinatively adapted to the locations of the intended users in the vertical plane. In [56], a fractional frequency reuse-based dual decomposition method is proposed to jointly optimize the resource block assignment, power allocation and downtilts adjustment. Some works on the simulation and field trial of 3D beamforming can be found in [57], [58]. In [59], a novel framework for coordinated user-specific tilting is proposed to enhance the throughput in MISO interference channels.

2.4 Limited Feedback for Beamforming

In order to adapt the transmit beamforming strategy to the dynamic wireless propagation environment, the knowledge of CSI is required at the BS. Since most of the existing cellular systems are operating in the FDD mode, limited feedback technique is usually applied to obtain and exploit the CSI at the transmitter [60]. It uses finite rate feedback channels to convey quantized observations of the channel from the user to the BS. A comprehensive overview of limited feedback in wireless communication systems is presented in [5]. Classical limited feedback schemes that are mainly based on the codebook design include covariance quantization, Grassmannian line packings, vector quantization (VQ) and random vector quantization (RVQ). The idea of covariance quantization is that given a codebook of possible covariance matrices, which is known to both the transmitter and receiver, the receiver will search for the codebook index that solves some functions of the channel, e.g., the ergodic rate maximization problem. Then the index corresponding to the optimal covariance matrix is fed back to the transmitter. Alternatively, the receiver can design and quantize the beamforming vector and send the designed vector back to the transmitter. The main idea is to restrict the beamforming vector to lie in a codebook. The entries of the codebook are chosen to maximize the minimum distance which is a function of the minimum angular separation between codebook lines. This problem is often referred to as the Grassmannian line packing problem. Another approach to construct the codebooks is to apply VQ techniques, in which the codebook is designed by formulating a distortion function and iteratively minimizing this distortion function to obtain local optimal solutions. Lastly, unlike the Grassmannian and VQ limited feedback approaches where the codebook is assumed to be fixed, the codebook vector in RVQ method is generated independently according

Fig. 2.3 System diagram of limited feedback.

to the stationary distribution of the channel. A general limited feedback procedure is demonstrated in Fig. 2.3.

In [61], the authors propose Grassmannian differential feedback strategy to reduce feedback overhead by exploiting the channel's correlation and Grassmannian structure. The RVQ method is fully investigated in [62], where the beamforming vector $\mathbf{w}$ is chosen by maximizing the SNR as follows:

$$\mathbf{w} = \arg \max_{\mathbf{f} \in \mathcal{F}} |\mathbf{h}\mathbf{f}|^2, \quad (2.2)$$

where $\mathbf{h}$ is the channel vector and the codebook $\mathcal{F}$ is randomly generated by selecting each column vector independently from the uniform distribution on the complex unit sphere. In [63], the SINR distributions and the sum rate under quantized CSI are derived. It is also shown that the CQI feedback with respect to the SINR is required to achieve both of the multiplexing and multiuser diversity gains. In [64], a limited feedback strategy for multicell beamforming is proposed, where beamforming vectors

are designed using a generalized eigenvector method to maximize the sum-rate. A feedback-bit allocation strategy that partitions the available bits between the desired and interfering channels is also developed. In [65], a new feedback approach based on predictive vector quantization is proposed to reduce feedback overhead by exploiting temporal correlation in the channels. In [66], a new partial CSI acquisition algorithm that models and tracks the variations between the dominant subspaces of channels at adjacent time instants is developed. Two partial CSI models are considered in [67], i.e., the shape feedback model and the limited feedback model. The authors derive the sum rate bounds for the ZF-DPC systems and channel inversion systems under these two feedback models. The limited feedback problem in distributed antenna systems, massive MIMO systems and relay networks is studied in [68], [69] and [70], respectively.

2.5 Beamforming in Secure Communications

Security and privacy are of much concern for future wireless communication technologies, due to the openness nature of the wireless propagation channels. The wireless network should have the ability to defend against malicious attacks and maintain services at an acceptable level. Conventionally, security is taken as a higher-layer issue to be solved by using cryptographic methods. However, cryptography may not be suitable for large-scale and dynamic networks. In the physical layer, the major form of attack is the eavesdropping in a wiretap channel. A general wiretap channel model is shown in Fig. 2.4. A transmitter sends confidential messages to a legitimate receiver. However, some unauthorized receivers can eavesdrop (or overhear) on the channel between the transmitter and legitimate receiver. Due to the eavesdropping, secrecy rate is introduced as a performance metric to indicate the secrecy of transmission [71]. The secrecy rate

Fig. 2.4 A general wiretap channel.

is defined as the difference between the capacity of main channel and the capacity of wiretap channel, which is given by

$$R_s = \log \left| \mathbf{I} + \mathbf{H}_M \mathbf{K} \mathbf{H}_M^\dagger \right| - \log \left| \mathbf{I} + \mathbf{H}_E \mathbf{K} \mathbf{H}_E^\dagger \right|, \quad (2.3)$$

where $\mathbf{H}_M$ and $\mathbf{H}_E$ denote the main channel and wiretap channel, respectively. The covariance matrix of the transmitted signal is denoted as $\mathbf{K}$. To maximize the secrecy rate, the transmitter needs to ensure that the eavesdropper cannot overhear the confidential message. Physical layer security, which exploits the physical characteristics of wireless channels, has emerged as a promising technique to protect the wireless communications against eavesdropping attacks [72]. It can be implemented by signal processing, communications and coding techniques. Transmit beamforming technique exploiting the channel knowledge has been shown to be an effective way to ensure secure and resilient transmission of confidential messages in wiretap channels and it has been widely used in the literature.

2.5 Beamforming in Secure Communications

One of the major techniques for physical layer security in the literature is the use of interference or artificial noise to confuse the eavesdropper [73], [74]. The transmitter can simultaneously transmit the information-bearing signal to the intended receiver and artificial noise to the eavesdroppers. Alternatively, the artificial noise can be transmitted by a friendly jammer or relay node [75]. In [76], cooperative transmission schemes are proposed from a game-theoretic perspective by taking transmitters as players and secrecy rates as utility functions, and an operating point is determined to balance network performance and fairness. In [77], iterative algorithms based on Taylor-series expansion are proposed to solve the power minimization and secrecy rate maximization problems for a MIMO wiretap channel with a single receiver and a single eavesdropper. In [78], SDP techniques are applied to solve the non-convex semiinfinite optimization problem maximizing the rate of the secondary link while avoiding harmful interference to the primary user in a multi-antenna cognitive radio network. In [79], a precoding scheme with full CSI and a postcoding scheme without full CSI are proposed to achieve secrecy rates. The beamformer design problem for the multiantenna amplify-and-forward relay in the relay-eavesdropper network is studied in [80]. In [81], a hybrid cooperative beamforming and jamming scheme based on semidefinite relaxation and randomization techniques is proposed to enhance the physical layer security of a two-way relay network. A regularized channel inversion (RCI) precoding approach is proposed in [82]. In [83], the authors investigate the design of the transmit power and the BS deployment density to make communications robust against malicious eavesdropping.

Chapter 3

Beamforming Design in Cooperative Multicell Systems

BS cooperation has been shown to be an effective way to dramatically improve the sum rate by enhancing the desired signal power and mitigating the interference [6]. The cooperation can be performed at two levels. Firstly, BSs are coordinated only at the beamforming level and it is termed as coordinate beamforming. Secondly, BS cooperation takes place at both signal and beamforming levels. By sharing the user data among BSs, the entire network forms a network MIMO system, in which users are served by cooperative BSs simultaneously rather than by a single one. It is well known that the latter cooperation approach achieves a better performance gain, however, with the cost of a much higher signal exchange overhead.

Most of the existing works consider global BS cooperation which leads to a heavy backhaul burden as there is a large amount of data/control traffic among BSs. Furthermore, the whole network is susceptible to central processor breakdown due to its highly centralized configuration. To deal with this problem, one can divide the network

into some clusters [6]. In [84], block diagonalization technique is used to coordinate the transmissions across multiple BSs in the same cluster. This strategy consists of a full intra-cluster coordination and a limited inter-cluster coordination. In [85], a linear precoding technique called soft interference nulling is proposed assuming that the coordination clusters for different users may overlap. In [86], an efficient algorithm is proposed to jointly design the BS clustering and linear beamformers in a heterogeneous network. In [87], BS clustering is considered where the coordination only happens within one cluster. Block diagonalization technique is applied to remove the intra-cluster interference while the inter-cluster interference remains.

In this chapter, we consider a clustered multicell system where the whole network is divided into some clusters and BSs within the same cluster work cooperatively to design their beamforming strategies. We assume a multicell MISO case with single active user per cell and uniform transmit power allocation is applied. As the objective function needs explicitly joint optimization, we adopt sub-optimal uplink-downlink duality to decouple the original problem into independent problems that could be solved locally. Then we use the SDR to relax and transform the original nonconvex optimization problem into a convex SDP that can be solved by existing algorithms.

3.1 System Description

We consider a downlink multicell cellular network with M cells where each BS is equipped with N_t transmit antennas, and each mobile user (MU) has a single receive antenna¹. The whole network is divided into some clusters. Each cluster contains V cells. There are I adjacent clusters that generate interference. The system model is illustrated

¹Note that all the notations are only applicable to this chapter.

Fig. 3.1 Clustered cooperative multicell cellular system with cluster size $V = 3$.

in Fig. 3.1. In each cluster there are K active users. There are two types of users in the network, namely, cluster-interior users and cluster-edge users. They are defined based on the distance between a user and its serving BS. Denote D as the distance between two adjacent BSs and d_k as the distance from the k th user to its own serving BS. Cluster-edge user is defined as: 1) $d_k \geq D/2$; 2) the user is close to the common edge of two adjacent clusters (not two adjacent cells in one cluster). Otherwise, the user is regarded as a cluster-interior user. We assume that all the users within the same cell are orthogonal to each other through intra-cell TDMA. Hence, the system can be viewed as a multicell interference channel model. We also assume that $K = V$, which means in a particular time slot there is one active user in each cell. On the other hand, as we apply universal frequency reuse where all the BSs in the system operate on the same frequency band, cluster-edge users may suffer severe interference from both neighboring clusters and their own cluster.

The received signal of user k in cluster c is given by

$$y_{k,c} = \sum_{j=1}^V \mathbf{h}_{jk,c}^H \mathbf{x}_{j,c} + \underbrace{\sum_{\hat{c}=1, \hat{c} \neq c}^I \sum_{i=1}^V \mathbf{h}_{ik,\hat{c} \rightarrow c}^H \mathbf{x}_{i,\hat{c}}}_{\text{inter-cluster interference}} + n_k \quad (3.1)$$

where $\mathbf{h}_{jk,c} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_{jk})$ represents the channel from BS $_j$ to MU $_k$ in cluster c , and $\mathbf{h}_{ik,\hat{c} \rightarrow c}$ represents the interfering channel from BS $_i$ in adjacent cluster $\hat{c}$ to MU $_k$ in cluster c . The covariance matrix $\mathbf{Q}_{jk}$ is defined as $\mathbf{Q}_{jk} \triangleq \mathbb{E}\{\mathbf{h}_{jk,c} \mathbf{h}_{jk,c}^H\}$. The propagation channel is assumed to be slowly-varying, flat and Rayleigh fading. $\mathbf{x}_{j,c} \in \mathbb{C}^{N_t \times 1}$ is the transmit signal at BS $_j$ in cluster c . $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ denotes the independent complex additive white Gaussian noise with zero mean and variance of σ_k^2 . $\mathbf{x}_{j,c}$ is given by

$$\mathbf{x}_{j,c} = \sum_{k=1}^K \mathbf{w}_{jk} s_k \quad (3.2)$$

where $\mathbf{w}_{jk}$ is the beamforming vector, and the transmission power is defined in terms of $\|\mathbf{w}_{jk}\|^2 = p_{jk}$. Let $s_k \sim \mathcal{CN}(0, 1)$ be the data symbol intended for MU $_k$, where $\mathbb{E}\{|s_k|^2\}$ is normalized to one. Each base station is subject to an individual transmit power constraint $\mathbb{E}\{\|\mathbf{x}_{j,c}\|^2\} = \sum_{k=1}^K p_{jk} \leq P_j$. We assume that all the data streams have uniform power allocation, i.e., $p_{jk} = p_{\hat{k}} = p$. This is a commonly made assumption to simplify the problem when we focus on the design of beamformers' directions [88]. We also assume that $N_t \geq K$ such that independent data streams can be transmitted to the users in each cluster. Substituting (3.2) into (3.1), we get:

$$y_{k,c} = \underbrace{\sum_{j=1}^V \mathbf{h}_{jk,c}^H \mathbf{w}_{jk} s_k}_{\text{desired signal}} + \underbrace{\sum_{\hat{k}=1, \hat{k} \neq k}^K \sum_{j=1}^V \mathbf{h}_{jk,c}^H \mathbf{w}_{j\hat{k}} s_{\hat{k}}}_{\text{intra-cluster interference}} + \underbrace{\sum_{\hat{c}=1, \hat{c} \neq c}^I \sum_{i=1}^V \mathbf{h}_{ik,\hat{c} \rightarrow c}^H \mathbf{x}_{i,\hat{c}}}_{\text{inter-cluster interference}} + n_k. \quad (3.3)$$

For the cluster-interior user, there is no inter-cluster interference, so their received signals only contain the first two terms of (3.3). Our analysis is based on the assumption that the data symbols intended for all users within the same cluster are shared by the BS in that cluster; the BSs are connected via backhaul links which are assumed to be error-free but have limited capacity, and the BSs are well synchronized. This enables a cooperative transmission technique to be employed where each receiver is served by several BSs simultaneously. On the other hand, BSs may also share the CSIT via backhaul links. BSs will design their beamforming vectors $\mathbf{w}_{jk}$ based on local full CSI, i.e., a perfect knowledge of vectors $\mathbf{h}_{jk,c}$ originating from the j th BS.

Assuming that single-user decoding is applied and the MU treats the interference as noise, the SINR at MU_k is given by

$$\text{SINR}_k = \frac{\left| \sum_{j=1}^V \mathbf{h}_{jk,c}^H \mathbf{w}_{jk} \right|^2}{\sum_{\substack{\hat{k}=1 \\ \hat{k} \neq k}}^K \left| \sum_{j=1}^V \mathbf{h}_{jk,c}^H \mathbf{w}_{j\hat{k}} \right|^2 + \sum_{\substack{\hat{c}=1 \\ \hat{c} \neq c}}^I \sum_{i=1}^V \left| \sum_{m=1}^K \mathbf{h}_{ik,\hat{c} \rightarrow c}^H \mathbf{w}_{im} \right|^2 + \sigma_k^2} \quad (3.4)$$

for $k = 1, 2, \dots, K$. The achievable instantaneous transmission rate at the k th MU is $R_k = \log_2(1 + \text{SINR}_k)$. The sum rate of all the users in one cluster is expressed as $R_s = \sum_k \log_2(1 + \text{SINR}_k)$. For the cooperative multi-antenna systems, there are several typical optimization goals. One design criterion that has been widely used is to minimize the transmit power at base stations [89]. While in the max-min-fair beamforming problem, the goal is to maximize the user's sum rate with the worst SINR [90]. In some existing works [27], [28], the authors attempted to maximize the weighted sum rate. In this chapter, we aim at designing a beamforming scheme to maximize the sum rate. The

problem is formulated as follows:

$$\max_{\mathbf{w}_{jk} \in \mathbb{C}^{N_t \times 1}} \sum_{k=1}^K R_k, \quad \text{s.t.} \quad \|\mathbf{w}_{jk}\|^2 = p_{jk}, \quad \forall j, k \quad (3.5)$$

The fixed uniform power allocation scheme is chosen in a way that it satisfies the per-BS power constraints. It is obvious that the design of beamformers affects the individual SINR at each MU and hence, the sum rate of the multicell system. The dependency of SINR_k on different beamformers $\mathbf{w}_{jk}$ implies that a joint optimization of all the users within the multicell system is required to maximize the sum rate. In the next section, we will remove this dependency through the sub-optimal uplink-downlink duality and solve the problem in a distributed manner.

We derive the achievable rate region for the clustered multicell scenario. This scenario is different from the interference channel as the data sharing enables MUs to be served by multiple BSs. Since the MUs treat the co-channel interference as noise rather than trying to decode it, the achievable rate region should be lower than the theoretical capacity region. In the information theory, the channel capacity is defined as the maximum mutual information that can be obtained during the communications. The rate achieved by a cluster-edge user k is given by

$$R_k = \log \frac{\left| \sigma_k^2 \mathbf{I} + \mathbf{h}_{k,c} \mathbf{T}_c \mathbf{h}_{k,c}^H + \sum_{\hat{c} \neq c}^I \mathbf{h}_{k,\hat{c} \rightarrow c} \mathbf{T}_{\hat{c}} \mathbf{h}_{k,\hat{c} \rightarrow c}^H \right|}{\left| \sigma_k^2 \mathbf{I} + \sum_{\hat{k} \neq k}^K \mathbf{h}_{k,c} \mathbf{W}_{\hat{k}} \mathbf{h}_{k,c}^H + \sum_{\hat{c} \neq c}^I \mathbf{h}_{k,\hat{c} \rightarrow c} \mathbf{T}_{\hat{c}} \mathbf{h}_{k,\hat{c} \rightarrow c}^H \right|} \quad (3.6)$$

where $\mathbf{h}_{k,c} = [\mathbf{h}_{1k,c}; \dots; \mathbf{h}_{V_k,c}]$ is a column vector that includes the local channels, $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$ with $\mathbf{w}_k = [\mathbf{w}_{1k}; \dots; \mathbf{w}_{V_k}]$ being a column vector that includes the individual beamformer, $\mathbf{T}_c = \mathbb{E}\{\mathbf{x}_{all,c} \mathbf{x}_{all,c}^H\}$ with $\mathbf{x}_{all,c} = [\mathbf{x}_{1,c}; \dots; \mathbf{x}_{V,c}]$. The rate achieved by a

cluster-interior user is the same as (3.6) except that there is no inter-cluster interference. The achievable sum-rate capacity of one cluster is defined as the convex union of all the user rates given by

$$\mathfrak{R} = \bigcup_{\{\mathbf{w}_{jk} \in \mathbb{C}^{N_t \times 1}, \|\mathbf{w}_{jk}\|^2 = p_{jk}, \forall j, k\}} (R_1, \dots, R_K). \quad (3.7)$$

Each cluster is a network MIMO system, which can be seen as a superposition of multiple broadcast channels. The rate tuples $(R_1, \dots, R_K)$ are achievable with feasible precoding strategies. The outer boundary of $\mathfrak{R}$ is known as the Pareto boundary [7], [91]. The rate tuples $(R_1, \dots, R_K)$ on the boundary are Pareto optimal if it is impossible to increase one of the rates without simultaneously decreasing the rate of any of other users. Moreover, it has been shown that the Pareto optimality can be achieved only when the BS uses the full transmit power, i.e., $\|\mathbf{w}_{jk}\|^2 = p_{jk}$ and $\sum_{k=1}^K p_{jk} = P_j$ [7], [91]. The optimal beamforming vectors are linear combinations of the MRT and ZF beamformers [91]. In general, we are looking for proper beamformers to approach Pareto boundary.

3.2 Semidefinite Relaxation-Based Algorithm

The objective rate function R_k is coupled by multiple beamformers that appear in both denominator and numerator of the SINR expression (3.4). The rate function is neither convex nor concave in $\mathbf{w}_{jk}$. Hence, the optimization problem in (3.5) is nonconvex and requires a complicated joint optimization across MUs as the SINR depends on the beamforming vectors of multiple users. In order to solve this problem efficiently, we apply the sub-optimal uplink-downlink duality to decouple various beamformers and transform the global optimization problem into a local one. Finally, we reformulate

3.2 Semidefinite Relaxation-Based Algorithm

the problem as a convex SDP problem by dropping the rank-one constraint and solve it accordingly. In this section, we assume full CSIT is available at all BSs.

The downlink is a network MIMO system. The direct dual of it is still in the form of network MIMO, which means the uplink dual problem is also a non-convex optimization problem that is intractable even if the global CSI is available. Instead, we seek an approximate design criterion by decomposing the network MIMO into a superposition of multiple broadcast channels and applying the uplink-downlink duality on the individual broadcast channel. We will use a suboptimal uplink-downlink duality for the downlink network MIMO to obtain a distributed optimization criterion [90]. Uplink-downlink duality refers to the fact that the same SINR target can be achieved in both downlink and uplink channels with the same set of beamforming vectors under the same per-BS power constraint, where the uplink channel is obtained by reversing the input and the output of the downlink [25], [90]. It implies that if per-BS sum-power constraint is considered for the virtual uplink, the uplink-downlink duality can be directly applied to solve the corresponding downlink optimization problem. Note that our goal is to achieve the performance that can approach the Pareto boundary. The interdependency among multiple beamformers can be removed with the aid of sub-optimal uplink-downlink duality, and the optimal beamforming vector can be obtained by the following theorem:

Theorem 3.1: Representing the downlink network MIMO as the superposition of some broadcast channels in each cluster and considering full uniform power allocation for beamformers, the Pareto optimal rate tuple can be approximately achieved by choosing beamforming vectors as

$$\mathbf{w}_{jk} = \arg \max_{\|\mathbf{w}\|^2=p} \frac{q_{jk} \mathbf{w}^H \mathbf{G}_{jk} \mathbf{w}}{\sum_{\hat{k} \neq k}^L q_{j\hat{k}} \mathbf{w}^H \mathbf{G}_{j\hat{k}} \mathbf{w} + \sigma_k^2}, \quad (3.8)$$

Beamforming Design in Cooperative Multicell Systems

where $\mathbf{G}_{jk} = \mathbf{h}_{jk,c} \mathbf{h}_{jk,c}^H$ for $1 \leq k \leq K$ and $p = P_j/K$. Duality variables q_{jk} can be interpreted as the transmit power of each user in the virtual uplink, and L is the number of channels originating from BS_{*j*} to users in the own cluster and adjacent clusters.

Proof. First, we decompose the downlink network MIMO into multiple broadcast channels. In the dual uplink multiple access channel, we treat each BS as an individual receiver rather than treating all the cooperative BSs in a cluster as a single receiver. Let $\rho_k = \log_2(1 + \text{SINR}_k^*)$ be an arbitrary Pareto optimal rate tuple. The rate tuple is achieved by choosing beamformers that solve the following optimization problem:

$$\begin{aligned} \max_{\mathbf{w}_{jk} \in \mathbb{C}^{N_t \times 1}, p_{jk} \geq 0} \quad & \min_k \frac{R_k}{\rho_k} \\ \text{s.t.} \quad & \|\mathbf{w}_{jk}\|^2 = p_{jk}, \forall j \text{ and } k, \end{aligned} \quad (3.9)$$

since the rate R_k cannot be increased without decreasing the rate of any other user. According to Lemma 1 in [25], the optimal solution to (3.9) yields the same R_k/ρ_k for all users. Here $\rho_1, \rho_2, \dots, \rho_K$ represent the target rate tuples of different users, which can be viewed as the weighting factors to consider heterogeneous user requirements. Larger ρ_k corresponds to higher rate R_k and vice versa. Each user's rate depends on the choice of beamformers and transmit power. The Pareto optimum can be achieved by maximizing the minimum weighted user rate with respect to the individual broadcast channel. Furthermore, as the user rate is monotonically increasing with the SINR, maximizing the minimum weighted R_k is equivalent to max-min of weighted SINR_k given by

$$\begin{aligned} \max_{\mathbf{w}_{jk} \in \mathbb{C}^{N_t \times 1}} \quad & \min_k \frac{\text{SINR}_k^{DL}}{\gamma_k} \\ \text{s.t.} \quad & \|\mathbf{w}_{jk}\|^2 = p_{jk}, \forall j \text{ and } k, \end{aligned} \quad (3.10)$$

3.2 Semidefinite Relaxation-Based Algorithm

where SINR_k^{DL} is given in (3.4), and γ_k is the SINR weighting factor that considers heterogeneous user requirements. Larger γ_k corresponds to higher SINR_k^{DL} , and vice versa. Based on [26], the above downlink max-min SINR problem is dual to the following virtual uplink min-max optimization problem:

$$\begin{aligned}
& \min_{\mathbf{v}} \max_{\mathbf{q}, \mathbf{w}, \alpha} \alpha \\
& \text{s.t. } \alpha \gamma_k \leq \text{SINR}_k^{UL} \\
& q_{jk} \geq 0, v_{jk} \geq 0, \|\mathbf{w}_{jk}\|^2 = p_{jk} \\
& \sum_{k=1}^K q_{jk} \leq \sum_{k=1}^K p_{jk}, \sum_{k=1}^K v_{jk} p_{jk} \leq \sum_{k=1}^K p_{jk}
\end{aligned} \tag{3.11}$$

where α is a slack variable and

$$\text{SINR}_k^{UL} = \frac{q_{jk} \mathbf{w}_{jk}^H \mathbf{G}_{jk} \mathbf{w}_{jk}}{\sum_{\hat{k} \neq k}^L q_{j\hat{k}} \mathbf{w}_{j\hat{k}}^H \mathbf{G}_{j\hat{k}} \mathbf{w}_{j\hat{k}} + v_{jk} \sigma_k^2} \tag{3.12}$$

is the virtual uplink SINR. The duality variables $\{q_{jk}\}_{k=1}^K$ are interpreted as the transmit power of each user in the virtual uplink, and $\{v_{jk}\}_{k=1}^K$ are the virtual noise variance coefficients at the BS. If we set $v_{jk} = 1$ for $k = 1, \dots, K$ and apply uniform power allocation, the above problem given by (3.11) can be further approximately transformed into

$$\begin{aligned}
& \max_{\mathbf{w}, \alpha} \alpha \\
& \text{s.t. } \alpha \gamma_k \leq \text{SINR}_k^{UL} \\
& \|\mathbf{w}_{jk}\|^2 = p, \quad \forall j \text{ and } k,
\end{aligned} \tag{3.13}$$

where

$$\text{SINR}_k^{UL} = \frac{q_{jk} \mathbf{w}_{jk}^H \mathbf{G}_{jk} \mathbf{w}_{jk}}{\sum_{\hat{k} \neq k}^L q_{j\hat{k}} \mathbf{w}_{j\hat{k}}^H \mathbf{G}_{j\hat{k}} \mathbf{w}_{j\hat{k}} + \sigma_k^2}. \tag{3.14}$$

Beamforming Design in Cooperative Multicell Systems

Consequently, the optimal beamforming vectors for the downlink problem can be found in a distributed way by solving the problem given in (3.13). When the optimum of (3.11) is achieved, v_{jk} are usually not equal to 1 for $k = 1, \dots, K$ and the non-uniform power allocation scheme should be applied. Thus, fixing the values of v_{jk} and q_{jk} leads to a sub-optimal uplink-downlink duality. The optimal beamformers should also maximize the virtual uplink SINR in (3.8) for each user. ■

As for the distributed optimization criterion, it is much easier to obtain the sub-optimal beamformer as compared with solving the joint optimization problem. At this point, we have transformed the original joint optimization problem given in (3.5) into a distributed problem given in (3.8) which could be solved locally.

