Spectral Efficiency Maximization of a Massive Multiuser MIMO System via Appropriate Power Allocation

Omid Saatlou

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of The Requirements

For The Degree of

Doctor of Philosophy(Electrical and Computer Engineering)

Concordia University

Montréal, Québec, Canada

April 2019

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Omid Saatlou

Entitled: Spectral Efficiency Maximization of a Massive Multiuser MIMO System Via Appropriate Power Allocation

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy (Electrical and Computer Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. Michelle Nokken

_____ External Examiner
Dr. Wasfy Boushra Mikhael

_____ External to Program
Dr. Chun-Yi Su

_____ Examiner
Dr. Yousef R. Shayan

_____ Examiner
Dr. Wei-Ping Zhu

_____ Thesis Co-Supervisor
Dr. M. Omair Ahmad

_____ Thesis Co-Supervisor
Dr. M.N.S. Swamy

Approved by _____
Dr. Rastko R. Selmic, Graduate Program Director

August 14, 2019 _____
Dr. Amir Asif, Dean
Gina Cody School of Engineering & Computer Science

Abstract

Spectral Efficiency Maximization of a Massive Multiuser MIMO System via Appropriate Power Allocation

Omid Saatlou, Ph.D.

Concordia University, 2019

Massive multiuser multiple-input multiple-output (MU-MIMO) systems are being considered for the next generation wireless networks in view of their ability to increase both the spectral and energy efficiencies. For such systems, linear detectors such as zero-forcing (ZF) and maximum-ratio combining (MRC) detectors on the uplink (UL) transmission have been shown to provide near optimal performance. As well, linear precoders such as ZF and maximum-ratio transmission (MRT) precoders on the downlink (DL) transmission offer lower complexity along with a near optimal performance in these systems.

One of the most challenging problems in massive MU-MIMO systems is obtaining the channel state information (CSI) at the transmitter as well as the receiver. In such systems, the base station (BS) obtains CSI using pilot sequences, which are transmitted by the users. Due to the channel reciprocity between the UL and DL channels in the time-division duplex (TDD) mode, BS employs CSI obtained to precode the data symbols in DL transmission. To accurately decode the received symbols in the DL transmission, the users also need to acquire CSI. In view of this, a beamforming training (BT) scheme has been proposed in the literature to obtain the estimates of CSI at each user. In this scheme, BS transmits a short pilot sequence to the users in a way such that each user estimates the effective channel gain.

Conventionally, the power of the pilot symbols has been considered equal to the power of data symbols for all the users. In this thesis, we pose and answer a basic question about the operation of a base station: How much the spectral efficiency could be improved if the transmit power allocated to the pilot and data symbols of each user are chosen in

some optimal fashion? In answering this question and in order to maximize the spectral efficiency for a given total energy budget, some methods of power allocation are proposed.

First, we derive a closed-form approximate expression for the achievable downlink rate for the maximum ratio transmission precoder based on small-scale fading in order to evaluate the spectral efficiency in the BT scheme. Then, we propose three methods of power allocation in order to maximize the spectral efficiency for a given total power budget among the users. In the first proposed method, we allocate equal pilot power as well as equal data power for all users in order to maximize the spectral efficiency. In the second proposed method, we allow for the allocation of different data powers among the users, whereas the pilot power for each user is kept the same and is specified. In the third method, we optimally allocate equal pilot power and a different data power for each user in such a way that the spectral efficiency is maximized. Numerical results are obtained showing that all the three proposed methods are superior to the existing methods in terms of spectral efficiency. In addition, they also show that the third proposed method of power allocation outperforms the other two proposed methods in terms of the spectral efficiency.

Next, we derive a closed-form approximate expression for the achievable downlink rate for the maximum ratio transmission precoder based on large-scale fading in order to evaluate the spectral efficiency in the BT scheme. Then, we propose four methods of power allocation in order to maximize the spectral efficiency for a given total power budget among the users. In the first method, power is allocated among the pilot and data symbols in such a way that the pilot power as well as the data power for each user is the same. In the second method, power is allocated among the data symbols of the various users, whereas the pilot power for each user is the same and is specified. In this method, the data power for each user is optimally determined to maximize the spectral efficiency. In the third method, power is allocated among the pilot and data symbols of the various users, whereas the pilot power for each user is the same but determined. In this method, the same pilot power along with the various data powers is optimized to maximize the spectral efficiency. Finally, in the fourth method, power is allocated optimally among each of the pilot and data symbols of the various users so as to maximize the spectral efficiency. Numerical results are obtained showing that the performance of

the first proposed method is approximately the same as that of the conventional approach. In addition, they also show that the second, third and fourth methods of power allocation yield similar performance in terms of spectral efficiency, and that the spectral efficiency of these methods is much superior to that of the first method or of the conventional method.

Finally, we investigate the spectral efficiency of massive MU-MIMO systems on an UL transmission with a very large number of antennas at the base station serving single-antenna users. A practical physical channel model is proposed by dividing the angular domain into a finite number of distinct directions. A lower bound on the achievable rate of the uplink data transmission is derived using a linear detector for each user and employed in defining the spectral efficiency. The lower bound obtained is further modified for the maximum-ratio combining and zero-forcing receivers. A power control scheme based on the large-scale fading is also proposed to maximize the spectral efficiency under the peak power constraint. Experiments are conducted to evaluate the lower bounds obtained and the performance of the proposed method. The numerical results show that the proposed power control method provides a spectral efficiency which is the same as that of the maximum power criterion using the ZF receiver. Further, the proposed method provides a spectral efficiency that is higher than that provided by the maximum power criterion using the MRC receiver.

To my parents

Acknowledgments

My most sincere thanks go to my supervisors Dr. M.N.S. Swamy and Dr. M. Omair Ahmad for their continuous support, patience and insightful guidance during different stages of my Ph.D. studies and writing of this thesis.

I am very grateful to my supervisors and Concordia University for the financial support that I received, which was very crucial for completing this research work. I would like to acknowledge the financial support from the School of Graduate Studies and Faculty of Engineering and Computer Science at Concordia University, NSERC, ReSMiQ and the supervisors for covering the costs related to this research. I acknowledge the support and help offered to me by the Department of Electrical and Computer Engineering, Concordia University.

Special thanks go to my mother for her patience, encouragement and continuous support for me. I would like also to thank my other family members who supported me and were available in times of need and eased the hardships of my life.

I wish to thank my friends and colleagues at Concordia University: Mohsen Ghaderi, Hamid nabati, Dr. Hamed Abdzadeh Ziabari, Dr. Shreyamsha Kumar Bidare, Ali Mohebbi, Alireza Esmaeilzehi and Mohammadreza Pourshahabi who walked by my side during the ups and downs of my research, shared my moments of distress and joy, and made the past five years one of the most memorable periods of my life.

Contents

List of Figures

List of Tables

List of Symbols

$\mathbf{x}^T$	Transpose of $\mathbf{x}$
A	Detector Matrix
$a(\phi)$	Steering Vector
C	Capacity
D_v	Diagonal Matrix
$E(.)$	Expected Value Operator
E_t	Consumed Energy of Each Users
F	Linear Detector Matrix for Physical Channels
$f(\phi)$	Probability Density Function of Arrival Angle
G	Propagation Matrix of Complex Valued Channel Coefficients
H	Channel Matrix
$H^\dagger$	Hermitian Transpose of Channel Matrix
H^*	Conjugate of Channel Matrix
H^Ψ	pseudo-inverse of Channel Matrix
$\hat{H}$	Estimate of H
I_{n_r}	Identity Matrix
$I(x;s)$	Mutual Information Operator
K	Number of Users
L	Number of Direction
M	Number of Antennas at BS
N_r	Number of Antennas at Receiver
N_t	Number of Antennas at Transmitter

p_d	Downlink Transmit Power
p_u	Uplink Transmit Power
p_p	Pilot Power
P_t	Total Power Budget
p^{max}	Maximum Available Power for Each User
R	Achievable Rate
$\widetilde{R}$	Lower Bound of Achievable Rate
S	Spectral Efficiency
T	Channel Coherence Interval
α	Power Normalization Factor
a_{ki}	Effective Channel Gain
$\hat{a}_{ki}$	Estimate of Effective Channel Gain
β_k	Shadow Fading Coefficient

List of Abbreviations

AWGN	Additive White Gaussian Noise
BS	Base Station
BT	Beamforming Training
CDF	Cumulative Distribution Function
CF	Cell Free
CRB	Cramer-Rao bound
CRC	Cyclic Redundancy Check
CSI	Channel State Information
DL	Down Link
EE	Energy Efficiency
FDD	Frequency Division Duplex
IID	Independent Identically Distributed
LTE	Long Term Evolution
MF	Matched Filter
MIMO	Multiple Input Multiple Output
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MU-MIMO	Multi User Multiple Input Multiple Output
OFDM	Orthogonal Frequency Division Multiplexing
PCP	Pilot Contamination Precoding
SE	Spectral Efficiency
SNR	Signal to Noise Ratio

TDD	Time Division Duplexing
UL	Up Link
ZF	Zero Forcing
WiFi	Wireless Fidelity

Chapter 1

Introduction

1.1 General

Massive Multiple-input multiple-output (MIMO) technology is being incorporated into emerging wireless broadband standards like long-term evolution (LTE) and 5G cellular networks which has been widely studied during the last decade. This technology can significantly improve the capacity and reliability of wireless systems by increasing the number of antennas at the receiver and transmitter [1, 2]. However, these are at the expense of the complexity of the hardware, the energy consumption of the signal processing at the receiver and the physical space needed to accommodate the antennas [3]. Initial researches focused on point-to-point MIMO links, where two devices with multiple antennas communicate with each other. However, in recent years, focus has shifted to more practical multi-user MIMO (MU-MIMO) systems where a base station (BS) serves multi users with single terminal antenna. In this technology, each BS is equipped with very large number of antennas, for e.g., 100 or more [4–6]. Massive MIMO requires a huge number of antennas simultaneously serving a much smaller number of terminals. Larger numbers of terminals can always be accommodated by combining very large MIMO technology with conventional time and frequency-division multiplexing via orthogonal frequency-division multiplexing (OFDM).

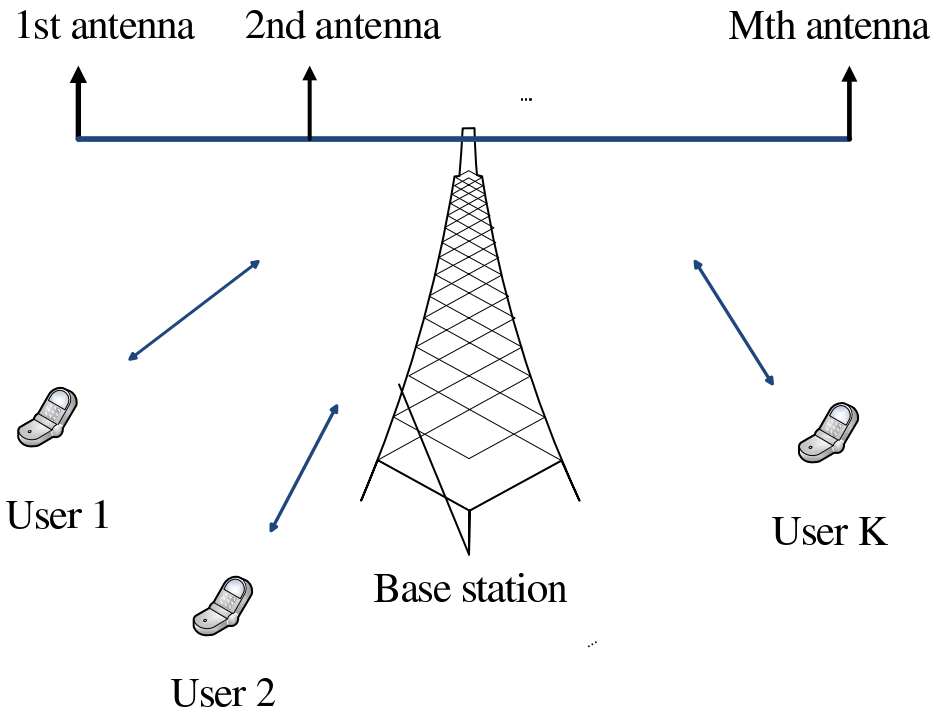

Figure 1.1: A massive MU-MIMO system.

Due to the multi-user diversity, the performance of the massive MU-MIMO systems is generally less sensitive to the propagation environment than in the point-to-point MIMO case. As a result, massive MU-MIMO has become an important part of communications standards, such as 802.11 (WiFi), 802.16 (WiMAX) and LTE. In addition, massive MU-MIMO is progressively being deployed throughout the world.

Fig. 1 shows a massive MU-MIMO system. In [1], it has been shown that the effects of the uncorrelated noise and small-scale fading are eliminated when the number of antennas at BS is very large. The number of the users per cell are independent of the size of the cell, and the required transmitted energy per bit decreases as the number of antennas in a massive MU-MIMO cell increases. Moreover, the performance of simple linear processing such as maximum-ratio combining (MRC) and zero-forcing (ZF) detections on the uplink (UL) channel, and maximum-ratio transmission (MRT) and zero-forcing (ZF) precoding on the downlink (DL) channel becomes near optimal in a massive MU-MIMO transmission. systems [7].

In [1], it has been shown that non-cooperative massive MIMO systems can achieve a data rate of 17 Mb/s for each of 40 users in a 20MHz channel in both the UL and DL transmissions with an overall spectral efficiency of 26.5 bps/Hz considering the realistic propagation assumptions.

1.2 Literature Review

The use of massive multiple-input multiple-output (MIMO) systems, where a number of users communicate with the base station (BS) with a very large number of antennas, is viable approach for achieving significant improvement in spectral efficiency (SE) [3, 5, 7–10]. It has been shown that by employing a very large number of antennas at BS, the interference among the users is canceled, the uncorrelated noise is eliminated and small-scale fading effects are averaged out [1]. In addition, linear detectors such as zero-forcing (ZF) and maximum-ratio combining (MRC) detectors on the uplink (UL) transmission have a near optimal performance along with an acceptable complexity in massive multiuser MIMO (MU-MIMO) systems. In these systems, linear precoders such as ZF and maximum-ratio transmission (MRT) precoders on the downlink (DL) transmission also offer lower complexity along with a near optimal performance [11]. Due to the aforementioned advantages, massive MIMO systems are studied for next generation of cellular networks [4, 6, 7, 12–18].

In the cellular networks such as 5G, all users occupy full time-frequency resources both in UL and DL transmissions. In a DL transmission, BS needs to ensure that each user receives only the data intended for it. In a UL transmission, BS requires to recover the individual signals transmitted by the users. In view of this, BS has to perform the huge amount of multiplexing and de-multiplexing signal processing, which is feasible using a massive number of antennas and having the channel state information (CSI) at BS [8], [4].

For DL transmission in a massive MIMO system, acquiring the channel state information (CSI) is one of the most challenging topics. BS requires CSI to transmit the precoded signal and the users also require CSI to decode the transmitted signal in DL

transmission [19]. This CSI can be estimated by received pilot signals or can be obtained through the feedback from the receiver to the transmitter. In frequency-division duplexing (FDD) operation, the users estimate CSI from DL pilots sent by BS and communicate the estimated CSI back to BS over a feedback channel [20]. This feedback is very costly in massive MU-MIMO, since the number of DL pilots is determined with the massive number of antennas at BS. On the other hand, in time-division duplexing (TDD) operation, BS estimates CSI from UL pilots sent by the users. Due to the channel reciprocity between UL and DL channels in TDD systems, once BS estimates the UL channel, it automatically has a valid estimate of the DL channel. Therefore, the pilot transmission is determined with the number of the users in TDD operation. Typically, the users are smaller than the antennas at BS in massive MIMO systems. As a result, CSI acquisition under TDD mode is more economical and preferable than FDD mode [8, 9].

In DL transmission under TDD operation, the users also need to obtain CSI to decode the received data symbols. To this end, one simple method is that BS transmits pilot sequences to the users so that each user estimates the DL channel based on the received pilot sequences. This overhead of channel estimation is not effective, since this method depends on the number of antennas at BS. In view of this, it is commonly assumed that the users are aware merely of the statistical properties of the channels and the perfect CSI is not available [21–23].

To solve this problem, in [24], the beamforming training (BT) scheme has been proposed to efficiently obtain estimates of CSI at each user in DL transmission for massive MU-MIMO systems. In the BT scheme, BS beamforms a pilot sequence such that each user is able to estimate the efficient channel gain using the minimum mean-square error (MMSE) channel estimation method. This channel estimation method depends only on the number of users. Thus, the BT scheme is preferable in DL transmission for massive MIMO systems. In [25], the BT scheme has been employed in association with the pilot contamination precoding (PCP) scheme to improve the spectral efficiency in a massive

MU-MIMO DL transmission.

One of the typical problems in wireless MIMO networks is to study as to how much training is required to estimate CSI. The effect of training sequences on the achievable rate has been investigated in [26] and [27]. In case of multiuser TDD MIMO systems, an attempt has been made in [28] to deal with the problem as to how much time should be spent in training for a given number of transmit antennas, number of receive antennas, and length of the channel coherence time. In addition, it has been shown in [29] that, by varying the transmit powers for the pilot and data sequences, the optimal number of pilot symbols is equal to the number of transmit antennas. In [30] and [31], the performance of channel estimation has been studied for different pilot symbol designs, where a lower bound on the achievable rate has been expressed as a function of the Cramer-Rao bound (CRB).

To improve the spectral efficiency of massive MU-MIMO systems, power control among the pilot sequences and payload signals is essential [32]. This power control strategy is employed in UL transmission for different purposes [32–37]. For instance, in [32], based on the large-scale fading coefficients, a power management method has been proposed by considering the channel hardening effect in massive MIMO systems. An optimal power control method that jointly optimizes the data and training signal power has been proposed in [33], in which the spectral efficiency is maximized for a given total power budget. In [34], a power control scheme based on the channel quality of each user has been proposed in order to maximize the minimum achievable rate of each user, where a MRC detector is used at BS. In [35], an optimal power control scheme over pilot and data power based on large-scale fading has been proposed to maximize the spectral efficiency, where a ZF receiver is employed. In [36], the data power among the users has been allocated in such a way that the sum rate is maximized in multi-cell massive MIMO systems. In [37], a power control strategy among the users has been proposed to maximize the spectral efficiency in a single cell massive MIMO system.

In DL transmission, power is allocated among the users by BS in order to improve the spectral efficiency of massive MU-MIMO systems [38–44]. For instance, a power allocation scheme among the users has been proposed in [42] in order to maximize the spectral efficiency under the total power constraint at BS. A power allocation among the various users with a regularized zero forcing (RZF) precoding has been studied in [38] to maximize the spectral efficiency. A method of power allocation among various users in conjunction with the BT scheme has been proposed in [39] based on the water-filling approach to maximize the spectral efficiency. In [40], a power allocation scheme between the pilot and data sequences has been proposed in order to improve the spectral efficiency, where the BT scheme is employed. In [41], a method of power allocation among each of the pilot and data symbols of all the users is proposed to maximize the spectral efficiency, where the total power budget per coherence interval for all users is given.

It is known that in practice, in UL transmission, the propagation environment and the geometry of the antenna arrays at BS have significant effect on the performance of massive MIMO systems [45, 46]. Most of the studies assume that the channel vectors for different users are independent or asymptotically orthogonal [8], [9], [23], [47]. However, in practice, the channel vectors for different users are generally correlated, or not asymptotically orthogonal, and are modeled by L-dimensional vectors, where L is the number of angular bins [48]. The reason for this is that the antennas are not sufficiently well separated or the propagation environment does not offer rich enough scattering. It has been shown in [48] that by increasing the number of antennas at BS, the system performance under a finite-dimensional channel model with L angular bins is the same as the performance under infinite-dimensional channel model with L antennas.

1.3 Motivation

Even though future wireless systems are likely to use an ever-increasing number of access points and new spectral bands, the need for maximizing the spectral efficiency in a given

band is never going to vanish. As mentioned in the above discussion, there exist a number of works on the spectral efficiency maximization in massive MIMO systems. However, none of them has considered a resource allocation scheme that jointly optimizes the pilot power, data power and duration of training in the BT scheme for a DL transmission. Since it is known that the joint optimization of the pilot power, data power and duration of training offers results in terms of the spectral efficiency maximization better than that offered by the optimization of just one of them [9], we are motivated to propose an optimal resource allocation that jointly optimizes the training duration in UL transmission, the training signal power in UL and DL transmissions, and the data signal power in DL transmission for a given total power budget in a coherence interval, where the BT scheme is employed.

In view of the discussion in the previous section, in this thesis, we employ the finite-dimensional channel model for the uplink transmission. Conventionally, it is assumed that the users transmit equal power in the uplink transmission. However, it should be noted that the spectral efficiency is not necessarily maximized in the conventional approach. Thus, an optimal transmit power of each user needs to be determined to maximize the spectral efficiency.

1.4 Objective

The objective of this thesis is to investigate the spectral efficiency of massive MU-MIMO systems in both UL and DL transmissions. The main focus of this study is on improving the spectral efficiency in massive MU-MIMO systems by allocating the power to the pilot and data transmitted symbols, given the total power budget per each coherence interval. Since the spectral efficiency is the main concern of this thesis, and its calculation using the lower bound on the achievable rate is computationally very intensive, in this thesis, we derive approximate expressions for the lower bound of achievable downlink rate, for both small scale and large scale fading.

In the first part of the thesis, we first derive a closed-form approximate expression for

the achievable DL rate for the maximum ratio transmission precoder based on small-scale fading in order to evaluate the spectral efficiency in the BT scheme. Then, we propose three methods of power allocation in order to maximize the spectral efficiency for a given total power budget among the users.

In the second part of the thesis, we first derive a closed-form approximate expression for the achievable DL rate for the maximum ratio transmission precoder based on large-scale fading in order to evaluate the spectral efficiency in the BT scheme. Then, we propose four methods of power allocation in order to maximize the spectral efficiency given the total power budget among the users.

The third part of the thesis deals with the spectral efficiency of massive MU-MIMO systems in an UL transmission with a very large number of antennas at the base station serving single-antenna users. A power control scheme based on large-scale fading is proposed to maximize the spectral efficiency under the peak power constraint.

1.5 Organization of the thesis

The thesis is organized as follows:

Chapter 2 provides the fundamental concepts of a massive MU-MIMO system. In addition, various channel estimation and precoding techniques are introduced for massive MU-MIMO systems.

Chapter 3 presents various power allocation schemes in order to maximize the spectral efficiency in DL channel employing small-scale fading. First, a closed-form approximate expression for the achievable DL rate for the MRT precoder is derived in order to evaluate the spectral efficiency. Then, three methods of power allocation are proposed in order to maximize the spectral efficiency given the total power budget among the users. Experiments are conducted to evaluate the performance of the proposed methods in terms of the spectral efficiency, and to determine as to which of the three methods provides the best performance.

Chapter 4 investigates the spectral efficiency in DL transmission employing large-scale fading. First, a closed-form approximate expression for the achievable DL rate for the MRT precoder is derived in order to evaluate the spectral efficiency. Then, four methods of power allocation are proposed in order to maximize the spectral efficiency given the total power budget for the users. Experiments are conducted to evaluate the performance of the proposed methods in terms of the spectral efficiency and to determine as to which of the four methods provides the best performance.

Chapter 5 presents a power control scheme based on large-scale fading in order to maximize the spectral efficiency in UL channel under the peak power constraint. First, an expression for the lower bound on the achievable rate of the uplink data transmission is derived using a linear detector for each user. This lower bound is further modified for the maximum-ratio combining and zero-forcing receivers. Then, a power control scheme based on large-scale fading is proposed to maximize the spectral efficiency under the peak power constraint. Experiments are conducted to evaluate the lower bounds obtained and the performance of the proposed method.

Chapter 6 concludes the thesis by summarizing and highlighting the work undertaken in the thesis. Some topics that could be undertaken following the ideas developed in this thesis are also briefly discussed in this chapter.

Chapter 2

Background

In this chapter, a brief introduction to point-to-point and massive MIMO systems is given. Then, the achievable rate of each user in UL and DL transmissions is described. Since our presented work relies on precoding techniques in DL transmission, the conventional precoding is described in detail.

2.1 Point-to-point MIMO

In order to investigate the capacity of massive MU-MIMO systems, first, we consider a point-to-point MIMO transmission, where the transmitter and receiver are equipped with n_t and n_r antennas, respectively. It is assumed that the MIMO channel is a narrow-band time-invariant and deterministic channel denoted by $H \in C^{n_r \times n_t}$. It is known that the OFDM schemes can convert a frequency-selective wide-band channel into multiple parallel flat-fading narrow-band channels [1]. A point-to-point MIMO link has the following mathematical description such that for each use of the channel, the received signals is given by

$$y = \sqrt{\rho} H x + w \ , \tag{2.1}$$

where y is n_r components vector of the transmitted signals, x is n_t components vector of the received signals, w is n_r components vector of the receiver noise, and ρ is signal to

noise ratio (SNR) of the link. To make sure that the expected total transmit power is unity we have

$$\mathbb{E}\left\{\left\|x^2\right\|\right\} = 1 \ , \tag{2.2}$$

where $\mathbb{E}(.)$ denotes the expected value. The components of the additive noise vector are independent and identically distributed (i.i.d) zero mean and unit variance complex-Gaussian random variables.

2.1.1 Achievable rate

Access to the channel state information (CSI) at the transmitter and the receiver has a substantial effect on the achievable rate of MIMO systems. Under the perfect knowledge of the channel matrix H at the transmitter, the mutual information or capacity is derived in bits-per-symbol as [49]

$$C = I(x;s) = \log_2(\det(I_{n_r} + \frac{\rho}{n_t} H H^{\dagger})), \tag{2.3}$$

where $I(x;s)$ indicates the mutual information operator, I_{n_r} denotes the $n_r \times n_t$ identity matrix, and $(.)^{\dagger}$ denotes the Hermitian transpose of the associated matrix [11]. We can rewrite (2.3) in form of singular values of the propagation matrix. To this end, we decompose the channel matrix H as

$$H = \varphi \ D_v \ \psi^{\dagger}, \tag{2.4}$$

where φ and ψ are unitary matrices of dimension $n_r \times n_r$ and $n_t \times n_t$, respectively. In addition, D_v denotes diagonal matrix with dimension of $n_r \times n_t$ which comprises the singular values of matrix H denoted. In this case, the capacity given by (2.3) is expressed

as

$$C = \sum_{l=1}^{\min(n_t,n_r)} \log_2(1 + \frac{\rho\ {v_l}^2}{n_t}) \tag{2.5}$$

where v_i is i^{th} singular value of matrix H. From (2.5), the capacity of the total system is equivalent to the combined achievable rate of parallel links when the l^{th} link has $\frac{\rho v_l^2}{n_t}$ SNR. The capacity is maximum when the singular values are equal and is minimum when all the singular values except one are zero. Thus, (2.5) can be bounded as

$$\log_2\left(1 + \frac{\rho\ Tr(GG^\dagger)}{n_t}\right) \leq C \leq \min(n_t n_r) \times \log_2\left(1 + \frac{\rho\ Tr(GG^\dagger)}{n_t \min(n_t n_r)}\right) \tag{2.6}$$

We assume that the propagation coefficient magnitude is equal to one, i.e., $Tr(GG^H) \approx n_t n_r$. Thus, we have

$$\log_2\left(1 + \rho\ n_t\right) \leq C \leq \min(n_t n_r) \times \log_2\left(1 + \frac{\rho \max(n_t n_r)}{n_t}\right) \tag{2.7}$$

The minimum capacity happens under the extreme keyhole propagation conditions while the maximum happens when the entries of the propagation matrix are i.i.d. random variables.

2.1.2 Limiting cases

In this subsection, we intend to deal with the effect of increasing the n_t and n_r on the achievable rate. It is known that at low SNRs, (2.3) becomes

$$C_{\rho\to 0} \approx \frac{\rho\ Tr(GG^\dagger)}{n_t \ln 2} \approx \frac{\rho\ n_r}{\ln 2} \tag{2.8}$$

In (2.7), capacity is dependent on only n_r. In other words, at low SNRs, the multiple transmit antennas have no value. In addition, if the number of transmitter antennas increases while keeping the number of receiver antennas constant, the capacity matches

the upper bound in (2.7), i.e., we have

$$C_{n_t >> \, n_r} \;\approx\; n_r \log_2(1+\rho) \tag{2.9}$$

In favorite propagation condition, we have [3]

$$\left(\frac{GG^{\dagger}}{n_t}\right)_{n_t >> n_r} \approx I_{n_r} \tag{2.10}$$

In (2.10), it is assumed that the row-vectors of the propagation matrix are asymptotically orthogonal. On the other hand, if the number of receiver antennas increases while the number of transmitter antennas are constant, we have

$$\left(\frac{GG^{\dagger}}{n_r}\right)_{n_r >> n_t} \approx I_{n_t} \tag{2.11}$$

Under this condition the capacity of the MIMO channel given by (4.8) is simplified to

$$C_{n_r >> n_t} \approx n_t \, \log_2(1+\frac{\rho \, n_r}{n_t}) \tag{2.12}$$

where the capacity again matches the upper bound (2.7). It should be noted that a very large number of transmitter or receiver antennas, combined with favorite propagation condition, constitutes a highly desirable scenario [7]. The results in (2.11) and (2.12) show the advantages of equipping the arrays in a MIMO link with a large number of antennas.

2.2 Multi-user MIMO

Multi user MIMO systems are considered as a cellular network, where the base station is equipped by M antennas that serves K single-terminal users in each cell. It is assumed that there are L cells. Therefore, the coefficient propagation of each user to each antenna

of the base station in L cells is denoted by $g_{i,k,l,n}$ which is defined as the channel coefficient from the kth user in the lth cell to the nth antenna of the ith BS. This coefficient consists of two parts, small-scale fading factor and large-scale fading (geometric attenuation). In this case, it can be written as

$$g_{i,k,l,n} = h_{i,k,l,n}\sqrt{d_{i,k,l}} \tag{2.13}$$

where $h_{i,k,l,n}$ and $d_{i,k,l}$ denote complex small-scale fading and large-scale fading coefficients, respectively. In this system model, it is assumed that small-scale fading coefficients are different for the different users for different antennas at each BS, but the large-scale fading coefficients are the same for different antennas at each BS and it depends on the distance of each user. In view of this, the channel matrix from all K users in the lth cell to the ith BS can be expressed as

$$G_{i,l} = H_{i,l}D^{\frac{1}{2}}{}_{i,l} \tag{2.14}$$

where $D_{i,l}$ is a $K \times K$ diagonal matrix that consists $d_{i,1,l}, ..., d_{i,K,l}$ elements. Typically, when the single-cell systems are considered, the cell and BS indices are dropped.

2.2.1 Uplink

In the uplink channel, when the perfect CSI is available at BS, the users only transmit their signal, but when CSI is not available at BS, the users send pilot signals along with transmitted data for estimating the channel at BS. In this section, we consider that BS is aware of CSI. In this case, the received signal vector at a single BS is given by [7]

$$y_u = \sqrt{p_u}Gx_u + n_u \tag{2.15}$$

where $x_u \in C^{K\times 1}$ is the transmitted signal vector from all the K users, $G \in C^{M\times K}$ is the uplink channel, $n_u \in C^{M\times 1}$ is a zero-mean noise vector with complex Gaussian distribution and identity covariance matrix, and p_u denotes the uplink transmit power. The transmitted symbol from kth user is the kth element of $x_u = \left[{x_1}^u, {x_2}^u, ..., {x_K}^u\right]$ with $E[|{x_k}^u|^2] = 1$. It is shown in [1] that the column channel vectors from different users are asymptotically orthogonal when the number of antennas at BS, M, grows to infinity. Thus, we have

$$G^{\dagger}G = D^{\frac{1}{2}}H^{\dagger}HD^{\frac{1}{2}} \approx MD \ . \tag{2.16}$$

This result is discussed in [50]. According to this result, the capacity of the MU-MIMO link becomes [7]

$$C = \sum_{K=1}^{K} \log_2(1 + \ M\, p_u d_k) \frac{\frac{bits}{s}}{HZ} \ . \tag{2.17}$$

The achievable rate of each user depends on the detectors that are employed. For example, (2.17) is derived when BS employs the simple matched filter (MF) detector. In this case, BS multiplies the received signal vector by $H^{\dagger}$, as

$$H^{\dagger}y_u = H^{\dagger}\left(\sqrt{p_u}Hx_u + n_u\right). \tag{2.18}$$

This technic does not color the noise since the channel vectors are asymptotically orthogonal when the number of antenna at BS grows to infinity [1].

2.2.2 Downlink

In TDD systems, the DL channel matrix is the conjugate transpose of the UL channel matrix [8]. The users are not aware of CSI while BS usually knows CSI based on the uplink pilot transmission. In this case, each user detects the corresponding signal data.

Let $y_d \in C^{K\times 1}$ be the received signal vector at K users. Then, the received signal vector can be written as

$$y_d = \sqrt{p_d}\ H^T x_d + n_d, \tag{2.19}$$

where $x_d \in C^{N\times 1}$ is the signal vector transmitted by BS, $n_d \in C^{K\times 1}$ is additive noise vector and p_d denotes the transmit power. To make sure that the transmitted symbol power is equal to one, we have $E\left\{\|x_d\|^2\right\} = 1$. In DL transmission, it is possible for BS to perform power allocation in order to maximize the sum rate [49]. With power allocation, the sum capacity for the system becomes

$$C = \max_{P} \log_2 \det\ (I_K\ +\ p_d\, M\, Q\, D)\frac{\frac{bits}{s}}{Hz}, \tag{2.20}$$

where Q is a positive diagonal matrix associated with the power allocation coefficients for each user, i.e., $(q_1\,, \ldots, q_K)$, where the power allocation coefficients satisfies $\sum_{k=1}^{K} q_k = 1$. It is shown in [3] that the simple MF detector can achieve the capacity of a massive MU-MIMO system when the number of antennas at BS, M, is much larger than the number of users, K, and grows to infinity, i.e., $M >> K$ and $M \to \infty$.

The CSI is not typically available at BS. Hence, the users transmit pilot symbols along with their uplink data symbols to BS to estimate the CSI. In this case, all users allocate τ symbols to pilot symbols per a coherence interval and BS receives the pilots and estimates the channel using minimum mean square error (MMSE) estimation.

2.3 Massive MIMO

Massive MIMO systems are a useful and scalable version of multi user MIMO systems [8]. There are three fundamental distictions between massive MIMO and conventional multi user MIMO. First, only base station learns G. Second, M is typically much larger than K,

although this does not have to be the case. Third, simple linear signal processing is used both on the uplink and on the downlink. These features render massive MIMO scalable with respect to the number of base station, M.

2.4 Channel estimation

As discussed in the previous subsection, BS needs to estimate CSI. Having CSI can be used in multi-user precoding in the DL transmission and multi-user detecting in the UL transmission. The resource, time or frequency required for channel estimation in a MU-MIMO system is proportional to the number of the transmit antennas and is independent of the number of the receive antennas [7]. When a massive MIMO system works in the frequency-division duplexing (FDD) mode, the frequency band of UL and DL channel are different. Hence, BS can not use the same CSI for both channels. Channel estimation for the UL channel is done at BS using different pilot sequences which are transmitted by the users. In the TDD mode, the time required for UL pilot transmission is independent of the number of antennas at BS. However, to get CSI for the DL channel in the FDD mode, a two-stage procedure is required. BS first transmits pilot symbols to all users, and then all users feed back estimated CSI (partial or complete) for the DL transmission to BS. The time required to transmit the DL pilot symbols is proportional to the number of antennas at BS [24], which is very large in massive MU-MIMO systems. For this reason, massive MU-MIMO systems are typically considered in TDD mode. Based on the assumption of channel reciprocity between UL and DL channel in the TDD mode, CSI is required to be estimated only in the UL channel [47].

According to the TTD protocol that is depicted in Fig.2, first, all the users in all the cells synchronously send UL data signals. Then, the users send pilot sequences. Base stations use these pilot sequences to estimate CSI for detecting the UL data. Also, base stations generate beamforming vectors for DL data transmission with the aid of CSI. One of the famous problems in this field is the pilot contamination. Due to limited channel

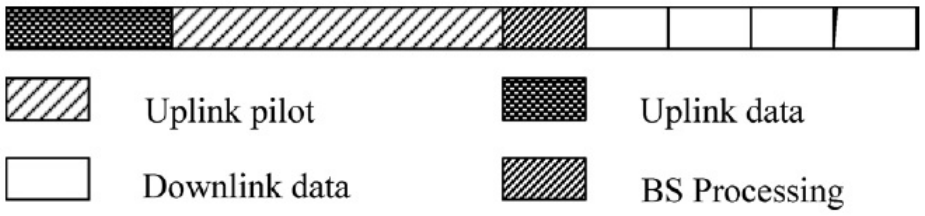

Figure 2.1: Multi-User MIMO TDD protocol.

coherence time, all the users in all the cells employ the same sequences of orthogonal pilot signals, which brings about the interference at the base stations [51–58]. Typically, the base stations employ the linear MMSE estimator in order to obtain CSI, but other methods, such as compressive sensing-based channel estimation approach which is a sub-optimal approach, are also used [59]. In addition, obtaining CSI at the receiver is a supportive method to improve the spectral efficiency in a massive MU-MIMO system. In this respect, a time-frequency training sequence design is developed in [60].

2.5 Precoding

As discussed in the previous subsection, precoding the DL data symbols has an important role on the capacity of the DL channels. In fact, precoding the DL data symbols for MIMO systems is done by linear and non-linear precoding techniques, such as dirty-paper-coding (DPC) [61–63], vector perturbation (VP) [64] and lattice-aided methods [65]. These schemes have better performance rather than linear precoding, but there are many complexity issues for implementation. On the other hand, with an increase in the number of antennas at the BS, linear precoders, such as MF and ZF, are shown to be near-optimal [3, 7]. Thus, it is more practical to use low-complexity linear precoding techniques in massive MIMO systems. Therefore, we mainly focus on linear precoding techniques.

2.5.1 Basic Precoding

In massive MU-MIMO systems, basic linear precoding methods are employed at BS [1]. The difference between the performance of MF and ZF precoders is discussed in [66]. When BS employs MF precoder, the transmitted signal from BS can be expressed as [66]

$$x_d{}^{MF} = \frac{1}{\sqrt{\alpha}} (H^T)^{\dagger} s_d = \frac{1}{\sqrt{\alpha}} H^* s_d, \tag{2.21}$$

where α is a power normalization factor and s_d is the data symbol [66]. When ZF is employed at BS, the transmitted signal from BS can be expressed as [66]

$$x_d{}^{ZF} = \frac{1}{\sqrt{\alpha}} (H^T)^{\Psi} s_d = \frac{1}{\sqrt{\alpha}} H^* (H^T H^*)^{-1} s_d \,, \tag{2.22}$$

where Ψ denotes the pseudo-inverse operator. The performance of ZF precoder outperforms the performance of MF precoder in the high SNR region, while the performance of MF precoder is better in the low SNR region [22]. The computational complexity of these precoders is discussed in [22]. This complexity is dependent of the number of the users. By decreasing the number of the users per each cell, the computational complexity of these precoders can be decreased.

2.6 Summary

In this chapter, we have presented the point to point MIMO, multi-user MIMO and massive MIMO systems in detail. In addition, we have introduced channel estimation and precoding techniques for MIMO systems.

Chapter 3

Spectral efficiency in downlink transmission for small-scale fading

In this chapter, we first investigate a massive MU-MIMO system in the TDD mode, where a base station is equipped with a very large number of antennas and serves single-antenna users simultaneously in the same frequency band. To accurately decode the data signal on the DL channel, BS and the users require the channel state information, which is obtained by the BT scheme employing the small-scale fading. Then, we propose various power allocation schemes in order to maximize SE.

3.1 System Model

The DL transmission is studied in a single-cell massive MU-MIMO system, where a BS with M antenna elements simultaneously communicates with K single antenna users as shown in Fig. 1. It is assumed that $M \gg K$ and BS employs the linear precoding technique before the DL transmission to all the users. In order to precode the data, BS requires CSI, which is obtained through the UL training. Due to the channel reciprocity between UL and DL channels in TDD operation, BS uses this CSI to precode the data in DL transmission.

3.1.1 Uplink channel estimation

The orthogonal pilot sequences of length τ_u symbols per coherence interval are simultaneously transmitted by all users in the cell. Since the pilot sequences are orthogonal, $\tau_u \geq K$. The pilot matrix of K users is denoted by $\mathbf{\Psi} = [\phi_1, \phi_2, ..., \phi_K] \in \mathbb{C}^{\tau_u \times K}$ with the orthogonality property $\mathbf{\Psi}^\dagger \mathbf{\Psi} = \mathbf{I_K}$, where ϕ_k denotes the pilot sequence of kth user and $(.)^\dagger$ denotes the Hermitian operation of the associated matrix.

Let $\mathbf{H} \in \mathbb{C}^{M \times K}$ be the channel matrix, where the elements of $\mathbf{H}$ are assumed to be independent Gaussian distributed with zero mean and unit variance. Thus, the $M \times \tau_u$ pilot matrix received at BS can be written as

$$\mathbf{Y_u} = \sqrt{\tau_u p_u}\ \mathbf{H}\ \mathbf{\Psi}^\dagger +\ \mathbf{N_u}, \tag{3.1}$$

where p_u and $\mathbf{N_u} \in \mathbb{C}^{M \times \tau_u}$ denote respectively, the average pilot transmission power of each user and the received noise matrix at BS. We assume that the elements of $\mathbf{N_u}$ are independent Gaussian distributed with zero mean and unit variance. Using the received pilot matrix given by (3.1), the minimum mean-square error (MMSE) estimate of $\mathbf{H}$ can be written as [67]

$$\hat{\mathbf{H}} = \frac{\tau_u p_u}{\tau_u p_u + 1}\mathbf{H}\ +\ \frac{\sqrt{\tau_u p_u}}{\tau_u p_u + 1}\tilde{\mathbf{N}}_\mathbf{u}, \tag{3.2}$$

where $\tilde{\mathbf{N}}_\mathbf{u} = \mathbf{N_u}\mathbf{\Psi}$ has the same distribution as $\mathbf{N_u}$. In this case, $\mathbf{H}$ is decomposed as

$$\mathbf{H} = \hat{\mathbf{H}} + \varepsilon, \tag{3.3}$$

where ε denotes the channel estimation error. Since MMSE estimation is employed, ε and $\hat{\mathbf{H}}$ are independent. In addition, ε and $\hat{\mathbf{H}}$ have i.i.d $\mathcal{CN}(0, \frac{1}{\tau_u p_u + 1})$ and $\mathcal{CN}(0, \frac{\tau_u p_u}{\tau_u p_u + 1})$ elements, respectively [24].

3.1.2 Downlink Transmission

BS first uses the channel estimate $\hat{\mathbf{H}}$ obtained in the previous subsection to precode the symbols, and then BS transmits the precoded symbols to the users in DL transmission. In view of this, let s_k be the symbol that BS transmits to the kth user, with $\mathbb{E}\{|s_k|^2\} = 1$ and $\mathbf{W} \in \mathbb{C}^{M\times K}$ be the linear precoding matrix. In this case, the $M \times 1$ transmit signal vector can be written as

$$\mathbf{x} = \mathbf{W}\mathbf{s}, \tag{3.4}$$

where $\mathbf{s} \triangleq [\sqrt{p_{d_1}}s_1, \sqrt{p_{d_2}}s_2, ..., \sqrt{p_{d_K}}s_K]^T$ and p_{d_k} denotes the average transmit power allocated to the k^{th} user. It is assumed that the power of precoding signal is unity, i.e., we have

$$\mathrm{E}\big[tr\big(\mathbf{W}\mathbf{W}^{\dagger}\big)\big] = 1, \tag{3.5}$$

The received vector in the DL transmission is given by

$$\mathbf{y} = \mathbf{H}^T\mathbf{W}\mathbf{s} + \mathbf{n}, \tag{3.6}$$

where $\mathbf{n}$ is a vector whose k^{th} element is additive noise at the k^{th} user that is denoted by $n_k \sim \mathcal{CN}(0, 1)$. In view of this, the received signal at the k^{th} user is given by

$$y_k = \sqrt{p_{d_k}}\, a_{kk}s_k + \sum_{i=1, i\neq k}^{K} \sqrt{p_{d_i}}a_{ki}s_i + n_k. \tag{3.7}$$

For accurately detecting the transmitted signal in DL transmission, each user needs to obtain CSI. A conventional method of channel estimation is that BS transmits pilot symbols in such a way that the users estimate the channel using minimum mean-square error (MMSE) estimation. This method is inefficient since the overhead on the aforementioned

channel estimation is proportional to M, which tends to infinity in a massive MU-MIMO system. To solve this problem, the BT scheme is employed to estimate a_{kk} for each user [24]. In the BT scheme, the channel estimation is proportional to K, which is much smaller than M. In the next subsection, we explain as to how to estimate a_{kk}.

3.1.3 Beamforming Training

In the BT scheme, BS beamforms the pilot sequences in DL transmission after channel estimation in the UL training. Then, the effective channel gain a_{ki} is estimated at each user by the received pilot sequences. We define $\mathbf{S}_p \in \mathbb{C}^{K\times\tau_d}$ to be a pilot matrix in the DL channel, where τ_d denotes the number of symbols for pilot sequences. Using this definition, the pilot matrix is given by

$$\mathbf{S}_p = \sqrt{\tau_d p_p}\ \mathbf{\Phi}. \tag{3.8}$$

where p_p and $\mathbf{\Phi}$ denote the power of each pilot symbol and the pilot sequence matrix in DL transmission, respectively. Since the pilot sequences are orthogonal, we have $\mathbf{\Phi}\mathbf{\Phi}^\dagger = \mathbf{I}_K$, which requires that $\tau_d \geqslant K$. In the BT method, using the precoding matrix $\mathbf{W}$, BS beamforms the pilot sequence for the users. In other words, the transmitted pilot matrix is $\mathbf{W}\mathbf{S}_p$. Thus, the received pilot matrix in DL transmission can be expressed as

$$\mathbf{Y}_p^T = \sqrt{\tau_d p_p}\ \mathbf{H}^T\mathbf{W}\mathbf{\Phi} + \mathbf{N}_p^T, \tag{3.9}$$

where $\mathbf{N}_p^T$ denotes the noise matrix in the received signal with i.i.d. $\mathcal{CN}(0,1)$ entries. To estimate the channel, we use the orthogonality of pilot sequences. In view of this, let $\tilde{\mathbf{Y}}_p^T \triangleq \mathbf{Y}_p^T\mathbf{\Phi}^\dagger$. In this case, we have

$$\tilde{\mathbf{Y}}_p^T = \sqrt{\tau_d p_p}\ \mathbf{H}^T\mathbf{W} + \tilde{\mathbf{N}}_p^T, \tag{3.10}$$

where $\tilde{\mathbf{N}}_p^T \triangleq \mathbf{N}_p^T \mathbf{\Phi}^\dagger$ has i.i.d $\mathcal{CN}(0,1)$ elements. By decomposing $\tilde{\mathbf{Y}}_p^T$ given by (4.12), we have

$$\tilde{\mathbf{y}}_{p,k}^T = \sqrt{\tau_d p_p}\ \mathbf{h}_k^T \mathbf{W} + \tilde{\mathbf{n}}_{p,k}^T = \sqrt{\tau_d p_p}\ \mathbf{a}_k^T + \tilde{\mathbf{n}}_{p,k}^T, \tag{3.11}$$

where $\tilde{\mathbf{y}}_{p,k}$ and $\tilde{\mathbf{n}}_{p,k}^T$ represent the kth columns of $\tilde{\mathbf{Y}}_p$ and $\tilde{\mathbf{N}}_p^T$, respectively and $\mathbf{a}_k \triangleq [a_{k1} a_{k2} ... a_{kK}]^T$. From (4.15), kth user estimates $\mathbf{a}_k$. Although the elements of $\mathbf{a}_k$ are correlated and should be jointly estimated, it has been shown in [24] that the performance loss due to independent estimation is negligible. As a result, a_{k1},...,a_{kK} are estimated independently. In view of this, the ith element of $\tilde{\mathbf{y}}_{p,k}$ is employed to estimate a_{ki} using MMSE channel estimation. In this case, the estimation of a_{ki} can be expressed as [67]

$$\hat{a}_{ki} = \mathbb{E}[a_{ki}] + \frac{\sqrt{\tau_d p_p}\ \mathrm{Var}(a_{ki})}{\tau_d p_p\ \mathrm{Var}(a_{ki}) + 1} \big(\tilde{y}_{p,ki} - \sqrt{\tau_d p_p}\ \mathbb{E}[a_{ki}]\big) \tag{3.12}$$

where $\tilde{y}_{p,ki}$ denotes the ith entry of $\tilde{\mathbf{y}}_{p,k}$ and $\mathrm{Var}(a_{ki})$ represents the variance of a_{ki}. This expression looks similar to (10) in [24]. However, in [24] the transmit power for pilot and data symbols are assumed to be the same in DL transmission. It is to be emphasized that in the present work, we distinguish between the pilot transmit power p_p and the data transmit power p_d in DL transmission.

We define ϵ_{ki} to be the channel estimation error. Since MMSE estimation is employed, the estimate $\hat{a}_{ki}$ and the estimation error ϵ_{ki} are uncorrelated. In view of this, the effective channel gain a_{ki} is given by

$$a_{ki} = \hat{a}_{ki} + \epsilon_{ki}\ . \tag{3.13}$$

Substituting (3.13) into (3.7), we have

$$y_k = \sqrt{p_{d_k}}\ \hat{a}_{kk} s_k + \sum_{i=1, i\neq k}^{K} \sqrt{p_{d_i}}\ \hat{a}_{ki} s_i + \sum_{i=1}^{K} \sqrt{p_{d_i}} \epsilon_{ki} s_i + n_k. \tag{3.14}$$

3.1.4 Achievable Downlink Rate

Employing an approach similar to that used in [48], it can be shown that the achievable DL rate (ADR) for kth user is lower bounded as

$$R_k = \mathbb{E}\left[\log_2\left(1+\frac{p_{d_k}|\hat{a}_{kk}|^2}{\sum_{i=1}^{K} p_{d_i}\mathbb{E}\{|\epsilon_{ki}|^2\}+\sum_{i\neq k}^{K} p_{d_i}|\hat{a}_{ki}|^2+1}\right)\right] \tag{3.15}$$

Even though this expression is similar to that obtained in [24], it is noted that the value of $\hat{a}_{ki}$ in the former expression is different from that in the latter in view of our distinguishing the pilot transmit power from the data transmit power.

It has been shown in [24] that the BT method with MRT precoding is more efficient than the BT method with ZF precoding. This is due to the fact that, with ZF, the randomness of the effective channel gain at each user is smaller than the one with MRT (with ZF, the channel gain becomes deterministic when the BS has perfect CSI) and hence, MRC has a higher advantage of using the channel estimate for the signal detection. In view of this, we have employed the MRT precoding in this thesis.