Over the past decade, the convex optimization technique has been widely used in wireless communication systems and signal processing applications. It can provide a local optimum which is also the global optimum [36]. Some researchers have applied the SDP to solve the downlink optimization problem [37], [40]. The SDP technique works surprisingly well to yield the optimum. The beamforming problem is often in the form of fractional quadratically constrained quadratic program (QCQP) which is nonconvex and NP-hard. SDR is a computationally efficient approximation technique for difficult optimization problems [38]. In particular, it can be applied to many nonconvex QCQP or fractional QCQP. The key point of SDR technique is to drop the rank-one constraint in the original problem to obtain a relaxed QCQP. The relaxed QCQP is a SDP, whose globally optimal solution may be found by available algorithms in polynomial time. The SDR makes use of the following fundamental observations:

$$\mathbf{W} = \mathbf{w}\mathbf{w}^H \iff \mathbf{W} \succeq 0, \text{rank}(\mathbf{W}) = 1 \quad (3.15)$$

3.2 Semidefinite Relaxation-Based Algorithm

and

$$\mathbf{w}^H \mathbf{G} \mathbf{w} = \text{tr}(\mathbf{w}^H \mathbf{G} \mathbf{w}) = \text{tr}(\mathbf{G} \mathbf{w} \mathbf{w}^H) = \text{tr}(\mathbf{G} \mathbf{W}). \quad (3.16)$$

Therefore, by defining $\mathbf{W} = \mathbf{w} \mathbf{w}^H$, we rewrite (3.8) in terms of $\mathbf{W}$ as follows:

$$\begin{cases} \mathbf{U}^* = \arg \max_{\mathbf{W} \in \mathbb{C}^{N_t \times N_t}} \frac{\text{tr}(\mathbf{G} \mathbf{W})}{\text{tr}(\mathbf{Q} \mathbf{W})} & (3.17a) \\ \text{s.t. } \text{tr}(\mathbf{W}) \leq P_{SDR} & (3.17b) \\ \mathbf{W} \succeq 0 & (3.17c) \\ \text{rank}(\mathbf{W}) = 1 & (3.17d) \end{cases}$$

where $\mathbf{Q} = \sum_{\hat{k} \neq k}^L \mathbf{G}_{j\hat{k}} + \mathbf{I}_{N_t} \cdot K \sigma_k^2 / P_j$, $P_{SDR} = P_j / K$, and we select the coefficient $q_{jk} = 1$ to account for strong interference. Problem **P1** explicitly reveals and isolates the nonconvex part of the problem. It can be observed from **P1** that the rank-one constraint in (3.17d) is not convex and is difficult to handle. The basic idea behind SDR is to approximate **P1** by removing the rank-one constraint while keeping the positive semidefinite constraint. In this way, we end up with the following semidefinite program

$$\begin{cases} \mathbf{W}^* = \arg \max_{\mathbf{W} \in \mathbb{C}^{N_t \times N_t}} \frac{\text{tr}(\mathbf{G} \mathbf{W})}{\text{tr}(\mathbf{Q} \mathbf{W})} & (3.18a) \\ \text{s.t. } \text{tr}(\mathbf{W}) \leq P_{SDR} & (3.18b) \\ \mathbf{W} \succeq 0 & (3.18c) \end{cases}$$

where $\mathbf{W}^*$ represents a global optimum solution of problem **P2**, which can be obtained by existing numerical algorithms in polynomial time. It has been shown that the solution to **P2** after dropping the rank-one constraint is very close to the optimal solution to **P1** [38]. Note that (3.18) is a quasi-convex optimization problem. The general approach to solve this type of problem is to transform it to the corresponding epigraph form by

introducing a new variable as suggested in [36]; the optimal solution can be obtained by the classical bisection method [92] in which a sequence of SDP problems need to be solved. Fortunately, we are able to obtain the global optimal solution to problem **P2** by solving just one SDP. By considering that the denominator of (3.18a) is always positive, we can convert the fractional SDP to an equivalent SDP. Precisely, let us define a new variable $\mathbf{Z} = s\mathbf{W}$, where $s > 0$ satisfies $\text{tr}(\mathbf{Q}s\mathbf{W}) = 1$. Then multiplying by s the numerator and the denominator of (3.18a), the fractional SDR problem **P2** can be turned into a convex SDP as follows:

$$\mathbf{P3} \left\{ \begin{array}{ll} \mathbf{Z}^* = \arg \max_{\mathbf{Z} \in \mathbb{C}^{N_t \times N_t}, s} \text{tr}(\mathbf{GZ}) & (3.19a) \\ \text{s.t. } \text{tr}(\mathbf{QZ}) = 1 & (3.19b) \\ \text{tr}(\mathbf{Z}) \leq s \cdot P_{SDR} & (3.19c) \\ \mathbf{Z} \succeq 0, s > 0. & (3.19d) \end{array} \right.$$

Proposition 3.1: The fractional quasi-convex problem **P2** has the same optimal objective value as the convex SDP in **P3**. Furthermore, if $\mathbf{Z}^*$ solves **P2**, $(\frac{\mathbf{Z}^*}{\text{tr}(\mathbf{ZQ}^*)}, \frac{1}{\text{tr}(\mathbf{ZQ}^*)})$ solves **P3**; if $(\mathbf{Z}^*, s^*)$ solves **P3**, then $\mathbf{Z}^*/s^*$ solves **P2**.

Proof. Suppose that $\mathbf{W}^*$ is the optimal solution of **P2**, $\text{opt}(\mathbf{P2})$ is the optimal value of **P2** and $\text{opt}(\mathbf{P3})$ is the optimal value of **P3**. It can be shown that $(\mathbf{W}^*/\text{tr}(\mathbf{QW}^*), s^*)$ with $s^* = 1/\text{tr}(\mathbf{QW}^*)$ is feasible for **P3**, and the objective function value at the feasible point is $\text{tr}(\mathbf{GW}^*)/\text{tr}(\mathbf{QW}^*) = \text{opt}(\mathbf{P2})$. Thus, we have $\text{opt}(\mathbf{P3}) \geq \text{opt}(\mathbf{P2})$. On the other hand, if $(\mathbf{Z}^*, s^*)$ is the optimal solution of **P3** with $s^* > 0$, then $\mathbf{Z}^*/s^*$ is feasible for **P2**. The objective function value of **P2** is given by $\text{tr}(\mathbf{GZ}^*)/\text{tr}(\mathbf{QZ}^*) = \text{tr}(\mathbf{GZ}^*) = \text{opt}(\mathbf{P3})$. Thus, we have $\text{opt}(\mathbf{P2}) \geq \text{opt}(\mathbf{P3})$; combined with the previous conclusion, it yields $\text{opt}(\mathbf{P2}) = \text{opt}(\mathbf{P3})$. ■

3.2 Semidefinite Relaxation-Based Algorithm

The objective function and constraints in **P3** are linear in $\mathbf{Z}$, and the optimum solution can be obtained by using an interior-point algorithm [92]. Once the solution $\mathbf{Z}^*$ is obtained, we need to check its rank. If $\mathbf{Z}^*$ is of rank one, the solution is the global optimum of the fractional SDP of **P1** since **P2** is a relaxation of **P1** by dropping the rank-one constraint. The optimal beamformer can then be obtained by performing the eigenvalue decomposition on $\mathbf{Z}^*$. On the other hand, if the solution is of more than rank one, we use Gaussian randomization to construct an approximate solution to (3.8). The idea of Gaussian randomization is to pick a random vector $\boldsymbol{\xi}_l \sim \mathcal{N}(\mathbf{0}, \mathbf{Z}^*)$ and project it to a feasible region [38]. The specific procedure is as follows:

$$\mathbf{w}^* = \arg \max_{\boldsymbol{\xi}_l, l=1, \dots, L} \frac{\boldsymbol{\xi}_l^H \mathbf{G} \boldsymbol{\xi}_l}{\boldsymbol{\xi}_l^H \mathbf{Q} \boldsymbol{\xi}_l}, \quad (3.20)$$

where L is the number of randomization. The optimal transmit beamformer will enhance the desired signal by the beam pattern main lobe and suppress the interfering signals by beam pattern nulls. In [17], the authors have proposed a generalized eigen decomposition approach to solve the problem. However, the solution to (3.8) is not unique due to the phase shifts in the beamformers $\mathbf{w}$. The solution is invariant to an angular rotation [64]. It is possible that the beamformers are selected such that the signals arriving at a given MU from multiple BSs will sum up destructively. Our proposed SDR-based algorithm solves the problem efficiently in the polynomial time in $\mathcal{O}(N_t^2 \log(1/\varepsilon))$ with ε being an accuracy parameter [92], which is comparable to the computational complexity of implementing eigenvalue decomposition. Furthermore, the SDR-based algorithm solves the problem in a different way from the eigen-beamforming approach in [17]. A global optimal solution can be obtained when a rank-one solution is obtained.

Fig. 3.2 Sum rate per cluster as function of SNR with $V = 3$, $N_t = 4$, $K = 3$.

3.3 Numerical Results and Discussions

In this section, we present some numerical results to evaluate the performance of the proposed SDR-based algorithm. In all simulations, spatially uncorrelated Rayleigh-fading channels are assumed and the noise variance is set to be $\sigma^2 = 1$. The system SNR is the average SNR a mobile user receives from a BS, without considering the co-channel interference. The optimal upper bound is generated by grouping all the BSs into one cluster and applying the DPC approach. The SDP problem **P3** is solved by using the optimization solver CVX [93].

The proposed SDR-based beamforming approach is compared with some existing methods in Fig. 3.2 in terms of the sum rate of a clustered system with $V = 3$. The

Fig. 3.3 Sum rate per cluster of the proposed algorithm with different number of transmit antennas for $V = 3$ and $K = 3$.

number of antennas at each BS is $N_t = 4$. The number of neighboring clusters that generate interference is $I = 4$. For simplicity, there are three active MUs in each cluster with two cluster interior users and one edge user. Sum rates of the proposed SDR based beamforming, eigen-beamforming approach in [17], ZF beamforming, MRT beamforming are compared. It can be observed that the proposed SDR-based beamforming scheme achieves almost the same performance as that of the eigen-beamforming approach, and it outperforms ZF and MRT across all SNRs. There is some performance loss compared with the sum-rate maximizing precoding due to the distributed optimization goal.

Figure 3.3 illustrates the sum rates of the proposed SDR-based beamforming, generalized eigen-beamforming and ZF beamforming schemes with perfect CSI for different

Fig. 3.4 Sum rate per cluster with different cluster size for $N_t = K$ and SNR = 10dB.

numbers of antennas with cluster size $V = 3$. With more transmit antennas, the system's DoF is increased, so is the sum-rate capacity. The proposed SDR-based scheme achieves almost the same performance with that of eigen-beamforming, and outperforms the conventional ZF beamforming scheme consistently.

In Fig. 3.4, we compare the sum-rate performance of various beamforming schemes for different cluster size V with SNR = 10dB. The number of transmit antennas is assumed to be $N_t = K$. The results are consistent with previous ones that the proposed SDR-based beamforming scheme has almost the same sum-rate performance with the generalized eigen-beamforming, and it outperforms the conventional ZF beamforming.

3.4 Conclusions

We have considered beamforming schemes in a cooperative clustered multicell multi-antenna system. The sum-rate maximization problem was transformed into a convex SDP by applying a sub-optimal uplink-downlink duality and the SDR. Numerical results have shown that the proposed SDR-based beamforming scheme can achieve good sum-rate performance improvement.

Chapter 4

Pilot Contamination and Beamforming in Massive MIMO Systems

Major market trends of the 5G network include higher system capacity, higher data rates, and the support of massive connectivity. The 5G network is expected to achieve a 1,000-fold increase in network capacity compared to the current LTE system [94]–[96]. Recently, massive MIMO technique has been considered as a candidate for 5G wireless standard [95], [96]. A key air interface technology to mitigate the interference and enhance the capacity is through beamforming. The network capacity can be significantly enhanced due to the large number of antenna arrays being deployed.

In this chapter, two problems in massive multi-antenna systems will be discussed, namely, the analysis of condition of pilot contamination mitigation and the design of 3D beamforming. To analyze the condition under which the pilot contamination can be mitigated, a 2D massive multi-antenna scenario is considered. The antenna elements are arranged in a uniform circular array (UCA) topology, and the AoAs of multipath components are assumed to follow the VMF distribution [97]. It is shown that the

interfering channels tend to fall into the null space of the covariance matrix of the desired user when the number of antennas increases. The pilot contamination effect can be mitigated by exploiting the property of channel covariance matrix.

In conventional multi-antenna systems, beamforming techniques usually operate only on the horizontal antenna pattern, and the vertical pattern is fixed. The signal may not be fully focused onto a particular user due to the fixed vertical pattern. The conventional 2D beamforming cannot take user's position into consideration. 3D beamforming, which enables fully dynamic adaptation of the vertical beam pattern, offers promising performance improvement. The signal strength can be increased by pointing the vertical main lobe directly to the receiver. The overall beamforming gain is fully exploited. In the second section of this chapter, the design of 3D beamforming is studied based on a realistic two-tier non-uniform user distribution. We derive the statistical distribution function of user's geographic position. Two practical 3D beamforming strategies are proposed. A region-partition 3D beamforming strategy, in which users in a certain partition are assigned a particular downtilt, is proposed when the user location information is available at BSs. When the location information is not available, we propose a statistic-based 3D beamforming strategy where the BS uses knowledge of statistical distribution of the user to determine the downtilt angle. We compare the proposed 3D beamforming schemes with user-specific beamforming where the BS directly points the vertical beam to the intended user.

4.1 Pilot Contamination Mitigation

4.1.1 System Description

A multi-cell system with L cells is considered¹. Each cell contains one BS with M antennas and K single-antenna user terminals (UTs). The system is assumed to operate in a TDD mode with universal frequency reuse. The CSI at BSs can be obtained through pilot-aided channel estimation. Non-orthogonal pilots are reused from cell to cell, leading to the pilot contamination from interfering UTs. The system model is demonstrated in Fig. 4.1. Since all UTs in the same cell transmit orthogonal pilot signals, one may consider a simple case where a single UT per cell transmits pilot signals to its target BS. The pilot signals used in the i -th cell are denoted as $\boldsymbol{\chi}_i = [\chi_{i1} \dots \chi_{i\tau}]^T$, where τ is the length of the sequence. Pilot signals are assumed to have equal power levels, i.e., $\mathbb{E}\{\|\boldsymbol{\chi}_i\|^2\} = \tau$, for $i = 1, \dots, L$. The $M \times \tau$ pilot signals received at the i -th BS are given by

$$\mathbf{Y}_i = \mathbf{g}_{ii}\boldsymbol{\chi}_i^T + \sum_{j=1, j \neq i}^L \mathbf{g}_{ji}\boldsymbol{\chi}_j^T + \mathbf{N}_i, \quad (4.1)$$

where $\mathbf{g}_{ji} \in \mathbb{C}^{M \times 1}$ represents the channel from the j -th cell user to the i -th BS, and $\mathbf{N}_i \in \mathbb{C}^{M \times \tau}$ is the independently and identically distributed (i.i.d.) Gaussian noise with its element distributed as $\mathcal{CN}(0, \sigma_n^2)$. The first term of the right hand side in (4.1) represents the desired pilot signals and the second term is the sum of interfering pilot signals from other cells. The propagation channel is assumed to be flat fading, and is given by $\mathbf{g}_{ji} = \mathbf{h}_{ji}\sqrt{\beta_{ji}}$, where $\mathbf{h}_{ji} \in \mathbb{C}^{M \times 1}$ represents the small-scale fading channel coefficient and β_{ji} is the large-scale fading coefficient accounting for path loss and shadow fading.

¹Note that all the notations are only applicable to Section 4.1.

Fig. 4.1 Multi-cell massive multi-antenna system with 3D spatial channel and pilot contamination effect.

Possible topologies for massive arrays include the uniform circular array (UCA), uniform rectangular array (URA) and concentric circular array (CCA) [98]. The UCA has the best spatial resolution, while the CCA achieves the deepest nulls. The selection of a topology depends on the specific requirements and hardware availability. A UCA topology is considered because a realistic 3D channel model can be applied. The coupling effects will be ignored. The 3D spatial channel model is shown in Fig. 4.2. There are M antenna elements on a planar circle. The angles θ and φ denote the co-elevation AoA and azimuth AoA, respectively. In most of the existing works, AoAs are modeled using uniform, Gaussian or Laplace distributions [99]. A more realistic model is applied, namely, the VMF distribution from the directional statistics [97]. In this model, the direction is represented by a point on the unit-radius Euclidean ball, and it can be written as a vector $\mathbf{\Omega} = [\sin\theta\cos\varphi \ \sin\theta\sin\varphi \ \cos\theta]^T$, with $0 \leq \theta \leq \pi$, $0 \leq \varphi \leq 2\pi$. It is assumed that scatterers are clustered [97], then the VMF pdf is given

4.1 Pilot Contamination Mitigation

are located in the space. The antenna arrays at BSs are assumed to be high enough such that the UTs are closely surrounded by all the scatterers.

A rich scattering situation with $P \gg 1$ number of paths is considered. The fading channel is given by

$$\mathbf{h}_{ji} = \frac{1}{\sqrt{P}} \sum_{p=1}^P h_{jip} \boldsymbol{\alpha}(\theta_{jip}, \varphi_{jip}), \quad (4.3)$$

where $h_{jip} \sim \mathcal{CN}(0, \delta_{ji}^2)$ is the complex channel gain. Note that $\boldsymbol{\alpha}(\theta_{jip}, \varphi_{jip})$ is a steering vector defined as $\boldsymbol{\alpha}(\theta_{jip}, \varphi_{jip}) = [e^{-i\frac{2\pi}{\lambda} R \sin \theta_{jip} \cos(\varphi_{jip} - \varphi_1)} \dots e^{-i\frac{2\pi}{\lambda} R \sin \theta_{jip} \cos(\varphi_{jip} - \varphi_M)}]^T$, where θ_{jip} and φ_{jip} denote the co-elevation and azimuth of the p -th path from the j -th cell user to the i -th BS, respectively, and $\varphi_m = 2\pi(m-1)/M$ defines the position of the m -th antenna element. Note that R is the radius of the array circle, λ is the wavelength and it is assumed that $R \leq \lambda$.

For the propagation channel coefficients $\mathbf{g}_{ji}$, the large-scale fading coefficient β_{ji} changes slowly and can be estimated over a long time interval. The small-scale fading coefficients $\mathbf{h}_{ji}$ model the fast fading and need to be estimated through pilot signals. The BS is assumed to use the minimum mean-square error (MMSE) channel estimation given by

$$\hat{\mathbf{h}}_{ii} = \sqrt{\beta_{ii}} \mathbf{R}_{ii} \left(\tau \sum_{j=1}^L \beta_{ji} \mathbf{R}_{ji} + \sigma_n^2 \mathbf{I}_M \right)^{-1} \boldsymbol{\Psi}^\dagger \mathbf{y}_i, \quad (4.4)$$

where $\mathbf{y}_i = \text{vec}\{\mathbf{Y}_i\}$, $\mathbf{R}_{ii}$ and $\mathbf{R}_{ji}$ are channel covariance matrices of $\mathbf{h}_{ii}$ and $\mathbf{h}_{ji}$, respectively. The pilot matrix is defined as $\boldsymbol{\Psi} = \boldsymbol{\chi}_i \otimes \mathbf{I}_M$, where $\otimes$ denotes the Kronecker product. The channel covariance matrix $\mathbf{R}_{ii}$ is given by [97]

$$\begin{aligned} \mathbf{R}_{ii} &= \frac{\delta_{ii}^2}{P} \sum_{p=1}^P \mathbb{E}\{\boldsymbol{\alpha}(\theta_{iip}, \varphi_{iip}) \boldsymbol{\alpha}^\dagger(\theta_{iip}, \varphi_{iip})\} \\ &= \int_{\varphi_{ii}} \int_{\theta_{ii}} \boldsymbol{\alpha}(\theta_{ii}, \varphi_{ii}) \boldsymbol{\alpha}^\dagger(\theta_{ii}, \varphi_{ii}) f_p(\theta_{ii}, \varphi_{ii}) d\theta_{ii} d\varphi_{ii} \end{aligned} \quad (4.5)$$

where $f_p(\theta_{ji}, \varphi_{ji})$ follows the VMF distribution in (4.2). Note that θ_{ji} and φ_{ji} denote the co-elevation and azimuth angles from the j -th cell user to the i -th BS, respectively. The MMSE estimation $\hat{\mathbf{h}}_{ii}$ is contaminated by the unwanted pilot signals.

4.1.2 Conditions for Pilot Contamination Mitigation

When the number of antennas increases, the existence of the null space of channel covariance matrix can assist us to deal with the pilot contamination problem. In [50], Yin *et al.* have shown that the pilot contamination effect can be mitigated if the multipath AoAs of the desired user do not overlap with AoAs of interfering users. Massive uniform linear array (ULA) is assumed to be deployed at the BS in [50]. In this section, the results of [50] are extended from the case of ULA to a scenario of 2D antenna arrays at the BS by applying a 3D channel model with VMF distribution.

A local scatterer multipath environment is considered [100]. The VMF distribution can be applied to model the distribution of the scatterers. Given the values of (θ_0, ϕ_0) and κ , the scatterers will be dropped in a region on the surface of the sphere. The region is defined as a circular surface that covers all these scatterers. The mean angles (θ_0, ϕ_0) correspond to the center of the circular surface. The radius of the circular surface is determined by the scatterer that is farthest to the center point of the circular surface. Since the dispersion factor κ indicates how dispersive the scatterers are distributed on the sphere, it indirectly determines the radius of the circular surface.

In local scatterer environment, scatterers are close to the UTs. This means that the scatterers can be modeled by VMF distribution with high dispersion κ . In this case, multipath signals from the desired user arrive at the reference BS in a region $[(\theta_{ii}^{\min}, \varphi_{ii}^{\min}), (\theta_{ii}^{\max}, \varphi_{ii}^{\max})]$ on the surface of sphere that is determined by the minimum and maximum values of the elevation angle and azimuth angle. The multipath AoA has

concentrated distribution with small deviation from the mean angle, because the UTs are closely surrounded by local scatterers. When any AoA $(\theta_{ji}, \varphi_{ji})$ of the j -th interfering user does not fall in the AoA range of the desired user, it must comply with one of the following non-overlapping conditions:

$$\begin{cases} \Upsilon = \theta_{ji} \in [\theta_{ii}^{\min}, \theta_{ii}^{\max}], \Lambda = \varphi_{ji} \notin [\varphi_{ii}^{\min}, \varphi_{ii}^{\max}] \\ \Upsilon = \theta_{ji} \notin [\theta_{ii}^{\min}, \theta_{ii}^{\max}], \Lambda = \varphi_{ji} \in [\varphi_{ii}^{\min}, \varphi_{ii}^{\max}] \\ \Upsilon = \theta_{ji} \notin [\theta_{ii}^{\min}, \theta_{ii}^{\max}], \Lambda = \varphi_{ji} \notin [\varphi_{ii}^{\min}, \varphi_{ii}^{\max}] \end{cases} \quad (4.6)$$

where (Υ, Λ) is the AoA pair of the interfering UT.

Lemma 4.1: For a clustering scattered 3D multipath structure where users are closely surrounded by local scatterers, if the VMF distribution $f_p(\theta, \varphi) = 0$ for $(\theta, \varphi) \notin [(\theta_{ii}^{\min}, \varphi_{ii}^{\min}), (\theta_{ii}^{\max}, \varphi_{ii}^{\max})]$, one has $\lim_{M \rightarrow \infty} \text{rank}(\mathbf{R}_{ii})/M \leq d_{ii}$, where $d_{ii} \triangleq |\cos(\gamma_{ii}^{\min}) - \cos(\gamma_{ii}^{\max})| \frac{R}{\lambda}$ with

$$\begin{aligned} \gamma_{ii}^{\min} &= \min_{\theta_{ii} \in [\theta_{ii}^{\min}, \theta_{ii}^{\max}], \varphi_{ii} \in [\varphi_{ii}^{\min}, \varphi_{ii}^{\max}]} \cos^{-1} \left(\frac{1}{2} [\sin(\theta_{ii} - \varphi_{ii} + \varphi_m) + \sin(\theta_{ii} + \varphi_{ii} - \varphi_m)] \right) \\ \gamma_{ii}^{\max} &= \max_{\theta_{ii} \in [\theta_{ii}^{\min}, \theta_{ii}^{\max}], \varphi_{ii} \in [\varphi_{ii}^{\min}, \varphi_{ii}^{\max}]} \cos^{-1} \left(\frac{1}{2} [\sin(\theta_{ii} - \varphi_{ii} + \varphi_m) + \sin(\theta_{ii} + \varphi_{ii} - \varphi_m)] \right). \end{aligned} \quad (4.7)$$

Proof. First, the real and imaginary parts of the element in $\mathbf{R}_{ii}$ are given by

$$\Re\{\mathbf{R}_{ii}(m, n)\} = \int_{\varphi_{ii}^{\min}}^{\varphi_{ii}^{\max}} \int_{\theta_{ii}^{\min}}^{\theta_{ii}^{\max}} \cos\left(\frac{2\pi}{\lambda} \sin \theta_{ii} \mathcal{B}_{ii}(m, n)\right) f_p(\theta_{ii}, \varphi_{ii}) d\theta_{ii} d\varphi_{ii} \quad (4.8)$$

and

$$\Im\{\mathbf{R}_{ii}(m, n)\} = \int_{\varphi_{ii}^{\min}}^{\varphi_{ii}^{\max}} \int_{\theta_{ii}^{\min}}^{\theta_{ii}^{\max}} \sin\left(\frac{2\pi}{\lambda} \sin \theta_{ii} \mathcal{B}_{ii}(m, n)\right) f_p(\theta_{ii}, \varphi_{ii}) d\theta_{ii} d\varphi_{ii} \quad (4.9)$$

where $\mathcal{B}_{ii}(m, n) = R \cos(\varphi_{ii} - \varphi_m) + R \cos(\varphi_{ii} - \varphi_n)$. Furthermore, the m th element in the steering vector $\boldsymbol{\alpha}(\theta_{ii}, \varphi_{ii})$ can be written as $e^{-i\frac{2\pi}{\lambda} R \sin \theta_{ii} \cos(\varphi_{ii} - \varphi_m)} = e^{-i\frac{2\pi}{\lambda} R \cos \gamma_{ii,m}}$ where $\gamma_{ii,m} = \cos^{-1}(\frac{1}{2}[\sin(\theta_{ii} - \varphi_{ii} + \varphi_m) + \sin(\theta_{ii} + \varphi_{ii} - \varphi_m)])$. Then $\boldsymbol{\alpha}(\theta_{ii}, \varphi_{ii})$ can be represented by $\boldsymbol{\alpha}(\gamma_{ii,\mathcal{A}})$ with $\mathcal{A} = [1 \dots M]$. Define

$$\mathbf{b}(x_{\mathcal{A}}) = \boldsymbol{\alpha}\left(\cos^{-1}\left(x_{\mathcal{A}} \frac{\lambda}{R}\right)\right) = [e^{-i2\pi x_1} \dots e^{-i2\pi x_M}]^T, \quad x_{\mathcal{A}} \in \left[-\frac{R}{\lambda}, \frac{R}{\lambda}\right]. \quad (4.10)$$

Based on Lemma 1 in [50], one knows that $\mathbf{b}(x_{\mathcal{A}}) = \boldsymbol{\rho}(2x_{\mathcal{A}}) \forall x_{\mathcal{A}} \in [x^{\min}, x^{\max}]$ in $[-1/2, 1/2]$ with $\boldsymbol{\rho}(x_{\mathcal{A}}) = [e^{-i\pi x_1} \dots e^{-i\pi x_M}]^T$, thus

$$\dim\{\text{span}[\mathbf{b}(x_{\mathcal{A}})]\} \sim (x^{\max} - x^{\min})M, \quad \forall x_{\mathcal{A}} \in [x^{\min}, x^{\max}] \quad (4.11)$$

for a large value of M , where $\dim\{\cdot\}$ is the dimension of the corresponding space. From (4.10), one can obtain

$$\text{span}\{\mathbf{R}_{ii}\} = \text{span}\left\{\int_{\gamma_{ii}^{\min}}^{\gamma_{ii}^{\max}} \mathbf{b}\left(\frac{R}{\lambda} \cos \gamma_{ii,\mathcal{A}}\right) \mathbf{b}^\dagger\left(\frac{R}{\lambda} \cos \gamma_{ii,\mathcal{A}}\right) f_p(\theta_{ii}, \varphi_{ii}) d\gamma_{ii,\mathcal{A}}\right\} \quad (4.12)$$

where $\gamma_{ii}^{\min}$ and $\gamma_{ii}^{\max}$ are given in (4.7).