3.1.4.1 MRT precoder

When BS uses the MRT precoder in the DL transmission, the precoding matrix is defined by

$$\mathbf{W} = \alpha_{MRT}\ \hat{\mathbf{H}}^*, \tag{3.16}$$

where $(.)^*$ denotes the conjugate operation of the associated matrix and α_{MRT} is a constant which is employed to satisfy the power constraint given by (3.5). Thus, we have [24]

$$\alpha_{MRT} = \sqrt{\frac{\tau_u p_u + 1}{MK\tau_u p_u}}. \tag{3.17}$$

Proposition 3.1 : Using MRT precoding technique, $\hat{a}_{ki}$ and $\mathbb{E}\{|\epsilon_{ki}|^2\}$ are given by

$$\begin{cases} \hat{a}_{ki} & = \frac{\sqrt{\tau_d p_p}}{\tau_d p_p + K} \tilde{y}_{p,ki} & i \neq k \\ \hat{a}_{kk} & = \frac{\sqrt{\tau_d p_p}}{\tau_d p_p + K} \tilde{y}_{p,kk} + \frac{K}{\tau_d p_p + K} \sqrt{\frac{\tau_u p_u M}{K(\tau_u p_u + 1)}} & i = k \\ \mathbb{E}\{|\epsilon_{ki}|^2\} & = \frac{1}{\tau_d p_p + K} & i \; \forall \; k \end{cases} \tag{3.18}$$

The proof of this proposition is given in Appendix A.

Substituting $\mathbb{E}\{|\epsilon_{ki}|^2\} = \frac{1}{\tau_d p_p + K}$ into (3.15), we obtain the lower bound of ADR for kth user as

$$R_k^{MRT} = \mathbb{E}\left[\log_2 \left(1 + \frac{p_{d_k} |\hat{a}_{kk}|^2}{\sum_{i \neq k}^{K} \frac{p_{d_i}}{\tau_d p_p + K} + \sum_{i \neq k}^{K} p_{d_i} |\hat{a}_{ki}|^2 + 1} \right) \right]. \tag{3.19}$$

3.1.5 Spectral Efficiency

The spectral efficiency S is defined by [24]

$$S = \frac{T - \tau_u - \tau_d}{T} \sum_{k=1}^{K} R_k, \tag{3.20}$$

where R_k is the lower bound on ADR for the kth user given by (3.19) for the MRT precoder, and T is the length of the coherence interval in DL transmission. The estimate for $\hat{a}_{ki}$ and $\hat{a}_{kk}$ depend on the precoder used. These estimates are given by (3.18) for the MRT precoder.

In the following section, we present various methods of power allocation and maximize the spectral efficiency. These are given below.

1. BS does not transmit any pilot symbols to the users and the data power of each user is the same, i.e.,

$\tau_d = 0$ and $p_{d_1} = p_{d_2} = ... = p_{d_K} = p_d$.

2. Pilot power as well as the data power for each user is the same, i.e.,

$p_{p_1} = ... = p_{p_K} = p_{d_1} = ... = p_{d_K} = p_d$.

3. Pilot power for each user is the same, with a similar statement holding true for the data power, i.e.,

$p_{p_1} = p_{p_2} = ... = p_{p_K} = p_p$.

$p_{d_1} = p_{d_2} = ... = p_{d_K} = p_d$.

4. Pilot power is the same for each user and is specified, while the data power for each user is determined to maximize SE, i.e.,

$p_{p_1} = p_{p_2} = ... = p_{p_K} = p_p$ is specified.

5. Pilot power is the same for each user, i.e.,

$p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$,

and we determine $p_p, p_{d_1}, ..., p_{d_K}$ such that SE is maximized.

3.2 Maximization of spectral efficiency

3.2.1 No pilot power and equal data power for all users

We assume that BS does not transmit any pilot symbols to the users and the data power for each user is the same, i.e., $\tau_d = 0$ and $p_{d_1} = p_{d_2} = ... = p_{d_K} = p_d$. Hence, from (3.19) we see that the lower bound of ADR is [22]

$$R_k^{MRT} = \log_2 \left(1 + \frac{M}{K} \frac{\tau_u p_u \; p_d}{(p_d + 1)(\tau_u p_u + 1)}\right). \tag{3.21}$$

In this method of power allocation, SE is obtained by substituting (3.21) in (3.20).

3.2.2 Equal pilot and data powers for each user

We now assume that the data power of each user is the same, i.e., $p_{p_1} = ... = p_{p_K} = p_p = p_{d_1} = ... = p_{d_K} = p_d$. Hence, from (3.19) the lower bound of ADR is obtained as [24]

$$R_k^{MRT} = \mathbb{E}\left[\log_2 \left(1 + \frac{p_d |\hat{a}_{kk}|^2}{\frac{Kp_d}{\tau_d p_d + K} + p_d \sum_{i \neq k}^{K} |\hat{a}_{ki}|^2 + 1}\right)\right]. \tag{3.22}$$

Again, in this method of power allocation, SE is obtained by substituting (3.22) in (3.20).

3.2.3 Equal pilot power and equal data power for all users

We assume that not only the pilot power for each user is the same, but also the data power for each user is the same, i.e.,$p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ and $p_{d_1} = p_{d_1} = ... = p_{d_K} = p_d$. Substituting these in (3.19), the lower bound of ADR in this case is given by

$$R_k^{MRT} = \mathbb{E}\left[\log_2 \left(1 + \frac{p_d |\hat{a}_{kk}|^2}{\frac{Kp_d}{\tau_d p_p + K} + p_d \sum_{i \neq k}^{K} |\hat{a}_{ki}|^2 + 1} \right) \right]. \tag{3.23}$$

We allocate the power among the pilot and data symbols in DL transmission given the energy budget in a coherence time in such a way that SE is maximized [68]. We also find the optimal value of τ_u in order to maximize SE. In view of this, first, we obtain a close approximation of the achievable rate given by (3.23) and then, we present our optimal power allocation method.

3.2.3.1 Approximation for the lower bound of achievable downlink rate

In order to obtain a close approximation of the achievable rate given by (3.23) for the MRT precoder, we use the following Lemma.

Lemma 3.1 : When X and Y are independent positive random variables, we have [69]

$$\mathbb{E}\left[\log_2 \left(1 + \frac{X}{Y} \right) \right] \approx \log_2 \left(1 + \frac{\mathbb{E}\{X\}}{\mathbb{E}\{Y\}} \right). \tag{3.24}$$

In [69], it has been shown that the right side expression of Lemma 3.1 lies between the lower and upper bounds of the left side expression, and hence, can be used as an approximation to the left side expression. In addition, it has been shown in [69] that as M becomes very large, this approximation becomes particularly accurate.

Employing Lemma 3.1, a tractable expression for the lower bound of the achievable

DL rate for the MRT precoder, given by (3.19), can be approximated as

$$R_k^{\mathrm{MRT}} \approx \log_2\left(1+\frac{p_d\ \mathbb{E}\{|\hat{a}_{kk}|^2\}}{\frac{K\ p_d}{\tau_d p_p+K}+p_d\sum_{i\neq k}^{K}\ \mathbb{E}\{|\hat{a}_{ki}|^2\}+1}\right), \tag{3.25}$$

Proposition 3.2 : Substituting (3.18) into the lower bound of the achievable DL rate given by (3.25), it can be shown that

$$R_k^{\mathrm{MRT}} \approx \log_2\left(1+\mathrm{SINR}_k^{\mathrm{MRT}}\right) = \tilde{R}_k^{\mathrm{MRT}}, \tag{3.26}$$

where

$$\mathrm{SINR}_k^{\mathrm{MRT}}=\frac{p_d\ [a\ \tau_d^2 p_p^2+b\ \tau_d p_p+c]}{\tau_d^2 p_p^2(d\ p_d+1)+\tau_d p_p(e\ p_d+f)+g(p_d+1)}\ , \tag{3.27}$$

and

$$\begin{aligned}
a &= \alpha_{MRT}^2\big(\frac{\tau_u p_u}{\tau_u p_u+1}\big)^2 M(M+1)+\alpha_{MRT}^2\frac{\tau_u p_u}{(\tau_u p_u+1)^2}M,\\
b &= \frac{2\ \tau_u p_u M}{(\tau_u p_u+1)}+1,\\
c &= \frac{K\ \tau_u p_u M}{(\tau_u p_u+1)},\\
d &= \frac{K-1}{K},\\
e &= 2K-1,\\
f &= 2K,\\
g &= K^2.
\end{aligned} \tag{3.28}$$

The proof of this proposition is given in Appendix B.

3.2.3.2 Optimal power allocation

We now present the proposed resource allocation scheme in order to maximize SE. It has been shown in [32] that allocating optimal powers for the training symbols and data symbols increases SE, where SE is a function of the energy per bit (EPB) defined as

$$\eta \triangleq \frac{\frac{\tau_d}{T}p_p + \frac{\tau_u}{T}p_u + (1 - \frac{\tau_d+\tau_u}{T})p_d}{S} \ . \tag{3.29}$$

It can be observed from (3.29) that when $p_p = p_u = p_d$ and $\tau_d = \tau_u$, we have $\eta = \frac{p_d}{S}$. Moreover, it can be observed from (3.20) that when the transmit power is reduced below a certain threshold, the bit energy increases. Hence, the minimum bit energy is obtained at a non-zero SE [33]. Operating below this SE is evidently inefficient. However, this regime can be operated by increasing the transmit power for training and reducing the transmit power for data. Motivated by these observations, we propose an optimal resource allocation to jointly select the training duration on UL transmission (τ_u), the training duration on DL transmission (τ_d), the training signal power on DL transmission (p_p), the training signal power on UL transmission (p_u), and the data signal power on DL transmission (p_d) in order to maximize SE for a given total energy budget spent in a coherence interval. In view of this, let the total transmit energy constraint at BS and each user be E_{td} and E_{tu}, respectively. Thus, we have

$$\tau_d p_p + (T - \tau_d - \tau_u)p_d \ \leq \ E_{td} \ , \tag{3.30}$$

and

$$\tau_u p_u \ \leq \ E_{tu} \ . \tag{3.31}$$

From (3.30), the channel estimate is degraded when $\tau_d p_p$ decreases, but the energy for the data transmission phase $(T-\tau_d-\tau_u)p_d$ is increased under the total energy constraint at BS.

Hence, SE may improve. Moreover, the accuracy of the channel estimate is improved by allocating more energy to the training transmission phase. However, less energy should be allocated to the data transmission phase to satisfy (3.30). Hence, SE may again improve. In addition, from (3.31), it is straightforward that total energy constraint at each user is allocated to the UL training transmission phase in order to improve SE. In view of this, there are optimal values of τ_u, τ_d, p_p, p_u, and p_d which maximize SE for given E_{td}, E_{tu}, and T. Mathematically speaking, we have the following problem.

$$\begin{aligned} \max_{p_u,p_d,p_p,\tau_u,\tau_d} \quad & S \\ \text{s.t.} \quad & \begin{cases} \tau_d p_p + (T-\tau_d-\tau_u)p_d \leq E_{td} \\ \tau_u p_u \leq E_{tu} \\ p_p \geq 0, p_d \geq 0, p_u \geq 0 \\ \tau_u \geq K, \tau_d \geq K \\ \tau_u + \tau_d \leq T \end{cases} \end{aligned} \tag{3.32}$$

Lemma 3.2 : The energy constraint given by (3.30) is satisfied with equality at the optimal solution.

Proof : Since the expressions for SINR given by (3.27) is monotonically increasing with p_p for a given p_d and vice versa, it can be observed from (3.20) that S is an increasing function of p_p when p_d is given. In addition, S is an increasing function of p_d when p_p is given. Hence, S is maximized when BS uses all the energy budget in one coherence interval, i.e., $\tau_d p_p + (T-\tau_d-\tau_d)p_d = E_{td}$. ■

Lemma 3.3 : The energy constraint given by (3.31) is satisfied with equality at the optimal solution.

Proof : Since SINRs given by (3.27) is monotonically increasing with p_u, it can be observed from (3.20) that S is an increasing function of p_u. Hence, S is maximized when

each user employs all energy budget in one coherence interval, i.e., $\tau_u p_u = E_{tu}$. ■

Remark 3.1 : It has been shown in [29] that when the transmit powers for pilot and data sequences are allowed to vary, the optimal number of training symbols is equal to the number of transmit antennas M. On the other hand, if the training and data powers are to be made equal, the optimal number of training symbols can be larger than the number of transmit antennas M. In massive MIMO systems, M is very large. Thus, it is ineffective that we optimally choose $\tau_d = M$. On the other hand, the BT scheme is employed to efficiently estimate the channel gain for each user in order to reduce the number of training symbols in DL transmission. In view of this. in this thesis, following the work in [24], we relax τ_d in the optimization problem given by (3.32) in such a way that the complexity of channel estimation is acceptable.

According to Lemma 3.1, Lemma 3.2, and Remark 3.1, the optimization problem given by (3.32) can be rewritten as

$$\begin{aligned} \max_{p_d, p_p, p_u, \tau_u} \quad & S \\ \text{s.t.} \quad & \begin{cases} \tau_d p_p + (T - \tau_d - \tau_u) p_d = E_{td} \\ \tau_u p_u = E_{tu} \\ p_p \geq 0, p_d \geq 0, p_u \geq 0 \\ K \leq \tau_u \leq T - \tau_d \end{cases} \end{aligned} \tag{3.33}$$

There are optimal values of τ_u, p_u, p_p and p_d which maximize SE. In the next subsections, we intend to find the optimal values of τ_u and p_u, and simplify the optimization problem given by (3.33). To this end, we first introduce the following theorem.

Theorem 3.1 : The function

$$g(x) = x \ \log_2\left(1+\frac{\beta_k}{\varsigma_k + \mu_k \ x}\right) \tag{3.34}$$

is a strictly increasing function in $x \in (0, \infty)$, where $\beta_k > 0$, $\varsigma_k > 0$, and $\mu_k > 0$.

Proof : When β_k, ς_k, and μ_k are positive, we have

$$g'(x)=\frac{-1}{\ln 2}\left(\frac{\beta_k\mu_k x}{(\varsigma_k+\mu_k x)(\beta_k+\varsigma_k+\mu_k x)}-\ln\big(1+\frac{\beta_k}{\varsigma_k+\ \mu_k x}\big)\right), \tag{3.35}$$

and

$$g''(x)=\frac{1}{\ln 2}\frac{-\beta_k\mu_k^2(\beta_k+2\varsigma_k)x-2\beta_k\varsigma_k\mu_k(\beta_k+\varsigma_k)}{(\varsigma_k+\mu_k x)^2(\beta_k+\varsigma_k+\mu_k x)^2}<0, \tag{3.36}$$

where $g'(x)$ and $g''(x)$ are first and second derivatives of $g(x)$, respectively. From (3.36), we conclude that $g'(x)$ is a strictly deceasing function in x since $g''(x) < 0$. As a result, $g'(x) > g'(\infty) = 0$ which implies that $g(x)$ is a strictly increasing function in $x \in (0, \infty)$. ■

To satisfy the first constraint of the optimization problem given by (3.33), we have $p_d = \frac{E_{td}-\tau_d p_p}{T-\tau_d-\tau_u}$. Using MRT precoder and substituting $p_d = \frac{E_{td}-\tau_d p_p}{T-\tau_d-\tau_u}$ into (3.26) and then, (3.26) into (3.20), we have

$$S(p_u, \tau_u, p_p, p_d) = \sum_{k=1}^{K} g_k(p_u, \tau_u, p_p, p_d) \tag{3.37}$$

where

$$g_k(p_u,\tau_u,p_p,p_d)=\left(1-\frac{\tau_d+\tau_u}{T}\right)\log_2\left(1+\frac{\beta_k}{\varsigma_k+\mu_k(1-\frac{\tau_d+\tau_u}{T})}\right), \tag{3.38}$$

and

$$\begin{aligned}
\beta_k &= \frac{(E_{td}-\tau_d p_p)}{T}\Big[a(\tau_d p_p)^2 + b(\tau_d p_p)+c\Big], \\
\varsigma_k &= \frac{(E_{td}-\tau_d p_p)}{T}\Big[d\ (\tau_d p_p)^2 + e\ (\tau_d p_p) + g\Big], \\
\mu_k &= \left((\tau_d p_p)^2 + f(\tau_d p_p) + g\right)\ . \qquad (3.39)
\end{aligned}$$

Proposition 3.3 : The optimal value of τ_u given by the optimization problem (3.33) is equal to K.

Proof : Let us assume that τ_u^*, $p_u^* = \frac{E_{tu}}{\tau_u^*}$, p_p^*, and $p_d^* = \frac{E_{td}-\tau_d p_p^*}{T-\tau_d-K}$ are the optimal solution of the optimization problem given by (3.33) satisfying the constraints where $\tau_u^* > K$. Then, we choose $\overline{\tau_u}=K$, $\overline{p_u} = \tau_u^* p_p^*/K$, $\overline{p_p}=p_p^*$, and $\overline{p_d}=\frac{E_{td}-\tau_d p_p^*}{T-\tau_d-K}$ satisfying the constraints of the optimization problem given by (3.33). Note that with this choice, we have $\overline{\tau_u p_u} = \tau_u^* p_u^* = E_{tu}$. Substituting $\overline{p_u}$, $\overline{\tau_u}$,$\overline{p_p}$, and $\overline{p_d}$ into (3.38) yields

$$g_k(\overline{p_u}, \overline{\tau_u}, \overline{p_p}, \overline{p_d}) = \left(1-\frac{\tau_d+K}{T}\right)\log_2\left(1+\frac{\overline{\beta_k}}{\overline{\varsigma_k} + \overline{\mu_k}(1-\frac{\tau_d+K}{T})}\right) \qquad (3.40)$$

where

$$\begin{aligned}
\overline{\beta_k} &= \frac{(E_{td}-\tau_d p_p^*)}{T}\Big[a(\tau_d p_p^*)^2 + b(\tau_d p_p^*)+c\Big], \\
\overline{\varsigma_k} &= \frac{(E_{td}-\tau_d p_p^*)}{T}\Big[d\ (\tau_d p_p^*)^2 + e\ (\tau_d p_p^*) + g\Big], \\
\overline{\mu_k} &= \left((\tau_d p_p^*)^2 + f(\tau_d p_p^*) + g\right)\ . \qquad (3.41)
\end{aligned}$$

Knowing $\overline{\tau_u p_u} = \tau_u^* p_u^*$ and using Theorem 1 and the fact $\tau_u^* > K$, we have

$$\begin{aligned}
g_k(\overline{p_u}, \overline{\tau_u}, \overline{p_p}, \overline{p_d}) &> \left(1-\frac{\tau_d+\tau_u^*}{T}\right)\log_2\left(1+\frac{\overline{\beta_k}}{\overline{\varsigma_k} + \overline{\mu_k}(1-\frac{\tau_d+\tau_u^*}{T})}\right) \\
&= g_k(p_u^*, \tau_u^*, p_p^*, p_d^*). \qquad (3.42)
\end{aligned}$$

Thus, from (3.37) and (3.42), we have

$$S(\overline{p_u}, \overline{\tau_u}, \overline{p_p}, \overline{p_d}) > S(p_u^*, \tau_u^*, p_p^*, p_d^*). \tag{3.43}$$

This contradicts the assumption and so, $\tau_u^* \leq K$. On the other hand, due to the orthogonality of the pilot sequences, we have $\tau_u^* \geq K$. As a result, $\tau_u^* = K$. ■

At this stage, we find the optimal value of p_u with help of the following lemma.

Lemma 3.4 : The optimal value of p_u in the optimization problem given by (3.33) is $p_u^* = E_{tu}/K$.

Proof : To satisfy the second constraint of the optimization problem given by (3.33), we have $\tau_u^* p_u^* = E_{tu}$. Since $\tau_u^* = K$, thus, $p_u^* = E_{tu}/K$. ■

According to Lemma 3.4 and Proposition 3.3, the optimization problem given by (3.33) can be rewritten as [40]

$$\begin{aligned} &\max_{p_d} && S\big|_{p_p = \frac{E_{td}}{\tau_d} - \left(\frac{T-\tau_d-K}{\tau_d}\right) p_d} \\ &\text{such that} && \left\{ 0 \leq p_d \leq \frac{E_{td}}{T-\tau_d-K} \right. . \end{aligned} \tag{3.44}$$

Lemma 3.5 : The objective function of the optimization problem given by (3.44) is concave with respect to p_d.

Proof : First, we substitute $p_p = \frac{E_{td}}{\tau_d} - \left(\frac{T-\tau_d-K}{\tau_d}\right) p_d$ into (3.26) which yields a concave function with respect to p_d in the range $0 \leq p_d \leq \frac{E_{td}}{T-\tau_d-K}$. Knowing that $\log_2(1+x)$ is a concave function, we conclude that $\log_2(1+\text{SINR}_k^{MRT})$ is also concave. Moreover, since the summation of concave functions is also concave, the proof of Lemma 3.5 is concluded. ■

As a result of Lemma 3.5, there is a global maximum point for the objective function given by (3.44). To obtain a globally optimal solution, any convex optimization scheme

can be employed. We employ the FMINCON function in MATLAB's optimization toolbox to derive the optimal solution of the optimization problem given by (3.44). It can be seen from (3.29) that when SE is maximized, EPB is minimized for a given E_{td}. As a result, this solution also provides the minimum value of EPB.

3.2.4 Experimental Results

We utilize the expressions presented in Sections 3.2.1, 3.2.2, and 3.2.3 to study the SE performance of the proposed optimal power allocation scheme. In all the experiments conducted, we define $SNR \triangleq \frac{E_{td}}{T}$. Since E_{td} is the total transmit energy spent in a coherence interval T and the noise variance is 1, SNR has the interpretation of average transmit signal-to-noise ratio (SNR). Thus, SNR is dimensionless. We also choose $\tau_u = \tau_d = K$, and $T = 200$ (corresponding to a coherence bandwidth of 200 KHz and a coherence time of 1 ms) in all examples.

3.2.4.1 Validation of the approximation for SE

In order to obtain SE for the MRT precoder, we first substitute the estimates of a_{ki} given by (3.18) in the expression for R_k given by (3.19). Then, we substitute (3.19) in the expression for SE given by (3.20). In order to obtain a_{ki}, 1000 Monte-Carlo simulations are carried out, where in the channel and noise matrices are generated for each snapshot. This process needs a large amount of calculation for a given SNR (for example, the run time for obtaining the SE for a given SNR is 21.76 seconds, using MATLAB software and a PC with Intel(R) Core(TM) i5 @ 2.7 GHz processor and 4 GB installed memory (RAM)). In order to validate the approximate expressions for SE for the MRT precoder, we substitute the lower bound of ADR given by (3.26) in the expression for SE given by (3.20). In this case, a small amount of calculation without any Monte-Carlo simulation is required (for example, the run time for obtaining the SE for a given SNR is 0.01 seconds, using MATLAB software and a PC with Intel(R) Core(TM) i5 @ 2.7 GHz processor and

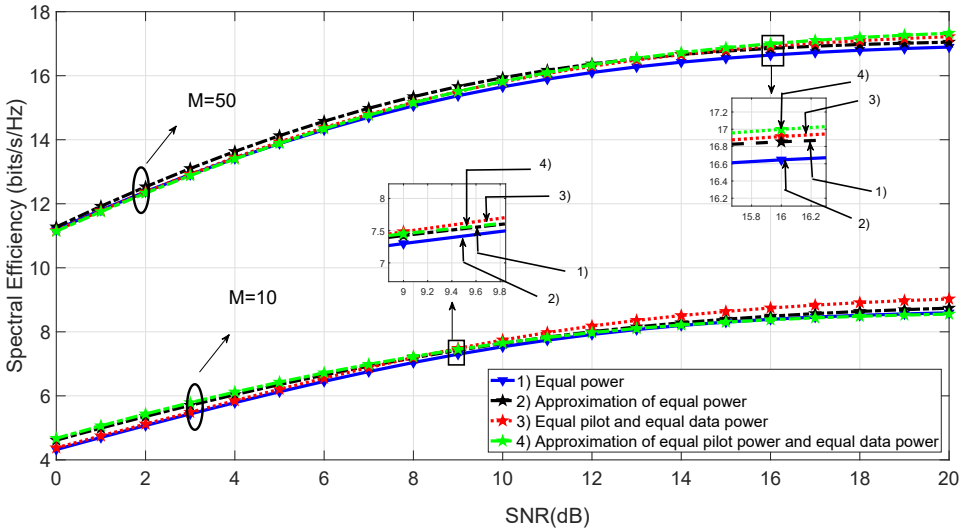

Figure 3.1: The spectral efficiency versus SNR

4 GB installed memory (RAM)).

Fig 3.1 shows SE versus SNR when $M = 10$ and $M = 50$ employing the MRT precoder. It can be seen from Fig 3.1 that SE obtained using the approximate expressions for $\tilde{R}_k^{\text{MRT}}$ is very close to that obtained using the actual one, for equal power allocation as well as for optimal power allocation when the MRT precoder is used. Hence, we can simply employ the approximated expressions $\tilde{R}_k^{\text{MRT}}$ given by (3.26) in order to obtain SE for the MRT precoder rather than using (3.19) that has a high complexity.

3.2.4.2 Results and Comparison

Fig 3.2 shows SE versus SNR when $M = 10$ and $M = 50$ employing the MRT precoder. It can be seen from this figure that the proposed method significantly improves SE with respect to the method provided in [22] and offers a slightly better SE than that provided by the method in [24]. This superiority in the performance is attributed to the optimal transmitted power p_d^* and p_p^*, which have been obtained in order to maximize SE.

Furthermore, Fig. 3.3 shows the variation of the ratio of the optimal pilot power p_p to the optimal transmitted data power p_d for the MRT precoder. We can see that in order to maximize SE, more power should be allocated to the data symbols at high SNR and less

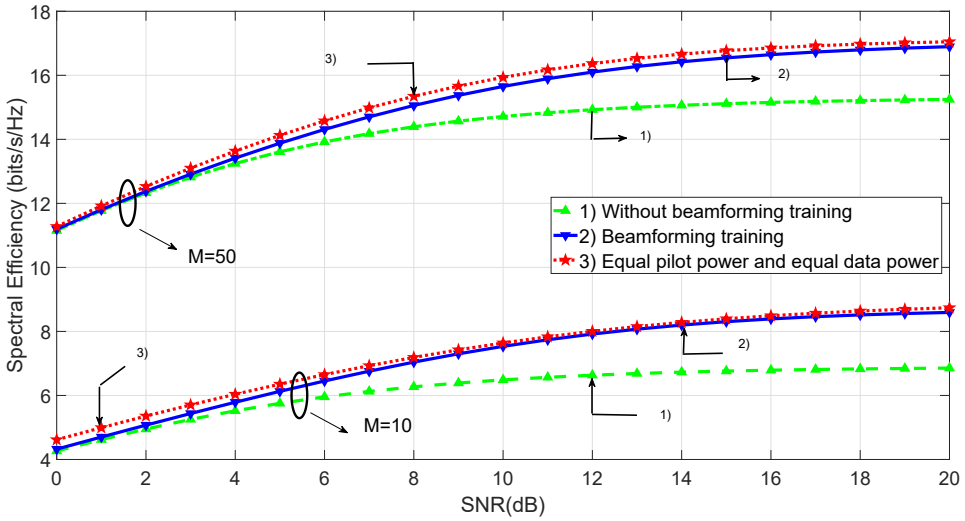

Figure 3.2: The spectral efficiency versus SNR of the proposed method and that of the methods using BT [24], without BT [22], and optimal power allocation for BT [40] with the number of BS antennas M=10 and M=50, where $E_{tu} = 6.9dB$ and $K = 5$

power at low SNR. It is also seen that the approximately half of the total energy budget is employed for DL training and the other half is employed for DL data transmission at low SNR.

As mentioned above, the proposed method of power allocation improves SE only slightly with that offered by the method in [24]. This motivated us to propose another method of power allocation in the following section in such a way that the power is allocated among all the data symbols of the users.

3.2.5 Equal pilot power for each user is same and is specified

We assume that the pilot power for each user is the same and is specified, i.e., $p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ is specified and the data power for each user, i.e., $p_{d_1}, p_{d_2}..., p_{d_K}$ is optimally selected to maximize SE. In order to maximize SE, a new method for power allocation among users is proposed based on the water-filling approach. Since maximizing SE is an NP-hard problem, an effective algorithm is also proposed to find local maximum points. The performance of the proposed power allocation method is verified by conducting

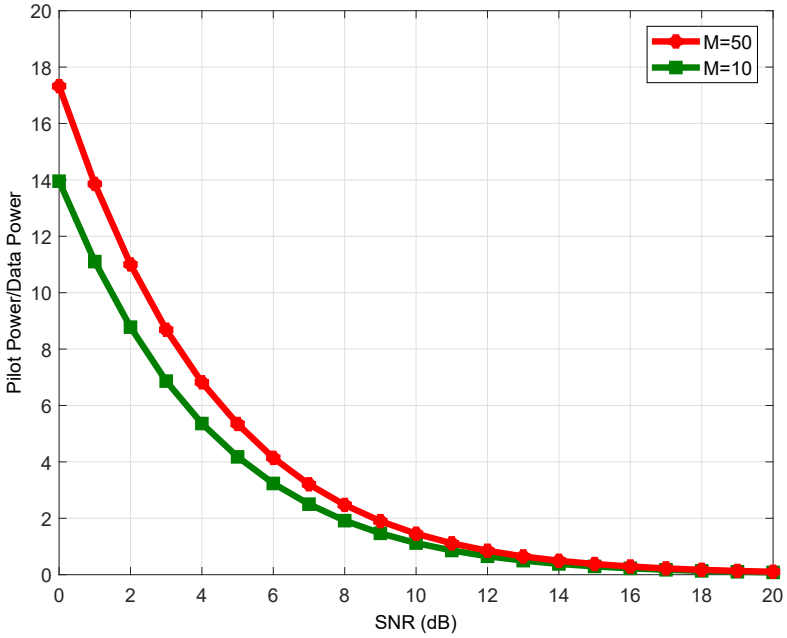

Figure 3.3: Ratio of the optimal pilot power to the optimal data power when $E_{tu} = 6.9dB$ and $K = 5$.

simulations and shown to be superior to other existing methods in terms of SE.

In this method, inspired by the water-filling power allocation method, larger power is allocated to the users with greater channel gains [70]. This strategy substantially increases the spectral efficiency compared to the case of allocating equal power. In view of this, let E_t be the total transmit power for BS per each channel use. Hence, the summation of the data transmit powers of the users is allowed not to exceed E_t and the data transmit power of each user can not be negative. Noting the above, the optimum power allocation strategy at BS can be formulated as the following optimization problem:

$$\max_{p_{d_1}, p_{d_2}, \ldots, p_{d_K}} \quad S$$
$$\text{such that} \quad \begin{cases} \sum_{k=1}^{K} p_{d_k} \leq E_t. \\ p_{d_k} \geq 0 \ , \ \forall k. \end{cases} \tag{3.45}$$

For a given $\sum_{i \neq k}^{K} p_{d_k}$, R_k is an increasing function of p_{d_k} as given by (3.19). As a result,

S is an increasing function of p_{d_k} as given by (3.20). Thus, S is maximized when the inequality of the total transmit power in (3.45) becomes the equality. Hence, the lower bound in (3.19) changes to

$$R_k = \mathbb{E}\Big[\log_2\big(1+\mathrm{SINR}_k\big)\Big], \tag{3.46}$$

where

$$\mathrm{SINR}_k = \frac{p_{d_k}\left|\frac{\sqrt{\tau_d p_p}}{\tau_d p_p+K}\tilde{y}_{p,k,k}+\frac{K}{\tau_d p_p+K}\sqrt{\frac{M\tau_u p_u}{K(\tau_u p_u+1)}}\right|^2}{\frac{E_t}{\tau_d p_p+K}+\sum_{i\neq k}^{K} p_{d_i}\left|\frac{\sqrt{\tau_d p_p}}{\tau_d p_p+K}\tilde{y}_{p,k,i}\right|^2+1}\ , \tag{3.47}$$

and the optimization problem given by (3.45) changes to

$$\begin{aligned} &\max_{p_{d_1},p_{d_2},\ldots,p_{d_K}} && S \\ &\text{such that} && \begin{cases} \sum_{k=1}^{K} p_{d_k} = E_t. \\ p_{d_k} \geq 0\ ,\ \forall k. \end{cases} \end{aligned} \tag{3.48}$$

Even under perfect CSI, power control problem to maximize the spectral efficiency is known to be an NP-hard problem [71]. As a result, we intend to find a local optimal solution for problem given by (3.48) along with an acceptable computational complexity. To this end, by using the epigraph form, let us equivalently reformulate the problem given by (3.48) as

$$\begin{aligned} &\max_{p_{d_k},\lambda_k} && \prod_k \lambda_k \\ &\text{such that} && \begin{cases} \sum_{k=1}^{K} p_{d_k} - E_t = 0\ , \\ 1+\mathrm{SINR}_k \geq \lambda_k\ ,\ \forall k \\ p_{d_k} \geq 0\ ,\ \forall k \end{cases} \end{aligned} \tag{3.49}$$

To obtain a valid Karush-Kuhn-Tucker (KKT) point (local maximum point) of the problem given by (3.52), we employ the general inner approximation algorithm [72]. In the problem of (3.52), only the constraints involving $SINRs$ are non-convex functions. Thus, we approximate $f(\mathbf{p}_k) = 1 + SINR_k$ by constructing a family of functions $\tilde{f}_i(\mathbf{p}_k)$ in each iteration i, where $\mathbf{p}_k = (p_{d_1}, ..., p_{d_K}, \lambda_k)$. This approximation is employed for every user k. Moreover, the approximated functions need to satisfy the following conditions [72]:

1. $f(\mathbf{p}_k) \leq \tilde{f}_i(\mathbf{p}_k),\ \forall \mathbf{p}_k$ in the feasible set,

2. $f(\mathbf{p}_k^{(i-1)}) \leq \tilde{f}_i(\mathbf{p}_k^{(i-1)})$, where $\mathbf{p}_k^{(i-1)}$ denotes the solution from the previous iteration and

3. $\nabla f(\mathbf{p}_k^{(i-1)}) = \nabla \tilde{f}_i(\mathbf{p}_k^{(i-1)})$.

We replace $f(\mathbf{p}_k)$ with $\tilde{f}_i(\mathbf{p}_k)$ in the i^{th} iteration, and solve the optimization problem given by (3.52). This algorithm converges to a KKT point of the problem given by (3.48).

The first condition ensures that the solution in each iteration is feasible for the problem (3.52). The second condition guarantees that the solution from the previous iteration is feasible for the current iteration. Thus, in each iteration, the objective value of the problem of (3.52) increases, since the solution from the previous iteration is a feasible point for the problem of (3.52) in the current iteration. The third condition ensures that the KKT conditions for problem of (3.52) are satisfied at the convergence. The objective value is monotonically increasing and bounded from the above. Hence, convergence of the algorithm is guaranteed.

To construct the approximated functions $\tilde{f}_i(\mathbf{p}_k)$, we employ the following Lemma

Lemma 3.6 : Let $g(x) = \sum_i m_i(x)$ be a posynomial. Then, for any α_i, we have [73]

$$g(x) \geq \tilde{g}(x) = \prod_i \left(\frac{m_i(x)}{\alpha_i}\right)^{\alpha_i} \tag{3.50}$$

In (3.52), the SINR constraints are not valid posynomial constraints since they are in the form $h(x)/g(x)$. In view of this, applying the above lemma on the denominator of SINR to replace $g(x)$ with $\tilde{g}(x)$ and leaving the numerator of SINR as a valid posynomial, we can make a valid posynomial constraint. In addition, by selecting $\alpha_i = m_i(x_0)/g(x_0)$, the aforementioned three conditions are satisfied. Employing this method for every SINR constraint in problem of (3.52) leads to a convex approximation of problem, which can be solved by convex optimization methods [74]. This procedure is repeated until convergence. To summarize the aforementioned procedure, we present Algorithm 3.1 to obtain a KKT point of the problem given by (3.52).

Algorithm 3.1 Successive convex optimization method for obtaining a KKT point of the optimization problem given by (3.52)

1: Initialize $i = 1$ and choose $P_k^{(0)}$ as the solution of the optimization problem (3.52) so that the constraints are satisfied.
2: Repeat.
3: Construct the i^{th} approximated optimization problem of (3.52) by employing Lemma 3.6 for every SINR constraint.
4: Solve the i^{th} approximated problem, which is a convex problem and obtain $P_k^{(i)}$ for every user k.
5: $i \leftarrow i + 1$.
6: Until convergence.
7: Return all $P_k^{(i)}$.

3.2.6 Experimental Results

The spectral efficiency of the proposed method in Section 3.2.5 is demonstrated through extensive experiments and is compared with that of the method provided in Section 3.2.3. For a fair comparison, we choose $\tau_u = \tau_d = K = 5$ and $p_u = 0dB$ in all the aforementioned schemes. We define $SNR = E_t/K$. Since E_t is the total transmit power for each channel use by K users and the noise variance is unity, SNR has the interpretation of average transmit signal-to-noise ratio (SNR) and is dimensionless.

Fig. 3.4 shows the spectral efficiency versus SNR when $T = 200$ (for example $1ms \times 200kHz$) for $M = 10$ and $M = 50$. It is seen from this figure that the method proposed in

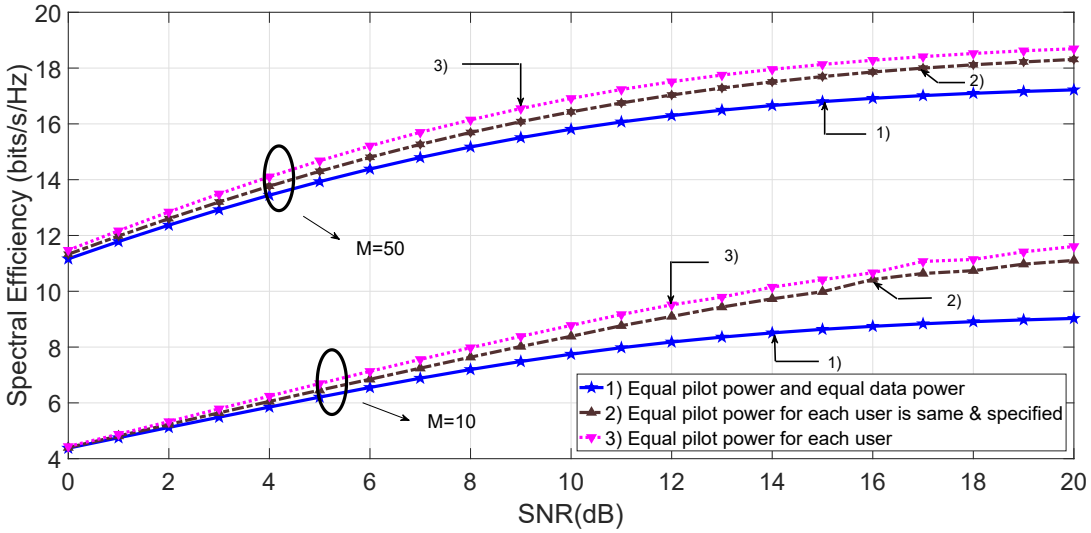

Figure 3.4: The spectral efficiency of the proposed methods and that of provided in Section 3.2.3 versus *SNR*, with the number of BS antennas M=10 and M=50, where $E_{tu} = 6.9dB$ and $K = 5$

Section 3.2.5 outperforms the scheme of Section 3.2.3 at all *SNR* values. The reason for this is that the optimal transmitted power among the users has been derived to maximize the spectral efficiency. It is also seen that the advantage of the proposed method become even more prominent at high *SNR*, since the channel estimate for each user is more accurate at high *SNR* [24]. In view of this, the spectral efficiency of the proposed method slightly improve at low *SNR* in comparison with that of Section 3.2.3.

Fig. 3.5 shows the optimal power allocated to the users versus *SNR* using the proposed method (note that for the y-axis in this figure a logarithmic scale has been plotted). It is also seen from this figure that approximately equal power has been allocated to the users at low *SNR*. As a result, at low *SNR*, the spectral efficiency of the proposed method is not as good as that of [24] where the power is equally allocated. However, it is seen from Fig. 3.4 that the proposed method outperforms that of [24] at high *SNR*.

By using this method of allocating power to the different users, SE is significantly improved as shown in Fig 3.4. This motivated us to present another method of power allocation in the following section in which the power is allocated not only to the data symbols but also to the pilot symbols.

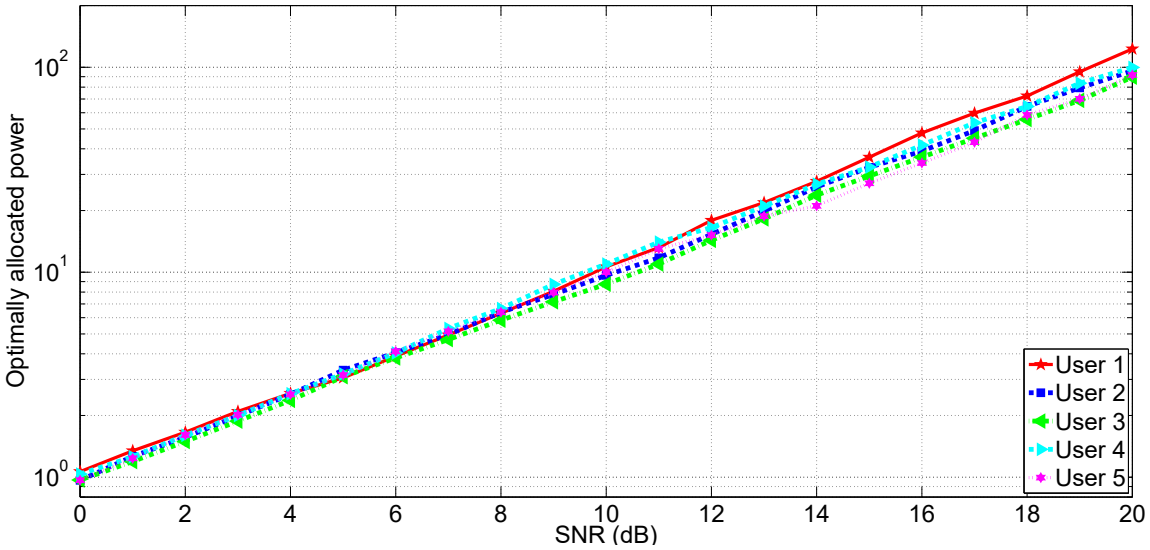

Figure 3.5: Optimal power allocated to users versus SNR when M=10.

3.2.7 Equal pilot power for each user

We optimally allocate equal pilot power and different data power for each user in such a way that SE is maximized, i.e., $p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ and $p_{d_1}, p_{d_2}..., p_{d_K}$ are optimally selected. In this case, the optimization problem is given by

$$\begin{aligned} &\max_{p_p, p_{d_1}, p_{d_2}, ..., p_{d_K}} \quad S \\ &\text{such that} \quad \begin{cases} \sum_{k=1}^{K} \tau_d p_p + (T - \tau_d - \tau_u) p_{d_k} = E_t. \\ p_p \geq 0 \ , p_{d_k} \geq 0 \ , \ \forall k. \end{cases} \end{aligned} \tag{3.51}$$

Using the epigraph form, let us equivalently reformulate the problem given by (3.51) as

$$\begin{aligned} &\max_{p_p, p_{d_k}, \lambda_k} \quad \prod_k \lambda_k \\ &\text{such that} \quad \begin{cases} \sum_{k=1}^{K} \tau_d p_p + (T - \tau_d - \tau_u) p_{d_k} = E_t \\ 1 + \text{SINR}_k \geq \lambda_k \ , \ \forall k \\ p_p \geq 0 \ , p_{d_k} \geq 0 \ , \ \forall k \end{cases} \end{aligned} \tag{3.52}$$

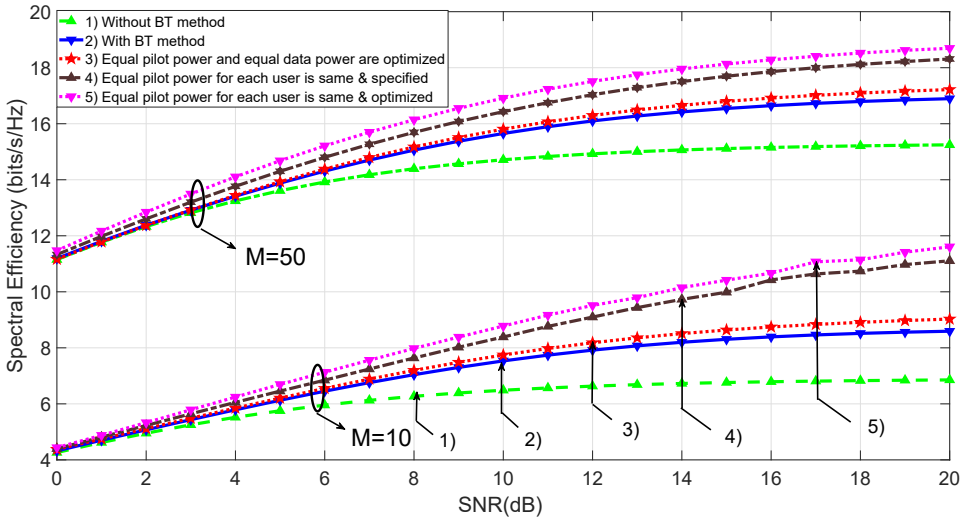

Figure 3.6: The spectral efficiency versus *SNR* of the proposed methods and that of the methods using BT [24] and without BT [22], with the number of BS antennas M=10 and M=50, where $E_{tu} = 6.9dB$ and $K = 5$

With the same approach employed in the previous section, this optimization problem is solved using Algorithm 3.1 in which $1 + \text{SINR}_k = f(\mathbf{p}_k)$ and $\mathbf{p}_k = (p_{d_1}, ..., p_{d_K}, p_p\ , \lambda_k)$.

3.2.8 Experimental Results

In Fig. 3.4, the spectral efficiency of the method proposed in Section 3.2.7 is demonstrated through extensive experiments and is compared with that of provided in Section 3.2.5 and Section 3.2.2. It can be seen from this figure that the method proposed in Section 3.2.7 outperforms the other methods.

Finally, Fig. 3.6 shows the spectral efficiency versus *SNR* for all the aforementioned schemes.

3.2.9 Summary

We have investigated the downlink transmission in a single multi-user massive MIMO system under time-division duplexing operation via a beamforming training method. We have proposed three methods of power allocation in order to maximize spectral efficiency.

The performance of the proposed methods have been demonstrated by conducting simulations and shown to be superior to that of the other two existing methods provided in Sections 3.2.1 and 3.2.2 in terms of spectral efficiency.

Conventionally, the transmit powers of the pilot and data symbols are considered to be equal in the beamforming training scheme. In the first proposed method, we have allocated equal pilot power and equal data power for all users in order to maximize the spectral efficiency.

Conventionally, the transmitted power at BS has been considered to be equal among the data symbols of the various users in the downlink transmission. In the second proposed method, inspired by the water-filling power allocation scheme, we have posed and answered a basic question about the operation of BS in downlink massive MU-MIMO systems as to how much is the improvement in the spectral efficiency if the power allocated to the data symbols of the various users are chosen optimally? In answering this question, we have found that the spectral efficiency can be significantly increased at high SNR, by optimally allocating the total transmit power to the users.

These two aforementioned methods motivated us to present a third method of power allocation in which the power is allocated not only to the data symbols but also to the pilot symbols of all the users (assuming all the pilot powers to be equal) in order to maximize the spectral efficiency. We have shown that this method of power allocation outperforms the two aforementioned methods in terms of spectral efficiency.

Chapter 4

Spectral efficiency for downlink transmission with large scale fading

In the previous chapter, three methods of power allocation between the data symbols and pilot symbols have been proposed to maximize the SE in the BT scheme for a given total power budget in a coherence interval for each user, assuming the channel between the users and the base station to have small scale fading. However, due to the channel hardening effect in massive MU-MIMO, allocating the power based on large-scale fading is more realistic rather than that based on small-scale fading. This case has been investigated for UL transmission in [32]. However, to the best of the author's knowledge, the power allocation problem that jointly determines the data power and pilot power allocation among the various users based on large-scale fading for the BT scheme for the downlink transmission has not yet been studied in the literature. This motivated the author to optimize the pilot power as well as the data power for each user to maximize SE, where a total power budget is given per coherence interval for all users. Intuitively, the SE obtained by this method should be superior to that of the methods provided in the previous chapter.

In view of this, in this chapter, first, we derive a lower bound on the achievable rate of each user based on large-scale fading in a single cell massive MU-MIMO DL system. Then, we propose various power allocation schemes among the pilot and data symbols of

the various users in order to maximize SE, where the total energy budget per coherence interval is given for all users. Since maximizing SE via jointly optimal pilot and data power allocation is an NP-hard optimization problem, we propose algorithms based on the general inner approximation algorithm to find the local maximum points.

4.1 System Model

Consider a single-cell massive MU-MIMO system for the DL transmission, where a BS with M multiple antennas serves K single antenna users in the same frequency band. In this system model, BS uses MRT precoding before the DL transmission. In view of this, BS needs to acquire CSI which is obtained through the UL training. Since there is a channel reciprocity between UL and DL channels in TDD operation, BS employs the obtained CSI to precode the data symbols in DL transmission.

4.1.1 Uplink Training

Let $\mathbf{G}$ be the $M \times K$ channel matrix between the users and BS, where the elements of each column of $\mathbf{G}$ are identical and independent distribution (i.i.d) with a Gaussian distribution having zero mean and β_k variance, where β_k represents the large-scale fading including path loss and shadowing for the kth user, i.e., $\mathbf{g}_k \sim \mathcal{CN}(0, \beta_k)$. To estimate G, the users simultaneously transmit orthogonal pilot sequences with length τ_u symbols per coherence interval to BS, where $\tau_u \geq K$. The pilot matrix of the users is denoted by $\mathbf{\Psi} = [\phi_1, \phi_2, ..., \phi_K] \in \mathbb{C}^{\tau_u \times K}$ with the orthogonality property $\mathbf{\Psi}^\dagger \mathbf{\Psi} = \mathbf{I_K}$, where ϕ_k denotes the pilot sequence of the kth user and $(.)^\dagger$ denotes the Hermitian operation of the associated matrix. In view of this, the $M \times \tau_u$ pilot matrix received at BS can be written as [32]

$$\mathbf{Y_u} = \sqrt{\tau_u p_u}\ \mathbf{G}\ \mathbf{\Psi}^\dagger +\ \mathbf{N_u}, \tag{4.1}$$

where p_u and $\mathbf{N_u} \in \mathbb{C}^{M\times\tau_u}$ denote, respectively, the transmit signal-to-noise ratio (SNR) of each pilot symbol and the received noise matrix at BS. Using the received pilot sequences and minimum mean-square error (MMSE) estimation, BS estimates $\mathbf{g}_k$ is given by [67]

$$\hat{\mathbf{g}}_k = \frac{\tau_u p_u \beta_k}{1+\tau_u p_u \beta_k}\mathbf{g}_k \; + \; \frac{\sqrt{\tau_u p_u \beta_k}}{1+\tau_u p_u \beta_k}\tilde{\mathbf{n}}_k, \tag{4.2}$$

where $\tilde{\mathbf{n}}_k \sim \mathcal{CN}(0, \mathrm{I}_M)$ is independent of $\mathbf{g}_k$. In this case, $\mathbf{g}_k$ can be decomposed as

$$\mathbf{g}_k = \hat{\mathbf{g}}_k + \epsilon_k, \tag{4.3}$$

where ϵ_k is the channel estimation error. Since MMSE channel estimation is employed, $\mathbf{g}_k$ and ϵ_k are independent. Moreover, $\hat{\mathbf{g}}_k$ and ϵ_k have i.i.d elements $C\mathcal{N}(0, \frac{\tau_u p_u \beta_k^2}{1+\tau_u p_u \beta_k})$ and $C\mathcal{N}(0, \frac{\beta_k}{1+\tau_u p_u \beta_k})$, respectively.