Interpreting the integral in (4.12) as a summation, each column of $\mathbf{R}_{ii}$ can be written as a specific (and finite) linear combination of the vector $\mathbf{b}(R \cos \gamma_{ii,\mathcal{A}}/\lambda)$. On the other hand, the space spanned by $\mathbf{b}(x_{\mathcal{A}})$ consists of infinite linear combinations of $\mathbf{b}(x_{\mathcal{A}})$. Thus

$$\text{span}\{\mathbf{R}_{ii}\} \subset \text{span}\{\mathbf{b}(x_{\mathcal{A}})\}, x_{\mathcal{A}} \in \left[\frac{R}{\lambda} \cos(\gamma_{ii}^{\max}), \frac{R}{\lambda} \cos(\gamma_{ii}^{\min})\right]. \quad (4.13)$$

Based on (4.11) and (4.13), one can obtain

$$\text{rank}\{\mathbf{R}_{ii}\} = \dim\{\text{span}\{\mathbf{R}_{ii}\}\} \leq \left|\cos(\gamma_{ii}^{\min}) - \cos(\gamma_{ii}^{\max})\right| \frac{R}{\lambda} M \quad (4.14)$$

for a large value of M , and the proof is complete. ■

According to Lemma 4.1, there exists a null space of the covariance matrix $\mathbf{R}_{ii}$ when M becomes very large. Any vector falling into this null space is strictly orthogonal to the columns of $\mathbf{R}_{ii}$.

Lemma 4.2: For multipath components with a specific AoA pair (Υ, Λ) that satisfies one of the non-overlapping conditions in (4.6), the channel corresponding to AoA (Υ, Λ) tends to fall into the null space of $\mathbf{R}_{ii}$ when M is very large.

Proof. First, a steering vector is defined as $\mathbf{v} \triangleq \frac{\boldsymbol{\alpha}(\Upsilon, \Lambda)}{\sqrt{M}}$, where (Υ, Λ) satisfies either $\Upsilon \neq \theta_{ii} \in [\theta_{ii}^{\min}, \theta_{ii}^{\max}]$ or $\Lambda \neq \varphi_{ii} \in [\varphi_{ii}^{\min}, \varphi_{ii}^{\max}]$, or both. Following that, one may have

$$\mathbf{v}^\dagger \mathbf{R}_{ii} \mathbf{v} = \frac{1}{M} \mathbb{E} \left\{ |\boldsymbol{\alpha}^\dagger(\Upsilon, \Lambda) \boldsymbol{\alpha}(\theta_{ii}, \varphi_{ii})|^2 \right\} = \frac{1}{M} \mathbb{E} \left\{ \left| \sum_{m=1}^M e^{i \frac{2\pi}{\lambda} R \mathcal{F}} \right|^2 \right\} \quad (4.15)$$

where $\mathcal{F} = \sin \Upsilon \cos(\Lambda - \varphi_m) - \sin \theta_{ii} \cos(\varphi_{ii} - \varphi_m)$. The above equation can be further transformed into

$$\int_{\varphi_{ii}^{\min}}^{\varphi_{ii}^{\max}} \int_{\theta_{ii}^{\min}}^{\theta_{ii}^{\max}} \left| \frac{1}{M} \sum_{m=1}^M e^{i \frac{2\pi}{\lambda} R \mathcal{F}} \right|^2 f_p(\theta_{ii}, \varphi_{ii}) d\theta_{ii} d\varphi_{ii}. \quad (4.16)$$

Since the sum of the exponential series in (4.16) gives a finite value, one can obtain

$$\lim_{M \rightarrow \infty} \left| \frac{1}{M} \sum_{m=1}^M e^{i \frac{2\pi}{\lambda} R \mathcal{F}} \right|^2 = 0. \quad (4.17)$$

Note that if the AoAs (Υ, Λ) overlap with (θ, φ) , $\mathcal{F} = 0$ and the result in (4.17) is not equal to zero. Then, $\lim_{M \rightarrow \infty} \mathbf{v}^\dagger \mathbf{R}_{ii} \mathbf{v} = 0$, which means the vector $\mathbf{v}$ is orthogonal to each column of the covariance matrix $\mathbf{R}_{ii}$. Hence, it falls into the null space of $\mathbf{R}_{ii}$, and the proof for Lemma 4.2 is complete. ■

Based on the two lemmas, multipath components with AOAs outside the AOA region of a desired user tend to fall into the null space of the user's covariance matrix. The channel of the interfering user will be orthogonal to the eigenvector matrix obtained by performing the eigenvalue decomposition on $\mathbf{R}_{ii}$. Then the MMSE channel estimate in (4.4) is identical to the perfect channel estimate where there is no pilot contamination [50]. When the users are closely surrounded by local scatterers, the AoAs of an interfering user are less likely to overlap with the AoA region of the desired user.

4.1.3 Beamforming of Large Antenna Arrays

The beamforming in massive MIMO systems is different from that in conventional systems in the sense that massive arrays can generate sharp beams pointing to a particular direction. We will apply the null-steering beamforming to maximize the desired signal power while minimizing the interference power [101]. The signal received at the k th intended UT in the l th cell is given by

$$x_{kl} = \sum_{j=1}^L \sum_{i=1}^K \mathbf{g}_{klj}^\dagger \mathbf{w}_{ij} s_{ij} + v_{kl} \quad (4.18)$$

where s_{ij} is the transmitted signal intended to the i th UT in the j th cell with unit power $\mathbb{E}\{\|s_{ij}\|^2\} = 1$, $\mathbf{w}_{ij} \in \mathbb{C}^{M \times 1}$ is the corresponding null-steering beamformer, $\mathbf{g}_{klj} \in \mathbb{C}^{M \times 1}$ is the channel coefficient from the j th BS to the k th UT in the l th cell, and $v_{kl} \sim \mathcal{CN}(0, \sigma_n^2)$ denotes the i.i.d. Gaussian noise component. The null-steering beamformer is given by [101]

$$\mathbf{w}_{kll}^{\text{ns}} = \frac{\sqrt{P_T}}{\|(\mathbf{I} - \Pi_{\hat{\mathbf{G}}})\hat{\mathbf{g}}_{kll}\|} \cdot (\mathbf{I} - \Pi_{\hat{\mathbf{G}}})\hat{\mathbf{g}}_{kll} \quad (4.19)$$

where P_T is the total transmission power, and $\hat{\mathbf{G}} = [\hat{\mathbf{g}}_{kl1} \dots \hat{\mathbf{g}}_{klj, j \neq l} \dots \hat{\mathbf{g}}_{klL}]$ with $\hat{\mathbf{g}}_{klj} = \hat{\mathbf{h}}_{klj} \sqrt{\beta_{klj}}$ for $k = 1 \dots K$. $\Pi_{\hat{\mathbf{G}}} \triangleq \hat{\mathbf{G}}(\hat{\mathbf{G}}^\dagger \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^\dagger$ is the orthogonal projection onto the

4.1 Pilot Contamination Mitigation

subspace spanned by the columns of $\hat{\mathbf{G}}$. The global CSI of the UT can be obtained by exchanging information through BS cooperation. The beam pattern of the antenna arrays is defined as

$$P(\xi_{kl}) \triangleq \frac{\beta_{kl}}{K} \left| \frac{\hat{\mathbf{h}}_{kl}^\dagger (\mathbf{I}_M - \Pi_{\hat{\mathbf{G}}}) \hat{\mathbf{h}}_{kl}}{\|(\mathbf{I}_M - \Pi_{\hat{\mathbf{G}}}) \hat{\mathbf{h}}_{kl}\|} \right|^2, \quad (4.20)$$

where ξ_{kl} represents the position of the k th UT in the l th cell. Substituting the MMSE estimator of (4.4) into (4.20), one can obtain

$$P(\xi_{kl}) = \frac{\beta_{kl}^2}{K} \left| \frac{\left\| \mathbf{R}_{kl} \left(\tau \sum_{i=1}^L \beta_{kli} \mathbf{R}_{kli} + \sigma_n^2 \mathbf{I}_M \right)^{-1} \boldsymbol{\Psi}^\dagger \mathbf{y}_{kl} \right\|^2 - \text{tr}(\Pi_{\hat{\mathbf{G}}} \hat{\mathbf{H}})}{\left\| (\mathbf{I} - \Pi_{\hat{\mathbf{G}}}) \mathbf{R}_{kl} \left(\tau \sum_{i=1}^L \beta_{kli} \mathbf{R}_{kli} + \sigma_n^2 \mathbf{I}_M \right)^{-1} \boldsymbol{\Psi}^\dagger \mathbf{y}_{kl} \right\|^2} \right|^2 \quad (4.21)$$

where the property $\hat{\mathbf{h}}_{kl}^\dagger \Pi_{\hat{\mathbf{G}}} \hat{\mathbf{h}}_{kl} = \text{tr}(\Pi_{\hat{\mathbf{G}}} \hat{\mathbf{h}}_{kl} \hat{\mathbf{h}}_{kl}^\dagger)$ is applied with $\hat{\mathbf{H}} = \hat{\mathbf{h}}_{kl} \hat{\mathbf{h}}_{kl}^\dagger$. The beam pattern is a function of the channel covariance matrix $\mathbf{R}$. The dispersion parameter κ plays an important role in determining the structure of the covariance matrix $\mathbf{R}$.

4.1.4 Numerical Results and Discussions

A 3-cell network is considered. The radius of the cell is 500 m, and there are $K = 15$ UTs in each cell. The number of multipath components is $P = 50$. The urban micro-cell model [102] for the large-scale fading is applied with $\beta_{kji}(\text{dB}) = PL - 10l \log_{10}(d/d_0)$, where PL is the path loss at the reference distance d_0 , l is the path-loss exponent, and d is the distance between the UT and the BS. In particular, $d_0 = 10$ m, $PL = -34.5$ dB and $l = 2.9$ are used. The operating frequency is set at 1 GHz.

The normalized channel estimation error for different directional dispersion factors κ is compared. The normalized channel estimation error at the i th BS is defined as

 Fig. 4.3 Estimation errors with different directional dispersion κ .

$\varsigma_i = 10\log_{10}\left(\sum_{j=1}^L \|\hat{\mathbf{h}}_{ji} - \mathbf{h}_{ji}\|^2 / \sum_{j=1}^L \|\mathbf{h}_{ji}\|^2\right)$. At each BS, the mean AoA pair of the desired user is $(\theta_0 = 45^\circ, \varphi_0 = 30^\circ)$, and the mean AoA pairs of the two interfering users are $(\theta_0^{\text{int}_1} = 60^\circ, \varphi_0^{\text{int}_1} = 90^\circ)$ and $(\theta_0^{\text{int}_2} = 36^\circ, \varphi_0^{\text{int}_2} = 150^\circ)$, respectively. The simulation results are shown in Fig. 4.3. The channel estimation error decreases as κ increases. As κ increases, the dispersion of AoA decreases. Small dispersion results in a highly non-overlapping area between AoA pairs, thereby mitigating the pilot contamination. On the other hand, as M becomes larger, the interfering channels tend to fall into the null space of $\mathbf{R}$. Note that the channel estimation error for the interference-free scenario is $\varsigma_i = 10\log_{10}\left(\sum_{i=1}^L \text{tr}\{\mathbf{R}_{ii}(\mathbf{I} + \tau\mathbf{R}_{ii}/\sigma_n^2)^{-1}\} / \sum_{i=1}^L \|\mathbf{h}_{ii}\|^2\right)$.

Fig. 4.4 Sum rates per cell with different directional dispersion and $K = 15$.

When M increases, the denominator increases faster than the numerator, thus the overall ratio decreases.

Next, the average sum rates for different values of κ are compared. The corresponding simulation results are shown in Fig. 4.4. The average sum rate is defined as $\mathcal{C} = \mathbb{E}\{\sum_{k=1}^K \log_2(1 + \text{SINR}_k)\}$, where SINR_k is the SINR at the k th UT. As κ increases, the AoA pairs between the desired user and interfering users have a larger non-overlapping region. Hence, the pilot contamination will have less impact on the channel estimation. More power can be radiated to the desired direction by the beamformer and the sum rate is enhanced accordingly.

Lastly, the azimuth beam patterns under different situations are investigated. It is assumed that there are nine active UTs. The beamformer focuses the power on

Fig. 4.5 Beam patterns with non-overlapping AoAs and perfect channel estimation.

Fig. 4.6 Beam patterns with imperfect channel estimation, overlapping AoAs between the desired and interfering UEs, $\kappa = 200$.

the direction of the desired UT and puts the nulls towards others. The desired UT is located with azimuth angle $\xi_d = 45^\circ$, and the locations of the other UTs are given by $\xi_{\text{int}} = [-110^\circ -88^\circ -10^\circ 19^\circ 68^\circ 98^\circ 120^\circ 158^\circ]$. For simplicity, it is assumed that the

Fig. 4.7 Beam patterns with imperfect channel estimation, overlapping AoAs between the desired and interfering UEs, $\kappa = 50$.

eight UTs have zero degree elevation angle. When the BSs have perfect CSI, a sharp beam is pointing to the direction of the desired UT as shown in Fig. 4.5. Meanwhile, there exist eight nulls that correspond to the eight unintended UTs. When the channel estimation is contaminated, the beam patterns are shown in Figs. 4.6 and 4.7 for $\kappa = 200$ and 50, respectively. When κ decreases, the channel estimation becomes less accurate as there are more overlaps between the AoAs. The main lobe becomes wider and the nulls deviate from the original positions.

The impact of the dispersion factor κ on the channel estimation has been investigated. In practice, it is determined by the wireless environment. One may place the BS antennas as high as possible, such that UTs are closely surrounded by local scatterers, which can be modeled by a relatively larger value of κ .

4.1.5 Conclusions

In this section, the condition under which the pilot contamination can be mitigated has been analyzed in a 2D massive multi-antenna system. The 3D VMF distribution to model the AoA of the multipath signals was analyzed. It has been shown that the pilot contamination can be mitigated if the AoA pairs of the desired user do not overlap with those of interfering users. Numerical results have shown that pilot contamination would have little impact on channel estimation when the UTs are closely surrounded by local scatterers.

4.2 3D Beamforming

4.2.1 System Description

We consider a multi-cell network with linear cellular array scenario, where L circular cells of radius R are arranged in a line [103]. Each cell has a BS equipped with $M \times N$ antennas in the URA topology, and each BS serves K single-antenna UTs². We assume that universal frequency reuse is applied and the system is operating in a TDD mode. We also assume that $MN \gg K$ as it is a typical scenario in massive multi-antenna systems.

In wireless multipath propagation channels, transmitted signals usually undergo scattering, reflection and refraction before they reach the receiver. All of these effects can be modeled by clustering of scatterers. We assume that the UTs are surrounded by many local scatterers that have small angular spread. The BS is assumed to be high enough so that the scatterers are only around UTs. The overall channel vector from the j th BS to the k th UT in the l th cell is given by $\mathbf{g}_{jkl} = \sqrt{\beta_{jkl}}\mathbf{h}_{jkl}$. The large-scale fading

²Note that all the notations are only applicable to Section 4.2.

Fig. 4.8 Two-dimensional massive multi-antenna network with multipath components of clustered scatterers.

factor is given by $\beta_{jkl} = \tau_{jkl}/d_{jkl}^\gamma$, where τ_{jkl} is a log-normal random variable, d_{jkl} is the distance between the k th UT in the l th cell and the j th BS, and γ is the path-loss exponent. The small-scale fading $\mathbf{h}_{jkl}$ is the superposition of multipath components. The local cluster channel model is shown in Fig. 4.8. Using the ray-based method for URA [9], the small-scale fading channel matrix is represented by [104]

$$\mathbf{H}_{jkl} = \frac{1}{\sqrt{P}} \sum_{p=1}^P \mathbf{a}(\theta_{kp}) \mathbf{a}^T(\theta_{kp}, \phi_{kp}), \quad (4.22)$$

where P is the total number of multipath components, θ_{kp} and ϕ_{kp} denote the elevation angle and azimuth angle of the p th multipath component of the k th UT, respectively.

The channel matrix is the sum of product between the horizontal steering vector and vertical steering vector. The $N \times 1$ steering vector with respect to the x -axis is expressed as $\mathbf{a}(\theta_{kp}, \varphi_{kp}) = \left[1 \ e^{-j\frac{2\pi}{\lambda}d_a \cos \theta_{kp} \cos \varphi_{kp}} \ \dots \ e^{-j\frac{2\pi}{\lambda}(N-1)d_a \cos \theta_{kp} \cos \varphi_{kp}} \right]^T e^{j\eta_{kp}}$, where d_a is the antenna spacing, λ is the carrier wavelength, and $e^{j\eta_{kp}}$ denotes the random phase of the path vector due to the small random perturbations of the UT's location [51]. η_{kp} is assumed to be i.i.d. uniformly distributed in $[-\Delta\eta, \Delta\eta]$, where $\Delta\eta$ is a small perturbation angle. The $M \times 1$ steering vector with respect to the z -axis can be expressed as $\mathbf{a}(\theta_{kp}) = \left[1 \ e^{-j\frac{2\pi}{\lambda}d_a \sin \theta_{kp}} \ \dots \ e^{-j\frac{2\pi}{\lambda}(M-1)d_a \sin \theta_{kp}} \right]^T e^{j\eta_{kp}}$. After vectorizing the channel matrix as $\mathbf{h}_{jkl} = \text{vec}(\mathbf{H}_{jkl})$, we obtain the small-scale fading vector $\mathbf{h}_{jkl}$. As we assume a rich scattering environment where $P \gg 1$, there are a significant number of multipath components in each of the resolvable angular directions. Thus, the entries of $\mathbf{h}_{jkl}$ can be well approximated by the i.i.d. Rayleigh fading model [9]. As β_{jkl} changes very slowly and can be assumed to be constant over many time intervals, the entries of $\mathbf{g}_{jkl}$ are i.i.d. complex Gaussian variables [45].

The received signal of the k th UT in the j th cell is given by

$$y_{kj} = \sqrt{p_j} \sum_{i=1}^K \mathbf{w}_{ji}^\dagger \mathbf{g}_{jkj} s_{ij} + \sum_{l \neq j}^L \sum_{i=1}^K \sqrt{p_l} \mathbf{w}_{li}^\dagger \mathbf{g}_{lkj} s_{il} + n_{kj} \quad (4.23)$$

where p_j represents the transmission power at the j th BS, s_{ij} is the transmitted signal distributed as $\mathcal{CN}(0, 1)$. Note that $\mathbf{w}_{ji} \in \mathbb{C}^{MN \times 1}$ denotes the MRT beamformer from the j th BS to the i th UT in the j th cell and $n_{kj} \sim \mathcal{CN}(0, 1)$ represents the complex Gaussian noise. The SINR of the k th UT in the j th cell is given by

$$\text{SINR}_{kj} = \frac{p_j \left| \mathbf{w}_{jkj}^\dagger \mathbf{g}_{jkj} \right|^2}{p_j \sum_{i \neq k}^K \left| \mathbf{w}_{ji}^\dagger \mathbf{g}_{jkj} \right|^2 + \sum_{l \neq j}^L \sum_{i=1}^K p_l \left| \mathbf{w}_{li}^\dagger \mathbf{g}_{lkj} \right|^2 + 1}. \quad (4.24)$$

In this case, the spatial capacity of the j th cell is defined as the ergodic sum capacity over the cell area:

$$C_j = \frac{B}{\pi R^2} \mathbb{E} \left\{ \sum_{k=1}^K \log_2 (1 + \text{SINR}_{kj}) \right\}, \quad (4.25)$$

where B represents the bandwidth of the system, R is the radius of the cell. The expectation is taken with respect to the distances between UTs and the BS, and the channel realizations. Hence, the spatial capacity depends on the user distribution and channel characteristics.

4.2.2 Distribution of User's Position and Downtilt

Most of the existing works consider the uniform user distribution due to the model's simplicity. However, the practical distribution of users may not behave in this way. Some public places such as malls and universities tend to be crowded with mobile users, while other places may have much less users. A non-uniform user-distribution model is more accurate in this case. Consequently, we consider a realistic two-tier clustered user distribution. In the first tier, some virtual points (VPs) are uniformly distributed in the cell. They correspond to the positions of places such as malls, buildings and universities in the real life. Then the UTs are located within the circular region centered at these VPs with uniform distribution in the second tier, and UTs in different clusters are independent. Note that the VP itself is not considered as the UT. Since user's distribution knowledge will be used for the beamforming design, the statistical distribution of UT's position is given in the following proposition.

Proposition 4.1: We assume a linear cellular array network where the users follow a two-tier clustered distribution as shown in Fig. 4.9. Let random variables $\mathcal{D}_{in}$ and $\mathcal{D}_o$ denote the UT's geographic position with respect to its own BS and the adjacent

Fig. 4.9 Demonstration of two-tier user distribution.

BS, respectively. Then, d and d_o are realizations of $\mathcal{D}_{in}$ and $\mathcal{D}_o$, respectively. Let r_1 and r_3 denote the distances from the tier-one VP to its own BS and the adjacent BS, respectively. The joint pdfs of $\mathcal{D}_{in}$ and $\mathcal{D}_o$ are given by:

$$f_{\mathcal{D}_{in}}(d, r_1) = \frac{2}{\mathcal{A}_c \mathcal{A}_2} \int_{-R}^R \left[\int_{-R}^R \sqrt{[R_2^2 - (z_3 - z_1)^2][R^2 - z_1^2]} dz_1 \right] \times \frac{1}{|d| \tan^{-1}(R_2/r_1) \sqrt{(d/z_3)^2 - 1}} dz_3 \quad (4.26)$$

and

$$f_{\mathcal{D}_o}(d_o, r_3) = \frac{2}{\mathcal{A}_c \mathcal{A}_2} \int_R^{3R} \left[\int_{-R}^R \sqrt{[R_2^2 - (2R - z_4 - z_1)^2][R^2 - z_1^2]} dz_1 \right] \times \frac{1}{|d_o| \tan^{-1}(R_2/r_3) \sqrt{(d_o/z_4)^2 - 1}} dz_4, \quad (4.27)$$

respectively, where $\mathcal{A}_c = \pi R^2$, $\mathcal{A}_2 = \pi R_2^2$, and R and R_2 denote the cell range and the radius of user cluster region, respectively.

Proof. As shown in Fig. 4.9, r_0 denotes the radius of the cell-hole in which there are no UTs [45]. The distances from the tier-one VP to its own BS and the adjacent BS are r_1 and r_3 , respectively. Let Z_1, Z_2, Z_3 and Z_4 denote four random variables, and z_1, z_2, z_3 and z_4 are realizations of them, respectively. First, we need to determine the pdf of Z_1 . The VPs have the distribution $f(r_1, \varphi_1) = r_1/\pi[R^2 - r_0^2]$ for $r_1 \in [r_0, R]$, as r_1 and φ_1 are independent. Since $f(z_1) = f(-z_1)$, we will concentrate on the case for $z_1 \geq 0$. The pdf of Z_1 can be derived as follows:

$$\begin{aligned}
 f_{Z_1}(z_1) &= -\frac{d}{dz_1} \Pr(Z > z_1) \\
 &= -\frac{d}{dz_1} \left[\frac{\frac{\pi R^2}{2} - R^2 \sin^{-1}\left(\frac{z_1}{R}\right) - z_1 R \sqrt{1 - \left(\frac{z_1}{R}\right)^2}}{\pi R^2} \right] \\
 &= \frac{1}{\pi R^2} \left(\frac{R}{\sqrt{1 - \left(\frac{z_1}{R}\right)^2}} + \sqrt{R^2 - z_1^2} - \frac{z_1^2}{\sqrt{R^2 - z_1^2}} \right) \\
 &= \frac{1}{\pi R^2} \left(\frac{R^2}{\sqrt{R^2 - z_1^2}} + \sqrt{R^2 - z_1^2} - \frac{z_1^2}{\sqrt{R^2 - z_1^2}} \right) \\
 &= \frac{2\sqrt{R^2 - z_1^2}}{\pi R^2}.
 \end{aligned} \tag{4.28}$$

In a similar way, the pdf of the random variable Z_2 is given by

$$\begin{aligned}
 f_{Z_2}(z_2) &= -\frac{d}{dz_2} \left[\frac{\frac{\pi R_2^2}{2} - R_2^2 \sin^{-1}\left(\frac{z_2}{R_2}\right) - z_2 \sqrt{R_2^2 - z_2^2}}{\pi R_2^2} \right] \\
 &= \frac{2\sqrt{R_2^2 - z_2^2}}{\pi R_2^2}.
 \end{aligned} \tag{4.29}$$

We have the relation $Z_3 = Z_1 - Z_2$. Thus the pdf of Z_3 can be derived as

$$f_{Z_3}(z_3) = \frac{4}{\mathcal{A}_c \mathcal{A}_2} \int_{-R}^R \sqrt{[R_2^2 - (z_3 - z_1)^2][R^2 - z_1^2]} dz_1, \quad (4.30)$$

where $\mathcal{A}_c = \pi R^2$ is the area of the cell and $\mathcal{A}_2 = \pi R_2^2$ is the area of the cluster. Similarly, we have

$$f_{Z_4}(z_4) = \frac{4}{\mathcal{A}_c \mathcal{A}_2} \int_{-R}^R \sqrt{[R_2^2 - (2R - z_4 - z_1)^2][R^2 - z_1^2]} dz_1. \quad (4.31)$$

On the other hand, the angles φ_3 and φ_4 should have uniform distributions with pdfs given by

$$f_{\Phi_3}(r_1) = \frac{1}{2 \tan^{-1}(R_2/r_1)}, \quad f_{\Phi_4}(r_3) = \frac{1}{2 \tan^{-1}(R_2/r_3)}. \quad (4.32)$$

Note that the distributions of φ_3 and φ_4 depend on the position of the tier-one VP and the cluster radius. Hence $Y_3 = \sec(\varphi_3)$ and $Y_4 = \sec(\varphi_4)$ have the pdfs given by

$$f_{Y_3}(y) = \frac{1}{2|y| \tan^{-1}(R_2/r_1) \sqrt{y^2 - 1}} \quad (4.33)$$

and

$$f_{Y_4}(y) = \frac{1}{2|y| \tan^{-1}(R_2/r_3) \sqrt{y^2 - 1}}. \quad (4.34)$$

We have the relation $d = z_3 Y_3$, where d is the product of Y_3 and Y_3 . Based on the probability theory [105], the distribution of d can be obtained as follows:

$$f(d, r_1) = \int_{-R}^R f_{Z_3}(z_3) \times f_{Y_3}\left(\frac{d}{z_3}\right) \frac{1}{|z_3|} dz_3. \quad (4.35)$$

Similarly, the distribution of d_o is given by

$$f(d_o, r_3) = \int_R^{3R} f_{Z_4}(z_4) \times f_{Y_4}\left(\frac{d_o}{z_4}\right) \frac{1}{|z_4|} dz_4. \quad (4.36)$$

By substituting (4.29) and (4.33) into (4.35), (4.30) and (4.34) into (4.36), followed by some parameter transformations, we can obtain the results in (4.26) and (4.27). Note that the final pdf is a joint distribution of the variables d and r_1 or d_o and r_3 , because the UT is distributed according to the position of its VP. ■

The radius of the cluster region, R_2 , is related to the properties of the specific places. Thus, R_2 is assumed to have some appropriate values such that all the UTs are located neither in the cell-hole area nor outside the cell itself. Moreover, different user clusters are allowed to have some overlapping regions.