4.1.2 Downlink Transmission

After the estimation of the channel, BS employs $\hat{\mathbf{g}}_k$ to linearly precode the data symbols. Let s_k be the symbol that is transmitted to the kth user with $\mathbb{E}\{|s_k|^2\} = 1$ and let $\mathbf{W} \in \mathbb{C}^{M\times K}$ be the linear precoding matrix. In this case, the $M \times 1$ transmit signal vector can be written as $\mathbf{x} = \mathbf{W}\mathbf{s}$, where $\mathbf{s} \triangleq [\sqrt{p_{d_1}}s_1, \sqrt{p_{d_2}}s_2, ..., \sqrt{p_{d_K}}s_K]^T$, $(.)^T$ denoting the transpose operator and p_{d_k} is the transmit SNR of each data symbol for the kth user. To satisfy the power constraint at BS, we have

$$\mathbb{E}\big[\mathrm{tr}\big(\mathbf{W}\mathbf{W}^\dagger\big)\big] = 1 \; . \tag{4.4}$$

Since it is shown in [24] that the MRT precoding is more efficient than the ZF precoding in the BT scheme, we employ the MRT precoding in this thesis. In view of this, $\mathbf{W}$ is

given by

$$\mathbf{W} = \alpha \hat{\mathbf{G}}^*, \tag{4.5}$$

where $(.)^*$ denotes the conjugate operator and α denotes a normalization constant for satisfying the transmit power constraint given by (4.4) which is obtained as $\mathbb{E}\big[\mathrm{tr}\big(\alpha^2 \hat{\mathbf{G}}^* \hat{\mathbf{G}}^T\big)\big]$=1. Equivalently, in this system model, using (4.2), α is obtained as

$$\alpha = \sqrt{\frac{1}{M\tau_u \; p_u \sum_{i=1}^{K} \beta_i^2/(1+\tau_u p_u \beta_i)}} \; . \tag{4.6}$$

In this case, the received vector in DL transmission at the users is given by

$$\mathbf{y} = \mathbf{G}^T \mathbf{W} \mathbf{s} + \mathbf{n}, \tag{4.7}$$

where $\mathbf{n}$ is a vector whose kth element is additive noise at the kth user that is denoted by $\mathbf{n}_k \sim \mathcal{CN}(0,1)$. Let us define $a_{ki} \triangleq \mathbf{g}_k^T \mathbf{w}_i$, where $\mathbf{w}_i$ is the ith column of $\mathbf{W}$. In view of this definition, the received signal at the kth user is given by

$$y_k = \sqrt{p_{d_k}} \; a_{kk} s_k + \sum_{i=1, i\neq k}^{K} \sqrt{p_{d_i}} a_{ki} s_i + n_k. \tag{4.8}$$

Since each user requires CSI to accurately detect the transmitted signal, we employ the BT scheme to estimate the effective channel gain a_{kk} at each user [24].

4.1.3 Beamforming Training

In the BT scheme, short pilot sequences are transmitted by BS in DL transmission [24]. Then, each user estimates the effective channel gain a_{ki} using the received pilot sequences and the MMSE channel estimation. Let $\mathbf{S}_p \in \mathbb{C}^{K \times \tau_d}$ be a pilot matrix in the DL transmission, where τ_d is the number of symbols for pilot sequences. In view of this, kth column of

the pilot matrix can be written as $\mathbf{S}_{p,k} = \sqrt{\tau_d p_{p_k}}\, \mathbf{\Phi}_k$, where p_{p_k} and $\mathbf{\Phi}_k$ denote SNR of each pilot symbol for kth user and kth column of pilot sequence matrix in DL transmission, respectively. Due to the orthogonality of the pilot sequences, we require that $\tau_d \geqslant K$. BS transmits the pilot matrix $\mathbf{WS}_p$ to the users. Hence, the received pilot matrix in DL transmission can be expressed as

$$\mathbf{Y}_p = \mathbf{G}^T \mathbf{W} \mathbf{S}_p + \mathbf{N}_p, \tag{4.9}$$

where $\mathbf{N}_p$ denotes the AWGN noise matrix whose elements are i.i.d. $\mathcal{CN}(0,1)$. To estimate a_{kk} at each user, we use the orthogonality of the pilot sequences. To this end, let $\tilde{\mathbf{Y}}_p \triangleq \mathbf{Y}_p \mathbf{\Phi}^\dagger$. Thus, we have

$$\tilde{\mathbf{Y}}_p = \mathbf{G}^T \mathbf{W} \mathbf{P} + \tilde{\mathbf{N}}_p, \tag{4.10}$$

where $\tilde{\mathbf{N}}_p \triangleq \mathbf{N}_p \mathbf{\Phi}^\dagger$ has the same distribution as $\mathbf{N}_p$ and $\mathbf{P} = [\, \sqrt{\tau_d p_{p_k}} \,]$ denotes the $K \times K$ diagonal matrix of the power allocation coefficients. Decomposing $\tilde{\mathbf{Y}}_p$ given by (4.9), we have

$$\tilde{\mathbf{y}}_{p,k} = \sqrt{\tau_d p_{p_k}}\, \mathbf{g}_k^T \mathbf{W} + \tilde{\mathbf{n}}_{p,k} = \sqrt{\tau_d p_{p_k}}\, \mathbf{a}_k^T + \tilde{\mathbf{n}}_{p,k}, \tag{4.11}$$

where $\tilde{\mathbf{y}}_{p,k}$ and $\tilde{\mathbf{n}}_{p,k}$ denote the kth columns of $\tilde{\mathbf{Y}}_p$ and $\tilde{\mathbf{N}}_p$, respectively and $\mathbf{a}_k \triangleq [a_{k1} a_{k2} ... a_{kK}]^T$. Using (4.10) and MMSE channel estimation, kth user estimates $\mathbf{a}_k$. In the BT scheme, a_{k1},...,a_{kK} are independently estimated by the kth user [24]. As a result, the estimation of a_{ki} can be written as [67]

$$\hat{a}_{ki} = \mathbb{E}[a_{ki}] + \frac{\mathrm{cov}(a_{ki}, \tilde{y}_{p,ki})}{\mathrm{cov}(\tilde{y}_{p,ki}, \tilde{y}_{p,ki})} \big(\tilde{y}_{p,ki} - \mathbb{E}[\tilde{y}_{p,ki}]\big) \tag{4.12}$$

where $\tilde{y}_{p,ki}$ denotes the ith element of $\tilde{\mathbf{y}}_{p,k}$. Let ε_{ki} be the channel estimation error. Employing MMSE channel estimation, the estimate $\hat{a}_{ki}$ and the estimation error ε_{ki} are

uncorrelated. Hence, a_{ki} can be written as

$$a_{ki} = \hat{a}_{ki} + \varepsilon_{ki} \tag{4.13}$$

Substituting (4.13) into (4.8), we have

$$y_k = \sqrt{p_{d_k}}\hat{a}_{kk}s_k + \sum_{i=1, i\neq k}^{K}\sqrt{p_{d_i}}\hat{a}_{ki}s_i + \sum_{i=1}^{K}\sqrt{p_{d_i}}\epsilon_{ki}s_i + n_k. \tag{4.14}$$

4.1.4 Achievable Downlink Rate

Following the work in [24] and using (4.14), a lower bound on the achievable DL rate for the kth user is obtained as

$$R_k = \mathbb{E}\left[\log_2\left(1 + \frac{p_{d_k}|\hat{a}_{kk}|^2}{\sum_{i=1}^{K} p_{d_i}\mathbb{E}\{|\varepsilon_{ki}|^2\} + \sum_{i\neq k}^{K} p_{d_i}|\hat{a}_{ki}|^2 + 1}\right)\right]. \tag{4.15}$$

To calculate the achievable DL rate for the kth user given by (4.15), we should generate the channel and compute $\hat{a}_{ki}$ several times. Then, we average over all realizations. To simplify these calculations, we obtain a close approximation of the achievable DL rate for the kth user employing Lemma 3.1.

Using Lemma 3.1, the achievable DL rate given by (4.15) can be approximated as

$$\hat{R}_k \approx \log_2\left(1 + \frac{p_{d_k}\ \mathbb{E}\{|\hat{a}_{kk}|^2\}}{\sum_{i=1}^{K} p_{d_i}\mathbb{E}\{\mathbb{E}\{|\epsilon_{ki}|^2\}\} + \sum_{i\neq k}^{K} p_{d_i}\mathbb{E}\{|\hat{a}_{ki}|^2\} + 1}\right). \tag{4.16}$$

Proposition 4.1 : Using MRT precoding, the achievable DL rate given by (4.16) can be written as

$$\hat{R}_k = \log_2(1 + \mathrm{SINR}_k), \tag{4.17}$$

where

$$\mathrm{SINR}_k = \frac{\frac{M\tau_u p_u p_{d_k}}{(\alpha\beta_k)^2}\gamma_{kk}^2 + \frac{p_{d_k}\gamma_{kk}^2}{\gamma_{kk} + \frac{1}{M\tau_u p_u \tau_d p p_k}}}{\frac{1}{M\tau_u p_u} + \sum_{i\neq k}^{K}\frac{p_{d_i}\gamma_{ki}^2}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p p_i}} + \sum_{i=1}^{K}\frac{\frac{p_{d_i}\gamma_{ki}}{M\tau_u p_u \tau_d p p_i}}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p p_i}}}, \tag{4.18}$$

and $\gamma_{ki} = \alpha^2 \frac{\beta_i^2 \beta_k}{1+\tau_u p_u \beta_i}$. The proof of Proposition 4.1 is given in Appendix C.

4.2 Spectral efficiency in two simple cases

The spectral efficiency, as defined in (3.20), is reproduced below for convenience

$$S = \frac{T - \tau_u - \tau_d}{T} \sum_{k=1}^{K} \hat{R}_k, \tag{4.19}$$

where T is the length of the coherence interval in DL transmission. We now consider the spectral efficiency for the following two cases.

4.2.1 No pilot power and equal data power for all users

We assume that BS does not transmit any pilot symbols to the users and the data power for each of the users is the same, i.e., $\tau_d = 0$ and $p_{d_1} = p_{d_1} = ... = p_{d_K} = p_d$. Hence, from (4.17) and (4.18) we see that the lower bound for ADR is

$$\hat{R}_k = \log_2(1 + \mathrm{SINR}_k), \tag{4.20}$$

where

$$\mathrm{SINR}_k = \frac{\frac{M\tau_u p_u p_d}{(\alpha\beta_k)^2}\gamma_{kk}^2}{\frac{1}{M\tau_u p_u} + \sum_{i=1}^{K} p_d \gamma_{ki}}, \tag{4.21}$$

In this method of power allocation, SE is obtained by substituting (4.20) in (4.19).

4.2.2 Equal pilot and data powers for each user

We assume that the pilot power as well as the data power for each user is the same, i.e., $p_{p_1} = ... = p_{p_K} = p_{d_1} = ... = p_{d_K} = p_d$. Hence, from (4.17) and (4.18) the lower bound for ADR is obtained as

$$\hat{R}_k = \log_2(1 + \text{SINR}_k) \tag{4.22}$$

where

$$\text{SINR}_k = \frac{\frac{M\tau_u p_u}{(\alpha\beta_k)^2}\gamma_{kk}^2 + \frac{\gamma_{kk}^2}{\gamma_{kk} + \frac{1}{M\tau_u p_u \tau_d p_d}}}{\frac{1}{M\tau_u p_u p_d} + \sum_{i\neq k}^{K}\frac{\gamma_{ki}^2}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p_d}} + \sum_{i=1}^{K}\frac{\frac{\gamma_{ki}}{M\tau_u p_u \tau_d p_d}}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p_d}}}, \tag{4.23}$$

Again, in this method of power allocation, SE is obtained by substituting (4.22) in (4.19).

4.3 Maximization of spectral efficiency

In the following section, we present a method of power allocation in such a way that SE is maximized. In this method, pilot power as well as data power for each user is optimized to maximize SE, i.e., we optimize $p_{p_1}, ..., p_{p_K}$, as well as $p_{d_1}, ..., p_{d_K}$ in order to maximize SE.

4.3.1 Joint pilot and data powers for each user

Conventionally, power is equally allocated among the pilot and data symbols in DL transmission, but in this thesis, we allow BS to allocate power among each of the pilot and data symbols of the various users in a given total power budget per coherence interval for

all users in order to maximize SE. Mathematically speaking, we have

$$\begin{aligned} \max_{p_{d_k}, p_{p_k}} \quad & \frac{T-\tau_u-\tau_d}{T} \sum_{k=1}^{K} \log_2(1+\mathrm{SINR}_k) \\ \text{s.t} \quad & \begin{cases} \sum_{k=1}^{K} \tau_d p_{p_k} + (T-\tau_d-\tau_u) p_{d_k} = P_{\mathrm{t}} \\ p_{p_k} \geq 0, p_{d_k} \geq 0, \forall k. \end{cases} \end{aligned} \tag{4.24}$$

where P_{t} is the total power budget spent in a coherence interval for all users. The optimization problem given by (4.24) is an NP-hard problem. Hence, using general inner approximation algorithm, we find a local optimal solution for the optimization problem given by (4.24). In view of this, with the aid of epigraph form, let us equivalently rewrite the optimization problem given by (4.24) as

$$\begin{aligned} \max_{p_{d_k}, p_{p_k}, \lambda_k} \quad & \prod_{k} \lambda_k \\ \text{s.t} \quad & \begin{cases} \sum_{k=1}^{K} \tau_d p_{p_k} + (T-\tau_d-\tau_u) p_{d_k} = P_{\mathrm{t}} \\ p_{p_k} \geq 0, p_{d_k} \geq 0, \forall k \\ 1+\mathrm{SINR}_k \geq \lambda_k \ , \forall k, \end{cases} \end{aligned} \tag{4.25}$$

where λ_k is a variable of the optimization problem for the kth user. In this optimization problem, only the constraints for SINR_k are not convex, which can be written as $\lambda_k - \mathrm{SINR}_k \leq 1 \ , \forall k$. After simplification, it can be shown that this constraint can be rewritten as $w(\mathbf{q}_k) = z(\mathbf{q}_k)/h(\mathbf{q}_k) \leq 1 \ , \forall k$, where $z(\mathbf{q}_k)$ and $h(\mathbf{q}_k)$ are posynomial functions and $\mathbf{q}_k = (\mathbf{p}_d, \mathbf{p}_p, \lambda_k)$, $\mathbf{p}_d = \{p_{d_1}, p_{d_2}, ..., p_{d_K}\}$, and $\mathbf{p}_p = \{p_{p_1}, p_{p_2}, ..., p_{p_K}\}$. Employing the general inner approximation algorithm [72], we intend to find an appropriate Karush-Kuhn-Tucker (KKT) point (local maximum point) for the optimization problem given by (4.25). In view of this, we approximate $w(\mathbf{q}_k)$ expressions in the ith iteration of the general inner approximation algorithm as $\hat{w}_i(\mathbf{q}_k)$ in such a way that $\hat{w}_i(\mathbf{q}_k)$ is a posynomial function.

Theorem 4.1 : To construct a posynomial function for the general inner approximation algorithm, it is required to satisfy the three following conditions [72].

1) $w(\mathbf{q}_k) \leq \hat{w}_i(\mathbf{q}_k),\ \forall \mathbf{q}_k,$

2) $w(\mathbf{q}_k^{(i-1)}) = \hat{w}_i(\mathbf{q}_k^{(i-1)})$, where $\mathbf{q}_k^{(i-1)}$ is the solution obtained from the previous iteration, and

3) $\nabla w(\mathbf{q}_k^{(i-1)}) = \nabla \hat{w}_i(\mathbf{q}_k^{(i-1)})$.

Since the objective value of the optimization problem given by (4.25) is bounded from above and increases monotonically in each iteration, convergence of the general inner algorithm is guaranteed.

Theorem 4.2 : Let $h(\mathbf{x}) = \sum_j n_j(\mathbf{x})$ be a posynomial function, where $n_j(\mathbf{x})$ and $\mathbf{x}$ are a monomial function and a set of variables, respectively. For any constant α_j, with the aid of the arithmetic-geometric mean inequality, we have [73]

$$h(\mathbf{x}) \geq \tilde{h}(\mathbf{x}) = \prod_j \left(\frac{n_j(\mathbf{x})}{\alpha_j}\right)^{\alpha_j}. \tag{4.26}$$

Moreover, for any fixed positive $\mathbf{x}_0$ and $\alpha_j = n_j(\mathbf{x}_0)/h(\mathbf{x}_0)$, $\tilde{h}(\mathbf{x})$ is the best local monomial approximation to $h(\mathbf{x_0})$ near $\mathbf{x_0}$ in the sense of first order Taylor approximation [73]. In addition, the three conditions given by Theorem 4.1 are satisfied by choosing $\alpha_j = n_j(\mathbf{x}_0)/h(\mathbf{x}_0)$ in Theorem 4.2 [13].

Lemma 4.1 : Replacing the denominator $h(\mathbf{q}_k)$ of the constraint for $w(\mathbf{q}_k)$ by $\tilde{h}(\mathbf{q}_k)$ and leaving the numerator, $z(\mathbf{q}_k)$, unchanged, the posynomial constraint $\hat{w}(\mathbf{q}_k) = z(\mathbf{q}_k)/\tilde{h}(\mathbf{q}_k)$ is constructed in each iteration of the general inner approximation algorithm [73].

Employing Lemma 4.1 leads to constructing a valid posynomial for every SINR constraint in the optimization problem given by (4.25). This results in a geometric programming (GP) optimization problem at each iteration of the general inner approximation algorithm since the objective function is a monomial and the constraints are posynomial [73]. This GP problem can be efficiently solved with any GP solver. In this thesis, we employ the ConVeX (CVX) package to solve this GP problem [75]. It is worth mentioning

that the GP optimization problem can be converted to a convex optimization problem by change of variable [73]. We repeat this procedure in the general inner approximation algorithm until convergence which is described in Algorithm 4.1.

Algorithm 4.1 Finding a KKT point of problem (4.25)

1: Initialize $i = 0$ and choose $\mathbf{q}_k^{(0)}$ as the solution of the optimization problem (4.25).
2: **while** $\mathbf{q}_k^{(i)}$ not converged **do**
3: $i \leftarrow i + 1$.
4: Setting $\alpha_j = n_j(\mathbf{q}_k^{(i-1)})/h(\mathbf{q}_k^{(i-1)})$ and using Lemma 4.1 for every SINR constraint, construct the ith approximated optimization problem given by (4.25)
5: The ith approximated optimization problem is a GP problem. Solve this GP problem using the ConVeX (CVX) package to obtain $\mathbf{q}_k^{(i)}$ for each k.
6: **end while**
7: Return all $\mathbf{q}_k^{(i)}$.

Employing Algorithm 4.1, we obtain the optimal values for $p_{p_1}, p_{p_2}, ..., p_{p_K}$, $p_{d_1}, p_{d_2}, ...,$ and p_{d_K}. Then, we substitute these optimal values into (4.19) in order to maximize the spectral efficiency.

In the following sections, we present three special cases of the proposed power allocation methods. These cases lead to various power allocation methods as discussed below.

1. Pilot power for each user is the same, with a similar statement holding true for the data power, i.e.,

$p_{p_1} = p_{p_2} = ... = p_{p_K} = p_p$.

$p_{d_1} = p_{d_2} = ... = p_{d_K} = p_d$.

2. Pilot power is the same for each user and is specified, while the data power for each user is determined to maximize SE, i.e.,

$p_{p_1} = p_{p_2} = ... = p_{p_K} = p_p$ is specified.

3. Pilot power is the same for each user, i.e.,

$p_{p_1} = p_{p_2} = ... = p_{p_K} = p_p$,

4.3.2 Equal pilot power and equal data power for all users

We assume that the pilot power for each user is the same as well as that for the data power for each user is the same, i.e.,$p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ and $p_{d_1} = p_{d_1} = ... = p_{d_K} = p_d$. Substituting these in (4.18), the lower bound of ADR in this case is given by

$$\hat{R}_k = \log_2(1 + \mathrm{SINR}_k), \tag{4.27}$$

where

$$\mathrm{SINR}_k = \frac{\frac{M\tau_u p_u}{(\alpha\beta_k)^2}\gamma_{kk}^2 + \frac{\gamma_{kk}^2}{\gamma_{kk} + \frac{1}{M\tau_u p_u \tau_d p_p}}}{\frac{1}{M\tau_u p_u p_d} + \sum_{i\neq k}^{K}\frac{\gamma_{ki}^2}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p_p}} + \sum_{i=1}^{K}\frac{\frac{\gamma_{ki}}{M\tau_u p_u \tau_d p_p}}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p_p}}}, \tag{4.28}$$

In this method of power allocation, we allocate the power among the pilot and data symbols in DL transmission given the energy budget in a coherence time in such a way that SE is maximized. Mathematically speaking, we have the following problem.

$$\begin{aligned} \max_{p_d, p_p} \quad & \frac{T - \tau_u - \tau_d}{T}\sum_{k=1}^{K}\log_2(1 + \mathrm{SINR}_k) \\ \text{s.t} \quad & \begin{cases} K\big(\tau_d p_p + (T - \tau_d - \tau_u)p_d\big) = P_{\mathrm{t}} \\ p_p \geq 0, p_d \geq 0. \end{cases} \end{aligned} \tag{4.29}$$

where P_{t} is the total power budget spent in a coherence interval for all users. Following the same procedure presented for solving the optimization problem given by (4.24), we obtain the optimal values p_p and p_d in the optimization problem given by (4.29). Then, we substitute these optimal values into (4.19) in order to maximize the spectral efficiency.

4.3.3 Equal pilot power for each user is same and is specified

We assume that the pilot power for each user is specified, i.e., $p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ is specified and the data power for each user, i.e., $p_{d_1}, p_{d_2}..., p_{d_K}$ is optimally selected to

maximize SE. Substituting these in (4.18), the lower bound of ADR in this case is given by

$$\hat{R}_k = \log_2(1 + \mathrm{SINR}_k), \tag{4.30}$$

where

$$\mathrm{SINR}_k = \frac{\frac{M\tau_u p_u p_{d_k}}{(\alpha\beta_k)^2}\gamma_{kk}^2 + \frac{p_{d_k}\gamma_{kk}^2}{\gamma_{kk} + \frac{1}{M\tau_u p_u \tau_d p_p}}}{\frac{1}{M\tau_u p_u} + \sum_{i\neq k}^{K}\frac{p_{d_i}\gamma_{ki}^2}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p_p}} + \sum_{i=1}^{K}\frac{\frac{p_{d_i}\gamma_{ki}}{M\tau_u p_u \tau_d p_p}}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p_p}}}, \tag{4.31}$$

In this method, we allocate the power among the data symbols in DL transmission for a given energy budget in a coherence time interval in such a way that SE is maximized. Mathematically speaking, we have the following problem.

$$\begin{aligned} \max_{p_{d_k}} \quad & \frac{T - \tau_u - \tau_d}{T}\sum_{k=1}^{K}\log_2(1 + \mathrm{SINR}_k) \\ \text{s.t} \quad & \begin{cases} \sum_{k=1}^{K}(T - \tau_d - \tau_u)p_{d_k} = P_{\mathrm{t}} - K\tau_d p_p \\ p_{d_k} \geq 0, \forall k. \end{cases} \end{aligned} \tag{4.32}$$

where P_{t} is the total power budget spent in a coherence interval for all users. Following the same procedure presented for solving the optimization problem given by (4.24), we obtain the optimal values p_{d_1},...,p_{d_K} in the optimization problem given by (4.32). Then, we substitute these optimal values into (4.19) in order to maximize the spectral efficiency.

4.3.4 Equal pilot power for each user

We now assume that the pilot power for each user is the same, i.e., $p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$. Substituting these in (4.18), the lower bound of ADR in this case is given by

$$\hat{R}_k = \log_2(1 + \mathrm{SINR}_k), \tag{4.33}$$

where

$$\mathrm{SINR}_k = \frac{\frac{M\tau_u p_u p_{d_k}}{(\alpha\beta_k)^2}\gamma_{kk}^2 + \frac{p_{d_k}\gamma_{kk}^2}{\gamma_{kk}+\frac{1}{M\tau_u p_u \tau_d p_p}}}{\frac{1}{M\tau_u p_u}+\sum_{i\neq k}^{K}\frac{p_{d_i}\gamma_{ki}^2}{\gamma_{ki}+\frac{1}{M\tau_u p_u \tau_d p_p}}+\sum_{i=1}^{K}\frac{\frac{p_{d_i}\gamma_{ki}}{M\tau_u p_u \tau_d p_p}}{\gamma_{ki}+\frac{1}{M\tau_u p_u \tau_d p_p}}}, \tag{4.34}$$

We now optimally allocate the equal pilot power p_p and the different data powers, i.e., $p_{d_1}, p_{d_2}..., p_{d_K}$ for each user in such a way that SE is maximized. Mathematically speaking, we have the following problem.

$$\begin{aligned} \max_{p_p, p_{d_k}} \quad & \frac{T-\tau_u-\tau_d}{T}\sum_{k=1}^{K}\log_2(1+\mathrm{SINR}_k) \\ \text{s.t} \quad & \begin{cases} K\tau_d p_p + \sum_{k=1}^{K}(T-\tau_d-\tau_u)p_{d_k} = P_{\mathrm{t}} \\ p_p \geq 0, p_{d_k} \geq 0, \forall k. \end{cases} \end{aligned} \tag{4.35}$$

where P_{t} is the total power budget spent in a coherence interval for all users. Following the same procedure presented for solving the optimization problem given by (4.24), we obtain the optimal values p_p, p_{d_1},...,p_{d_K} in the optimization problem given by (4.35). Then, we substitute these optimal values into (4.19) in order to maximize the spectral efficiency.