Let random variable Φ denote the downtilt angle, then ϕ is a realization of Φ . The downtilt depends on user's position and is defined as $\phi = \tan^{-1}((h_{BS} - h_{UT})/d)$ with $\phi \in (0, \pi/2)$, where h_{BS} and h_{UT} denote the height of BS and UT, respectively. The downtilt is a strictly decreasing function of the distance d . According to the probability theory [106], the distribution of Φ can be derived as

$$\begin{aligned} f_{\Phi}(\phi, r_1) &= 1 - f_{\mathcal{D}_{in}}\left(\frac{h_{BS} - h_{UT}}{\tan \phi}, r_1\right) \\ &= 1 - \frac{2}{\mathcal{A}_c \mathcal{A}_2} \int_{-R}^R \left[\int_{-R}^R \sqrt{[R_2^2 - (z_3 - z_1)^2][R^2 - z_1^2]} dz_1 \right] \times \\ &\quad \frac{1}{|\tan((h_{BS} - h_{UT})/\phi)| \tan^{-1}(R_2/r_1) \sqrt{(\tan((h_{BS} - h_{UT})/\phi)/z_3)^2 - 1}} dz_3. \end{aligned} \quad (4.37)$$

4.2.3 Region-Partition Beamforming

In massive antenna array networks, antennas are to be deployed in the 2D plane or in the 3D space. For a 2D URA antenna topology shown in Fig. 4.8, there are two types of AoA of multipath propagation, i.e., the elevation angle θ and the azimuth angle φ . The beamformer should be adapted to these AoAs accordingly. We propose two practical

3D beamforming strategies. The objective is to control the antenna downtilt so that the vertical beams can adapt to user's location.

Field trials have been conducted to investigate the performance of 3D beamforming [53], [57], [58]. User-specific beamforming (direct steering) is to point the vertical main lobe towards the UT being served. However, it is not yet available in practice due to the specific hardware constraints [57], [58]. Hence, we take it as a theoretically optimal scheme. In the following, two practical 3D beamforming strategies are proposed.

Essentially, the purpose of the beamforming is to match the signal well with the channel characteristics. Based on its nature, the 3D beamforming tends to be a location-dependent technique. The location information can be obtained through either the AoA estimation or the user positioning technique in cellular networks [107]. It is appropriate to assume that BSs have perfect CSI containing both the AoAs information and channel coefficients. Then, the 3D MRT beamformer is considered as a function of the antenna downtilt ϕ , which is given by

$$\begin{aligned}\mathbf{w}_{jkl}(\phi_k) &= \text{vec}(\mathbf{H}_{jkl}(\phi_k)) \\ &= \frac{1}{\sqrt{P}} \text{vec} \left(\boldsymbol{\zeta}(\phi_k) \sum_{p=1}^P \mathbf{a}^T(\theta_{kp}, \phi_{kp}) \right),\end{aligned}\tag{4.38}$$

where $\boldsymbol{\zeta}(\phi_k) = [1 \ e^{-j\frac{2\pi}{\lambda}d_a \sin \phi_k} \dots e^{-j\frac{2\pi}{\lambda}(M-1)d_a \sin \phi_k}]^T$. Then we normalize it as $\mathbf{w}_{jkl}/\|\mathbf{w}_{jkl}\|$ to obtain the 3D MRT beamformer. For the optimal direct steering strategy, the downtilt angle is selected as $\phi = \theta$. In practice, however, ϕ can only be adjusted in a step of 1° [58].

In a cell with radius $R = 500$ m, antenna array height of $h_{\text{BS}} = 40$ m and UT height of $h_{\text{UT}} = 1.5$ m, the cumulative distribution function (CDF) of the downtilt is shown in Fig. 4.10. The simulations match well with analytical results obtained from Proposition

Fig. 4.10 CDF of antenna downtilt for the two-tier clustered distribution, $h_{BS} = 40m$ and $h_{UT} = 1.5m$.

4.1, and most of the downtilt angles are less than 30° . On the other hand, the downtilt cannot be too small as there will be strong inter-cell interference. After calculating the half-power beam width (HPBW) of the vertical beam pattern, we confine the downtilt in the range $[6^\circ 30^\circ]$. The downtilt ϕ_k for the specific UT is determined in the following steps:

1. Assuming the antenna space is $d_a = \lambda/2$, the corresponding HPBW of the vertical beam pattern is given by $HPBW^\circ = 318.96/\pi M$ [10].
2. For $\Phi = [6^\circ, 7^\circ, \dots, 30^\circ]$, calculate the inner and outer downtilt coverage radius associated with each downtilt as $R_{in/out,i} = (h_{BS} - h_{UT})/\tan(\Phi_i \pm HPBW^\circ/2)$, where i represents the index of Φ .

3. For the k th UT in the j th cell with $k = 1, \dots, K$, the downtilt ϕ_k assigned to it is selected as

$$\Phi_i = \arg \min_i \left| d_{jkj} - \frac{R_{\text{in},i} + R_{\text{out},i}}{2} \right|. \quad (4.39)$$

In a nutshell, the cell is partitioned into some regions, each of which is associated with a particular downtilt angle. After calculating the distances from a UT to the center of each region, the UT will be put into the region with the shortest distance. Thus, we call it the region-partition beamforming strategy.

4.2.4 Statistic-Based Beamforming

In some situations, the BS only has the distribution information of the UT's position instead of the exact position information. The downtilt needs to be designed based on the statistical knowledge accordingly. For the 3D MRT beamforming in the vertical plane, the objective is to minimize the difference between ϕ_k and $\theta_{0,k}$, where $\theta_{0,k}$ denotes the elevation angle of the line-of-sight (LOS) propagation for the k th UT. Thus, we adopt the MMSE method to determine the downtilt as follows:

$$\arg \min_{\phi_k} \mathbb{E} [(\phi_k - \theta_{0,k})^2], \text{ for } k = 1, \dots, K. \quad (4.40)$$

The well-known optimal result is $\phi_k^* = \mathbb{E}\{\theta_{0,k}\}$. As we have obtained the pdf of user's position in a two-tier clustered distribution model in Section 4.2.2, and $\theta_{0,k} = \tan^{-1}((h_{\text{BS}} - h_{\text{UT}})/d_{jkj})$, the expectation of $\theta_{0,k}$ can be obtained after some mathematical manipulations. The statistical knowledge of the two-tier user distribution may play a vital role in the design of statistic-based 3D beamforming strategy.

Fig. 4.11 Spatial capacity comparison among different beamforming strategies.

4.2.5 Numerical Results and Discussions

We consider a three-cell network and conduct simulations to evaluate the performance of the proposed 3D beamforming strategies. The radius of the cell is $R = 500$ m, and the cell-hole radius is $r_0 = 30$ m. The BS height is assumed to be $h_{BS} = 40$ m, and all the UTs have the height of 1.5 m. Users are distributed according to the proposed two-tier clustered distribution. In all simulation settings, we assume a total system bandwidth of $B = 20$ MHz. The carrier frequency is set at 3 GHz. The path-loss exponent is $\gamma = 3.7$, and the shadow-fading standard deviation is $\sigma_{shad} = 8.0$ dB for the urban macro-cell [102]. The antenna arrays are in the URA topology with the antenna spacing $d_a = \lambda/2$.

In Fig. 4.11, the performance of spatial capacity of different 3D beamforming strategies is evaluated. We consider a 20×10 URA topology at BSs. In each cell,

Fig. 4.12 BER for different beamforming strategies.

there are five clusters and each cluster contains six UTs according to the two-tier UT distribution, resulting in $K = 30$ for each cell. Since our objective is to compare the performance of different beamforming schemes, we assume that the BSs have perfect CSI. The user-specific beamforming is taken as the optimal strategy. As for the region-partition beamforming, the downtilt for each UT is determined according to the process in Section 4.2.3. BSs use UT's distribution information to implement the statistic-based beamforming. The statistic-based beamformer is obtained according to both (4.38) and (4.40). For the 2D beamforming, the downtilt is fixed as $\phi = \tan^{-1}(38.5/250) \approx 8.8^\circ$ for all the beamformers. It can be observed that the proposed 3D beamforming strategies perform much better than the conventional 2D beamforming. The performance of the region-partition beamforming is very close to that of user-specific beamforming. As the

performance loss of the statistic-based beamforming is acceptable, it can be used as an alternative when the UT's location information is not available.

The performance of bit error rate (BER) per user of different beamforming schemes is compared in Fig. 4.12. The simulation settings remain the same. The proposed region-partition beamforming strategy achieves nearly the same BER with user-specific beamforming scheme. Both of the proposed 3D beamforming strategies have better BER performance than that of the conventional 2D beamforming.

4.2.6 Conclusions

We have derived the pdf of a two-tier clustered user distribution model. We have also investigated the performance of different 3D beamforming strategies with 2D massive antenna deployment. The effectiveness of the proposed strategies have been validated by numerical results.

Chapter 5

Limited Feedback for Beamforming

Most of the multi-antenna techniques require the channel state information at the transmitter (CSIT) to improve throughput and reliability. In an FDD system, CSIT can be obtained over finite-rate feedback links. The assumption of the perfect CSIT is almost impossible to be realized at the BSs, because of the finite bandwidth of feedback links. There has been some work on limited feedback in massive MIMO systems [69]. Conventional limited feedback approaches, either based on channel quantization or Grassmannian line packing, do not scale to massive MIMO systems because a large number of antennas yields high computational complexity.

Recently, the compressive sensing (CS) technique [110]–[112] has been applied to signal processing and communication systems, such as ultra-wideband signal processing, radar detection, and the channel estimation [113], [114]. In particular, some researchers have applied the CS to reduce the feedback in conventional multi-antenna systems [115]–[118]. CS exploits the sparsity of a signal to reduce the sampling rate and the amount of data generated. If a signal is sparse in a particular transformation domain, it can be reconstructed exactly from a small number of linear measurements with a high

accuracy. Various recovery algorithms with good robustness and low computational complexity have been proposed to reconstruct the signal. As for limited feedback schemes for massive multi-antenna systems, CS is more efficient than some conventional quantization approaches. If the long channel vector can be represented in a sparse form, users may compress and send it to BS via finite-rate feedback links.

In this chapter, we propose a novel efficient limited feedback scheme based on the CS technique for a massive multi-antenna system. The CSI of downlink channels is compressed through linear measurement and fed back to the BS. The CS measurement is modelled as a quantization procedure that generates quantization errors. Numerical results show that the proposed novel limited feedback scheme outperforms the conventional VQ and Grassmannian line packing schemes.

5.1 System Description

5.1.1 Downlink Signal Model

We consider the downlink of a single cell large-scale MISO network. The BS is equipped with N transmit antennas and there are K mobile users (MUs)¹. For simplicity, we assume that each MU has one receive antenna. The received signal of user k can be represented as

$$y_k = \sqrt{p_k} \gamma_k \mathbf{h}_k^H \mathbf{w}_k s_k + \underbrace{\sum_{j \in \mathcal{U}, j \neq k} \sqrt{p_j} \gamma_k \mathbf{h}_k^H \mathbf{w}_j s_j}_{\text{inter-user interference}} + n_k \quad (5.1)$$

¹Note that all the notations are only applicable to this chapter.

Limited Feedback for Beamforming

where $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ represents the channel from the BS to MU_k. $s_k \sim \mathcal{CN}(0, 1)$ is the data symbol intended for MU_k, where $\mathbb{E}\{|s_k|^2\} = 1$. The transmit power allocated to user k is denoted as p_k , and $\mathbf{w}_k \in \mathbb{C}^{N \times 1}$ is the beamforming vector that has unit norm $\|\mathbf{w}_k\|^2 = 1$. Uniform power allocation strategy is assumed. The selected set of users is given by $\mathcal{U} = \{\pi(1), \dots, \pi|\mathcal{U}|\}$. $n_k \sim \mathcal{CN}(0, 1)$ is the independent complex additive white Gaussian noise. γ_k denotes the path loss from the BS to MU_k, and $p_k \gamma_k$ is the corresponding receive power. We consider a spatially-correlated flat and Rayleigh fading channel [119]. Since each MU is assumed to have a single antenna, we only need to consider correlation at the transmit antenna array. The correlated channel vector is given by

$$\mathbf{h}_k = \mathbf{R}_t^{\frac{1}{2}} \mathbf{h}_{iid} \quad (5.2)$$

for $k = 1, \dots, K$, where the vector $\mathbf{h}_{iid} \in \mathbb{C}^{N \times 1}$ consists of i.i.d. complex Gaussian elements with zero mean and unit variance. The matrix $\mathbf{R}_t \in \mathbb{R}^{N \times N}$ is the correlation matrix at the transmitter, which follows the Clarke model widely used in the literature [120]. The entry of $\mathbf{R}_t$ is given by

$$r_{ij} = J_0\left(\frac{2\pi d_{ij}}{\lambda}\right), \quad (5.3)$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind and d_{ij} is the distance between the i th and j th antenna. For the massive multi-antenna system, the antenna array can be dense. Thus, it is appropriate to assume that the channel has relatively high correlation. We also make an ideal assumption that each user has full knowledge of downlink channels through perfect channel estimation, since our focus is on the part of channel quantization. This is a common assumption made in the literature of limited feedback [119]. Note that in [119], the authors have discussed the CS limited feedback

Fig. 5.1 Massive MISO downlink model with limited feedback, user scheduling and beamforming.

approach in point-to-point massive multi-antenna systems with perfect feedback link condition. We extend their work by considering multi-user massive multi-antenna systems over a noisy feedback link.

We assume that single-user decoding (SUD) is applied, so the receiver will treat inter-user interference (IUI) as noise rather than decoding it. Then the SINR at the k th MU is given by

$$\text{SINR}_k = \frac{p_k \gamma_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{1 + \sum_{j \in \mathcal{U}, j \neq k} p_j \gamma_k |\mathbf{h}_k^H \mathbf{w}_j|^2} \quad (5.4)$$

for $k = 1, 2, \dots, K$. The achievable instantaneous transmission rate at MU_k is $R_k = \log_2(1 + \text{SINR}_k)$. The sum rate of the active users in the massive MISO system is given by

$$R_s = \sum_k \log_2(1 + \text{SINR}_k). \quad (5.5)$$

5.1.2 User Scheduling and Beamforming

As shown in [46], if the number of active users is much smaller than the number of transmit antenna, a simple linear precoding scheme such as ZF beamforming is able to approach the optimal performance. Therefore, we assume that the BS selects $|\mathcal{U}| \ll N$ out of K users. As finding the optimal user set requires an exhaustive search, we apply the semi-orthogonal user selection (SUS) method presented in [63]. For any user k , the direction of its channel $\tilde{\mathbf{h}}_k = \mathbf{h}_k / \|\mathbf{h}_k\|$ is quantized to $\hat{\mathbf{h}}_k$. The i th user is selected from the candidate set

$$\mathcal{S}_i = \{k : |\hat{\mathbf{h}}_k^H \hat{\mathbf{h}}_{\pi(j)}| \leq \varepsilon, j = 1, \dots, i-1\} \quad (5.6)$$

as

$$\pi(i) = \arg \max_{k \in \mathcal{S}_i} \|\mathbf{h}_k\|^2, \quad (5.7)$$

where ε is a design parameter reflecting the orthogonality between quantized channels. In other words, the user scheduler will select those MUs that are ε -orthogonal to each other. Further details of the SUS method can be found in [63]. The system model is shown in Fig. 5.1.

After the user set is selected, we apply ZF beamforming at the BS. The idea of ZF beamforming is to force the IUI to be zero. The unit-norm ZF beamforming vector for the k th user is given by [7]

$$\mathbf{w}_k = \frac{\left(\mathbf{I} - \sum_{l=1}^{o_k} \Pi_{\mathbf{e}_k^{(l)}} \right) \hat{\mathbf{h}}_k}{\left\| \left(\mathbf{I} - \sum_{l=1}^{o_k} \Pi_{\mathbf{e}_k^{(l)}} \right) \hat{\mathbf{h}}_k \right\|}, \quad k \in \mathcal{U} \quad (5.8)$$

5.2 Compressive Sensing Limited Feedback

where $\mathbf{e}_k^{(1)}, \dots, \mathbf{e}_k^{(o_k)}$ is an orthogonal basis of $\text{span}(\bigcup_{\bar{k} \neq k} \{\hat{\mathbf{h}}_{\bar{k}}\})$ and o_k is the dimension of the orthogonal base. In this way, the k th beamforming vector lies in the null-spaces of all the other channels. Since the channels are independent of each other with high probability, ZF beamformer in (5.8) ensures perfect IUI suppression, i.e., $\hat{\mathbf{h}}_j^H \mathbf{w}_k = 0$ for $j \in \mathcal{U}$ and $j \neq k$. Then the achievable sum rate of the system under ZF beamforming and ideal SUD is given by

$$R_s = \sum_{k \in \mathcal{U}} \log \left| 1 + \frac{p_k \gamma_k \mathbf{h}_k^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_k}{1 + \sum_{j \in \mathcal{U}, j \neq k} p_j \gamma_k \mathbf{h}_k^H \mathbf{w}_j \mathbf{w}_j^H \mathbf{h}_k} \right|. \quad (5.9)$$

The instantaneous sum rate is an important metric for cellular networks because it reflects the total network throughput. Maximizing the sum rate by applying proper preprocessing schemes is an ongoing research topic. Since we focus on the design of a limited feedback strategy, we will use the SUS and ZF beamforming to facilitate our analysis.

5.2 Compressive Sensing Limited Feedback

5.2.1 Background Knowledge

Over the past several years, CS has been widely applied to signal and image processing areas [110], [112]. CS exploits a sparse structure present in the signal. The basic idea of CS is that if we capture the non-zero components of the sparse signal through very few measurements/samples, we are able to recover the exact signal with high accuracy. Basically, the CS theory involves three major steps: sparse representation, measurements taking and signal recovery via the l_1 -norm minimization.

Limited Feedback for Beamforming

Consider an $N \times 1$ non-sparse and discrete-time signal $\mathbf{x}$, it can be represented in terms of an orthonormal basis as follows:

$$\mathbf{x} = \mathbf{\Psi}\mathbf{u} \quad (5.10)$$

where $\mathbf{x}$ is transformed to $\mathbf{u}$ by the $N \times N$ orthonormal matrix $\mathbf{\Psi}$ with $u_i = \langle \mathbf{x}, \boldsymbol{\psi}_i \rangle = \boldsymbol{\psi}_i^T \mathbf{x}$. Note that u_i is the i th element of $\mathbf{u}$ and $\boldsymbol{\psi}_i$ is the i th column of matrix $\mathbf{\Psi}$. The signal $\mathbf{x}$ is called S -sparse if it can be represented by a set of coefficients u_i where only S coefficients are non-zero with $S \ll N$. The orthonormal basis $\mathbf{\Psi}^T$ is called the dictionary or transformation basis. Typical examples of dictionary include discrete cosine transform (DCT) matrix and discrete Fourier transform (DFT) matrix.

The measurement is taken through random projection at a rate lower than the Nyquist rate. This can be realized by multiplying the signal with a measurement matrix. Consider an $M \times N$ measurement matrix $\boldsymbol{\Phi} = [\phi_1, \dots, \phi_N]$ with $M \ll N$ and $M > S$. The noisy measurement can be obtained by

$$\mathbf{f} = \boldsymbol{\Phi}\mathbf{\Psi}\mathbf{u} + \mathbf{z} = \boldsymbol{\Theta}\mathbf{u} + \mathbf{z} \quad (5.11)$$

where $\boldsymbol{\Theta} = \boldsymbol{\Phi}\mathbf{\Psi}$ is the compression matrix, and $\mathbf{z} \in \mathbb{R}^{M \times 1}$ is an error term. According to the CS theory, when $\boldsymbol{\Theta}$ satisfies the so-called restricted isometry property (RIP) [121], one can reconstruct the signal with high accuracy from the incomplete measurement $\mathbf{f}$. Popular measurement matrices that satisfy the RIP condition are Gaussian and Bernoulli matrices who have the i.i.d. Gaussian and Bernoulli entries, respectively. They are typically incoherent to most transforms and are also physically realizable. We use a Gaussian random matrix as the measurement matrix $\boldsymbol{\Phi}$ since it is easy to generate, and it is universal in the sense that $\boldsymbol{\Theta} = \boldsymbol{\Phi}\mathbf{\Psi}$ will be i.i.d Gaussian and thus have the RIP

5.2 Compressive Sensing Limited Feedback

with high probability regardless of the choice of Ψ [111]. The number of measurement satisfies

$$M \geq CS \log(N/S) \quad (5.12)$$

where C is a positive constant. On the other hand, the measurements of real-world application are often inaccurate, e.g., contaminated by noise, hence one cannot obtain the measurement with arbitrary precision. Typically, it is perturbed by a small noise vector with $\mathbb{E}\{\|\mathbf{z}\|^2\} = \kappa^2$. To reconstruct the signal $\mathbf{x}$ from $\mathbf{f}$, the CS decoder needs to solve the following convex optimization problem:

$$\begin{aligned} \min_{\mathbf{u}} \quad & \|\mathbf{u}\|_{l_1} \\ \text{s.t.} \quad & \|\Phi\Psi\mathbf{u} - \mathbf{f}\|^2 \leq \kappa^2, \end{aligned} \quad (5.13)$$

where the l_1 -norm is defined as $\|\mathbf{u}\|_{l_1} = \sum_i |u_i|$. Following that, $\mathbf{x}$ will be recovered using the relation $\mathbf{x} = \Psi\mathbf{u}$. Generally, there are two main types of recovery algorithms. The first one is based on greedy pursuits such as compressive sampling Matching Pursuit (CoSaMP) [122]. Another scheme is based on the convex optimization technique such as interior-point methods. It has been shown in [121] that the solution $\hat{\mathbf{u}}$ to the above convex programming is a good approximation to signal $\mathbf{u}$ with $\|\mathbf{u} - \hat{\mathbf{u}}\|^2 \leq D\kappa^2$, where D is a small constant. Both CoSaMP and convex programming algorithms have good stability in the presence of small perturbations.

5.2.2 Compressive Sensing for Limited Feedback

In this subsection, we propose a novel limited feedback scheme based on the CS theory. Assuming that the MUs have perfect channel estimation, they will compress the CSI and feed back it to the BS. Since we have very limited rate on the reverse links, the

Limited Feedback for Beamforming

feedback accuracy will be determined by the feedback rate. We model the feedback process with a quantization effect. The received signal of the CS decoder at the BS will be contaminated by quantization noise plus the link noise, which affects the recovery quality.

Conventional channel quantization approaches are based on the VQ [5]. The channel is normally decomposed into channel direction information (CDI) and channel quality information (CQI). The CQI is defined in terms of either the channel gain or SINR, and it is often directly fed back without quantization since it is just a scalar. For quantizing the CDI, a classical way is through RVQ [62], where the receiver will search across a codebook $\mathcal{C}_k$ for $k = 1, \dots, K$ that is generated randomly with uniform distribution to find some codeword vectors with the smallest angular separation from the CDI as follows:

$$\begin{aligned}\hat{\mathbf{h}}_k &= \arg \max_{\mathbf{c}_i \in \mathcal{C}_k} |\tilde{\mathbf{h}}_k^H \mathbf{c}_i|^2 \\ &= \arg \max_{\mathbf{c}_i \in \mathcal{C}_k} \cos^2(\angle(\tilde{\mathbf{h}}_k, \mathbf{c}_i))\end{aligned}\tag{5.14}$$

where $\tilde{\mathbf{h}}_k = \mathbf{h}_k / \|\mathbf{h}_k\|$ is the CDI. With a larger size of the codebook, it is more likely to pick a vector that is close to the CDI. However, the computational complexity can be very high due to large number of antennas, which makes the approach impractical for the massive multi-antenna system.

We treat the channel $\mathbf{h}_k$ as a signal vector and transform it into its sparse representation. In signal and image processing, high correlation of the original signal data means that the data are dominated by low-frequency components in the frequency domain. The data can be transformed into frequency domain with few low-frequency components through a certain transformation such as DCT [123], where only few low-frequency components are able to approximately represent the original data. The sparsity in fre-

5.2 Compressive Sensing Limited Feedback

quency domain depends on the level of correlation of the original signal data. As the channels are assumed to have high correlation, it is appropriate to use the 2D-DCT matrix as the dictionary basis [119]. The measurement taken by DCT yields a random vector $\mathbf{u}_k = \mathbf{\Psi}^T \mathbf{h}_k$, which is a complex Gaussian vector distributed as $\mathcal{CN}(\mathbf{0}, 2\mathbf{C}^{-1})$ [9] with pdf given by

$$f(\mathbf{z}) = \frac{1}{\pi^n \det(2\mathbf{C}^{-1})} \exp(-\frac{1}{2} \mathbf{z}^H \mathbf{C} \mathbf{z}), \quad (5.15)$$

where $\mathbf{C}$ is a positive definite matrix.

We assume that there is no delay for reverse links and the BS has a perfect estimation of the reverse channel, which is a common assumption in the literature [62]. We model the CS measurement as a quantization process, where the quantization noise is determined by the feedback rate B as in the quantization theory. The received noisy CSI measurements at the BS are given by

$$\mathbf{f}_k = \mathcal{Q}(\mathbf{\Theta} \mathbf{u}_k) + \mathbf{n}_k = \mathbf{\Theta} \mathbf{u}_k + \mathbf{e}_k + \mathbf{n}_k \quad (5.16)$$

for $k \in \mathcal{U}$. $\mathcal{Q}(\cdot)$ is either the uniform or non-uniform quantization operator, $\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is the Gaussian noise of the feedback link, and $\mathbf{e}_k \in \mathbb{C}^{M \times 1}$ is i.i.d. measurement noise with element $e_i \sim \mathcal{N}(0, \sigma^2)$. The measurement noise power is limited by $\|\mathbf{e}_k\|^2 \leq \varepsilon^2$, where ε^2 is the quantization error. In the rate-distortion theory, ε^2 can also be taken as the distortion rate of the quantizer [112]. Define the quantization interval to be $\Delta = L/2^B$, where L is the interval of magnitude of quantizer input. Let $R = \frac{1}{M} \log_2(L/\Delta) = B/M$ be the rate of the quantizer codebook, the distortion rate $D(R)$ can be approximated as [112]

$$D(R) \approx \frac{M\Delta^2}{12} = \frac{ML^2}{12} 2^{-2R} \quad (5.17)$$

Limited Feedback for Beamforming

for the uniform quantization (UQ), and

$$D(R) \approx \frac{\pi\sqrt{3}M}{2}\sigma_z^2 2^{-2R} \quad (5.18)$$

for the Lloyd-max non-uniform quantization (NUQ), where σ_z^2 is the variance of quantizer input. We need to determine the interval of magnitude L and the variance σ_z^2 . Since the Gaussian matrix Φ is incoherent to the matrix Ψ , the entries of Θ are also Gaussian distributed. We know that the quantizer input $\Theta \mathbf{u}_k$ is the product of two independent Gaussian vectors. It has been shown that the distribution of the product of two independent zero-mean complex Gaussian variables $X \sim \mathcal{N}(0, \sigma_x^2)$ and $Y \sim \mathcal{N}(0, \sigma_y^2)$ is given by [124]

$$f_{XY}(z) = \frac{1}{\pi\sigma_x^2\sigma_y^2} K_0\left(\frac{z}{\sigma_x\sigma_y}\right) \quad (5.19)$$

where $K_0(r)$ is the modified Bessel function of the second kind. Note that almost all values of $K_0(r)$ lie within the range $[-4, 4]$. Therefore, the magnitude interval of the quantizer input is $L = 8/\sigma_x\sigma_y$. The variance of Z is given by: $\text{var}(Z) = E\{Z^2\} - (E\{Z\})^2 = E\{X^2\}E\{Y^2\} = \text{var}(X)\text{var}(Y)$, hence $\sigma_z^2 = \sigma_x^2\sigma_y^2$. At each BS, the CS decoder will recover the original CSI via solving the following l_1 minimization problem

$$\begin{aligned} \hat{\mathbf{u}}_k &= \arg \min \|\mathbf{u}_k\|_{l_1} \\ \text{s.t. } &\|\Theta \mathbf{u}_k - \mathbf{f}_k\|^2 \leq \varepsilon^2 + 1 \end{aligned} \quad (5.20)$$

with $\varepsilon^2 = ML^2 2^{-2R}/12$ for uniform quantization, and $\varepsilon^2 = \pi\sqrt{3}\sigma_z^2 2^{-2R}M/2$ for Lloyd-max non-uniform quantization as given in (5.17) and (5.18), respectively.

We use the iterative recovery algorithm CoSaMP proposed in [122] to solve the optimization problem. The CoSaMP is a greedy pursuit algorithm in which an approx-

5.2 Compressive Sensing Limited Feedback

imation to the optimum value is built one step at a time by making locally optimal choice at each step. Further explanation of the algorithm itself can be found in [122]. For the optimization problem in (5.20), the CoSaMP algorithm produces a $2s$ -sparse approximation $\hat{\mathbf{u}}$ that satisfies

$$\|\mathbf{u} - \hat{\mathbf{u}}\| \leq C \cdot \max \left\{ \eta, \frac{1}{\sqrt{s}} \|\mathbf{u} - \mathbf{u}_s\|_{l_1} + \sqrt{\varepsilon^2 + 1} \right\} \quad (5.21)$$

where η is a given precision parameter, and $\mathbf{u}_s$ is the best s -sparse approximation to $\mathbf{u}$. It has been shown that the CoSaMP algorithm is able to produce a good approximation to the original signal $\mathbf{u}$ with small distortion [122]. Thus, the recovered channels are close to the original CSI and the BS can obtain the CSIT with high quality.

The proposed CS-based limited feedback scheme is different from the conventional VQ-based approaches. For the RVQ approach, the quantized CSI is obtained indirectly by choosing the codeword that has the smallest angular separation from the CDI. In the proposed scheme, we directly handle the channel vectors through compression, measurement and reconstruction. Furthermore, the size of quantization codebook for RVQ is $N \times 2^B$, and the computational complexity can be high for a massive multi-antenna system. In contrast, the proposed CS-based limited feedback scheme is more efficient as it exploits the nature of the CS technique. The CSIT can be well recovered with high accuracy by exploiting the sparsity of channels with only a small number of measurements given in (5.12). Consequently, the proposed limited feedback scheme is able to yield more precise CSIT than the conventional VQ approach at the same feedback rate.

5.3 Sum-Rate Lower Bound

When the imperfect CSIT is available at the BS through limited feedback, the sum-rate performance becomes worse than the performance where the BS has perfect CSIT. A sum-rate lower bound can indicate how much performance degradation the imperfect CSIT brings. In this section, we first derive a lower bound on the SINR in (5.4) under the case where the channels of the scheduled users are ε -orthogonal, imperfect CSIT is available at the BS and ZF beamforming is applied. A lower bound of the sum rate is obtained accordingly by assuming the uniform power allocation strategy. We decompose the channel vector as

$$\mathbf{h}_k = \|\mathbf{h}_k\| \tilde{\mathbf{h}}_k = \|\mathbf{h}_k\| (\cos \theta_k \hat{\mathbf{h}}_k + \sin \theta_k \mathbf{e}_k), \quad k \in \mathcal{U} \quad (5.22)$$

where $\hat{\mathbf{h}}_k$ is the quantized channel obtained from limited feedback scheme and $\mathbf{e}_k$ is a unit vector orthogonal to $\hat{\mathbf{h}}_k$. Denote θ_k as the angle between $\tilde{\mathbf{h}}_k$ and $\hat{\mathbf{h}}_k$, and ϕ_k as the angle between $\hat{\mathbf{h}}_k$ and $\mathbf{w}_k$. We can obtain $\cos \theta_k = |\tilde{\mathbf{h}}_k^H \hat{\mathbf{h}}_k|$ and $\sin \theta_k = |\tilde{\mathbf{h}}_k^H \mathbf{e}_k|$. The relationship between these vectors is illustrated in Fig. 5.2. As the ZF beamforming vectors are chosen by the BS based on the feedback CSI, they will satisfy

$$|\hat{\mathbf{h}}_k^H \mathbf{w}_j| = 0 \quad \forall j \neq k, (j, k) \in \mathcal{U}. \quad (5.23)$$

Based on the channel decomposition in (5.22) and ZF beamforming in (5.23), the instantaneous SINR is given by

$$\text{SINR}_k = \frac{p_k \gamma_k \|\mathbf{h}_k\|^2 |\cos \theta_k \hat{\mathbf{h}}_k + \sin \theta_k \mathbf{e}_k|^2}{1 + \sum_{j \neq k} p_j \gamma_k \|\mathbf{h}_k\|^2 |\sin \theta_k \mathbf{w}_j^H \mathbf{e}_k|^2}, \quad k \in \mathcal{U}. \quad (5.24)$$

Fig. 5.2 Channel decomposition.

Following that, we will derive a lower bound on the SINR for the k th user.

Theorem 5.1: For the downlink broadcast channels, where CS limited feedback scheme is applied. The SINR at the k th active user under transmit ZF beamforming and ε -orthogonal user selection is lower bounded by

$$\text{SINR}_k \geq \frac{p_k \gamma_k \|\mathbf{h}_k\|^2 \cos^2(\theta_k + \phi)}{1 + \sum_{j \neq k} p_j \gamma_j \|\mathbf{h}_k\|^2 \sin^2 \theta_k}, \quad (k, j) \in \mathcal{U} \quad (5.25)$$

where ϕ is a constant given by $\phi = \cos^{-1} \sqrt{\frac{1-(N-1)\varepsilon}{1-(N-2)\varepsilon}}(1 + \varepsilon)$.

Limited Feedback for Beamforming

Proof. First, we derive the lower bound on the numerator of (5.24). Based on the fact that $|a + b| \geq ||a| - |b||$, we can obtain

$$\begin{aligned}
 |\mathbf{h}_k^H \mathbf{w}_k|^2 &= \|\mathbf{h}_k\|^2 \left| \cos \theta_k \hat{\mathbf{h}}_k^H \mathbf{w}_k + \sin \theta_k \mathbf{e}_k^H \mathbf{w}_k \right|^2 \\
 &\stackrel{(a)}{\geq} \|\mathbf{h}_k\|^2 \left| \cos \theta_k \hat{\mathbf{h}}_k^H \mathbf{w}_k - \sin \theta_k \mathbf{e}_k^H \mathbf{w}_k \right|^2 \\
 &\stackrel{(b)}{\geq} \|\mathbf{h}_k\|^2 \left| \cos \theta_k \cos \phi_k - \sin \theta_k \sin \phi_k \right|^2 \\
 &\stackrel{(c)}{\geq} \|\mathbf{h}_k\|^2 \left| \cos \theta_k \cos \phi - \sin \theta_k \sin \phi \right|^2 \\
 &= \|\mathbf{h}_k\|^2 \cos^2(\theta_k + \phi)
 \end{aligned} \tag{5.26}$$

where the constant $\phi = \cos^{-1} \sqrt{\frac{1-(N-1)\varepsilon}{1-(N-2)\varepsilon}}(1 + \varepsilon)$ is given in [63]. Equality holds in (a) when $\cos \theta_k \hat{\mathbf{h}}_k^H \mathbf{w}_k$ and $\sin \theta_k \mathbf{e}_k^H \mathbf{w}_k$ have opposite phase. The equality in (b) holds when $\mathbf{w}_k$, $\hat{\mathbf{h}}_k$ and $\mathbf{e}_k$ are in the same plane, which yields $\mathbf{e}_k^H \mathbf{w}_k = \sin \phi_k$. The proof of (c) can be found in [63].

Next, let us find an upper bound on the term $|\mathbf{w}_j^H \mathbf{e}_k|$ in the denominator in (5.24). $\mathbf{w}_j$ should be orthogonal to $\hat{\mathbf{h}}_k$ for $j \neq k$ as indicated in (5.23). $|\mathbf{w}_j^H \mathbf{e}_k|$ can be viewed as the projection of the unit vector $\mathbf{e}_k$ onto the ZF beamforming vector $\mathbf{w}_j$. Assume that the angle between these two vectors is ϕ_{jk} , we have $|\mathbf{w}_j^H \mathbf{e}_k| = |\cos \phi_{jk}|$. As $\mathbf{w}_j$ rotates around $\hat{\mathbf{h}}_k$ with fixed angle $\phi_j = \pi/2$, the maximum value of $|\cos \phi_{jk}|$ occurs when $\mathbf{e}_k$ and $\mathbf{w}_j$ are either in the same or in opposite directions. The upper bound is given by

$$|\mathbf{w}_j^H \mathbf{e}_k|^2 \leq |\cos(\pi)|^2 = 1. \tag{5.27}$$

By substituting the results in (5.26) and (5.27) into (5.24), the lower bound in (5.25) can be proved. ■

5.4 Numerical Results and Discussions

As we assume that uniform power allocation is applied as $p_k = p$ for $k \in \mathcal{U}$, the SINRs of different active users are independent with each other. The lower bound of the sum rate can be obtained as

$$R_s \geq \sum_k \log \left| 1 + \frac{\gamma_k \|\mathbf{h}_k\|^2 \cos^2(\theta_k + \phi)}{1/p + (S_{\mathcal{U}} - 1) \gamma_k \|\mathbf{h}_k\|^2 \sin^2 \theta_k} \right|, \quad k \in \mathcal{U} \quad (5.28)$$

where $S_{\mathcal{U}}$ is the size of $\mathcal{U}$. As shown in (5.25), the SINR is dependent on the channel gain $\|\mathbf{h}_k\|$ and θ_k . The tightness of the lower bound depends on the quality of channel quantization, which is reflected by θ_k . Since θ_k indicates how close the feedback channel $\hat{\mathbf{h}}_k$ is to the original channel $\mathbf{h}_k$, it can be seen as an indication of the feedback accuracy. As the CoSaMP algorithm in (5.21) produces a small distortion, θ_k can be a small angle in the proposed scheme. Particularly, in the high SNR region where the noise is negligible relative to interference, the values of numerator and denominator of (5.25) will be at a comparable level. Any small change of θ_k will affect the lower bound significantly. Thus, the feedback accuracy becomes a major constraint factor in the high SNR region.

5.4 Numerical Results and Discussions

In this section, we present some numerical results to evaluate the performance of our proposed CS limited feedback (CS-LF) scheme. Within a single cell, we randomly generate $K = 20$ MUs according to uniform distribution. In [63], the parameter ε is set to be 0.25 for the conventional broadcast channels. Since we assume that there is a large amount of transmit antennas in the simulations, the BS may select more users each time. Thus, we use $\varepsilon = 0.314$ as the user orthogonality threshold. The entries of measurement matrix Φ are generated using real Gaussian variables with zero mean and unit variance,

Fig. 5.3 Sum rates versus receive power with $N = 30$, $B = 8$ and compression ratio $\rho = 0.2$.

and the constant C in (5.12) is set to be 1.3 [116], [117]. For the correlated channel in (5.3), we assume that $d_{ij} = 0.3\lambda$, and receivers have perfect channel estimation. On the other hand, we apply the “urban micro-cell” model for the path-loss fading [102]. The path-loss fading factor γ_k is given by

$$\gamma_k(\text{dB}) = R - 10l \log_{10} \left(\frac{d}{d_0} \right), \quad (5.29)$$

where R is the path-loss at the reference distance d_0 , l is the path-loss exponent whose value is typically between 2.7 and 3.5 for the urban micro-cell model. In particular, we choose $d_0 = 10$ m, $R = -34.5$ dB and $l = 2.9$ in our simulations, and the shadowing effect will not be considered.

5.4 Numerical Results and Discussions

The sum rates for different receive power levels are shown in Fig. 5.3. The transmit power of each BS is increased so that the corresponding receive power varies from -5 dBm to 35 dBm. The optimal upper bound is generated by randomly generating 5,000 beamforming vectors and averaging over the highest data rate for each channel realization. The number of linear measurements for CS-LF satisfies the condition in (5.12). We select $M = 6$ so that the minimum sampling rate is achieved. The compression ratio is defined as $\rho = M/N = 0.2$. We fix the number of feedback bits to be 8, and model the measurement as NUQ because it yields a smaller quantization error. For the Grassmannian beamforming (BF) [108], the Grassmannian codebook $\mathcal{F}$ has the following form [109]

$$\mathcal{F} = \left\{ \mathbf{f}_k | \mathbf{f}_k = \frac{1}{\sqrt{N}} \left[e^{j\frac{2\pi}{L}ku_1}, \dots, e^{j\frac{2\pi}{L}ku_N} \right]^T \right\}, k = 0, \dots, L-1 \quad (5.30)$$

where $L = 2^B$ and $0 \leq u_i < L$ are the integer design parameters and should be optimized as follows

$$\mathbf{u} = \arg \max_{u_i} \min_{1 \leq k < L} \sqrt{1 - |\mathbf{f}_0^\dagger \mathbf{f}_k|^2} \quad (5.31)$$

where $\mathbf{u} = [u_1, u_2, \dots, u_N]^T$ and $\sqrt{1 - |\mathbf{f}_0^\dagger \mathbf{f}_k|^2}$ is called the chordal distance. We transform the codebook form in (5.30) to

$$\mathcal{F} = \left\{ \mathbf{f}_k | \mathbf{f}_k = \frac{1}{\sqrt{N}} \left[e^{jk\alpha_1}, \dots, e^{jk\alpha_N} \right]^T \right\}, k = 0, \dots, L-1 \quad (5.32)$$

where $0 \leq \alpha_i < 2\pi$ are the design parameters and should be optimized as

$$\boldsymbol{\alpha} = \arg \max_{\alpha_i} \min_{1 \leq k < L} \sqrt{1 - |\mathbf{f}_0^\dagger \mathbf{f}_k|^2} \quad (5.33)$$

Limited Feedback for Beamforming

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_N]^T$. Usually, exhaustive computer search or random search is applied since it is very difficult to get the optimum through analytical method. We randomly generate 5,000 sets of $\boldsymbol{\alpha}$ with α_i generated from a uniform distribution in $[0, 2\pi)$. The set $\boldsymbol{\alpha}$ that yields the maximum chordal distance in (5.33) will be selected and the sub-optimal Grassmannian codebook $\mathcal{F}$ is constructed accordingly. The quantization codebook for RVQ is randomly generated as in [62]. It is observed that the proposed CS-LF scheme outperforms Grassmannian BF and RVQ schemes consistently. In the high receive power region, the gaps between them become larger. Thus, the CS-LF scheme can feed back more precise channel information as compared to the Grassmannian BF and RVQ approaches at the same feedback rate. Moreover, the sum rate lower bound derived in (5.25) for the proposed scheme is shown to be tight across all receive power levels.

Next, we compare the sum rates between different schemes by varying the number of feedback bits B . The simulation results are shown in Fig. 5.4. The number of transmit antennas is $N = 30$ and the receive power is fixed at 15 dBm. The compression ratio remains to be 0.2. We can observe that both the NUQ and UQ models are better than the conventional Grassmannian BF and RVQ approaches at the same feedback rate. For feedback rates higher than 6, the sum rate of the proposed CS-LF scheme becomes flat. The reason is that when $B \geq 6$, the quantization errors in (5.17) and (5.18) become very small. The feedback CSI is perturbed by the noise that is much lower than the signal itself. In this situation, the input to the CS decoder is almost the same as the measurement. The sum rate lower bound for the proposed scheme is shown to be tight.

In Fig. 5.5, we compare the sum rate performance by varying the number of transmit antennas N . The number of feedback bits is fixed at 6 and the receive power is still fixed at 15 dBm. For different N , the number of CS measurements should satisfy the condition

Fig. 5.4 Sum rates for different feedback rates with $N = 30$, $\rho = 0.2$ and the receive power is set at 15 dBm.

in (5.12). In the simulation, we choose the minimum valid value of ρ because it is sufficient to recover the CSI with high accuracy according to the CoSaMP algorithm. For the Grassmannian BF and RVQ approaches, if we increase the number of antennas by one, we need to generate one more row with 2^B elements for the quantization codebook, which will increase the system complexity at both BS and users. For the CS-LF scheme, even if we increase the number of antennas, we may not need to increase the number of measurements accordingly. Therefore, we can make full use of the sparsity of signals to reduce the sampling rate and data feedback to the BS.

We investigate the impact of compression ratio on the CS-LF scheme in Fig. 5.6, where the number of transmit antennas is given by $N = 30$, the number of feedback bits

Fig. 5.5 Sum rate for different antenna numbers with $B = 6$ and the receive power is set at 15 dBm.

is $B = 8$ and the receive power is set at 15 dBm. We choose NUQ to model the CS measurement procedure. According to the condition in (5.12), the compression ratio ρ should vary from 0.2 to 1. With a higher ratio, more precise CSI can be recovered, at the cost of increasing the burden of feedback links. We observe that the minimum compression ratio of 0.2 is good enough as compared with the Grassmannian BF and RVQ approaches. In particular, the sum rate of CS-LF is much closer to the full CSIT case than the other two approaches in the high receive power region. Therefore, the proposed CS-LF scheme can provide BSs with higher quality CSI feedback at a relatively low feedback rate.

Fig. 5.6 Sum rates with different compression ratio ρ for $N = 30$ and $B = 8$.

We compare the performance of different schemes when the factor $\alpha \in [0, 1]$ is varied. In this particular setup, the factor α is given by the ratio of the path loss at distance d over the path loss at distance D , where d is the distance from a user to the BS and D is radius of the cell. The factor α reflects how close users are to the cell edge. The number of antennas is given by $N = 30$, the number of feedback bits is fixed at 8, the compression ratio is set at $\rho = 0.26$ and the receive power is fixed at 15 dBm. The simulation results are shown in Fig. 5.7. It can be observed that the CS-LF scheme achieves higher sum rate than the conventional Grassmannian BF and RVQ methods across all values of α .

Lastly, the average BER results per user of the proposed CS-LF scheme and the conventional approaches are compared in Fig. 5.8. The number of antennas is given

Fig. 5.7 Sum rates versus different factor α with $N = 30$, $B = 8$ and the receive power is set at 15 dBm.

by $N = 15$, the number of feedback bits is fixed at 8, and the compression ratio is set at $\rho = 0.33$. It can be observed that the CS-LF schemes have better BER performance than the Grassmannian BF and RVQ approaches. We can also observe that with a large number of transmit antennas, the diversity can be increased significantly. As a result, the users are able to receive the correct information with high probability.

5.5 Conclusions

We have considered the CSI acquisition problem in a massive MISO system and proposed a novel limited feedback strategy based on the compressive sensing technique. The

Fig. 5.8 Average BER for different schemes with $N = 15$, $B = 8$ and $\rho = 0.33$.

CS measurements were modelled as a quantization procedure, where the quantization noise was added to the feedback CSI. We have also derived a tight lower bound on the sum rate, assuming that ZF beamforming was applied and that the active users were semi-orthogonal. Numerical results have shown that the proposed CS-LF scheme has significant sum rate and BER improvement as compared with the Grassmannian beamforming and RVQ approaches.

Chapter 6

Distributed Beamforming for Secure Communications

Due to the broadcast nature of the wireless medium, wireless transmissions over MIMO channels are constantly vulnerable to attacks. The primary form of attack at the physical layer in wireless communications is eavesdropping, in which an unauthorized recipient intercepts the communication between a transmitter and a legitimate receiver. This problem, referred to as the wiretap channel, was first introduced and characterized by Wyner [125]. Transmitting confidential messages securely in a MIMO wiretap channel for the network resilience still remains a challenging task. An information-theoretic performance metric indicating the amount of information transmitted to the intended receiver reliably and confidentially is called secrecy rate [71]. Consequently, the typical design objective is usually to find the proper beamformer to maximize the secrecy sum rate of all the legitimate receivers to protect the wireless transmission. A lot of research efforts regarding the application of transmit beamforming have been made to improve the secrecy rate. Knowing the full or partial channel knowledge, the transmitter is able

to strike a perfect balance between ensuring the secure transmission and interfering the eavesdroppers.

Most of the existing works consider the secure wireless communications in a simple wiretap channel scenario where there are only a single transmitter, a single receiver and a single eavesdropper. In a more complicated system consisting of multiple transmitters, receivers and eavesdroppers, there exists inter-user interference and the secrecy-rate function is coupled across all the users. Thus, finding an optimal transmission scheme to maximize the secrecy sum rate requires a joint optimization of high computational complexity. In this chapter, we focus on the parallel and distributed optimization algorithm with relatively low complexity to achieve secure transmission in a multicell multi-antenna interference wiretap channel. We propose an iterative optimization algorithm by using a decomposition framework based on successive convex approximation (SCA). The existing framework is for the case of real-valued vectors, and is not proper for the secrecy-rate maximization problem that has complex-valued matrix variables. We extend the framework from the case of real-valued vectors to the case of complex-valued matrices by applying techniques from complex derivatives. The framework preserves the convexity of the objective function and linearizes the non-convex part. All the BSs solve in parallel a sequence of decoupled convex optimization problems to obtain the optimal transmit covariance matrices in the current iteration. Then the covariance matrices are updated in a gradient-based manner.

6.1 System Description

We consider downlink transmissions in a K -cell interference wiretap channel model as shown in Fig. 6.1. In each cell, a BS with M antennas transmits confidential messages to

Fig. 6.1 Multicell multi-antenna interference wiretap channel model.

its users, each of which is equipped with N antennas¹. We assume that multiple-access techniques are applied such that the users in each cell are strictly orthogonal to each other. Thus, there will be one user being served in each cell in a particular resource block. Furthermore, there also exist some multi-antenna eavesdroppers who overhear the transmission between the BSs and the legitimate users. The eavesdroppers may or may not collude with each other [126]. In the colluding scenario, all the eavesdroppers cooperate to form a virtual antenna array with multiple antennas. It represents the worst-case scenario from the perspective of secure communications. In the non-colluding scenario, the eavesdroppers do not cooperate with each other. We consider the worst-case of colluding eavesdropper, where all eavesdroppers can perform joint processing and they can be seen as a single eavesdropper with N_e antennas [127], and the non-colluding scenario will not be discussed. The received signals at the i th legitimate user and the

¹Note that all the notations are only applicable to this chapter.

eavesdropper are given by

$$\mathbf{y}_i = \mathbf{H}_{ii}\mathbf{x}_i + \sum_{j \neq i} \mathbf{H}_{ji}\mathbf{x}_j + \mathbf{n}_i \quad (6.1)$$

and

$$\mathbf{y}_e = \mathbf{G}_i\mathbf{x}_i + \sum_{j \neq i} \mathbf{G}_j\mathbf{x}_j + \mathbf{n}_e \quad (6.2)$$

respectively. The channel between the i th BS and its intended user, denoted as $\mathbf{H}_{ii} \in \mathbb{C}^{N \times M}$, is called the main channel, while $\mathbf{H}_{ji} \in \mathbb{C}^{N \times M}$ is the interference channel from the j th BS to the i th user in the other cell for $i = 1, \dots, K$. The channel coefficients between the i th BS and the colluding eavesdropper are denoted by $\mathbf{G}_i \in \mathbb{C}^{N_e \times M}$. The transmitted signal vector at the i th BS is $\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{W}_i)$, and the transmit covariance matrix is subject to some power constraints as $\text{Tr}\{\mathbf{W}_i\} \leq P_t$. The complex Gaussian noise components at the i th user and the colluding eavesdropper are denoted by $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ and $\mathbf{n}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, respectively. Note that all the important symbols used in this chapter are summarized in Table 6.1.

All the channels are assumed to be quasi-static and i.i.d. flat Rayleigh fading. It is also assumed that full CSI knowledge is available at the BSs, which is a common assumption widely used in the literature [76], [77], [81], [128], [129]. The assumption on the known CSI of the colluding eavesdropper corresponds to the situation where the eavesdropper is a known node that transmits and receives non-confidential information in other time slots [81], [128]. Alternatively, the BS may detect the eavesdropper from the local oscillator power leaked from its radio frequency front end [130]. After detecting the leakage signal, the CSI can be obtained by using channel estimation at the BS. Note that the similar assumption has been made in [131].

Table 6.1 Definition of symbols.