4.4 Numerical Results

We now study the relative performance of the various methods discussed in Sections 4.2 and 4.3 by conducting a number of experiments. In all the experiments, we set $T = 200$, $\tau_u = \tau_d = K$, $p_u = 0$ dB, and the power per symbol to be $10dB$ on an average. Thus, the total available power for DL transmission is $P_{\mathrm{t}} = 10 \times (T - \tau_u) \times K = 10(200 - K)K$. Moreover, we assume that no user is closer to BS than $r_h = 100$ meters and the radius of the cell is $R = 1000$ meters. We also assume that the path loss exponent is $v = 3.8$ and z_k is the log-normal random variable with standard deviation $\delta_{shadow} = 8$ dB. In view of

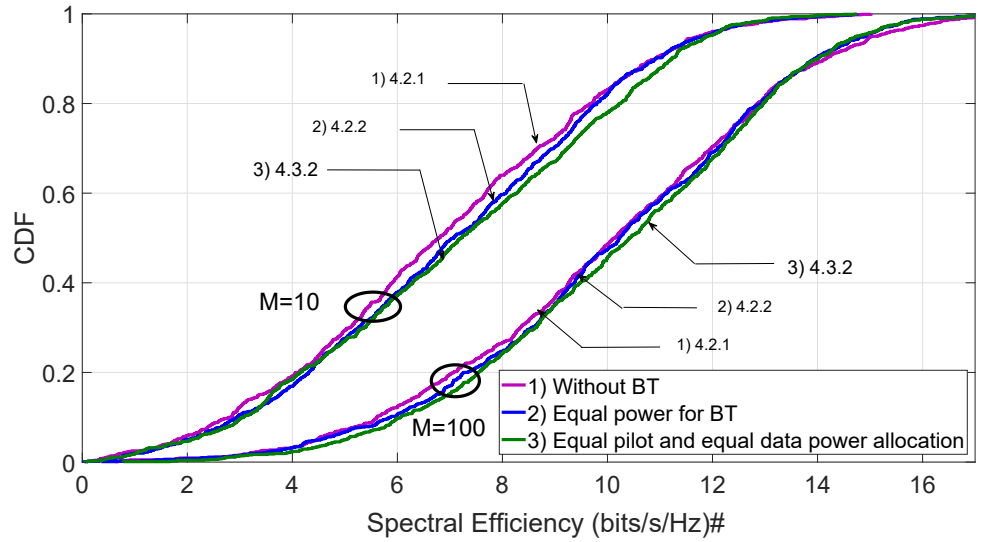

Figure 4.1: CDF of the spectral efficiency ($K = 5$).

this, the large-scale fading coefficient for the kth user given by $\beta_k = z_k/(d_k/r_h)^v$, where d_k is the distance between BS and the kth user. All methods are run for 1000 Monte-Carlo simulations, using MATLAB software and a PC with Intel(R) Core(TM) i5 @ 2.7 GHz processor and 4 GB installed memory (RAM). In each snapshot, the users are randomly located in the cell so that the large-scale fading coefficient for each user changes.

Fig. 4.1 shows the cumulative distribution function (CDF) of SE when $M = 10$ and $M = 100$ for the following three methods. 1) The method discussed in Section 4.2.1, where BS does not transmit any pilot symbols to the users and the data power for each of the users is the same, i.e., $\tau_d = 0$ and $p_{d_1} = p_{d_1} = ... = p_{d_K} = p_d$. 2) The method described in Section 4.2.2, where the pilot power as well as the data power for each user is the same, i.e., $p_{p_1} = ... = p_{p_K} = p_{d_1} = ... = p_{d_K} = p_d$. 3) The method proposed in Section 4.3.2, where the pilot power for each user is the same as well as that for the data power for each user is the same, i.e.,$p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ and $p_{d_1} = p_{d_1} = ... = p_{d_K} = p_d$. It can be seen from this figure that the method proposed in Section 4.3.2 improves SE very slightly over that offered by those of sections 4.2.1 and 4.2.2, where no optimization is required.

Fig. 4.2 shows the cumulative distribution function (CDF) of SE when $M = 10$ and $M = 100$ for following four methods. 1) The method proposed in 4.3.2, where the pilot power for each user is the same as well as that for the data power for each user is the same, i.e.,$p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ and $p_{d_1} = p_{d_1} = ... = p_{d_K} = p_d$. 2) The method proposed in Section 4.3.3, where the pilot power for each user is specified, i.e., $p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ is specified and the data power for each user, i.e., $p_{d_1}, p_{d_2}..., p_{d_K}$ is optimally selected to maximize SE. 3) The method proposed in Section 4.3.4, where the pilot power for each user is the same, i.e., $p_{p_1} = p_{p_1} = ... = p_{p_K} = p_p$ and p_p along with the data power for each user, i.e., $p_{d_1}, p_{d_2}..., p_{d_K}$ is optimally selected to maximize SE. 4) The method proposed in Section 4.3.1, where the pilot power as well as the data power for each user is optimized to maximize SE, i.e., $p_{p_1}, ..., p_{p_K}$, as well as $p_{d_1}, ..., p_{d_K}$ are optimized in order to maximize SE. It can be seen from this figure that SEs of the methods proposed in Sections 4.3.1, 4.3.3, and 4.3.4 are approximately the same, and very much superior to that of the method proposed in Section 4.3.2. This superiority in the performance can be attributed to the optimal power allocation to the data symbols for each of the users at BS in DL transmission, where a total power budget P_{t} is given for the users. As expected, it can be seen from this figure that when the number of antennas at BS increases, the performance of the proposed power allocation methods proposed in Sections 4.3.1, 4.3.3, and 4.3.4 significantly improves in terms of SE. This indicates that these power allocation methods are particularly suitable for massive MU-MIMO systems, where the number of antennas at BS is very large.

Next, we increase K to study the performance of the proposed methods in terms of SE. Fig. 4.3 shows CDF of SE for the power allocation methods proposed in Sections 4.3.1, 4.3.3, and 4.3.4 as well as for the equal power allocation method proposed in Section 4.2.2 when $K = 5$ and $K = 10$. It can be seen from this figure that when K increases, SE decreases irrespective of the method used.

In the optimization problem given by (4.25), the average run time for obtaining the

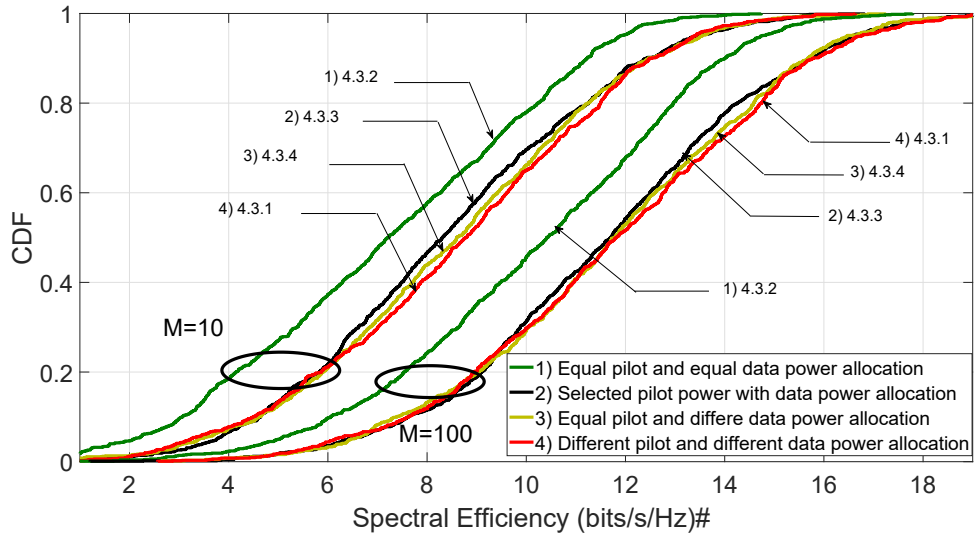

Figure 4.2: CDF of the spectral efficiency ($K = 5$).

solution is 1.086548 seconds averaged over all the 1000 snapshots when $K = 5$. The corresponding run times to obtain solutions for the problems given by (4.29), (4.32), and (4.35) are 0.546985, 0.731824, and 0.785823 seconds, respectively. These are shown in Tables 4.1 and 4.2 for $K = 5$. The corresponding run times to obtain solutions for the problems given by (4.25), (4.29), (4.32), and (4.35) when $K = 10$ are 2.3903, 0.7548, 1.0977 and 1.3359, respectively. These are shown in Tables 4.3 and 4.4.

Tables 4.1 and 4.2 show a comparison among the power allocation methods proposed in Sections 4.3.1, 4.3.3, 4.3.4 and 4.3.2, when $K = 5$; the values of the power are presented in Watts and dB, respectively. Tables 4.3 and 4.4 give the corresponding results when $K = 10$. It can be seen from these tables that optimizing the power of the data symbols is much more important than optimizing the power of the pilot symbols in order to maximize SE.

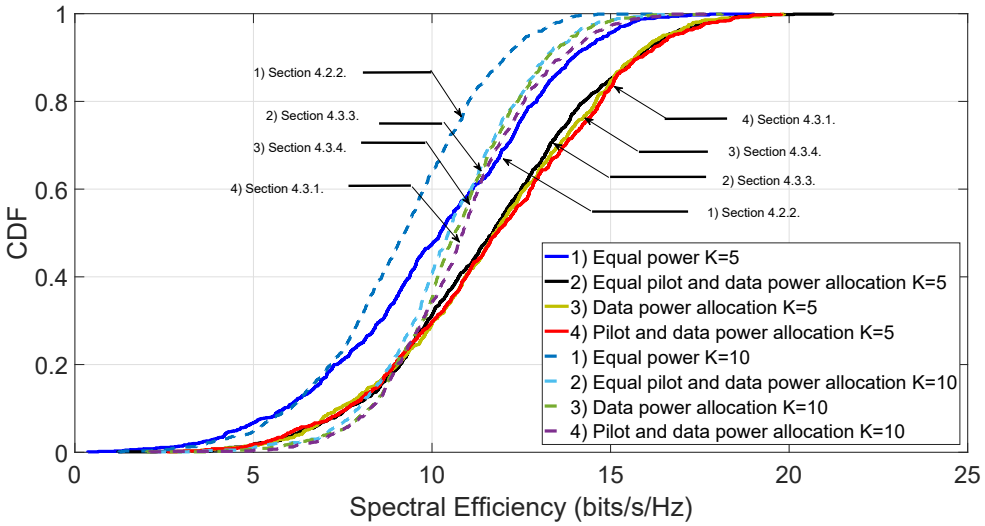

Figure 4.3: CDF of the spectral efficiency ($M = 100$).

4.5 Summary

In this chapter, we have investigated spectral efficiency of a massive MU-MIMO downlink system based on large-scale fading. In order to maximize the spectral efficiency for a given total power budget, we have proposed the four following methods of power allocation.

1) Power is allocated among the pilot and data symbols in such a way that the pilot power as well as the data power for each user is the same. The pilot power and the data power are optimally selected in order to maximize the spectral efficiency. 2) Power is allocated among the data symbols of the various users, where the pilot power for each user is same and is specified. In this method, the data power for each user is optimally selected to maximize the spectral efficiency. 3) Power is allocated among the pilot and data symbols of the various users, where the pilot power for each user is the same. In this method, the same pilot power along with the various data powers is optimized to maximize the spectral efficiency. 4) Power is allocated among each of the pilot and data symbols of the various users. In this method, the pilot power as well as the data power for each user is optimized to maximize SE.

Numerical results have shown that methods 2, 3,and 4 offer similar performance in

Table 4.1: Comparison among the proposed power allocation methods ($K = 5$)

Proposed methods		Method of Section 4.3.1		Method of Section 4.3.4		Method of Section 4.3.3		Method of Section 4.3.2	
Users	β_k	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}
User 1	0.001	1.2425	10.3347	0.7813	10.2270	10	9.6453	0.2925	10.2555
User 2	0.006	1.0361	10.5380	0.7813	10.4152	10	10.8772	0.2925	10.2555
User 3	0.188	0.7456	10.2383	0.7813	10.2763	10	10.3342	0.2925	10.2555
User 4	0.307	0.3927	10.0992	0.7813	10.2089	10	9.2512	0.2925	10.2555
User 5	1.049	0.4895	10.0027	0.7813	10.0739	10	9.9426	0.2925	10.2555
Total pilot power in watts		3.90×5=19.53		3.90×5=19.54		50×5=250		1.46×5=7.31	
Total data power in watts		51.21×190=9730.5		51.2×190=9728.2		50.05×190=9509.5		51.27×190= 9742.7	
Total power P_t in watts		9750		9750		9750		9750	
Run time (Secs)		1.0865		0.7858		0.7318		0.5469	

terms of SE with approximately the same run time. However, the SE of these methods is much better than that of method 1, even though the run time for the former methods are slightly more than that of latter method. This indicates that in a downlink transmission, where the total power budget is given, optimizing the power of data symbols is much more important than optimizing the power of pilot symbols in order to maximize the spectral efficiency.

Table 4.2: Comparison among the proposed power allocation methods ($K = 5$)

Proposed methods		Method of Section 4.3.1		Method of Section 4.3.4		Method of Section 4.3.3		Method of Section 4.3.2	
Users	β_k	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}
User 1	0.001	0.94dB	10.14dB	-1.07dB	10.09dB	10dB	9.84dB	-5.33dB	10.10dB
User 2	0.006	0.15dB	10.22dB	-1.07dB	10.17dB	10dB	10.36dB	-5.33dB	10.10dB
User 3	0.188	-1.27dB	10.10dB	-1.07dB	10.11dB	10dB	10.14dB	-5.33dB	10.10dB
User 4	0.307	-4.05dB	10.04dB	-1.07dB	10.07dB	10dB	9.66dB	-5.33dB	10.10dB
User 5	1.049	-3.10dB	10.00dB	-1.07dB	10.03dB	10dB	9.97dB	-5.33dB	10.10dB
Total power P_t		39.89dB		39.89dB		39.89dB		39.89dB	
Run time (Secs)		1.0865		0.7858		0.7318		0.5469	

Table 4.3: Comparison among the proposed power allocation methods ($K = 10$)

Proposed methods		Method of Section 4.3.1		Method of Section 4.3.4		Method of Section 4.3.3		Method of Section 4.3.2	
Users	β_k	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}
User 1	0.0001	4.3785	10.4238	1.0368	10.658	10	10.3872	0.7662	10.5130
User 2	0.0020	1.1421	10.5427	1.0368	10.268	10	10.1376	0.7662	10.5130
User 3	0.0145	4.1520	10.0812	1.0368	10.523	10	9.6635	0.7662	10.5130
User 4	0.0321	3.1922	10.5082	1.0368	10.414	10	9.7715	0.7662	10.5130
User 5	0.0455	0.8523	10.3123	1.0368	10.522	10	9.8624	0.7662	10.5130
User 6	0.0566	2.3715	10.3867	1.0368	10.474	10	10.2285	0.7662	10.5130
User 7	0.1228	0.8020	10.6219	1.0368	10.338	10	10.3365	0.7662	10.5130
User 8	0.3230	3.0381	10.5210	1.0368	10.582	10	9.6128	0.7662	10.5130
User 9	0.4847	0.6244	10.5110	1.0368	10.728	10	9.7427	0.7662	10.5130
User 10	0.8733	0.2314	10.4921	1.0368	10.473	10	9.2573	0.7662	10.5130
Total pilot power in watts		20.78×10=207.845		10.368×10=103.68		100×10=10^3		7.662×10=76.620	
Total data power in watts		104.40×180=18792		104.98×180=18896		100×180=18×10^3		105.13×180=18923	
Total power P_t in watts		19000		19000		19000		19000	
Run time (Secs)		2.3903		1.3359		1.0977		0.7548	

Table 4.4: Comparison among the proposed power allocation methods ($K = 10$)

Proposed methods		Method of Section 4.3.1		Method of Section 4.3.4		Method of Section 4.3.3		Method of Section 4.3.2	
Users	β_k	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}	p_{p_k}	p_{d_k}
User 1	0.0001	6.41dB	10.18dB	0.15dB	10.27dB	10dB	10.16dB	-1.15dB	10.21dB
User 2	0.0020	0.57dB	10.22dB	0.15dB	10.11dB	10dB	10.05dB	-1.15dB	10.21dB
User 3	0.0145	6.18dB	10.03dB	0.15dB	10.22dB	10dB	9.85dB	-1.15dB	10.21dB
User 4	0.0321	5.04dB	10.21dB	0.15dB	10.17dB	10dB	9.89dB	-1.15dB	10.21dB
User 5	0.0455	-0.69dB	10.13dB	0.15dB	10.22dB	10dB	9.93dB	-1.15dB	10.21dB
User 6	0.0566	3.75dB	10.16dB	0.15dB	10.20dB	10dB	10.09dB	-1.15dB	10.21dB
User 7	0.1228	-0.95dB	10.26dB	0.15dB	10.14dB	10dB	10.14dB	-1.15dB	10.21dB
User 8	0.3230	4.82dB	10.22dB	0.15dB	10.24dB	10dB	9.82dB	-1.15dB	10.21dB
User 9	0.4847	-2.04dB	10.21dB	0.15dB	10.30dB	10dB	9.88dB	-1.15dB	10.21dB
User 10	0.8733	-6.35dB	10.20dB	0.15dB	10.20dB	10dB	9.66dB	-1.15dB	10.21dB
Total power P_t		42.78dB		42.78dB		42.78dB		42.78dB	
Run time (Secs)		2.3903		1.3359		1.0977		0.7548	

Chapter 5

Uplink Transmission

In this chapter, we investigate the spectral efficiency of massive MU-MIMO systems with a very large number of antennas at a base station serving single antenna users in a uplink transmission. A practical physical channel model is proposed by dividing the angular domain into a finite number of distinct directions. A lower bound on the achievable rate of uplink data transmission is derived using a linear detector for each user. The obtained lower bound is further modified for the maximum-ratio combining and zero-forcing receivers. A power control scheme based on large-scale fading is also proposed to maximize the spectral efficiency under peak power constraint, where the MRC or ZF receiver is employed at BS.

5.1 Power Allocation with Finite-Dimensional Channel

It is known that in practical applications, complexity of the propagation environment and properties of the antenna arrays at BS have significant effect on the performance of massive MIMO systems. In addition, the channel vectors for different users are not asymptotically orthogonal and have been modeled by an L-dimensional vector, where L is the number of angular bins. It has been shown in [48] that by increasing the number of antennas at BS

under a finite-dimensional model with L angular bins, the performance of massive MIMO systems is almost the same as its performance under uncorrelated channel model with L antennas. Another important practical consideration is that the dimension of the physical channel should be finite [45, 46].

In view of this, in this section, we propose a finite-dimensional channel model and derive the achievable rate of a multi-user massive MIMO system on the uplink channel. A lower bound on the achievable rate is derived and employed in defining the SE. It should be noted that due to employing a finite-dimensional uplink channel and consequently increasing the multiuser interferences, conventional approaches for equal power control among the users may not be used and thus, an optimal transmit power of each user needs to be determined and used to maximize the SE based on the large-scale fading coefficient [10].

5.1.1 Channel Model

A BS in a multi-user massive MIMO network is assumed to be equipped with M antennas and serves K single-antenna users as shown in Fig 5.1. Since the propagation environment and properties of the antenna arrays have significant impact on the performance of massive MIMO systems [48], we assume that the dimension of the physical channel L is finite and the angular bin is divided into a number of directions such that $L \leq M$. Moreover, since the distance between BS and each user is much larger than the distance among the antennas, the angles of arrival at the M antennas of BS, $\phi_l \in [-\pi/2, \pi/2], l = 1, ..., L$, are considered to be equal for each user. In this case, each direction is associated with an $M \times 1$ array as

$$\mathbf{a}(\phi_l) = \frac{1}{\sqrt{L}}[e^{-jf_1(\phi_l)}, e^{-jf_2(\phi_l)}, ..., e^{-jf_M(\phi_l)}]^T, \tag{5.1}$$

where $(.)^T$ denotes the transpose operator and $f_m(\phi)$ is a function of ϕ. The channel vector from the kth user to BS is combination of L steering vectors, i.e., $\sum_{l=1}^{L} g_{kl}\mathbf{a}(\phi_l)$, where g_{kl} is the propagation coefficient from the kth user to BS associated with the lth direction

of arrival [48]. To normalize the channel, a $\frac{1}{\sqrt{L}}$ factor is used in (5.1). We now define our channel model $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_K]$ as a $L \times K$ matrix having $\mathbf{g}_k \triangleq [g_{k1}, g_{k2}, ..., g_{kL}]^T$ columns. Each column represents the propagation coefficients from the kth user to BS. The channel matrix between the K users and BS is then expressed as

$$\mathbf{T} = \mathbf{A}\ \mathbf{G}, \tag{5.2}$$

where $\mathbf{A} \triangleq [\mathbf{a}(\phi_1), ..., \mathbf{a}(\phi_L)]$ is a full rank $M \times L$ known matrix and $\mathbf{G}$ represents the propagation channel matrix, whose g_{kl}th element is given by

$$g_{kl} = h_{kl}\ \sqrt{\beta_k} \qquad l = 1, 2, \ldots, L\ , \tag{5.3}$$

where h_{kl} is the fast fading coefficient associated with the lth direction from the kth user to BS, having zero-mean and unit variance Gaussian distribution, and $\sqrt{\beta_k}$ is the large-scale fading coefficient. This coefficient models the geometrical attenuation (path loss) and shadow fading that is constant over l. In view of the fact that the distances between the users and BS are much larger than the those between the antennas, $\sqrt{\beta_k}$ will be constant over time intervals. The propagation channel matrix is then written as

$$\mathbf{G} = \mathbf{H}\mathbf{D}^{1/2}, \tag{5.4}$$

where $\mathbf{H} = [h_{kl}]$ denotes the $L \times K$ matrix of fast fading coefficients between BS and the K users and $\mathbf{D} = [\beta_k]$ denotes the $K \times K$ diagonal matrix of large-scale coefficients.

In the uplink transmission, the users transmit data in the same time-frequency resource. Thus, the $M \times 1$ received vector at the BS can be written as

$$\mathbf{y} = \ \mathbf{T}\ \mathbf{x}\ +\ \mathbf{n}, \tag{5.5}$$

where $\mathbf{x}$ denotes the data transmitted by K users and $\mathbf{n} \sim \mathcal{CN}(0, 1)$ is an additive white

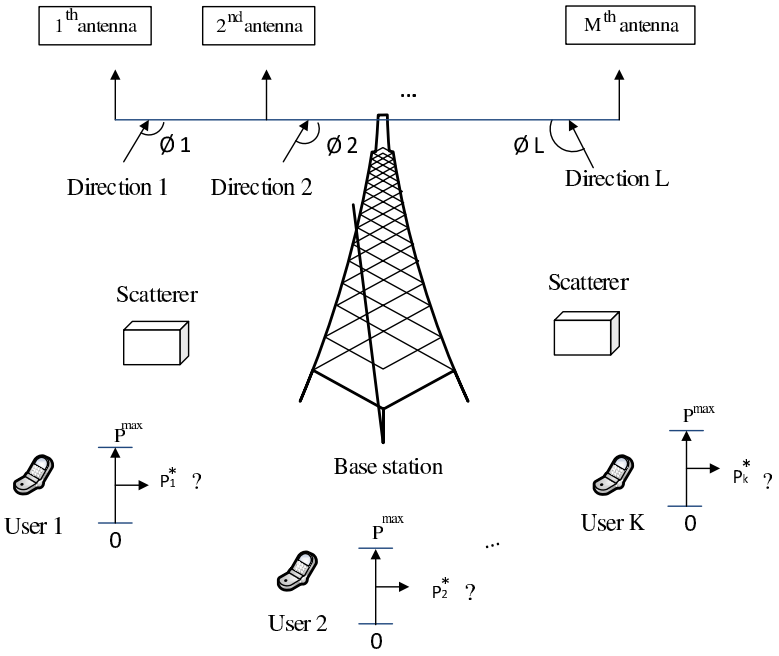

Figure 5.1: System model along with the proposed method diagram.

Gaussian noise. The kth element of $\mathbf{x}$ is $\sqrt{p_k}\ x_k$, where p_k and x_k are the average transmitted power of the kth user and the kth transmitted symbol, respectively. It is worth mentioning that wideband channels handled by orthogonal frequency division multiplexing (OFDM) over restricted intervals can be applied to the signal model (5.5).

5.1.2 Problem Formulation

In this subsection, we present the proposed power control method to maximize the SE in the cell. First, the lower bound of achievable rates on the uplink channel is derived, where uniform linear array is considered at BS. For such a uniform linear array, $f_m(\phi_l)$ is given by $f_m(\phi_l) = \frac{2\pi(m-1)d}{\lambda} \sin\phi_l\ ,\ m = 1, 2, ..., M$ [48]. In this case, the response vector in (5.1) can be rewritten as [48]

$$\mathbf{a}(\phi_l) = \frac{1}{\sqrt{L}}[1, e^{-j2\pi\frac{d}{\lambda}\sin\phi_l}, ..., e^{-j2\pi\frac{(M-1)d}{\lambda}\sin\phi_l}]^T, \tag{5.6}$$

where d and λ are the antenna spacing and the carrier wavelength, respectively. In this respect, we have

$$\frac{1}{M}\ \mathbf{a}^{\dagger}(\phi_l)\mathbf{a}(\phi_q) = \frac{1}{ML}\sum_{m=0}^{M-1} e^{j2\pi\frac{d}{\lambda}(\sin\phi_l - \sin\phi_q)m}\ , \tag{5.7}$$

where $(.)^{\dagger}$ denotes the Hermitian operator. When the number of antennas at BS (M) is very large and $l \neq q$, we have

$$\frac{1}{M}\ \mathbf{a}^{\dagger}(\phi_l)\mathbf{a}(\phi_q) = \frac{1}{ML}\ \frac{1 - e^{j2\pi\frac{d}{\lambda}(\sin\phi_l - \sin\phi_q)M}}{1 - e^{j2\pi\frac{d}{\lambda}(\sin\phi_l - \sin\phi_q)}} \simeq 0. \tag{5.8}$$

If $l = q$, $\frac{1}{M}\mathbf{a}^{\dagger}(\phi_l)\mathbf{a}(\phi_q)\ =\ \frac{1}{L}$. As a result, for large M, we have

$$\frac{1}{M}\mathbf{A}^{\dagger}\mathbf{A}\ \simeq\ \frac{1}{L}\mathbf{I}_L\ . \tag{5.9}$$

5.1.2.1 Achievable Uplink Rate

We now derive the lower bounds on the achievable uplink rate for each user when BS is aware of their channel coefficients (g_{kl}). To this end, BS needs a linear detector depending on the channel matrix to separate the received signal into streams. Letting $\mathbf{V}$ be an $M \times K$ linear detector matrix, beamforming of the received signal is defined as [11]

$$\mathbf{r} = \mathbf{V}^{\dagger}\mathbf{y}. \tag{5.10}$$

Substituting (5.5) in (5.10), the received vector can be written as

$$\mathbf{r} = \mathbf{V}^{\dagger}\ \mathbf{T}\ \mathbf{x}\ +\ \mathbf{V}^{\dagger}\mathbf{n}. \tag{5.11}$$

To separate the transmitted signal of kth user and obtain its achievable rate, (5.11) can be expressed as

$$\mathbf{r}_k = \sqrt{p_k}\ \mathbf{v}_k^{\dagger}\mathbf{t}_k x_k + \sum_{i=1, i\neq k}^{K} \sqrt{p_i}\ \mathbf{v}_k^{\dagger}\mathbf{t}_i x_i + \mathbf{v}_k^{\dagger}\mathbf{n}, \tag{5.12}$$

where $\mathbf{v}_k$ and $\mathbf{t}_k$ represent the kth columns of the matrices $\mathbf{V}$ and $\mathbf{T}$, respectively. The last two terms on the right hand side of (5.12) are the interference and noise terms, respectively. Spectral Efficiency for a single-cell multiuser MIMO system is given by [11]

$$S = \sum_{k=1}^{K} \tilde{R}_k, \tag{5.13}$$

where $\tilde{R}_k$ is the lower bound of the ergodic achievable uplink rate of the kth user, R_k, which is given by [11]

$$R_k = \mathbb{E}\Bigg\{ \log_2 \left(1 + \frac{p_k |\mathbf{v}_k^{\dagger}\mathbf{t}_k|^2}{\sum_{i=1, i\neq k}^{K} p_i\ |\mathbf{v}_k^{\dagger}\mathbf{t}_i|^2 + \parallel \mathbf{v}_k \parallel^2} \right) \Bigg\}, \tag{5.14}$$

At this stage, two linear detectors, namely, ZF and MRC, are employed to derive $\tilde{R}_k$.

Lemma 5.1 : Using ZF detector, the interference term cancels out and $\tilde{R}_k$ is obtained as

$$\tilde{R}_k^{zf} = \log_2 \left(1 + p_k \times \frac{M}{L}(L-K)\beta_k \right) . \tag{5.15}$$

Proof : See Appendix D. ■

Lemma 5.2 : Using MRC detector, $\tilde{R}_k$ is obtained as

$$\tilde{R}_k^{mrc} = \log_2 \left(1 + \frac{\beta_k (L-1) p_k}{\sum_{i=1, i\neq k}^{K} \beta_i\ p_i + \frac{L}{M}} \right) . \tag{5.16}$$

Proof : See Appendix E. ■

Using MRC detector, the lower bound of achievable rate of each user depends on the transmit power of the other users due to the interference term. Hence, when users consume their peak power for transmission, the interference term may be increased and the spectral efficiency given by (5.13) may be diminished. Although the interference has an impact on the achievable rate of each user in MRC receiver, this receiver is still being employed in multi-user massive MIMO systems. This is due to the fact that MRC detector is used in a distributed fashion, independently at each antenna unit and offers less per-symbol complexity [4].

5.1.2.2 Optimal Power Control

We now present a power control scheme for the users to maximize the SE given in (5.13). To this end, we find the optimal transmitted power for each user. Let p^{max} be the peak power that each user can transmit in one symbol. Then, the optimal power control strategy for each user can be formulated as

$$\begin{aligned} &\max_{p_1,p_2,...,p_K} \sum_{k=1}^{K} \tilde{R}_k^C \\ \text{s.t.} \quad &0 < p_k \leq p^{max} \; ; \qquad k = 1, 2, ..., K. \end{aligned} \tag{5.17}$$

where $C \in \{mrc, zf\}$ corresponds to MRC or ZF detectors. The given constraints guarantee that the power control function is positive and do not exceed p^{max}.

Theorem 5.1. Using ZF receiver, (5.17) can be maximized when all users transmit their maximum powers, i.e., $p_k = p_k^{max}$.

Proof : Since there is no interference in ZF receiver obtained in (5.15), the proof is straightforward. ■

Theorem 5.2. Using MRC receiver, (5.17) is NP-hard [71], and can be reformulated

as

$$\max_{p_1,p_2,...,p_K} \prod_{k=1}^{K} \left(\frac{1}{M} + \gamma_k(p_1, p_2, ..., p_K) \right) \tag{5.18}$$
$$\text{s.t.} \qquad 0 < p_k \leq p^{max} \ ; \qquad k = 1, 2, ..., K.$$

where $\gamma_k(p_1, p_2, ..., p_K) = \frac{\beta_k(L-1)p_k}{M \sum_{i=1, i\neq k}^{K} \beta_i p_i + L}$.

This optimization problem is still an NP-hard problem. We solve (5.18) using an alternative optimization approach as

$$\max_{p_1,p_2,...,p_K} \prod_{k=1}^{K} \gamma_k(p_1, p_2, ..., p_K) \tag{5.19}$$
$$\text{s.t.} \qquad 0 < p_k \leq p^{max} \ ; \qquad k = 1, 2, ..., K.$$

For large M, the optimal solutions of (5.19) and (5.18) are the same.

Proof : Let y^* and y_{opt} be the optimal solution of (5.19) and (5.18), respectively. As a result, considering (5.19) and (5.18), $y^* \leq y_{opt}$. Defining $u_k \triangleq \max_{p_1,p_2,...,p_K} \{\gamma_k(p_1, p_2, ..., p_K)\}$ with $0 \leq p_k \leq p^{max}$, we have

$$u_k = \frac{p^{max} \beta_k (L-1)}{L}. \tag{5.20}$$

Thus, for any given indices of γ_{j_i}, $i = 1, ..., K$, we have

$$\gamma_{j_1}(p_1, p_2, ..., p_K) ... \gamma_{j_i}(p_1, p_2, ..., p_K) \leq u_1 u_2 ... u_K \ . \tag{5.21}$$

Expanding (5.18) and using (5.21), we have

$$y^* \leq y_{opt} \leq y^* + Q(M), \tag{5.22}$$

where

$$Q(M) \triangleq \frac{1}{M^K} + \sum_{i=1}^{K-1} \frac{1}{M^{K-i}} \binom{K}{i} \prod_{j=1}^{i} u_j. \tag{5.23}$$

When M is very large, $Q(M)$ in (5.23) approaches zero, and hence, $y^* \simeq y_{opt}$. ■

The optimization problem given by (5.19) is a standard geometric programming (GP) problem, since the objective and subject functions are posynomials [73]. This GP problem can be easily solved by using standard numerical optimization packages. In this thesis, we employ ConVeX (CVX) package to solve the optimization problem given by (5.19) [75]. It is worth mentioning that the optimization problem given by (5.19) can be converted to a convex optimization problem by change of variable [73].

It should also be mentioned that by increasing M when using MRC receiver, the interference may not be negligible compared to the noise, i.e., $\sum_{i=1}^{K} \beta_i p_i \gg \frac{L}{M}$. As a result, the maximum transmitted power is no longer a valid solution for (5.17). In view of this, the proposed power control method at BS is indispensable to inform the amount of transmit power to the users.

5.1.3 Numerical Results

Experiments are conducted to evaluate the performance of the proposed power control method and verify the expressions obtained for lower achievable bounds. We first describe a scenario corresponding to the impact of peak power p^{max} and the dimension of the physical channel L on the performance of the system when using MRC receiver with fixed β_k. Moreover, we evaluate the lower bounds obtained for MRC and ZF receivers in this scenario. We then compare the SE of a massive MIMO system with proposed and peak power control scheme for MRC and ZF receivers in a scenario where β_k changes. In both the scenarios, we consider the number of users $K = 10$ in a hexagonal single-cell massive MIMO system with a radius of $R = 500$ meters. We also consider that the users are not closer to BS than $d_h = 100$ meters and the path loss exponent is $v = 3.8$. In this case,

the large-scale fading (β_k) follows the log-normal distribution with standard deviation $\delta_{shadow} = 8\ dB$, i.e., $\beta_k = z_k/(d_k/d_h)^v$, where z_k is the log-normal random variable and d_k is the distance between the kth user and BS. Since the noise variance is 1, we define signal-to-noise ratios (SNR) to be equal to peak power, i.e., SNR $= p^{max}$.

5.1.3.1 Scenario I

In this scenario, we investigate the effect of SNR on the performance of the system. Following $\beta_k = z_k/(d_k/d_h)^v$, the large-scale fading coefficients are generated as β_1=0.0006, β_2=0.0033, β_3=0.0063, β_4=0.0074, β_5=0.0105, β_6=0.0204, β_7=0.0270, β_8=0.0413, β_9=0.0584 and β_{10}=0.0908. Fig. 5.2 shows numerically evaluated values (NEV) of SE using (5.14) versus SNR for both ZF and MRC receivers, where the peak power is allocated to each of the users. Fig. 5.2 also shows SE obtained using (5.15) and (5.16) for different values of SNR, employing ZF and MRC receivers, respectively. It is seen from this figure that the lower bounds obtained in Lemmas 5.1 and 5.2 in terms of SE are pretty accurate for ZF and MRC receivers, respectively. Fig. 5.3 shows SE obtained using the proposed method in terms of SNR employing MRC receiver. It is seen from this figure that the performance of the method using the proposed power control scheme is superior to that of the conventional peak power criterion, in terms of providing a higher SE. In addition, it is also seen that by increasing L, the performance of the proposed power control scheme can be significantly increased. Fig. 5.4 shows SE obtained using the proposed method in terms of SNR employing ZF receiver, where the proposed power control is equal to the peak power control. It can be seen from this figure that by increasing SNR and L, SE also increases.

In order to illustrate the superiority of the proposed power control method, we present the achievable rate of each user obtained by using this method and compare it with that provided by the maximum power control strategy, when M=300 and L=200. Employing the MRC receiver and assuming SNR=10dB, the optimal power and the optimal rate of the users are obtained as p_1^*=10dB, p_2^*=10dB, p_3^*=10dB, p_4^*=10dB, p_5^*=10dB, p_6^*=10dB,

p_7^*=8.8084dB, p_8^*=6.9626dB, p_9^*=5.4580dB, p_{10}^*=3.5412dB, and R_1^*=0.6845, R_2^*=2.1329, R_3^*=2.9194, R_4^*=3.1304, R_5^*=3.6075, R_6^*=4.5822, R_7^*=4.5914, R_8^*=4.5914, R_9^*=4.5914, R_{10}^*=4.5914, respectively. Thus, SE obtained is S^*=35.4225. However, using maximum power control strategy, we have $p_1^{\max}$=$p_2^{\max}$=$p_3^{\max}$=$p_4^{\max}$=$p_5^{\max}$=$p_6^{\max}$=$p_7^{\max}$=$p_8^{\max}$=$p_9^{\max}$=$p_{10}^{\max}$=10dB, and $R_1^{\max}$=0.4431, $R_2^{\max}$=1.5820, $R_3^{\max}$=2.2754, $R_4^{\max}$=2.4666, $R_5^{\max}$=2.9042, $R_6^{\max}$=3.8074, $R_7^{\max}$=4.2155, $R_8^{\max}$=4.8683, $R_9^{\max}$=5.4387, $R_{10}^{\max}$=6.2424, resulting in SE given by $S^{\max}$=34.2436. This clearly shows that the proposed power control method provides a higher SE when the MRC receiver is employed.

Employing the ZF receiver, the optimal power that is the same as peak power and the optimal rate of the users are obtained as p_1^*=$p_1^{\max}$ =10dB, p_2^*=$p_2^{\max}$ =10dB, p_3^*=$p_3^{\max}$ =10dB, p_4^*=$p_4^{\max}$ =10dB, p_5^*=$p_5^{\max}$ =10dB, p_6^*=$p_6^{\max}$ =10dB, p_7^*=$p_7^{\max}$ =10dB, p_8^*=$p_8^{\max}$ =10dB, p_9^*=$p_9^{\max}$ =10dB, p_{10}^*=$p_{10}^{\max}$ =10dB, and R_1^*=$R_1^{\max}$= 1.4383, R_2^*=$R_2^{\max}$ =3.3792, R_3^*=$R_3^{\max}$ = 4.2445, R_4^*=$R_4^{\max}$ =4.4653, R_5^*=$R_5^{\max}$ =4.9507, R_6^*=$R_6^{\max}$ =5.8861, R_7^*=$R_7^{\max}$ =6.2845, R_8^*=$R_8^{\max}$ =6.8912, R_9^*=$R_9^{\max}$ =8.0211, R_{10}^*=$R_{10}^{\max}$ =14.2640, respectively. Thus, SE obtained is S^*=59.8250.

Table 5.1 shows a comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 10dB$, and $K = 10$. It can be seen from this table that the proposed power control method provides a higher SE when the MRC receiver is employed. However, the performance of the proposed power allocation method is the same as maximum power control when the ZF receiver is employed. In addition, it can be seen from this table that the performance of the ZF receiver outperforms that of the MRC receiver in terms of SE.

Next, the performance of the proposed power allocation method is studied with respect to the variation of SNR. Tables 5.1, 5.3, 5.4, 5.5 and 5.6 give the corresponding results when $M = 300$, $L = 200$, $K = 10$ and SNR is $-5dB$, $5dB$, $10dB$, $20dB$ and $30dB$, respectively. It can be seen from these tables that with increasing SNR, SE significantly improves using the proposed power control method when the MRC receiver is employed. In

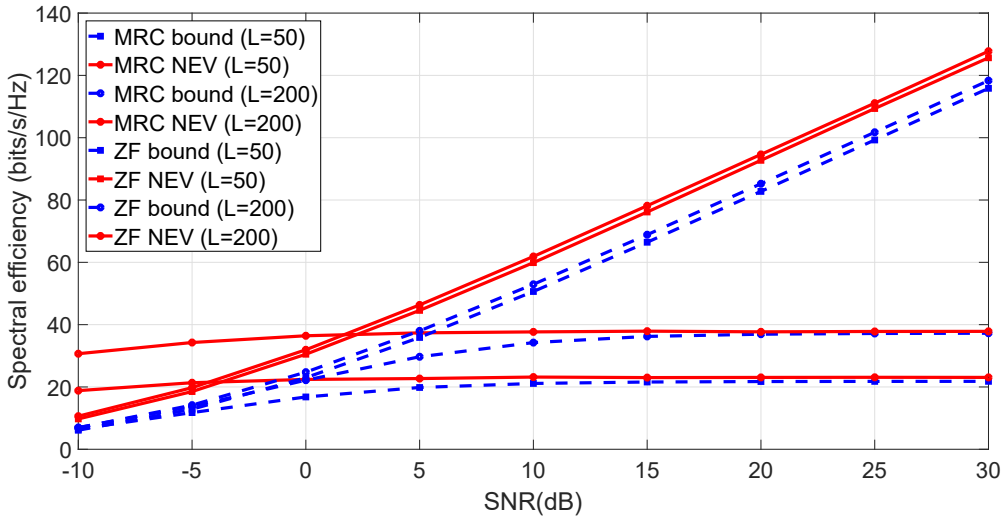

Figure 5.2: Numerically evaluated values of SE along with those obtained using (5.15) and (5.16) for different values of SNR when ZF and MRC receivers are employed and peak power is allocated to each user with $M = 300$.

addition, it can be seen that the performance of the ZF receiver significantly outperforms that of the MRC receiver in terms of SE.

Then, the performance of the proposed power allocation method is studied with respect to the variation of K. Tables 5.7, 5.8, 5.9 and 5.10 give the corresponding results when $M = 300$, $L = 200$, $SNR = 10dB$ and K is 5, 10, 15 and 20, respectively. It can be seen from these tables that with increasing K, SE significantly improves. It can be also seen that the advantage of using the proposed power control method becomes more evident as K increases when the MRC or ZF receiver is employed.

5.1.3.2 Scenario II

We now compare the SE of a massive MIMO system with proposed and peak power control scheme for both receivers. To this end, we set up $\beta_k = z_k/(d_k/d_h)^v$ and set SNR $= 0dB$. We run 1000 Monte-Carlo simulations for which the users are randomly located in the cell so that the large-scale fading β_k changes. Fig. 5.5 shows the cumulative distribution function values of the SE, with using the proposed and peak power control schemes. It is seen from this figure that using the ZF detector results in the highest value for SE, as

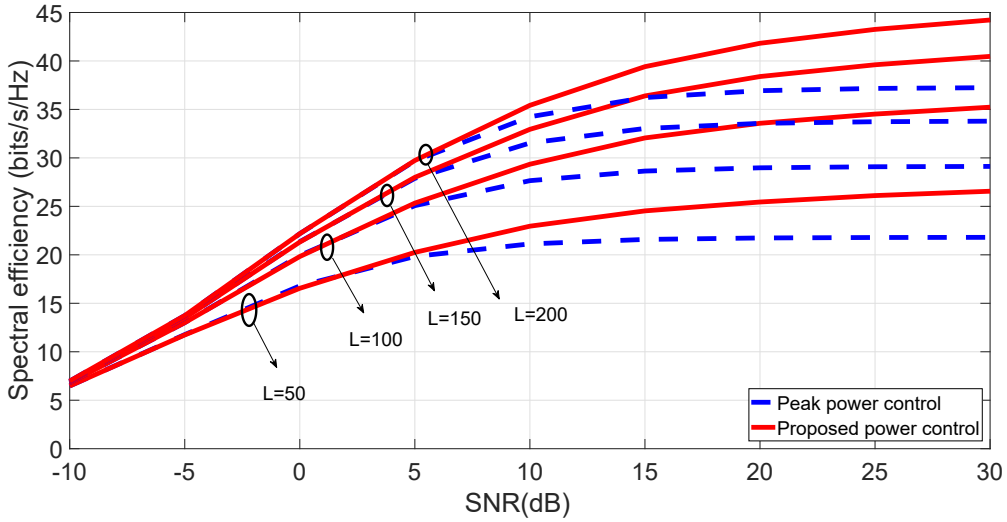

Figure 5.3: SE obtained using the proposed method as well as the peak power criterion for each user, when MRC receiver is employed and $M = 300$.

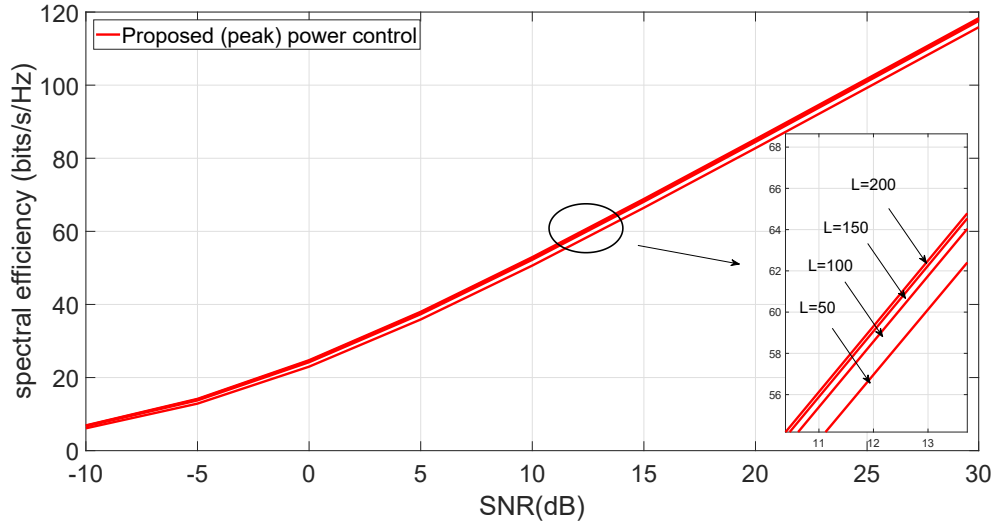

Figure 5.4: SE obtained using the proposed method (peak power criterion) for each user, when ZF receiver is employed and $M = 300$.

Table 5.1: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 10dB$, and $K = 10$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0006	10.0000dB	3.9719	10dB	0.0238	10dB	1.4383
User 2	0.0033	10.0000dB	4.0838	10dB	0.1264	10dB	3.3792
User 3	0.0063	10.0000dB	4.1970	10dB	0.2325	10dB	4.2445
User 4	0.0074	10.0000dB	4.3115	10dB	0.2695	10dB	4.4653
User 5	0.0105	10.0000dB	4.4272	10dB	0.3691	10dB	4.9507
User 6	0.0204	10.0000dB	4.5444	10dB	0.6482	10dB	5.8861
User 7	0.0270	8.8084dB	4.5477	10dB	0.8086	10dB	6.2845
User 8	0.0413	6.9626dB	4.5477	10dB	1.1055	10dB	6.8912
User 9	0.0908	3.5412dB	4.5477	10dB	1.8278	10dB	8.0211
User 10	6.9028	-15.2682dB	4.5477	10dB	12.2904	10dB	14.2640
Total Power		18.7425dB		20dB		20dB	
SE			35.4225		17.7018		59.8250

expected. It is also seen that by increasing L, SE significantly improves with proposed and peak power control schemes for both ZF and MRC receivers. It is also seen that the use of MRC detector with the proposed power control scheme provides a substantially higher SE as compared to that with the peak power criterion. Fig. 5.6 shows the probability density function (pdf) of the optimally allocated power to each user when MRC detector is employed. It can be seen from this figure that the various users have approximately the same pdf.

5.1.4 Summary

In this chapter, we have investigated the spectral efficiency of a single-cell massive MIMO system where a channel model is considered by dividing the angular domain into a finite number of distinct directions. First, we have derived the achievable rate of each user with ZF and MRC receivers. Then, we have proposed a power control strategy among the

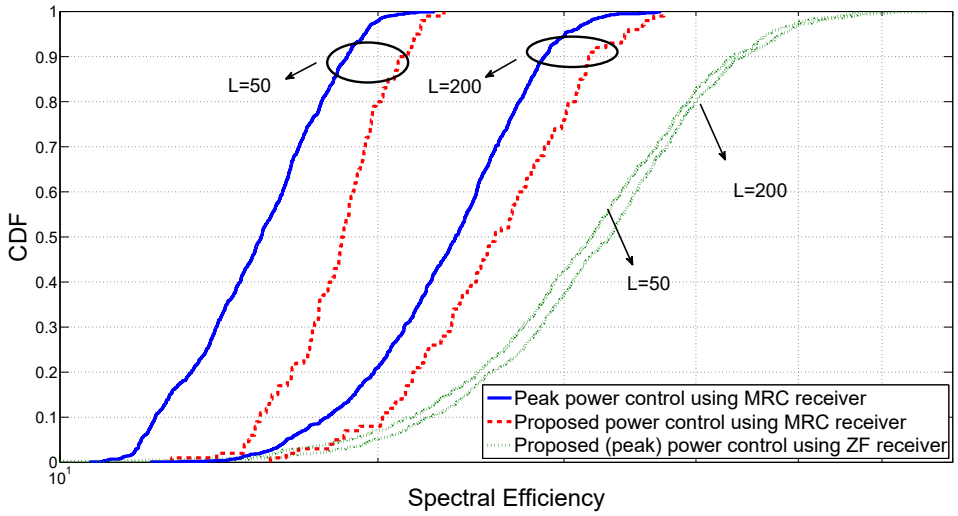

Figure 5.5: Spectral efficiency obtained with the proposed and peak power control for MRC and ZF receivers, when $M = 500$ and SNR $= 0dB$.

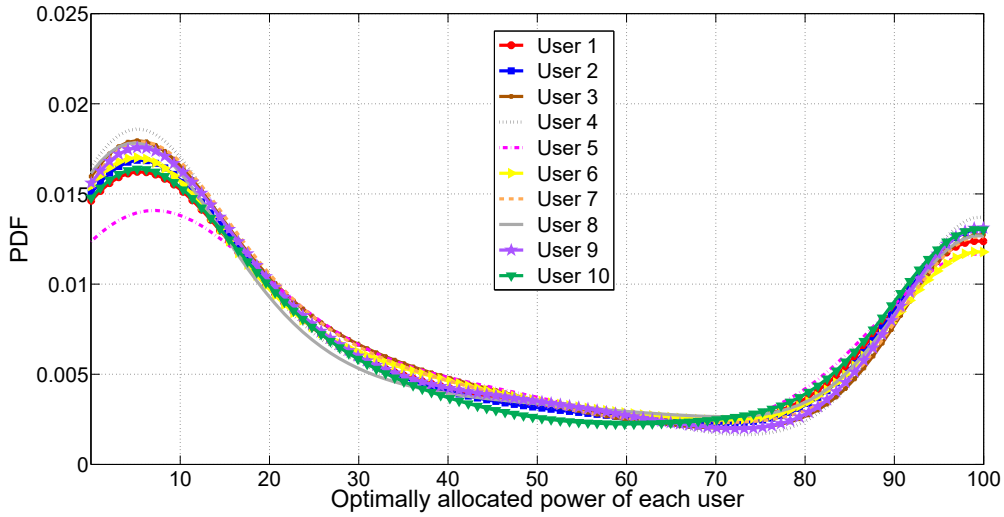

Figure 5.6: pdf of the optimally allocated power to each user, when SNR $= 20dB$, $M = 300$, and $L = 200$.