M	Number of antennas of the BS
N	Number of antennas of the user
N_e	Total number of antennas of eavesdroppers
K	Total number of legitimate users
$\mathbb{C}$	Complex number fields
$\mathbf{H}_{ii} \in \mathbb{C}^{N \times M}$	Channel matrix from the i th BS to its intended user
$\mathbf{H}_{ji} \in \mathbb{C}^{N \times M}$	Channel matrix from the j th BS to the i th user
$\mathbf{G}_i \in \mathbb{C}^{N_e \times M}$	Channel matrix from the i th BS to the eavesdropper
P_t	Transmit power constraint at the BS
$\mathbf{W}_i \in \mathbb{C}^{M \times M}$	Transmit covariance matrix at the i th BS
γ^n	Step size at the n th iteration
ε	Termination accuracy

From the perspective of information theory, the secrecy rate is defined as the difference between the information rate of main channel and the information rate of the wiretap channel. Taking into consideration the inter-cell interference, the achievable secrecy sum rate of the multicell interference wiretap channel is given by [82], [71]

$$\mathcal{R} \triangleq \sum_{i=1}^K \left[\log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{j \neq i} \mathbf{H}_{ji} \mathbf{W}_j \mathbf{H}_{ji}^\dagger \right)^{-1} \mathbf{H}_{ii} \mathbf{W}_i \mathbf{H}_{ii}^\dagger \right| - \log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{j \neq i} \mathbf{G}_j \mathbf{W}_j \mathbf{G}_j^\dagger \right)^{-1} \mathbf{G}_i \mathbf{W}_i \mathbf{G}_i^\dagger \right| \right]^+, \quad (6.3)$$

where $\mathbf{W}_i \succeq \mathbf{0}$, for $i = 1, \dots, K$, is the transmit covariance matrix at the i th BS, which needs to be properly designed to guarantee secure transmission. Formally, the secrecy sum-rate optimization problem can be formulated as

$$\mathcal{P}: \max_{\mathbf{W}} \mathcal{R} \quad \text{s.t.} \quad \text{Tr}\{\mathbf{W}_i\} \leq P_t, \mathbf{W}_i \succeq \mathbf{0}. \quad (6.4)$$

It can be observed that the problem $\mathcal{P}$ is non-convex, and the objective function is coupled by all the users and eavesdroppers. Hence, finding the optimal covariance matrices requires centralized optimization across all the users. It is intractable to obtain

the global optimal solution and the corresponding algorithm has high computational complexity. Recent research efforts have been concentrated on finding the suboptimal solution efficiently via distributed algorithms. The main difficulty lies in approximating the non-convex objective function and performing the optimization in a distributed manner with limited signaling among the users. Existing distributed optimization algorithms usually include SCA algorithms, subgradient methods and primal-dual decomposition methods. In the following, based on a decomposition framework, we propose an efficient parallel and distributed algorithm that falls into the category of the SCA method.

6.2 A Decomposition Framework Based on SCA

Recently, some advances in the parallel optimization of multi-user systems offer a new decomposition framework for the parallel and distributed optimization of non-convex sum-utility functions arising from the design of wireless communication systems [132]–[133]. Consider the SCA-based approach for optimization of a multiuser system consisting of I coupled users. The objective is to properly select each user's strategy so that a sum-utility function is optimized. In some cases, the objective function associated with each user is the difference between two convex functions. It is referred to as difference convex (DC) programming, and the system design is formulated as

$$\begin{aligned} \min_{\mathbf{x}} \quad & U(\mathbf{x}) = \sum_{i=1}^I [f_i(\mathbf{x}) - g_i(\mathbf{x})] \\ \text{s.t.} \quad & \mathbf{x}_i \in \mathcal{X}_i \quad \forall i = 1, \dots, I, \end{aligned} \tag{6.5}$$

where $\mathbf{x} \triangleq (\mathbf{x}_i)_{i=1}^I$ denotes the real strategy profile for all the users, $\mathcal{X}_i \in \mathbb{C}$ denotes the strategy set. Note that $\mathbf{x}_{-i} \triangleq (\mathbf{x}_j)_{j \neq i}$ is the real-valued strategy vector of all users except

the i th user. The functions f_i and g_i are convex and continuously differentiable on $\mathcal{X}_i$. The idea of the decomposition is to preserve the convex structure of $U(\mathbf{x})$ while partially linearizing the non-convex part. The resultant convexified function is separable in the users' strategies $\mathbf{x}_i$. Denoting the strategy profile at the n th iteration as $\mathbf{x}^n \triangleq (\mathbf{x}_i^n)_{i=1}^I$ and $\mathbf{x}_{-i}^n \triangleq (\mathbf{x}_j^n)_{j \neq i}$, we have

$$\begin{aligned} f_i(\mathbf{x}) &= f_i(\mathbf{x}_i, \mathbf{x}_{-i}^n) + [f_i(\mathbf{x}_i, \mathbf{x}_{-i}) - f_i(\mathbf{x}_i, \mathbf{x}_{-i}^n)] \\ &\approx f_i(\mathbf{x}_i, \mathbf{x}_{-i}^n) + \sum_{j \neq i} \nabla_{\mathbf{x}_j} f_j(\mathbf{x}^n)^T (\mathbf{x}_j - \mathbf{x}_j^n), \end{aligned} \quad (6.6)$$

where the Taylor-series expansion at $\mathbf{x}^n$ has been applied. Similarly, the linearization of $g_i(\mathbf{x})$ at $\mathbf{x}^n$ is given by

$$g_i(\mathbf{x}) \approx g_i(\mathbf{x}^n) + \sum_{j=1}^I \nabla_{\mathbf{x}_j} g_j(\mathbf{x}^n)^T (\mathbf{x}_j - \mathbf{x}_j^n). \quad (6.7)$$

The linearization in (6.6) and (6.7) leads to a convex approximation of $U(\mathbf{x})$. Then all the users will iteratively solve a sequence of decoupled convex optimization problems in parallel given by [132]

$$\hat{\mathbf{x}}(\mathbf{x}^n) = \arg \min_{\mathbf{x}_i \in \mathcal{X}_i} \tilde{U}(\mathbf{x}; \mathbf{x}^n) = \sum_{i=1}^I \tilde{U}_i(\mathbf{x}_i; \mathbf{x}^n), \quad (6.8)$$

where

$$\begin{aligned} \tilde{U}_i(\mathbf{x}_i; \mathbf{x}^n) &= f_i(\mathbf{x}_i, \mathbf{x}_{-i}^n) + \sum_{j \neq i} \nabla_{\mathbf{x}_j} f_j(\mathbf{x}^n)^T (\mathbf{x}_j - \mathbf{x}_j^n) - \left[g_i(\mathbf{x}^n) + \sum_{j=1}^I \nabla_{\mathbf{x}_j} g_j(\mathbf{x}^n)^T (\mathbf{x}_j - \mathbf{x}_j^n) \right] \\ &\quad + \frac{\tau_i}{2} \triangle \mathbf{x}_i^T \mathcal{H}_i(\mathbf{x}^n) \triangle \mathbf{x}_i \end{aligned} \quad (6.9)$$

is the convex approximation of the original function $U(\mathbf{x})$. The optimal users' strategies in the n th iteration are denoted by $\hat{\mathbf{x}}(\mathbf{x}^n)$. The second term on the right-hand side (RHS) of (6.9) corresponds to the partial linearization on $f_i(\mathbf{x})$. The third term (in square bracket) represents the linearization of $g_i(\mathbf{x})$, which removes the non-convex part of the sum-utility function $U(\mathbf{x})$. The last term is a proximal-like regularization term, where τ_i is a given non-negative constant, and $\Delta \mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_i^n$. The matrix $\mathcal{H}_i(\mathbf{x}^n)$ is a positive definite matrix, the exact expression of which is problem dependent. The purpose of the last term is to relax the convergence condition or improve the convergence speed. Within such a decomposition framework, each function $\tilde{U}_i(\mathbf{x}_i; \mathbf{x}^n)$ depends only on $\mathbf{x}_i$ given the values of $\mathbf{x}^n$ and is convex in $\mathbf{x}_i$. Hence, it is separable in users' strategy profile $\mathbf{x}_i$.

6.3 Extension of the Framework

The decomposition framework mentioned in the previous subsection is generally applicable for real-valued vector variables. However, in wireless multi-antenna communication systems, we often deal with the optimization problem whose variables are complex-valued vectors or matrices. It is natural to consider the extension of the framework to the case of complex-valued variables. In practical applications, the real-valued function of complex variables can be written as $f : z \in \mathbb{C} \rightarrow f(z) \in \mathbb{R}$. If f is considered as a function of two independent real-valued scalar variables $z_R = \Re(z)$ and $z_I = \Im(z)$, the derivative and conjugate derivative of f can be formally defined as [134]

$$\frac{\partial f}{\partial z} = \frac{1}{2} \left(\frac{\partial f(z)}{\partial z_R} - j \frac{\partial f(z)}{\partial z_I} \right) \quad (6.10)$$

and

$$\frac{\partial f}{\partial z^*} = \frac{1}{2} \left(\frac{\partial f(z)}{\partial z_R} + j \frac{\partial f(z)}{\partial z_I} \right), \quad (6.11)$$

respectively, where $j = \sqrt{-1}$ is the complex operator. Subsequently, the definition of derivatives of the function f with respect to its complex variables is introduced. When f is a scalar function of complex matrix $\mathbf{Z}$, the derivative and conjugate derivative of f at $\mathbf{Z}_0$ are defined as [134]

$$\nabla_{\mathbf{Z}} f(\mathbf{Z}_0) = \left. \frac{\partial f(\mathbf{Z})}{\partial \mathbf{Z}} \right|_{\mathbf{Z}=\mathbf{Z}_0} \quad (6.12)$$

and

$$\nabla_{\mathbf{Z}^*} f(\mathbf{Z}_0) = \left. \frac{\partial f(\mathbf{Z})}{\partial \mathbf{Z}^*} \right|_{\mathbf{Z}=\mathbf{Z}_0}, \quad (6.13)$$

respectively. The derivatives of the function $f(\mathbf{Z})$ are obtained in an element-wise manner. Thus, both $\nabla_{\mathbf{Z}} f(\mathbf{Z}_0)$ and $\nabla_{\mathbf{Z}^*} f(\mathbf{Z}_0)$ are matrices having the same size with $\mathbf{Z}$. Alternatively, we can re-arrange the elements of derivative matrices in a row vector, and define $\mathcal{D}_{\mathbf{Z}} f(\mathbf{Z}_0)$ and $\mathcal{D}_{\mathbf{Z}^*} f(\mathbf{Z}_0)$ as

$$\mathcal{D}_{\mathbf{Z}} f(\mathbf{Z}_0) = \left. \frac{\partial f(\mathbf{Z})}{\partial \text{vec}(\mathbf{Z})^T} \right|_{\mathbf{Z}=\mathbf{Z}_0} = \text{vec}(\nabla_{\mathbf{Z}} f(\mathbf{Z}_0))^T \quad (6.14)$$

and

$$\mathcal{D}_{\mathbf{Z}^*} f(\mathbf{Z}_0) = \left. \frac{\partial f(\mathbf{Z})}{\partial \text{vec}(\mathbf{Z}^*)^T} \right|_{\mathbf{Z}=\mathbf{Z}_0} = \text{vec}(\nabla_{\mathbf{Z}^*} f(\mathbf{Z}_0))^T, \quad (6.15)$$

respectively. Based on (6.14) and (6.15), the higher-order derivatives of f can be derived accordingly.

Essentially, the framework in (6.9) mainly comes from the Taylor-series expansion. The linearization part corresponds to the first-order Taylor expansions of $f_i(\mathbf{x})$ and $g_i(\mathbf{x})$. The regularization term varies from case to case. When it is dependent on the user's

strategy, one may choose the matrix $\mathcal{H}_i(\mathbf{x}^n)$ as the Hessian of $f_i(\mathbf{x})$ [133]. In other cases, $\mathcal{H}_i(\mathbf{x}^n)$ is a positive definite matrix satisfying $\mathcal{H}_i(\mathbf{x}^n) - c_{\mathcal{H}_i} \mathbf{I} \succeq \mathbf{0}$ for some $c_{\mathcal{H}_i} \geq 0$. To extend the decomposition framework to the complex case, the (second-order) Taylor expansion of the real-valued function f is shown in the following Lemma.

Lemma 6.1: Given a real-valued function with complex-valued matrix variable $f(\mathbf{Z}) : \mathbf{Z} \in \mathbb{C}^{m \times n} \rightarrow f(\mathbf{Z}) \in \mathbb{R}$, the second-order Taylor-series expansion at the point $\mathbf{Z}$ can be expressed as

$$f(\mathbf{Z} + \Delta \mathbf{Z}) = f(\mathbf{Z}) + 2\langle \Delta \mathbf{Z}, \nabla_{\mathbf{Z}} f(\mathbf{Z}) \rangle + \frac{1}{2} \text{vec}([\Delta \mathbf{Z}, \Delta \mathbf{Z}^*])^\dagger \mathcal{H}_{\mathbf{Z}\mathbf{Z}^*} f(\mathbf{Z}) \text{vec}([\Delta \mathbf{Z}, \Delta \mathbf{Z}^*]), \quad (6.16)$$

where the inner product is defined as $\langle \mathbf{A}, \mathbf{B} \rangle \triangleq \Re\{\text{Tr}(\mathbf{A}^\dagger \mathbf{B})\}$. The Hessian of $f(\mathbf{Z})$ is given by

$$\mathcal{H}_{\mathbf{Z}\mathbf{Z}^*} f(\mathbf{Z}) = \begin{bmatrix} \nabla_{\mathbf{Z}\mathbf{Z}^*}^2 f(\mathbf{Z}) & \nabla_{\mathbf{Z}^* \mathbf{Z}^*}^2 f(\mathbf{Z}) \\ \nabla_{\mathbf{Z}\mathbf{Z}}^2 f(\mathbf{Z}) & \nabla_{\mathbf{Z}^* \mathbf{Z}}^2 f(\mathbf{Z}) \end{bmatrix}. \quad (6.17)$$

Proof. To obtain the first-order Taylor expansion, we need to first introduce a real-valued coordinate representation of $f(\mathbf{Z})$. There exists a transformation between the complex space $\mathbb{C}^{m \times n}$ and the real space $\mathbb{R}^{2mn}$, which is shown as follows:

$$\mathbf{Z} \Longleftrightarrow \boldsymbol{\zeta} = \begin{bmatrix} \text{vec}(\Re(\mathbf{Z})) \\ \text{vec}(\Im(\mathbf{Z})) \end{bmatrix} \in \mathbb{R}^{2mn \times 1}. \quad (6.18)$$

Under the transformation given in (6.18), the real-valued coordinate representation of $f(\mathbf{Z})$ is denoted as $f_r(\boldsymbol{\zeta}) = f(\mathbf{Z})$. Then the connection between the Jacobian of $f_r(\boldsymbol{\zeta})$ and the Jacobian of $f(\mathbf{Z})$ can be established. By definition, the Jacobian of $f_r(\boldsymbol{\zeta})$ is

given by

$$\begin{aligned}\mathcal{D}_{\boldsymbol{\zeta}} f_r(\boldsymbol{\zeta}) &= \left[\frac{\partial f_r(\boldsymbol{\zeta})}{\partial \text{vec}(\Re(\mathbf{Z}))^T}, \frac{\partial f_r(\boldsymbol{\zeta})}{\partial \text{vec}(\Im(\mathbf{Z}))^T} \right] \\ &= \left[\mathcal{D}_{\boldsymbol{\zeta}_R} f_r(\boldsymbol{\zeta}), \mathcal{D}_{\boldsymbol{\zeta}_I} f_r(\boldsymbol{\zeta}) \right].\end{aligned}\quad (6.19)$$

According to definitions in (6.10)–(6.11) and (6.14)–(6.15), for any $\mathbf{Z} \iff \boldsymbol{\zeta} \in \mathbb{R}^{2mn}$, we have

$$\mathcal{D}_{\mathbf{Z}} f(\mathbf{Z}) = \frac{\partial f(\mathbf{Z})}{\partial \text{vec}(\mathbf{Z})^T} = \frac{1}{2} \left[\mathcal{D}_{\boldsymbol{\zeta}_R} f_r(\boldsymbol{\zeta}) - j \cdot \mathcal{D}_{\boldsymbol{\zeta}_I} f_r(\boldsymbol{\zeta}) \right] \quad (6.20)$$

and

$$\mathcal{D}_{\mathbf{Z}^*} f(\mathbf{Z}) = \frac{\partial f(\mathbf{Z})}{\partial \text{vec}(\mathbf{Z}^*)^T} = \frac{1}{2} \left[\mathcal{D}_{\boldsymbol{\zeta}_R} f_r(\boldsymbol{\zeta}) + j \cdot \mathcal{D}_{\boldsymbol{\zeta}_I} f_r(\boldsymbol{\zeta}) \right]. \quad (6.21)$$

At this point, the desired relationship among $\mathcal{D}_{\boldsymbol{\zeta}} f_r(\boldsymbol{\zeta})$, $\mathcal{D}_{\mathbf{Z}} f(\mathbf{Z})$ and $\mathcal{D}_{\mathbf{Z}^*} f(\mathbf{Z})$ is provided.

Following that, the first-order Taylor expansion can be derived. Given a real-valued continuously differentiable function f , the first-order Taylor expansion of $f(\mathbf{Z}) = f_r(\boldsymbol{\zeta})$ at $\mathbf{Z}$ is calculated as

$$\begin{aligned}f(\mathbf{Z} + \Delta \mathbf{Z}) - f(\mathbf{Z}) &= f_r(\boldsymbol{\zeta} + \Delta \boldsymbol{\zeta}) - f_r(\boldsymbol{\zeta}) \\ &\stackrel{(a)}{=} \mathcal{D}_{\boldsymbol{\zeta}} f_r(\boldsymbol{\zeta}) \cdot \Delta \boldsymbol{\zeta} \\ &\stackrel{(b)}{=} 2 \Re \{ \mathcal{D}_{\mathbf{Z}} f(\mathbf{Z}) \text{vec}(\Delta \mathbf{Z}) \} \\ &\stackrel{(c)}{=} 2 \langle \Delta \mathbf{Z}, \nabla_{\mathbf{Z}} f(\mathbf{Z}) \rangle,\end{aligned}\quad (6.22)$$

where we have applied the relation in (6.20) in step (b) and the property $\text{vec}(\mathbf{A})^T \text{vec}(\mathbf{B}) = \text{Tr}(\mathbf{A}^T \mathbf{B})$ in step (c). Thus, the proof of the first-order Taylor expansion of a real-valued function with complex-valued matrix variables is complete.

6.3 Extension of the Framework

Next, when computing the second-order Taylor expansion of $f(\mathbf{Z})$, the differentiation of $\mathcal{D}_{\boldsymbol{\zeta}_R} f_r(\boldsymbol{\zeta})$ is needed. Based on (6.20) and (6.21), we have

$$\mathcal{D}_{\boldsymbol{\zeta}_R} f_r(\boldsymbol{\zeta}) = \mathcal{D}_{\mathbf{Z}} f(\mathbf{Z}) + \mathcal{D}_{\mathbf{Z}^*} f(\mathbf{Z}). \quad (6.23)$$

The differentiations of $\mathcal{D}_{\mathbf{Z}} f(\mathbf{Z})$ and $\mathcal{D}_{\mathbf{Z}^*} f(\mathbf{Z})$ are given by

$$d\mathcal{D}_{\mathbf{Z}} f(\mathbf{Z}) = \nabla_{\mathbf{Z}}(\nabla_{\mathbf{Z}} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}) + \nabla_{\mathbf{Z}^*}(\nabla_{\mathbf{Z}} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}^*) \quad (6.24)$$

and

$$d\mathcal{D}_{\mathbf{Z}^*} f(\mathbf{Z}) = \nabla_{\mathbf{Z}}(\nabla_{\mathbf{Z}^*} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}) + \nabla_{\mathbf{Z}^*}(\nabla_{\mathbf{Z}^*} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}^*), \quad (6.25)$$

respectively. Utilizing the results in (6.24) and (6.25), the second-order differentiation of $f(\mathbf{Z})$ can be found as

$$\begin{aligned} d^2 f(\mathbf{Z}) &= \text{vec}(\Delta \mathbf{Z})^\dagger \cdot [\nabla_{\mathbf{Z}}(\nabla_{\mathbf{Z}} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}) + \nabla_{\mathbf{Z}^*}(\nabla_{\mathbf{Z}} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}^*)] + \\ &\quad \text{vec}(\Delta \mathbf{Z}^*)^\dagger \cdot [\nabla_{\mathbf{Z}}(\nabla_{\mathbf{Z}^*} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}) + \nabla_{\mathbf{Z}^*}(\nabla_{\mathbf{Z}^*} f(\mathbf{Z})) \cdot \text{vec}(\Delta \mathbf{Z}^*)] \\ &= [\text{vec}(\Delta \mathbf{Z}), \text{vec}(\Delta \mathbf{Z}^*)]^\dagger \begin{bmatrix} \nabla_{\mathbf{Z}}(\nabla_{\mathbf{Z}^*} f(\mathbf{Z})) & \nabla_{\mathbf{Z}^*}(\nabla_{\mathbf{Z}^*} f(\mathbf{Z})) \\ \nabla_{\mathbf{Z}}(\nabla_{\mathbf{Z}} f(\mathbf{Z})) & \nabla_{\mathbf{Z}^*}(\nabla_{\mathbf{Z}} f(\mathbf{Z})) \end{bmatrix} \begin{bmatrix} \text{vec}(\Delta \mathbf{Z}) \\ \text{vec}(\Delta \mathbf{Z}^*) \end{bmatrix} \\ &= \text{vec}([\Delta \mathbf{Z}, \Delta \mathbf{Z}^*])^\dagger \mathcal{H}_{\mathbf{Z}\mathbf{Z}^*} f(\mathbf{Z}) \text{vec}([\Delta \mathbf{Z}, \Delta \mathbf{Z}^*]). \end{aligned} \quad (6.26)$$

The matrix $\mathcal{H}_{\mathbf{Z}\mathbf{Z}^*} f(\mathbf{Z})$ represents the complex Hessian of $f(\mathbf{Z})$. Based on the results in (6.22) and (6.26), the second-order Taylor-series expansion of a real-valued function with complex-valued matrix variables can be obtained as shown in (6.16), which completes the proof of Lemma 6.1. ■

It is worth mentioning that the purpose of the regularization term in the decomposition framework is to relax the convergence condition or improve the convergence behavior. The choice of the matrix $\mathcal{H}_i(\mathbf{x}^n)$ in (6.9) is problem dependent. It can be the Hessian of $f(\mathbf{x})$, or it can be any positive definite matrix satisfying $\mathcal{H}_i(\mathbf{x}^n) - c_{\mathcal{H}_i} \mathbf{I} \succeq \mathbf{0}$ for some $c_{\mathcal{H}_i} \geq 0$ [133]. It is known that if $\mathcal{H}(\mathbf{x}_u)$ is positive definite, then $\mathbf{x}_u$ is a local minimum of $f(\mathbf{x})$ [92]. A simple choice of $\mathcal{H}_i(\mathbf{x}^n)$ is $\mathcal{H}_i(\mathbf{x}^n) = \mathbf{I}$, which leads to an approximation framework for the pricing-based optimization algorithm.

6.4 Distributed Optimization Algorithm

The sum-rate optimization problem of $\mathcal{P}$ in (6.4) is non-convex, and the objective function is coupled by users' covariance matrices. Thus, a distributed method is preferred to obtain the optimum solution efficiently. As $\mathcal{P}$ falls into the category of the DC programming, we develop a distributed and iterative optimization approach that involves solving a sequence of decoupled convex optimization problems in parallel to solve the problem $\mathcal{P}$. The algorithm is based on the extension of the decomposition framework to real-valued functions with complex-valued matrix variable shown in Lemma 6.1. The optimal covariance matrices can be found according to a gradient-based method. Given the values of $\mathbf{W}^n = (\mathbf{W}_i^n)_{i=1}^K$ in the current iteration, the feasible covariance matrices $\mathbf{W}^{n+1}$ in the next iteration are given by

$$\mathbf{W}^{n+1} = \mathbf{W}^n + \gamma^n \mathbf{D}^n, \quad (6.27)$$

where $\gamma^n \in (0, 1]$ is the step size that needs to be properly chosen to guarantee convergence. The searching directions are $\mathbf{D}^n = \widehat{\mathbf{W}}^n - \mathbf{W}^n$ with $\widehat{\mathbf{W}}^n$ being the solution of the

following convex optimization problem

$$\widehat{\mathbf{W}}^n = \max_{\mathbf{W}} \tilde{\mathcal{R}} \quad \text{s.t.} \quad \text{Tr}\{\mathbf{W}_i\} \leq P_i, \mathbf{W}_i \succeq \mathbf{0}, \quad (6.28)$$

where $\tilde{\mathcal{R}}$ is a concave approximation of the original sum-rate function. Notice that maximizing a concave function $\theta(x)$ can be re-formulated equivalently as a problem of minimizing the corresponding convex function $-\theta(x)$. Thus, the convex optimization technique contains the concave maximization as well [92]. Since we are dealing with a maximization problem, we need to find a concave approximation of the original objective function.

According to the decomposition framework, $\tilde{\mathcal{R}} = \sum_{i=1}^K \tilde{R}_i(\mathbf{W}_i; \mathbf{W}^n)$ where $\tilde{R}_i(\mathbf{W}_i; \mathbf{W}^n)$ is given by

$$\begin{aligned} \tilde{R}_i(\mathbf{W}_i; \mathbf{W}^n) = & \log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{j \neq i} \mathbf{H}_{ji} \mathbf{W}_j^n \mathbf{H}_{ji}^\dagger \right)^{-1} \mathbf{H}_{ii} \mathbf{W}_i \mathbf{H}_{ii}^\dagger \right| + \langle \Pi_1(\mathbf{W}^n), \Delta \mathbf{W}_i \rangle \\ & - \log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{j \neq i} \mathbf{G}_j \mathbf{W}_j^n \mathbf{G}_j^\dagger \right)^{-1} \mathbf{G}_i \mathbf{W}_i^n \mathbf{G}_i^\dagger \right| - \langle \Pi_2(\mathbf{W}^n), \Delta \mathbf{W}_i \rangle \\ & + \frac{\tau_i}{2} \text{vec}([\Delta \mathbf{W}_i, \Delta \mathbf{W}_i^*])^\dagger \mathcal{H}_i(\mathbf{W}^n) \text{vec}([\Delta \mathbf{W}_i, \Delta \mathbf{W}_i^*]), \end{aligned} \quad (6.29)$$

where $\Delta \mathbf{W}_i = \mathbf{W}_i - \mathbf{W}_i^n$. Given the values of $(\mathbf{W}_j^n)_{j \neq i}^K$, the first term on the RHS of (6.29) is concave in $\mathbf{W}_i$. It preserves the concave structure of the first term on the RHS in (6.3). The pricing matrix $\Pi_1(\mathbf{W}^n)$ is defined as

$$\Pi_1(\mathbf{W}^n) \triangleq \sum_{j \neq i} \nabla_{\mathbf{W}_i} \rho_j(\mathbf{W})|_{\mathbf{W}=\mathbf{W}^n}, \quad (6.30)$$

where

$$\rho_j(\mathbf{W}) = \log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{l \neq j} \mathbf{H}_{lj} \mathbf{W}_l \mathbf{H}_{lj}^\dagger \right)^{-1} \mathbf{H}_{jj} \mathbf{W}_j \mathbf{H}_{jj}^\dagger \right| \quad (6.31)$$

is non-concave in $(\mathbf{W}_i)_{i \neq j}$. By defining the interference-plus-noise covariance matrix as

$$\mathbf{R}_j = \mathbf{I} + \sum_{l \neq j} \mathbf{H}_{lj} \mathbf{W}_l \mathbf{H}_{lj}^\dagger, \quad (6.32)$$

the pricing matrix $\Pi_1(\mathbf{W}^n)$ can be differentiated as follows:

$$\begin{aligned} \nabla_{\mathbf{W}_i} \rho_j(\mathbf{W}) &= \frac{\partial}{\partial \mathbf{W}_i} \log \left| \mathbf{I} + \mathbf{R}_j^{-1} \mathbf{H}_{jj} \mathbf{W}_j \mathbf{H}_{jj}^\dagger \right| \\ &= -\text{Tr} \left\{ \mathbf{R}_j \cdot \mathbf{R}_j^{-1} \cdot \frac{\partial \mathbf{R}_j}{\partial \mathbf{W}_i} \cdot \mathbf{R}_j^{-1} \right\} + \\ &\quad \text{Tr} \left\{ \left(\mathbf{R}_j + \mathbf{H}_{jj} \mathbf{W}_j \mathbf{H}_{jj}^\dagger \right)^{-1} \frac{\partial}{\partial \mathbf{W}_i} \left(\mathbf{R}_j + \mathbf{H}_{jj} \mathbf{W}_j \mathbf{H}_{jj}^\dagger \right) \right\} \\ &= -\mathbf{H}_{ij}^\dagger \left[\mathbf{R}_j^{-1} - \left(\mathbf{R}_j + \mathbf{H}_{jj} \mathbf{W}_j \mathbf{H}_{jj}^\dagger \right)^{-1} \right] \mathbf{H}_{ij}. \end{aligned} \quad (6.33)$$

Note that we have applied the fact that $\partial \log |\mathbf{X}| = \text{Tr} \{ \mathbf{X}^{-1} \partial \mathbf{X} \}$, $\partial \mathbf{X}^{-1} = -\mathbf{X}^{-1} (\partial \mathbf{X}) \mathbf{X}^{-1}$ and $\frac{\partial}{\partial \mathbf{X}} \text{Tr} \{ \mathbf{X} \mathbf{A} \} = \mathbf{A}^T$ [135]. Hence, the pricing matrix that only depends on the values of $\mathbf{W}^n$ can be obtained as

$$\Pi_1(\mathbf{W}^n) = \sum_{j \neq i} -\mathbf{H}_{ij}^\dagger \left[\mathbf{R}_j^{-1} - \left(\mathbf{R}_j + \mathbf{H}_{jj} \mathbf{W}_j \mathbf{H}_{jj}^\dagger \right)^{-1} \right] \mathbf{H}_{ij} \Big|_{\mathbf{W}=\mathbf{W}^n}. \quad (6.34)$$

The second term on the RHS of (6.29) corresponds to the partial linearization of the first term on the RHS of (6.3). The partial linearization eliminates the non-concave part and makes it separable in users' covariance matrices $\mathbf{W}_i$. The third term together with the fourth term is the linearization of the second term on the RHS of (6.3), which removes the non-concavity of the second term on the RHS of (6.3).

6.4 Distributed Optimization Algorithm

Next, the pricing matrix $\Pi_2(\mathbf{W}^n) = \Pi_2(\mathbf{W})|_{\mathbf{W}=\mathbf{W}^n}$ in (6.29) can be obtained in a similar way with $\Pi_2(\mathbf{W})$ given by

$$\begin{aligned}\Pi_2(\mathbf{W}) &= \sum_{j \neq i} \frac{\partial}{\partial \mathbf{W}_i} \log \left| \mathbf{I} + \mathbf{T}_j^{-1} \mathbf{G}_j \mathbf{W}_j \mathbf{G}_j^\dagger \right| \\ &= \sum_{j \neq i} -\mathbf{G}_i^\dagger \mathbf{T}_j^{-1} \mathbf{G}_i + \mathbf{G}_i^\dagger \left(\mathbf{T}_j + \mathbf{G}_j \mathbf{W}_j \mathbf{G}_j^\dagger \right)^{-1} \mathbf{G}_i \\ &= \sum_{j \neq i} -\mathbf{G}_i^\dagger \left[\mathbf{T}_j^{-1} - \left(\mathbf{T}_j + \mathbf{G}_j \mathbf{W}_j \mathbf{G}_j^\dagger \right)^{-1} \right] \mathbf{G}_i,\end{aligned}\tag{6.35}$$

where $\mathbf{T}_j$ is defined as

$$\mathbf{T}_j = \mathbf{I} + \sum_{i \neq j} \mathbf{G}_i \mathbf{W}_i \mathbf{G}_i^\dagger.\tag{6.36}$$

The last term on the RHS of (6.29) is a proximal regularization term, the purpose of which is to guarantee convergence to a stationary solution of the original sum-rate optimization problem. The parameter τ_i is a given non-negative constant. The positive definite matrix $\mathcal{H}_i(\mathbf{W}^n)$ may be set as an identity matrix, or it may be considered as the Hessian matrix given by

$$\begin{aligned}\mathcal{H}_i(\mathbf{W}) &= \frac{\partial}{\partial \text{vec}([\mathbf{W}_i, \mathbf{W}_i^*])^T} \left(\frac{\partial \delta_i(\mathbf{W})}{\partial \text{vec}([\mathbf{W}_i^*, \mathbf{W}_i])^T} \right)^T \\ &= \begin{bmatrix} \nabla_{\mathbf{W}_i \mathbf{W}_i^*}^2 \delta_i(\mathbf{W}) & \nabla_{\mathbf{W}_i^* \mathbf{W}_i^*}^2 \delta_i(\mathbf{W}) \\ \nabla_{\mathbf{W}_i \mathbf{W}_i}^2 \delta_i(\mathbf{W}) & \nabla_{\mathbf{W}_i^* \mathbf{W}_i}^2 \delta_i(\mathbf{W}) \end{bmatrix},\end{aligned}\tag{6.37}$$

where $\delta_i(\mathbf{W})$ corresponds to the first term on the RHS of (6.3).

The distributed algorithm to solve problem $\mathcal{P}$ is formally described in Algorithm 6.1. The termination accuracy is denoted as ε . At each iteration, the BS is required to communicate the pricing matrix and $\mathbf{W}_{-i}^n$ with limited signaling exchange. Through the BS coordination, each BS cannot be too aggressive in choosing the respective

Distributed Beamforming for Secure Communications

Algorithm 6.1 Distributed algorithm for solving the secrecy sum-rate optimization problem in a multicell interference wiretap channel

- 1: **Data:** $n_{\max}, \varepsilon, \gamma^n \in (0, 1]$, set $n = 0$
 - 2: **while** $(\tilde{\mathcal{R}}^{n+1} - \tilde{\mathcal{R}}^n) \leq \varepsilon$ or $n < n_{\max}$ **do**
 - 3: Compute $\widehat{\mathbf{W}}^n$ by solving (6.28);
 - 4: Set $\mathbf{W}^{n+1} = \mathbf{W}^n + \gamma^n (\widehat{\mathbf{W}}^n - \mathbf{W}^n)$;
 - 5: Update the step size based on (6.38);
 - 6: $n \leftarrow n + 1$.
 - 7: **end while**
-

beamformer. As the algorithm is based on SCA for DC program, the analysis of convergence can be found in [132]. Generally, the convergence of the distributed algorithm is guaranteed by using either a constant step size or a diminishing step size [132]. We apply the latter rule in which the step size varies in a diminishing manner as

$$\gamma^{n+1} = \gamma^n (1 - \varsigma \gamma^n), \quad (6.38)$$

where $\varsigma \in (0, 1)$ is a given constant that determines the behavior of the step size. In each iteration, a sequence of decoupled convex optimization problems is solved in parallel by using some well-known convex optimization software packages such as CVX [93]. The iteration will be terminated either when the sum-rate difference between two consecutive iterations is smaller than ε which is a prescribed termination accuracy, or when it reaches the maximum iteration number. It is reasonable to expect that the proposed algorithm performs better than the conventional gradient algorithms because the structure of the original sum-rate function is fully explored.

6.4.1 Analysis of Computational Complexity

The proposed parallel and distributed algorithm requires limited signaling exchange among BSs. In this subsection, we analyze the complexity in implementing the proposed distributed algorithm in a multicell interference wiretap channel. The computational complexity is analyzed in one round, where the update of all users is considered. Each iteration of the algorithm mainly involves the computation of the pricing matrices $\Pi_1(\mathbf{W}^n)$ and $\Pi_2(\mathbf{W}^n)$. At each iteration, the interference-plus-noise covariance matrix $\mathbf{R}_j$ needs to be computed at the BS. As shown in (6.32), $\mathbf{R}_j$ is approximately equal to the summation of K components, where the matrix multiplication in each component yields a complexity of $\mathcal{O}(N^2M)$. Thus, the complexity of computing $\mathbf{R}_j$ is $\mathcal{O}(KN^2M)$. The computation of $\nabla_{\mathbf{W}_i} \rho_j(\mathbf{W})$ in (6.33) involves the matrix inversion and the summation of matrix multiplication, which yields a complexity of $\mathcal{O}(NM^2 + N^3)$. As the computation of $\Pi_1(\mathbf{W}^n)$ in (6.34) involves the summation of approximately K matrix multiplications, the complexity of computing $\Pi_1(\mathbf{W}^n)$ is therefore given by $\mathcal{O}(KNM^2)$. Similarly, the complexity of calculating $\Pi_2(\mathbf{W}^n)$ can be analyzed and obtained in the same way. Let us denote $L = \max\{N, N_e\}$. Then the overall complexity of computing the pricing matrices is given by $\mathcal{O}(KM^2L + KML^2 + L^3)$. Moreover, the computation of the two logarithmic terms in (6.29) yields a complexity of $\mathcal{O}(KL^2M + L^3)$. For simplicity, we choose $\mathcal{H}_i(\mathbf{W}^n)$ to be an identity matrix. Hence, the complexity of computing the regularization term $\|\triangle \mathbf{W}_i\|^2$ is $\mathcal{O}(M^3)$. In summary, in order to update all the transmit covariance matrices across all K BSs, the computational complexity of the proposed algorithm is given by $\mathcal{O}(K^2(ML^2 + M^2L) + K(L^3 + M^3))$.

6.5 Numerical Results and Discussions

In this section, numerical results on the performance of the proposed distributed algorithm for solving the secrecy sum-rate optimization problem are presented. For simplicity, we consider a 3-cell interference wiretap channel scenario. The numbers of antennas at the BS and the user are given by $M = 4$ and $N = 2$, respectively. The colluding eavesdropper has $N_e = 3$ antennas. The channel coefficients of all the channel matrices are i.i.d. complex Gaussian distributed as $\mathcal{CN}(0, 1)$. The channels are corrupted by additive white Gaussian noise with zero mean and unit variance. It is assumed that all the BSs have perfect CSI and equivalent power budget. Furthermore, the convex optimization problem in Algorithm 6.1 is numerically solved by using the commercially available optimization solver CVX [93].

In Fig. 6.2, the convergence behavior of the proposed distributed algorithm is evaluated in terms of the secrecy sum rate versus the number of iterations. In the simulations, the transmit power constraint at each BS is set as $P_t = 5\text{W}$. For the distributed optimization method described in Algorithms 6.1, the initial step size is set as $\gamma^0 = 0.8$, the matrix $\mathcal{H}_i(\mathbf{W}^n)$ is selected as an identity matrix, and the maximum iteration number is set as 50. We compare the proposed algorithm with the method proposed in [136], termed weighted minimum mean-square error (WMMSE) method, and the method proposed in [137], termed distributed weighted sum-rate maximization (DWSM) method. The WMMSE method is a linear transceiver design algorithm for weighted sum-rate maximization that is based on iterative minimization of weighted MSE. DWSM is a distributed weighted sum-rate maximization algorithm for MIMO ad-hoc networks under per-node transmit-power constraints. The particular description of these two algorithms are listed in Appendix A. The secrecy sum rates are obtained by averaging over 300

Fig. 6.2 Average secrecy sum rates versus number of iterations with $M = 4$, $N = 2$, $N_e = 3$, and $K = 3$.

independent channel realizations. It can be observed that the proposed algorithm shows better convergence behavior. By applying the proposed algorithm, the stationary optimal secrecy sum rate can be achieved with fewer iterations than those of the DWSM and WMMSE methods. Moreover, the optimal secrecy sum rate achieved by the proposed algorithm is better than those achieved by the WMMSE and DWSM methods.

Next, the convergence behavior of the proposed algorithm with respect to the iteration error, defined as $|\tilde{\mathcal{R}}^{n+1} - \tilde{\mathcal{R}}^n|$, is illustrated in Fig. 6.3. The system parameters are set as follows: $K = 3$, $M = 4$, $N = 2$, $N_e = 3$, $P_t = 5W$, and $\gamma^0 = 0.8$. The algorithm is implemented by using the regularization term $\mathcal{H}_i(\mathbf{W}^n)$ as an identity matrix and a Hessian matrix, respectively. For both cases, the iteration error descends quickly during

Fig. 6.3 Iteration error versus number of iterations with $M = 4$, $N = 2$, $N_e = 3$, and $K = 3$.

the first ten iterations and then converges to a value in the order of 10^{-3} . A stationary optimum solution can be achieved by the proposed distributed algorithm. It can also be observed that the convergence performance of these two cases are almost the same. The convergence speed of using Hessian matrix is a little faster than that of using an identity matrix. Since generating the Hessian matrix is complicated, it's proper to use an identity matrix for $\mathcal{H}_i(\mathbf{W}^n)$ due to its simplicity.

In Fig. 6.4, the secrecy sum rates of the proposed algorithm, the DWSM method, the WMMSE method, and the RCI precoding [82] are compared by varying the transmit power P_t at the BS. RCI is a linear precoding scheme to achieve a tradeoff among desired signal power, interference, and the information leakage with a regularization

Fig. 6.4 Average secrecy sum rates versus P_t with $M = 4$, $N = 2$, $N_e = 3$, and $K = 3$.

parameter. The form of RCI precoding matrix is given in Appendix A. The system parameters are set as follows: $K = 3$, $M = 4$, $N = 2$, $N_e = 3$, $\gamma^0 = 0.8$, and $\varepsilon = 10^{-3}$. It can be observed that the secrecy sum rate achieved by the proposed algorithm is higher than those achieved by the WMMSE and DWSM methods across all the values of P_t . Furthermore, the proposed scheme performs much better than the RCI precoding scheme.

The CPU running time per user of the three algorithms versus the number of users is compared in Table 6.2. All the algorithms are computed using the same computer platform and they have the same termination criterion. The time to complete one iteration is proportional to the computational complexity of the algorithm. Specifically, the complexities of the DWSM and WMMSE methods are given by $\mathcal{O}(K^2(L^2M +$

Table 6.2 Average CPU time versus number of users.

Number of users	5	10	15	20	25	30
CPU time (s): proposed algorithm	1.24	2.19	3.15	4.11	5.06	6.02
CPU time (s): DWSM method	1.37	2.48	3.6	4.72	5.82	6.94
CPU time (s): WMMSE method	2.26	4.47	6.71	8.95	11.17	13.41

$LM^2 + L^3) + KM^3)$ and $\mathcal{O}(K^2(L^2M + LM^2 + M^3) + KL^3)$, respectively [136]. It can be observed that the proposed algorithm has relatively shorter CPU computation time than the DWSM method and it significantly outperforms the WMMSE method. In particular, the gap becomes bigger when the number of users increases. Hence, the proposed algorithm is more suitable for the dense network in which a large number of users are distributed.

Lastly, the convergence speed of the proposed algorithm is compared with those of the DWSM and WMMSE methods for different number of users in Fig. 6.5. The system parameters are set as follows: $M = 4$, $N = 2$, $N_e = 4$, and $P_t = 5\text{W}$. The average number of iterations is taken over 200 independent channel realizations. All the algorithms are terminated when the iteration error between two consecutive iterations becomes smaller than 10^{-3} . It is observed that the proposed algorithm has higher convergence speed. The gap between our proposed scheme and the WMMSE method is about 15 iterations consistently. Note that in the DWSM method, there is only one user updating its strategy at each iteration. Hence, the number of iterations increases significantly as the number of users increases. Essentially, the better convergence behavior of the proposed algorithm results from the decomposition framework that preserves the inherent concavity while linearizing the non-concave part of the objective function.

Fig. 6.5 Average number of iterations versus number of users with $M = 4$, $N = 2$, $N_e = 4$, and $P_t = 5W$.

6.6 Conclusions

In this chapter, the secrecy sum-rate optimization problem in a multi-antenna interference wiretap channel model has been considered. We have proposed a distributed algorithm by using a decomposition framework based on SCA to solve a sequence of decoupled convex optimization problems in parallel to obtain the optimal transmit covariance matrices. We have extended the framework from the case of real-valued vectors to the case where the variables are complex-valued matrices. The proposed algorithm has been shown to outperform the existing distributed optimization algorithms in terms of the convergence behavior.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In this thesis, we have investigated the design of beamforming in various wireless MIMO systems, i.e., conventional cellular networks, massive MIMO systems and interference wiretap channels. Beamforming has been recognized as a key air interface technology in the current LTE system and the next generation communication system. By creating the desired radiation pattern of an antenna array, the inter-user interference can be mitigated and the system throughput can be enhanced dramatically.

In the current cellular networks, CoMP techniques have been applied because of the performance benefits they can bring. The problem of finding the optimal beamformer is usually formulated as an optimization problem according to the specific design objective. However, it is intractable to find the optimal solution directly. We proposed an SDR-based beamforming approach in cooperative clustered multicell systems. The original sum-rate maximization problem was transformed into subproblems through uplink-downlink duality which can be solved by using SDP technique. For massive MIMO

systems, pilot contamination is a major bottleneck for the system performance. We have applied the VMF distribution to model the 3D channel and analyzed the condition under which pilot contamination can be mitigated by exploiting the property of the channel covariance matrix. We have also proposed region-partition 3D beamforming strategy where users in a partition are assigned a particular downtilt, and statistic-based 3D beamforming when the BS only knows the statistical distribution of the user. Furthermore, limited feedback problem has been discussed when the BS cannot acquire perfect CSIT. We have proposed a novel limited feedback scheme based on the theory of compressive sensing, where the channel was compressed through linear measurement and fed back to the BS. At the BS, the channel can be recovered with high accuracy by solving a convex optimization problem. Compared with conventional VQ-based indirect approaches where CSIT is chosen from some codebooks, the proposed CS-based scheme handles the channel vector directly. Lastly, beamforming has been recognized as an effective approach to defend against the eavesdropping. We have proposed a distributed optimization method for secure communication in interference wiretap channels. The iterative algorithm was based on a decomposition framework that preserves the convexity of the original objective function and linearizing the non-convex part. The BSs can simultaneously solve a sequence of convex optimization problems until the optimal beamformers are obtained.

7.2 Future Work

In the thesis, we have discussed beamforming problem in various wireless communication systems. However, there are still some challenging issues that need to be addressed and further explored. Some possible future research works include:

Conclusions and Future Work

1. In practice, the BS is difficult to acquire perfect CSI knowledge. The transmit beamformer should be able to adapt to changes of wireless environment. Hence, robust beamforming strategies are required against the channel uncertainty in various wireless communication systems. This involves the application of robust optimization techniques which have not been covered by this thesis [92], [138].
2. The 5G network in the future tends to be dense and heterogeneous [139]. Both macro cell and small cell will coexist in the communication system. In order to mitigate the interference, the beamforming strategies in different tiers need to be coordinated and carefully designed.
3. Physical layer security of wireless communication in the 5G network remains elusive. Technical challenges driven by massive MIMO techniques, heterogeneous networks and millimeter wave communications need to be addressed to fulfill the confidential data transmission [140].

Appendix A

A.1 WMMSE Algorithm

The WMMSE method is a linear transceiver design algorithm for weighted sum-rate maximization that is based on iterative minimization of weighted MSE [136]. After tailoring the WMMSE approach for the secrecy sum-rate maximization problem, the optimization problem in (6.4) is equivalent to the following

$$\begin{aligned} \min_{\mathbf{V}, \mathbf{U}, \mathbf{A}} \quad & \sum_{i=1}^K \text{Tr}(\mathbf{A}_i \mathbf{E}_i) - \log |\mathbf{A}_i| \\ \text{s.t.} \quad & \text{Tr}\{\mathbf{V}_i \mathbf{V}_i^\dagger\} \leq P_t, \end{aligned} \tag{A.1}$$

where $\mathbf{V}_i$ is the transmit beamforming matrix such that $\mathbf{W}_i = \mathbf{V}_i \mathbf{V}_i^\dagger$. The MSE matrix $\mathbf{E}_i$ can be written as

$$\begin{aligned} \mathbf{A}_i = & \left[\mathbf{U}_i^\dagger (\mathbf{H}_{ii} - \mathbf{G}_i) - \mathbf{I} \right] \mathbf{V}_i \mathbf{V}_i^\dagger \left[\mathbf{U}_i^\dagger (\mathbf{H}_{ii} - \mathbf{G}_i) - \mathbf{I} \right]^\dagger \\ & + \sum_{j \neq i} \mathbf{U}_i^\dagger (\mathbf{H}_{ji} - \mathbf{G}_j) \mathbf{V}_i \mathbf{V}_i^\dagger (\mathbf{H}_{ji} - \mathbf{G}_j)^\dagger \mathbf{U}_i, \end{aligned} \tag{A.2}$$

where $\mathbf{U}_i$ is defined as the receive beamformer given by:

$$\mathbf{U}_i = \left[\sum_{j=1}^K (\mathbf{H}_{ji} - \mathbf{G}_j) \mathbf{V}_j \mathbf{V}_j^\dagger (\mathbf{H}_{ji} - \mathbf{G}_j)^\dagger + \mathbf{I} \right]^{-1} \mathbf{H}_{ii} \mathbf{V}_i. \quad (\text{A.3})$$

The matrix $\mathbf{A}_i$ is a weight matrix for the i th receiver, and the optimal $\mathbf{A}_i$ in each iteration is given by

$$\mathbf{A}_i = \left(\mathbf{I} - \mathbf{U}_i^\dagger \mathbf{H}_{ii} \mathbf{V}_i \right)^{-1}. \quad (\text{A.4})$$

Next, the optimal transmit beamformer in each iteration can be updated as

$$\mathbf{V}_i^{\text{opt}} = \left[\sum_{j=1}^K (\mathbf{H}_{ji} - \mathbf{G}_j)^\dagger \mathbf{U}_j \mathbf{A}_j \mathbf{U}_j^\dagger (\mathbf{H}_{ji} - \mathbf{G}_j) + \mu_i \mathbf{I} \right]^{-1} (\mathbf{H}_{ii} - \mathbf{G}_i)^\dagger \mathbf{U}_i \mathbf{A}_i. \quad (\text{A.5})$$

where $\mu_i > 0$ should be chosen such that the power constraint is satisfied. Let $\mathbf{D} \mathbf{\Lambda} \mathbf{D}^\dagger$ be the eigen-decomposition of $\sum_{j=1}^K (\mathbf{H}_{ji} - \mathbf{G}_j)^\dagger \mathbf{U}_j \mathbf{A}_j \mathbf{U}_j^\dagger (\mathbf{H}_{ji} - \mathbf{G}_j)$ and $\mathbf{\Phi} = \mathbf{D}^\dagger (\mathbf{H}_{ii} - \mathbf{G}_i)^\dagger \mathbf{U}_i \mathbf{A}_i \mathbf{U}_i^\dagger (\mathbf{H}_{ii} - \mathbf{G}_i) \mathbf{D}$. Let $[\mathbf{\Phi}]_{mm}$ denote the m th diagonal element of $\mathbf{\Phi}$. The optimal value of μ_i can be computed as

$$\sum_{m=1} \frac{[\mathbf{\Phi}]_{mm}}{([\mathbf{\Lambda}]_{mm} + \mu_i)^2} = P_t, \quad (\text{A.6})$$

where μ_i^* can be obtained using one dimensional search techniques such as bisection method. By plugging μ_i^* into $\mathbf{V}_i^{\text{opt}}$, we get $\mathbf{V}_i^{\text{opt}}$ in each iteration. The matrices $\mathbf{U}_i$, $\mathbf{V}_i$ and $\mathbf{A}_i$ will be updated iteratively until some criterion is satisfied. The WMMSE algorithm for secrecy sum-rate maximization problem is summarized in Algorithm A.1.

Algorithm A.1 WMMSE algorithm for solving the secrecy sum-rate optimization problem in a multicell interference wiretap channel

- 1: Initialize $\mathbf{V}_i$ as $\text{Tr}(\mathbf{V}_i \mathbf{V}_i^\dagger) \leq P_t$
 - 2: **Repeat**
 - 3: $\mathbf{A}_i \leftarrow \mathbf{A}_i, \quad i = 1, \dots, K$
 - 4: $\mathbf{U}_i \leftarrow \left[\sum_{j=1}^K (\mathbf{H}_{ji} - \mathbf{G}_j) \mathbf{V}_j \mathbf{V}_j^\dagger (\mathbf{H}_{ji} - \mathbf{G}_j)^\dagger + \mathbf{I} \right]^{-1} \mathbf{H}_{ii} \mathbf{V}_i, \text{ for } i = 1, \dots, K$
 - 5: $\mathbf{A}_i \leftarrow \left(\mathbf{I} - \mathbf{U}_i^\dagger \mathbf{H}_{ii} \mathbf{V}_i \right)^{-1}, \text{ for } i = 1, \dots, K$
 - 6: $\mathbf{V}_i \leftarrow \left[\sum_{j=1}^K (\mathbf{H}_{ji} - \mathbf{G}_j)^\dagger \mathbf{U}_j \mathbf{A}_j \mathbf{U}_j^\dagger (\mathbf{H}_{ji} - \mathbf{G}_j) + \mu_i^* \mathbf{I} \right]^{-1} (\mathbf{H}_{ii} - \mathbf{G}_i)^\dagger \mathbf{U}_i \mathbf{A}_i, \forall i$
 - 7: **Until** $|\sum_j \log |\mathbf{A}_j| - \sum_j \log |\mathbf{A}'_j|| \leq \varepsilon$
-

A.2 DWSM Algorithm

DWSM is a distributed weighted sum-rate maximization algorithm for MIMO ad hoc networks under per-node transmit-power constraints [137]. To solve the sum-rate maximization problem, each link in the network needs to solve a convex optimization problem per iteration, requiring only local channel knowledge from the interfering links. The secrecy sum-rate maximization problem in (6.4) can be rewritten as

$$\begin{aligned} \max_{\mathbf{W}} \quad & \sum_{i=1}^K \left[\log \left| \mathbf{I} + \mathbf{B}_i^{-1} (\mathbf{R}_i^{-1} \mathbf{H}_{ii} - \mathbf{S}_i^{-1} \mathbf{G}_i) \mathbf{W}_i (\mathbf{H}_{ii}^\dagger - \mathbf{G}_i^\dagger) \right| \right]^+ \\ \text{s.t.} \quad & \text{Tr}\{\mathbf{W}_i\} \leq P_t, \quad \mathbf{W}_i \succeq \mathbf{0}, \end{aligned} \quad (\text{A.7})$$

where $\mathbf{R}_i = \mathbf{I} + \sum_{j \neq i} \mathbf{H}_{ji} \mathbf{W}_j \mathbf{H}_{ji}^\dagger$, $\mathbf{S}_i = \mathbf{I} + \sum_{j \neq i} \mathbf{G}_j \mathbf{W}_j \mathbf{G}_j^\dagger$, and $\mathbf{B}_i = \mathbf{I} + \mathbf{S}_i^{-1} \mathbf{G}_i \mathbf{W}_i \mathbf{G}_i^\dagger$. The negative partial derivative of the secrecy rate at the i th user is given by

$$\mathbf{C}_i = \sum_{j \neq i} \mathbf{G}_j^\dagger \left[\mathbf{S}_j^{-1} \mathbf{B}_j^{-1} - \left(\mathbf{B}_j + (\mathbf{R}_j^{-1} \mathbf{H}_{jj} - \mathbf{S}_j^{-1} \mathbf{G}_j) \mathbf{W}_j (\mathbf{H}_{jj}^\dagger - \mathbf{G}_j^\dagger) \right)^{-1} \right] \mathbf{G}_j \Big|_{\mathbf{W}_i = \bar{\mathbf{W}}_i}. \quad (\text{A.8})$$

By factorizing $\mathbf{W}_i = \mathbf{F}_i \mathbf{F}_i$, the above maximization problem can be approximated by a set of K per-link problems and rewritten as

$$\max_{\mathbf{F}_i} \log \left| \mathbf{I} + \mathbf{F}_i^\dagger \left(\mathbf{H}_{ii}^\dagger - \mathbf{G}_i^\dagger \right) \mathbf{B}_i^{-1} \left(\mathbf{R}_i^{-1} \mathbf{H}_{ii} - \mathbf{S}_i^{-1} \mathbf{G}_i \right) \mathbf{F}_i \right| - \text{tr} \left\{ \mathbf{F}_i^\dagger (\mathbf{C}_i + v_i \mathbf{I}) \mathbf{F}_i \right\}, \quad (\text{A.9})$$

where v_i is the Lagrange multiplier corresponding to the power constraint. The Gauss-Seidel iteration can be adopted to obtain the optimal $\mathbf{F}_i$ and thus $\mathbf{W}_i$ accordingly [137].

The iterative algorithm is listed in Algorithm A.2 below.

Algorithm A.2 DWSM algorithm for solving the secrecy sum-rate optimization problem in a multicell interference wiretap channel

- 1: Initialize $\mathbf{W}_i = \mathbf{0}$, $\forall i$
 - 2: Repeat until convergence
 - 3: **for** $i = 1, \dots, K$ **do**
 - 4: Compute $\mathbf{C}_i$, set $\hat{v}_i = 0$, $\bar{v}_i = v_{i,\max}$
 - 5: Repeat until convergence
 - 6: $v_i = (\hat{v}_i + \bar{v}_i)/2$,
 - 7: Compute the generalized eigenmatrix $\mathbf{F}_i$ and eigenvalues $\boldsymbol{\Sigma}_i$ of the pair of
 - 8: matrices $\left(\mathbf{H}_{ii}^\dagger - \mathbf{G}_i^\dagger \right) \mathbf{B}_i^{-1} \left(\mathbf{R}_i^{-1} \mathbf{H}_{ii} - \mathbf{S}_i^{-1} \mathbf{G}_i \right)$ and $(\mathbf{C}_i + v_i \mathbf{I})$,
 - 9: Compute the normalized eigenmatrix $\bar{\mathbf{F}}_i$ of $\mathbf{F}_i$,
 - 10: Compute $\boldsymbol{\Sigma}_i^{(1)} = \bar{\mathbf{F}}_i^\dagger \left(\mathbf{H}_{ii}^\dagger - \mathbf{G}_i^\dagger \right) \mathbf{B}_i^{-1} \left(\mathbf{R}_i^{-1} \mathbf{H}_{ii} - \mathbf{S}_i^{-1} \mathbf{G}_i \right) \bar{\mathbf{F}}_i$,
 - 11: Compute $\mathbf{P}_i = (\boldsymbol{\Sigma}_i - \mathbf{I})^+ / \boldsymbol{\Sigma}_i^{(1)}$,
 - 12: If $\text{tr}\{\mathbf{P}_i\} > P_t$, set $\hat{v}_i = v_i$; otherwise set $\bar{v}_i = v_i$
 - 13: Set $\mathbf{W}_i = \bar{\mathbf{F}}_i \mathbf{P}_i \bar{\mathbf{F}}_i^\dagger$.
 - 14: **end for**
-

A.3 RCI Precoding

The regularized channel inversion (RCI) is a linear precoding scheme to achieve a tradeoff among desired signal power, interference, and the information leakage with a regularization parameter α for the multi-user MIMO downlink [82]. The RCI precoding

matrix at the i th BS is given by

$$\mathbf{Q}_i = \mathbf{H}_{ii}^\dagger \left(\mathbf{H}_{ii} \mathbf{H}_{ii}^\dagger + \alpha \mathbf{I} \right)^{-1}, \quad (\text{A.10})$$

where $\alpha = K / (3P_t + 1 + \sqrt{3P_t + 1})$.

Author's Publications

Journal Publications

- [1] Z. Zhang, K. C. Teh, and K. H. Li, "Performance analysis of two-dimensional massive antenna arrays for future mobile networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5400–5405, Nov. 2015.
- [2] Z. Zhang, K. C. Teh, and K. H. Li, "A semidefinite relaxation approach for beamforming in cooperative clustered multicell systems with novel limited feedback scheme," *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1740–1748, May 2014.
- [3] Z. Zhang, K. C. Teh, and K. H. Li, "Application of compressive sensing to limited feedback strategy in large-scale multiple-input single-output cellular networks," *IET Commun.*, vol. 8, no. 6, pp. 947–955, Apr. 2014.
- [4] Z. Zhang, K. C. Teh, and K. H. Li, "Distributed optimization for resilient transmission of confidential information in interference channels," submitted to *IEEE Trans. Veh. Technol.*, under review.
- [5] Z. Zhang, K. C. Teh, and K. H. Li, "Study of 3D beamforming strategies in cellular networks with clustered user distribution," submitted to *IEEE Trans. Veh. Technol.*, under review.

- [6] Z. Zhang, K. C. Teh, and K. H. Li, "Network throughput optimization: Parallel and distributed optimization with new decomposition framework," submitted to *IEEE Veh. Technol. Mag.*, under review.

Conference Publications

- [1] Z. Zhang, K. C. Teh, and K. H. Li, "Analysis of beam pattern and spatial capacity of massive antenna array networks with non-uniform user distribution," in Proc. *International Conference on Signal Processing & Communication System (ICSPCS)*, Gold Coast, Australia, Dec. 2014, pp. 1–6.
- [2] Z. Zhang, K. C. Teh, and K. H. Li, "Semidefinite relaxation based beamforming in clustered cooperative multicell MISO systems," in Proc. *IEEE International Conference on Communications (ICC'13)*, Budapest, Hungary, Jun. 2013, pp. 5635–5639.
- [3] Z. Zhang, K. C. Teh, and K. H. Li, "Novel channel covariance matrix feedback strategy in coordinated multicell MISO networks," in Proc. *IEEE International Conference on Information, Communications and Signal Processing (ICICS'13)*, Tainan, Dec. 2013, pp. 1–5.
- [4] Z. Zhang, K. C. Teh, and K. H. Li, "Downlink beamforming in clustered cooperative multicell MIMO systems," in Proc. *IEEE International Conference on Communication Systems (ICCS'12)*, Singapore, Dec. 2012, pp. 524–527.

Bibliography

- [1] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, “From theory to practice: An overview of MIMO space-time coded wireless systems,” *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.
- [2] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, “An overview of MIMO communications—A key to Gigabit wireless,” *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.
- [3] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, “Transmit beamforming and power control for cellular wireless systems,” *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Aug. 1998.
- [4] R. Mudumbai, D. R. Brown, U. Madhow, and H. V. Poor, “Distributed transmit beamforming: Challenges and recent progress,” *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 102–110, Feb. 2009.
- [5] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, “An overview of limited feedback in wireless communication systems,” *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1341–1364, Aug. 2008.
- [6] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, “Multi-cell MIMO cooperative networks: A new look at interference,” *IEEE J. Select. Areas Commun.*, vol. 28, no. 9, pp. 1380–1407, Dec. 2010.

- [7] E. Bjornson, R. Zakhour, D. Gesbert, and B. Ottersten, “Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI,” *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, Aug. 2010.
- [8] H. Dahrouj and W. Yu, “Coordinated beamforming for the multi-cell multi-antenna wireless system,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.
- [9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge Univ. Press, 2005.
- [10] J. Litva and T. K.-Y. Lo, *Digital Beamforming in Wireless Communications*. Artech House Publishers, 1996.
- [11] G. Caire and S. Shamai, “On the achievable throughput of a multi-antenna gaussian broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
- [12] 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.
- [13] H. Weingarten, Y. Steinberg, and S. Shamai, “The capacity region of the Gaussian multiple-input multiple-output broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [14] X. Shang, B. Chen, and M. J. Gans, “On the achievable sum rate for MIMO interference channels,” *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4313–4320, Sep. 2006.

Bibliography

- [15] 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, Jun. 2010.
- [16] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [17] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1711–1721, May 2007.
- [18] X. Wang and X. Zhang, "Linear transmission for rate optimization in MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3247–3257, Oct. 2010.
- [19] A. J. Tenenbaum and R. S. Adve, "Linear processing and sum throughput in the multiuser MIMO downlink," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2652–2662, May 2009.
- [20] S. Shi, M. Schubert, and H. Boche, "Rate optimization for multiuser MIMO systems with linear processing," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 4020–4030, Aug. 2008.
- [21] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct. 1999.
- [22] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
- [23] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292–5296, Oct. 2008.

- [24] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced," *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 26–34, Jun. 2010.
- [25] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
- [26] S. He, Y. Huang, L. Yang, A. Nallanathan, and P. Liu, "A multi-cell beamforming design by uplink-downlink max-min SINR duality," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2858–2867, Aug. 2012.
- [27] H. J. Choi, S. H. Park, S. R. Lee, and I. Lee, "Distributed beamforming techniques for weighted sum-rate maximization in MISO interfering broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1314–1320, Apr. 2012.
- [28] P. C. Weeraddana, M. Codreanu, M. Latvaaho, and A. Ephremides, "Multicell MISO downlink weighted sum-rate maximization: A distributed approach," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 556–570, Feb. 2013.
- [29] W. W. L. Ho, T. Q. S. Quek, S. Sun, and R. W. Heath, "Decentralized precoding for multicell MIMO downlink," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1798–1809, Jun. 2011.
- [30] B. L. Ng, J. S. Evans, S. V. Hanly, and D. Aktas, "Distributed downlink beamforming with cooperative base stations," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5491–5499, Dec. 2008.
- [31] Y. Huang, G. Zheng, M. Bengtsson, K.-K. Wong, L. Yang, and B. Ottersten, "Distributed multicell beamforming with limited intercell coordination," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 728–738, Feb. 2011.

Bibliography

- [32] C.-B. Chae, I. Hwang, R. W. Heath, and V. Tarokh, “Interference aware-coordinated beamforming in a multi-cell system,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3692–3703, Oct. 2012.
- [33] L. Venturino, N. Prasad, and X. Wang, “Coordinated linear beamforming in downlink multi-cell wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1451–1461, Apr. 2010.
- [34] D. H. N. Nguyen and T. Le-Ngoc, “Multiuser downlink beamforming in multicell wireless systems: A game theoretical approach,” *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3326–3338, Jul. 2011.
- [35] S. He, Y. Huang, S. Jin, and L. Yang, “Coordinated beamforming for energy efficient transmission in multicell multiuser systems,” *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4961–4971, Dec. 2013.
- [36] Z.-Q. Luo and W. Yu, “An introduction to convex optimization for communications and signal processing,” *IEEE J. Select. Areas Commun.*, vol. 24, no. 8, pp. 1426–1438, Aug. 2006.
- [37] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, “Convex optimization-based beamforming,” *IEEE Trans. Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.
- [38] Z.-Q. Luo and T.-H. Chang, “SDP relaxation of homogeneous quadratic optimization: Approximation bounds and applications,” in *Convex Optimization in Signal Processing and Communications*. Cambridge, U. K.: Cambridge Univ. Press, 2010, ch. 4, pp. 117–162.
- [39] Z.-Q. Luo, W.-K. Ma, A. So, Y. Ye, and S. Zhang, “Semidefinite relaxation of quadratic optimization problems,” *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

- [40] Y. Huang and D. P. Palomar, "Rank-constraint separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.
- [41] C. Shen, T.-H. Chang, K.-Y. Wang, Z. Qiu, and C.-Y. Chi, "Distributed robust multicell coordinated beamforming with imperfect CSI: An ADMM approach," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2988–3003, Jun. 2012.
- [42] 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.
- [43] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [44] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [45] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [46] J. Hoydis, S. 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, Feb. 2013.
- [47] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.

Bibliography

- [48] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [49] R. R. Muller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.
- [50] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- [51] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with interference in distributed large-scale MIMO systems: A statistical approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 942–953, Oct. 2014.
- [52] Y.-H. Nam, B. L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.
- [53] H. Halbauer, S. Saur, J. Koppenborg, and C. Hoek, "3D beamforming: Performance improvement for cellular networks," *Bell Labs Tech. Journal*, vol. 18, no. 2, pp. 37–56, Sep. 2013.
- [54] W. Lee, S.-R. Lee, H.-B. Kong, and I. Lee, "3D beamforming designs for single user MISO systems," in *Proc. IEEE Globecom 13*, Dec. 2013, pp. 4019–4024.
- [55] N. Seifi, J. Zhang, R. W. Heath, T. Svensson, and M. Coldrey, "Coordinated 3D beamforming for interference management in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5396–5410, Oct. 2014.
- [56] W. Zhang, Y. Wang, and P. Li, "Coordinated resource allocation with vertical beamforming in 3D MIMO-OFDMA networks," in *Proc. ICC'14*, Jun. 2014, pp. 4765–4770.

- [57] S. Saur and H. Halbauer, “Exploring the vertical dimension of dynamic beam steering,” in *Proc. Int. Workshop on Multi-Carrier Syst. and Sol. (MC-SS)*, May 2011, pp. 1–5.
- [58] H. Halbauer, J. Koppenborg, J. Holfeld, M. Danneberg, M. Grieger, and G. Fettweis, “Field trial evaluation of 3D beamforming in a multicell scenario,” in *Proc. Int. ITG Workshop on Smart Antennas (WSA’13)*, Mar. 2013, pp. 1–7.
- [59] N. Seifi, M. Coldrey, and M. Viberg, “Throughput optimization for MISO interference channels via coordinated user-specific tilting,” *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1248–1251, Aug. 2012.
- [60] D. J. Love, R. W. Heath, W. Spantipach, and M. L. Honig, “What is the value of limited feedback for MIMO channels,” *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 54–59, Oct. 2004.
- [61] O. E. Ayach and R. W. Heath, “Grassmannian differential limited feedback for interference alignment,” *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6481–6494, Dec. 2012.
- [62] C. K. Au-Yeung and D. J. Love, “On the performance of random vector quantization limited feedback beamforming in a MISO system,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 458–462, Feb. 2007.
- [63] T. Yoo, N. Jindal, and A. Goldsmith, “Multi-antenna downlink channels with limited feedback and user selection,” *IEEE J. Select. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, Sep. 2007.
- [64] R. Bhagavatula and R. W. Heath, “Adaptive limited feedback for sum-rate maximizing beamforming in cooperative multicell systems,” *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 800–811, Feb. 2011.

Bibliography

- [65] R. Bhagavatula and R. W. Heath, “Predictive vector quantization for multicell cooperation with delayed limited feedback,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2588–2597, Jun. 2013.
- [66] J. Yang and D. B. Williams, “Transmission subspace tracking for MIMO systems with low-rate feedback,” *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1629–1639, Aug. 2007.
- [67] P. Ding, D. J. Love, and M. D. Zoltowski, “Multiple antenna broadcast channels with shape feedback and limited feedback,” *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3417–3428, Jul. 2007.
- [68] H. Lee, E. Park, H. Park, and I. Lee, “Bit allocation and pairing methods for multi-user distributed antenna systems with limited feedback,” *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2905–2915, Aug. 2014.
- [69] J. Choi, Z. Chance, D. J. Love, and U. Madhow, “Noncoherent Trellis coded quantization: A practical limited feedback technique for massive MIMO systems,” *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5016–5029, Dec. 2013.
- [70] E. Karamad, B. Khoshnevis, and R. S. Adve, “Quantization and bit allocation for channel state feedback in relay-assisted wireless networks,” *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 327–339, Jan. 2013.
- [71] F. Oggier and B. Hassibi, “The secrecy capacity of the MIMO wiretap channel,” in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT’08)*, Toronto, Canada, Jul. 2008, pp. 524–528.
- [72] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, “Principles of physical layer security in multiuser wireless networks: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, 2014.

- [73] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.
- [74] S. Luo, J. Li, and A. P. Petropulu, "Uncoordinated cooperative jamming for secret communications," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 7, pp. 1081–1090, Jul. 2013.
- [75] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. L. Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, to be published.
- [76] S. A. A. Fakoorain and A. L. Swindlehurst, "MIMO interference channel with confidential messages: Achievable secrecy rates and precoder design," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 640–649, Sep. 2011.
- [77] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1678–1690, May 2014.
- [78] Y. Pei, Y. C. Liang, K. C. Teh, and K. H. Li, "Secure communication in multi-antenna cognitive radio networks with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1683–1693, Apr. 2011.
- [79] C.-H. Lin, S.-H. Tsai, and Y.-P. Lin, "Secure transmission using MIMO precoding," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 5, pp. 801–813, May 2014.
- [80] X. Wang, K. Wang, and X.-D. Zhang, "Secure relay beamforming with imperfect channel side information," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2140–2155, Jun. 2013.

Bibliography

- [81] H.-M. Wang, M. Luo, Q. Yin, and X.-G. Xia, “Hybrid cooperative beamforming and jamming for physical-layer security of two-way relay networks,” *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 12, pp. 2007–2020, Dec. 2013.
- [82] G. Geraci, M. Egan, J. Yuan, A. Razi, and I. B. Collings, “Secrecy sum rates for multi-user MIMO regularized channel inversion precoding,” *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3472–3482, Nov. 2012.
- [83] G. Geraci, H. S. Dhillon, J. G. Andrews, J. Yuan, and I. B. Collings, “Physical layer security in downlink multi-antenna cellular networks,” *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 2006–2021, Jun. 2014.
- [84] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath, “Networked MIMO with clustered linear precoding,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1910–1921, Apr. 2009.
- [85] C. T. K. Ng and H. Huang, “Linear precoding in cooperative MIMO cellular networks with limited coordination clusters,” *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1446–1454, Dec. 2010.
- [86] M. Hong, R. Sun, H. Baligh, and Z. Luo, “Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 226–240, Feb. 2013.
- [87] R. Corvaja, J. J. G. Fernandez, and A. G. Armada, “Mean achievable rates in clustered coordinated base station transmission with block diagonalization,” *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3483–3493, Aug. 2013.
- [88] P. Kerret and D. Gesbert, “Degree of freedom of the network MIMO channel with distributed CSI,” *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6806–6824, Nov. 2012.

- [89] A. Tolli, H. Pennanen, and P. Komulainen, “Decentralized minimum power multi-cell beamforming with limited backhaul signaling,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 570–580, Feb. 2011.
- [90] Y. Huang, G. Zheng, M. Bengtsson, K.-K. Wong, L. Yang, and B. Ottersten, “Distributed multicell beamforming design approaching Pareto boundary with max-min fairness,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2921–2933, Aug. 2012.
- [91] E. A. Jorswieck, E. G. Larsson, and D. Danev, “Complete characterization of the Pareto boundary for the MISO interference channel,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292–5296, Oct. 2008.
- [92] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.
- [93] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming,” 2015. Available online: <http://cvxr.com/cvx/>
- [94] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, “5G on the horizon: Key challenges for the radio-access network,” *IEEE Veh. Technol. Mag.*, vol. 8, no. 3, pp. 47–53, Jul. 2013.
- [95] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [96] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. Zhang, “What will 5G be?,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

Bibliography

- [97] K. Mammasis, R. W. Stewart, and J. S. Thompson, "Spatial fading correlation model using mixtures of Von Mises Fisher distributions," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2046–2055, Apr. 2009.
- [98] P. Ioannides and C. A. Balanis, "Uniform circular and rectangular arrays for adaptive beamforming applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 4, pp. 351–354, Sep. 2005.
- [99] J. A. Tsai, R. M. Buehrer, and B. D. Woerner, "The impact of AOA energy distribution on the spatial fading correlation of linear antenna array," in *Proc. Veh. Technol. Conf.*, May 2002, vol. 2, pp. 933–937.
- [100] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. D. Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.
- [101] K. Zarifi, S. Affes, and A. Ghayeb, "Collaborative null-steering beamforming for uniformly distributed wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1889–1903, Mar. 2010.
- [102] F. Khan, *LTE for 4G Mobile Broadband: Air Interface Technologies and Performance*. Cambridge Univ. Press, 2009.
- [103] J. Zhang, X. Ge, Z. Li, G. Mao, and Y. Yang, "Analysis of the uplink maximum achievable rate with location-dependent intercell signal interference factors based on linear Wyner model," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4615–4628, Nov. 2013.
- [104] S. K. Yong and J. S. Thompson, "A three-dimensional spatial fading correlation model for uniform rectangular arrays," *IEEE Antenna Wireless Propagation Lett.*, vol. 2, no. 2, pp. 182–185, Feb. 2003.

- [105] V. K. Rohatgi, *An Introduction to Probability Theory and Mathematical Statistics*. Wiley. 1976.
- [106] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*. Oxford Univ. Press, 3rd edition. 2001.
- [107] D. Liu, M.-C. Lee, C.-M. Pun, and H. Liu, “Analysis of wireless localization in nonline-of-sight conditions,” *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1484–1492, May 2013.
- [108] R.-A. Pitaval, H.-L. Maattanen, K. Schober, O. Tirkkonen, and R. Wichman, “Beamforming codebooks for two transmit antenna systems based on optimum Grassmannian packings,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6591–6602, Oct. 2011.
- [109] D. Love, R. W. Heath, and T. Strohmer, “Grassmannian beamforming for multiple-input multiple-output wireless systems,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Dec. 2008.
- [110] E. J. Candes and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Feb. 2008.
- [111] R. G. Baraniuk, “Compressive sensing [Lecture notes],” *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
- [112] V. K. Goyal, A. K. Fletcher, and S. Rangan, “Compressive sampling and lossy compression,” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 48–52, Feb. 2008.
- [113] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, “Compressed channel sensing: A new approach to estimating sparse multipath channels,” *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

Bibliography

- [114] S. L. H. Nguyen and A. Ghrayeb, “Compressive sensing-based channel estimation for massive multiuser MIMO systems,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, Apr. 2013, pp. 2890–2895.
- [115] X. L. Liu, C. Luo, W. Hu, and F. Wu, “Compressive broadcast in MIMO systems with receive antenna heterogeneity,” in *Proc. IEEE INFOCOM*, Orlando, US, Mar. 2012, pp. 3011–3015.
- [116] K. Kim, S. Jang, D. Kim, and J. Seo, “An efficient feedback scheme using compressive sensing for MIMO broadcast channel with random beamforming,” in *Proc. International Conference on Computer and Automation Engineering (ICCAE)*, Singapore, Feb. 2010, pp. 743–747.
- [117] R. N. Netto and C. C. Cavalcante, “Quantization and noise impact over feedback reduction of MIMO systems using compressive sensing,” in *Proc. International Symposium on Wireless Communication Systems (ISWCS)*, Paris, France, Aug. 2012, pp. 396–400.
- [118] S. T. Qaseem and T. Y. Al-Naffouri, “Compressive sensing for reducing feedback in MIMO broadcast channels,” in *Proc. IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 2010, pp. 1–5.
- [119] P. H. Kuo, H. T. Kung, and P. A. Ting, “Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Paris, France, Apr. 2012, pp. 492–497.
- [120] R. Iqbal, T. D. Abhayapala, and T. A. Lamahewa, “Generalised Clarke model for mobile-radio reception,” *IET Commun.*, vol. 3, no. 4, pp. 644–654, Apr. 2009.

- [121] E. J. Candes, J. K. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Commun. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
- [122] D. Needell and J. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, Mar. 2009.
- [123] N. Ahmed, T. Natarajan, and K. R. Rao, “Discrete cosine transform,” *IEEE Trans. Computers*, vol. 23, no. 1, pp. 90–93, Jan. 1974.
- [124] N. O’Donoughue and J. M. F. Moura, “On the product of independent complex Gaussians,” *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1050–1063, Mar. 2012.
- [125] A. D. Wyner, “The wire-tap channel,” *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
- [126] J. Yang, I.-M. Kim, and D. I. Kim, “Optimal cooperative jamming for multiuser broadcast channel with multiple eavesdroppers,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2840–2852, Jun. 2013.
- [127] G. Geraci, S. Singh, J. G. Andrews, J. Yuan, and I. B. Collings, “Secrecy rates in broadcast channels with confidential messages and external eavesdroppers,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2931–2943, May 2014.
- [128] S. A. A. Fakoorian and A. L. Swindlehurst, “Solutions for the MIMO Gaussian wiretap channel with a cooperative jammer,” *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 5013–5022, Oct. 2011.
- [129] S. Jeong, K. Lee, H. Huh, and J. Kang, “Secure transmission in downlink cellular network with a cooperative jammer,” *IEEE Wireless Commun. Lett.*, vol. 2, no. 4, pp. 463–466, Aug. 2013.

Bibliography

- [130] A. Mukherjee and A. L. Swindlehurst, “Detecting passive eavesdroppers in the MIMO wiretap channel,” in *Proc. IEEE Int. Conf. Acous., Speech and Signal Process. (ICASSP’12)*, Tokyo, Japan, Mar. 2012, pp. 2809–2812.
- [131] C. Wang, H. Wang, D. W. K. Ng, X. Xia, and C. Liu, “Joint beamforming and power allocation for secrecy in peer-to-peer relay networks,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3280–3293, Jun. 2015.
- [132] A. Alvarado, G. Scutari, and J.-S. Pang, “A new decomposition method for multiuser DC-programming and its applications,” *IEEE Trans. Signal Process.*, vol. 62, no. 11, pp. 2984–2998, Jun. 2014.
- [133] G. Scutari, F. Facchinei, L. Lampariello, and P. Song, “Parallel and distributed methods for nonconvex optimization,” in *Proc. IEEE ICASSP’14*, Florence, Italy, May 2014, pp. 840–844.
- [134] A. Hjørungnes, *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge Univ. Press, 2011.
- [135] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Tech. Univ. Denmark, 2007.
- [136] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, “An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel,” *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
- [137] S.-J. Kim and G. B. Giannakis, “Optimal resource allocation for MIMO ad hoc cognitive radio networks,” *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.
- [138] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton Univ. Press, 2009.

- [139] “5G wireless communications systems: Prospects and challenges part 2,” *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014.
- [140] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, “Safeguarding 5G wireless communication networks using physical layer security,” *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.