Table 5.2: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = -5dB$, and $K = 10$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0006	-5.0000dB	0.0647	-5dB	0.0186	-5dB	0.0760
User 2	0.0033	-5.0000dB	0.3250	-5dB	0.0993	-5dB	0.3756
User 3	0.0063	-5.0000dB	0.5684	-5dB	0.1841	-5dB	0.6487
User 4	0.0074	-5.0000dB	0.6485	-5dB	0.2140	-5dB	0.7372
User 5	0.0105	-5.0000dB	0.8529	-5dB	0.2950	-5dB	0.9607
User 6	0.0204	-5.0000dB	1.3631	-5dB	0.5274	-5dB	1.5052
User 7	0.0270	-5.0000dB	1.6259	-5dB	0.6641	-5dB	1.7796
User 8	0.0413	-5.0000dB	2.0735	-5dB	0.9226	-5dB	2.2394
User 9	0.0908	-5.0000dB	3.0347	-5dB	1.5751	-5dB	3.1990
User 10	6.9028	-18.8182dB	4.6794	-5dB	9.2147	-5dB	9.2834
Total Power		4.5624dB		5dB		5dB	
SE			15.2361		13.7149		20.8049

users to maximize the spectral efficiency. Experiments have been conducted to assess the performance of the proposed method in terms of the spectral efficiency. The numerical results show that the proposed power control method provides a spectral efficiency which is the same as that of the conventional maximum power criterion using the ZF receiver. Further, the proposed method provides a spectral efficiency which is higher than that of the conventional maximum power criterion using the MRC receiver. In addition, it has been shown that the ZF receiver outperforms the MRC receiver in terms of spectral efficiency. In spite of this fact, the MRC receiver is still being employed in multi-user massive MIMO systems, due to the fact that the MRC receiver is used in a distributed fashion, independently at each antenna unit and offers less per-symbol complexity.

Table 5.3: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 5dB$, and $K = 10$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0006	5.0000dB	0.3645	5dB	0.0233	5dB	0.6236
User 2	0.0033	5.0000dB	1.3738	5dB	0.1240	5dB	1.9906
User 3	0.0063	5.0000dB	2.0217	5dB	0.2282	5dB	2.7394
User 4	0.0074	5.0000dB	2.2031	5dB	0.2647	5dB	2.9391
User 5	0.0105	5.0000dB	2.6214	5dB	0.3627	5dB	3.3872
User 6	0.0204	5.0000dB	3.4936	5dB	0.6379	5dB	4.2769
User 7	0.0270	5.0000dB	3.8889	5dB	0.7963	5dB	4.6630
User 8	0.0413	5.0000dB	4.5196	5dB	1.0901	5dB	5.2563
User 9	0.0908	1.8695dB	4.6228	5dB	1.8070	5dB	6.3721
User 10	6.9028	-16.9399dB	4.6228	5dB	11.6812	5dB	12.6032
Total Power		14.2905dB		15dB		15dB	
SE			29.7322		17.0154		44.8515

Table 5.4: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 10dB$, and $K = 10$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0006	10.0000dB	3.9719	10dB	0.0238	10dB	1.4383
User 2	0.0033	10.0000dB	4.0838	10dB	0.1264	10dB	3.3792
User 3	0.0063	10.0000dB	4.1970	10dB	0.2325	10dB	4.2445
User 4	0.0074	10.0000dB	4.3115	10dB	0.2695	10dB	4.4653
User 5	0.0105	10.0000dB	4.4272	10dB	0.3691	10dB	4.9507
User 6	0.0204	10.0000dB	4.5444	10dB	0.6482	10dB	5.8861
User 7	0.0270	8.8084dB	4.5477	10dB	0.8086	10dB	6.2845
User 8	0.0413	6.9626dB	4.5477	10dB	1.1055	10dB	6.8912
User 9	0.0908	3.5412dB	4.5477	10dB	1.8278	10dB	8.0211
User 10	6.9028	-15.2682dB	4.5477	10dB	12.2904	10dB	14.2640
Total Power		18.7425dB		20dB		20dB	
SE			35.4225		17.7018		59.8250

Table 5.5: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 20dB$, and $K = 10$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0006	20.0000dB	1.6589	20dB	0.0240	20dB	4.1779
User 2	0.0033	20.0000dB	3.7524	20dB	0.1274	20dB	6.5706
User 3	0.0063	19.5431dB	4.5516	20dB	0.2343	20dB	7.4963
User 4	0.0074	18.8442dB	4.5516	20dB	0.2716	20dB	7.7272
User 5	0.0105	17.3247dB	4.5516	20dB	0.3719	20dB	8.2300
User 6	0.0204	14.4403dB	4.5516	20dB	0.6526	20dB	9.1859
User 7	0.0270	13.2229dB	4.5516	20dB	0.8138	20dB	9.5897
User 8	0.0413	11.3771B	4.5516	20dB	1.1121	20dB	10.2022
User 9	0.0908	7.9557dB	4.5516	20dB	1.8366	20dB	11.3381
User 10	6.9028	-10.8537dB	4.5516	20dB	12.6466	20dB	17.5859
Total Power		26.8977dB		30dB		30dB	
SE			41.8245		18.0907		92.1037

Table 5.6: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 30dB$, and $K = 10$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0006	30.0000dB	3.3712	30dB	0.0240	30dB	7.4263
User 2	0.0033	26.1027dB	4.5395	30dB	0.1275	30dB	9.8788
User 3	0.0063	23.2944dB	4.5395	30dB	0.2345	30dB	10.8110
User 4	0.0074	22.5955dB	4.5395	30dB	0.2718	30dB	11.0430
User 5	0.0105	21.0759dB	4.5395	30dB	0.3721	30dB	11.5476
User 6	0.0204	18.1915dB	4.5395	30dB	0.6531	30dB	12.5056
User 7	0.0270	16.9742dB	4.5395	30dB	0.8143	30dB	12.9099
User 8	0.0413	15.1283dB	4.5395	30dB	1.1127	30dB	13.5230
User 9	0.0908	11.7070dB	4.5395	30dB	1.8375	30dB	14.6595
User 10	6.9028	-7.1024dB	4.5395	30dB	12.6875	30dB	20.9078
Total Power		33.2106dB		40dB		40dB	
SE			44.2267		18.1350		125.2124

Table 5.7: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 10dB$ and $K = 5$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0004	10.0000dB	0.6847	10dB	0.3901	10dB	1.1177
User 2	0.0009	10.0000dB	1.2460	10dB	0.7655	10dB	1.8610
User 3	0.0028	10.0000dB	2.4143	10dB	1.6754	10dB	3.2001
User 4	0.0380	9.0234dB	5.9243	10dB	5.1524	10dB	6.8093
User 5	0.1480	3.1187dB	5.9243	10dB	8.0863	10dB	8.7612
Total Power		16.0246dB		16.9897dB		16.9897dB	
SE			16.1935		16.0697		21.7492

Table 5.8: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 10dB$, and $K = 10$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0006	10.0000dB	3.9719	10dB	0.0238	10dB	1.4383
User 2	0.0033	10.0000dB	4.0838	10dB	0.1264	10dB	3.3792
User 3	0.0063	10.0000dB	4.1970	10dB	0.2325	10dB	4.2445
User 4	0.0074	10.0000dB	4.3115	10dB	0.2695	10dB	4.4653
User 5	0.0105	10.0000dB	4.4272	10dB	0.3691	10dB	4.9507
User 6	0.0204	10.0000dB	4.5444	10dB	0.6482	10dB	5.8861
User 7	0.0270	8.8084dB	4.5477	10dB	0.8086	10dB	6.2845
User 8	0.0413	6.9626dB	4.5477	10dB	1.1055	10dB	6.8912
User 9	0.0908	3.5412dB	4.5477	10dB	1.8278	10dB	8.0211
User 10	6.9028	-15.2682dB	4.5477	10dB	12.2904	10dB	14.2640
Total Power		18.7425dB		20dB		20dB	
SE			35.4225		17.7018		59.8250

Table 5.9: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 10dB$, and $K = 15$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0001	10.0000dB	0.1547	10dB	0.0042	10dB	0.3533
User 2	0.0002	10.0000dB	0.2945	10dB	0.0084	10dB	0.6369
User 3	0.0009	10.0000dB	1.0166	10dB	0.0375	10dB	1.8063
User 4	0.0015	10.0000dB	1.4389	10dB	0.0619	10dB	2.3681
User 5	0.0018	10.0000dB	1.6121	10dB	0.0740	10dB	2.5838
User 6	0.0040	10.0000dB	2.4930	10dB	0.1596	10dB	3.5969
User 7	0.0043	10.0000dB	2.5814	10dB	0.1709	10dB	3.6929
User 8	0.0189	8.0488dB	3.9675	10dB	0.6360	10dB	5.7401
User 9	0.0385	4.9588dB	3.9675	10dB	1.0920	10dB	6.7527
User 10	0.0482	3.9829dB	3.9675	10dB	1.2744	10dB	7.0742
User 11	0.1814	-1.7730dB	3.9675	10dB	2.6887	10dB	8.9784
User 12	0.7618	-8.0050dB	3.9675	10dB	4.7045	10dB	11.0464
User 13	1.3224	-10.4002dB	3.9675	10dB	5.6139	10dB	11.8418
User 14	1.3394	-10.4557dB	3.9675	10dB	5.6363	10dB	11.8602
User 15	3.0180	-13.9838dB	3.9675	10dB	7.3171	10dB	13.0320
Total Power		19.1939dB		21.7609dB		21.7609dB	
SE			41.3313		29.4794		91.3641

Table 5.10: Comparison between the proposed power control method and maximum power control for MRC and ZF receivers when $M = 300$, $L = 200$, $SNR = 10dB$, and $K = 20$

	Fading	MRC				ZF	
	β_k	p_k^*	R_k^*	p_k^{max}	R_k^{max}	$p_k^* = p_k^{max}$	$R_k^* = R_k^{max}$
User 1	0.0001	10.0000dB	0.1355	10dB	0.0042	10dB	0.3448
User 2	0.0015	10.0000dB	1.3146	10dB	0.0084	10dB	2.3363
User 3	0.0026	10.0000dB	1.8449	10dB	0.0375	10dB	3.0036
User 4	0.0027	10.0000dB	1.8850	10dB	0.0619	10dB	3.0514
User 5	0.0030	10.0000dB	1.9990	10dB	0.0740	10dB	3.1859
User 6	0.0034	10.0000dB	2.1387	10dB	0.1596	10dB	3.3477
User 7	0.0042	10.0000dB	2.3845	10dB	0.1709	10dB	3.6253
User 8	0.0052	8.0488dB	2.6446	10dB	0.6360	10dB	3.9107
User 9	0.0055	4.9588dB	2.7147	10dB	1.0920	10dB	3.9864
User 10	0.0072	3.9829dB	3.0615	10dB	1.2744	10dB	4.3533
User 11	0.0075	-1.7730dB	3.1154	10dB	2.6887	10dB	4.4094
User 12	0.0156	-8.0050dB	3.5450	10dB	4.7045	10dB	5.4303
User 13	0.0226	-10.4002dB	3.5450	10dB	5.6139	10dB	5.9547
User 14	0.0291	-10.4557dB	3.5450	10dB	5.6363	10dB	6.3142
User 15	0.1052	-13.9838dB	3.5450	10dB	7.3171	10dB	8.1550
User 16	0.1253	-13.9838dB	3.5450	10dB	7.3171	10dB	8.4065
User 17	0.1709	-13.9838dB	3.5450	10dB	7.3171	10dB	8.8531
User 18	0.1716	-13.9838dB	3.5450	10dB	7.3171	10dB	8.8590
User 19	0.7282	-13.9838dB	3.5450	10dB	7.3171	10dB	10.9419
User 20	1.7121	-13.9838dB	3.5450	10dB	7.3171	10dB	12.1748
Total Power		19.2043dB		23.0103dB		23.0103dB	
SE			55.1432		34.2744		110.6441

Chapter 6

Conclusion and Scope for Further Work

This chapter concludes the thesis by summarizing and highlighting the work undertaken therein. We also briefly discuss some topics that could be undertaken following the ideas developed in this thesis.

6.1 Conclusion

In this dissertation, we have investigated the spectral efficiency of multi-user massive MIMO systems in downlink and uplink transmissions. In order to maximize the spectral efficiency, we have proposed a number of methods for power allocation given the total power budget.

In downlink transmission, we have investigated the spectral efficiency under time-division duplexing operation via a beamforming training method. Conventionally, in the beamforming training method, the power of the pilot symbols is assumed to be equal to that of the data symbols for all the users. In this dissertation, we have proposed several methods of allocating power so that, in each case, the spectral efficiency is maximized under the assumption of both small scale and large scale fading.

In uplink transmission, we have investigated the spectral efficiency of a single-cell massive MIMO system where the channel vectors for different users are generally correlated, or not asymptotically orthogonal and can be modelled as a finite number of distinct directions.

The investigation concerning the spectral efficiency in the downlink transmission has been considered in the first two parts of the thesis, where as the investigation in the uplink transmission has been carried out in the third part.

In the first part of the dissertation, we have derived a closed-form approximate expression for the achievable downlink rate using the maximum ratio transmission precoder based on small-scale fading in order to evaluate the spectral efficiency. Then, we have proposed three methods of power allocation in order to maximize the spectral efficiency. In the first method, we have posed and answered a fundamental question about the operation of BS in a downlink transmission: how much is the improvement in the spectral efficiency if the average transmit power allocated to the pilot and data symbols are chosen optimally? In answering this question, we have demonstrated that in order to maximize the spectral efficiency for a given total power budget, the spectral efficiency is remarkably improved at high signal to noise ratio by allocating the optimal power to the pilot and data symbols in downlink transmission. In the second method, inspired by the water-filling power allocation scheme, we have posed and answered a basic question about the operation of BS in downlink massive MU-MIMO systems as to how much is the improvement in the spectral efficiency if the power allocated to the data symbols of the various users are chosen optimally? In answering this question, we have found that the spectral efficiency can be significantly increased at high signal to noise ratio by optimally allocating the total transmit power to data symbols of the various users. In the third method, we have allocated power not only to the data symbols but also to the pilot symbols of all the users, where all the pilot powers for the various users are equal. We have shown that all these three proposed methods are superior to other existing methods in terms of spectral

efficiency. In addition, we have shown that the third proposed method of power allocation outperforms the first two aforementioned methods in terms of spectral efficiency.

In the second part of the dissertation, we have derived a closed-form approximate expression for the achievable downlink rate using the maximum ratio transmission precoder based on large-scale fading in order to evaluate the spectral efficiency. Then, we have proposed four schemes of power allocation in order to maximize the spectral efficiency. In the first scheme, we have allocated power among the pilot and data symbols in such a way that the pilot power as well as the data power for each user is the same. The pilot power and the data power are optimally selected in order to maximize the spectral efficiency. In the second scheme, we have allocated power among the data symbols of the various users, whereas the pilot power for each user is the same and is specified. In this scheme, the data power for each user is optimally determined to maximize the spectral efficiency. In the third scheme, we have allocated power among the pilot and data symbols of the various users, whereas the pilot power for each user is the same but determined. In this scheme, the same pilot power along with the various data powers is optimized to maximize the spectral efficiency. In the forth scheme, we have optimally allocated power among each of the pilot and data symbols of the various users so as to maximize the spectral efficiency. We have shown that the second, third and forth schemes offer similar performance in terms of spectral efficiency with approximately the same run time. However, the spectral efficiency of these schemes is much superior to that of the first scheme or that of the existing schemes, even though the run times for the second, third and fourth proposed methods are slightly more than that of the first method. This indicates that in a downlink transmission, where the total power budget is given, optimizing the power of data symbols is much more important than optimizing the power of the pilot symbols in order to maximize the spectral efficiency.

In the third part of the dissertation, first we have derived the achievable uplink rate using zero-forcing and maximum ratio combining receivers based on large-scale fading in

order to evaluate the spectral efficiency. Then, we have proposed a power control strategy among the users to maximize the spectral efficiency. We have shown that the proposed power control method provides a spectral efficiency which is the same as that of the maximum power criterion using the zero-forcing receiver. Further, the proposed method provides a spectral efficiency that is higher than that provided by the maximum power criterion using the maximum ratio combining receiver. In addition, it has been shown that the zero-forcing receiver outperforms the maximum ratio combining receiver in terms of spectral efficiency. In spite of this fact, the maximum ratio combining receiver is still being employed in multi-user massive MIMO systems, due to the fact that the maximum ratio combining receiver is used in a distributed fashion independently at each antenna unit and offers less per-symbol complexity.

6.2 Scope for Further Work

While the research work undertaken in this dissertation has focused on maximizing the spectral efficiency in a massive multi-user MIMO system, there are a number of additional studies that can be undertaken along the lines developed in this dissertation. Some of the possible studies are as follows:

6.2.1 Multi-Cell Systems

We have investigated the spectral efficiency maximization of single-cell massive MIMO systems in this dissertation. However, massive MIMO systems can be employed in multi-cell scenarios, where the activities in the various cells occur synchronously and there is no cell-to-cell cooperation [8]. In multi-cell scenarios, the system model should first be refined to take into account the inter-cell interference. Then, the proposed method can be further investigated to maximize the spectral efficiency by using this refined model.

6.2.2 Multiple-Antenna Terminals

Massive MIMO systems, in contrast to point-to-point MIMO systems, work very well with only single-antenna users. However, massive MIMO systems can employ multiple-antenna users as well. A multiple-antenna user could enjoy throughput in proportion to the number of antennas that the user possesses without requiring exponentially growing signal to interference noise ratios. In addition, the users and the base station could identify the subspace in which the interference is contained, and then operate in the orthogonal interference-free subspace [8]. It would be interesting to investigate the possibility of maximizing the spectral efficiency of such systems, just as has been done in this thesis for the case of single-antenna user.

6.2.3 Massive MIMO with FDD Operation

We have explained in Chapter 3 that a massive MIMO system with time-division duplexing (TDD) mode is superior to frequency-division duplexing (FDD) mode. However, for systems with symmetric traffic and delay-sensitive applications, a massive MIMO system with FDD mode is generally considered to be more effective. Thus, there is a substantial interest in the study of a massive MIMO system with FDD mode [76–79]. In view of this, the possibility of maximizing the spectral efficiency of such systems could also be investigated.

6.2.4 Cell-Free Massive MIMO

Cell-Free (CF) massive MIMO is an alternative topology for massive MIMO networks, where a large number of single-antenna access points (APs) are distributed over the coverage area. There are no cells, but all users are jointly served by the APs using network MIMO methods. It has been shown in [80–88] that CF massive MIMO inherits the basic properties of cellular massive MIMO, namely, channel hardening and favorable

propagation. Maximizing the spectral efficiency in CF massive MIMO systems using the approaches presented in this dissertation would be worth exploring.

References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[2] T. L. Marzetta, "Multi-cellular wireless with base stations employing unlimited numbers of antennas," in *Proc. UCSD Information Theory and Applications Workshop*, Feb 2010.

[3] 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 larg arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

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

[5] E. Bjornson, E. G. Larsson and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[6] M. Agiwal, A. Roy and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, Feb. 2016.

[7] L. Lu, G. Y. Li, A. L. swindlehurst, A. Ashikhmin and R. Zhang, "An overview of massive MIMO benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–754, Oct. 2014.

[8] T. L. Marzetta, E. G. Larsson, H. Yang and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge University Press, 2016.

[9] A. Chockalingam and B. S. Rajan, *Large MIMO Systems*. Cambridge, U.K.: Cambridge University Press, 2014.

[10] O. Saatlou, M. Omair Ahmad and M. N. S. Swamy, "Spectral efficiency maximization of multiuser massive MIMO systems with finite-dimensional channel via control of users power," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 7, pp. 883–887, Jul. 2018.

[11] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[12] C. Xu, Y. Hu, C. Liang, J. Ma and L. Ping, "Massive MIMO, non-orthogonal multiple access and interleave division multiple access," *IEEE Access*, vol. 5, pp. 14 728–14 748, Jul. 2017.

[13] E. L. Bengtsson, F. Rusek, S. Malkowsky, F. Tufvesson, P. C. Karlsson and O. Edfors, "A simulation framework for multiple-antenna terminals in 5G massive MIMO systems," *IEEE Access*, vol. 5, pp. 26 819–26 831, Nov. 2017.

[14] J. Li, D. Wang, P. Zhu and X. You, "Energy efficiency optimization of distributed massive MIMO systems under ergodic QoS and Per-RAU power constraints," *IEEE Access*, vol. 7, pp. 5001–5013, Dec. 2019.

[15] W. Cheng, H. Zhang, L. Liang, H. Jing and Z. Li, "Orbital-angular-momentum embedded massive MIMO: Achieving multiplicative spectrum-efficiency for mmWave communications," *IEEE Access*, vol. 6, pp. 2732–2745, Dec. 2018.

[16] J. Zhang, Y. Wei, E. Bjornson, Y. Han and S. Jin, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," *IEEE Access*, vol. 6, pp. 55 302–55 314, Sep. 2018.

[17] O. Saatlou, M. Omair Ahmad and M.N.S. Swamy, "Max-Min fairness power control for massive MU-MIMO systems with finite-dimensional channel model," in *Proc. IEEE Canadian Conference on Electrical Computer Engineering (CCECE)*, May 2018, pp. 1–4.

[18] O. Saatlou, M. Omair Ahmad and M.N.S. Swamy, "Power control for massive multiuser MIMO systems with finite-dimensional channel," in *Proc. IEEE International New Circuits and Systems Conference (NEWCAS)*, Jun. 2018, pp. 18–21.

[19] A. Khansefid and H. Minn, "Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4849–4864, Sep. 2015.

[20] Y. Wang, C. Li, Y. Huang, D. Wang, T. Ban and L. Yang, "Energy-efficient optimization for downlink massive MIMO FDD systems with transmit-side channel correlation," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7228–7243, Sep. 2016.

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

[22] H. Yang and T. L. Marzetta, "Performance of conjugate and zeroforcing beamforming in large-scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[23] J. Jose, A. Ashikhmin, T. L. Marzetta and S. Vishwanath, "Pilot contamination problem in multi-cell TDD systems," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2009, pp. 2184–2188.

[24] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. Allerton Conference on Communication, Control, and Computing*, Oct. 2013, pp. 293–298.

[25] J. Zuo, J. Zhang, C. Yuen, W. Jiang and W. Luo, "Multi-cell multiuser massive MIMO transmission with downlink training and pilot contamination precoding," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6301–6314, Aug. 2016.

[26] J. Baltersee, G. Fock and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 12, pp. 2358–2368, Dec. 2001.

[27] X. Ma, L. Yang and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 453–456, Mar. 2005.

[28] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. Asilomar Conference on Signals, Systems and Computers (ACSSC)*, Oct. 2006, pp. 359–363.

[29] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[30] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Proceedings (ICASSP)*, May 2001, pp. 2109–2112.

[31] L. Berriche, K. Abed-Meraim and J.C. Belfiore, "Cramer-Rao bounds for MIMO channel estimation," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2004, pp. 397–400.

[32] H. V. Cheng, E. Bjornson and E. G. Larsson, "Optimal pilot and payload power control in single-cell massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2363–2378, May 2017.

[33] H. Q. Ngo, M. Matthaiou and E. G. Larsson, "Massive MIMO with optimal power and training duration allocation," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 605–608, Dec. 2014.

[34] H. V. Cheng, E. Bjornson and E. G. Larsson, "Uplink pilot and data power control for single cell massive MIMO systems with MRC," in *Proc. IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2004, pp. 397–400.

[35] O. Saatlou, M. Omair Ahmad and M.N.S. Swamy, "Spectral efficiency maximization of single cell massive multiuser MIMO systems via optimal power control with ZF receiver," in *Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–5.

[36] Q. Zhang, S. Jin, M. McKay, D. Morales-Jimenez and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5941–5955, Nov. 2015.

[37] A. Khansefid and H. Minn, "Asymptotically optimal power allocation for massive MIMO uplink," in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec. 2014, pp. 627–631.

[38] S. Zarei, W. Gerstacker, R. R. Mller and R. Schober, "Low-complexity linear precoding for downlink large-scale MIMO systems," in *Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2014, pp. 1119–1124.

[39] O. Saatlou, M. Omair Ahmad and M.N.S. Swamy, "Spectral efficiency maximization for massive multiuser MIMO downlink TDD systems via data power allocation with MRT precoding," in *Proc. IEEE Vehicular Technology Conference (VTC-Fall)*, Sep. 2017, pp. 1–5.

[40] O. Saatlou, M. Omair Ahmad and M.N.S. Swamy, "Optimal power allocation for massive MU-MIMO downlink TDD systems," in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, May 2016, pp. 1–4.

[41] O. Saatlou, M. Omair Ahmad and M.N.S. Swamy, "Joint data and pilot power allocation for massive MU-MIMO downlink TDD systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 3, pp. 512–516, Mar. 2019.

[42] Q. Zhang, S. Jin, M. McKay, D. M. Jimenez and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5941–5955, Nov. 2015.

[43] L. Zhao, K. Li, K. Zheng and M. Omair Ahmad, "An analysis of the tradeoff between the energy and spectrum efficiencies in an uplink massive MIMO-OFDM system," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 3, pp. 291–295, Mar. 2015.

[44] J. Chen, Z. Zhang, H. Lu, J. Hu, and G. E. Sobelman, "An intra-iterative interference cancellation detector for large-scale MIMO communications based on convex optimization," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 11, pp. 2062–2072, Nov. 2016.

[45] R. R. Muller, "A random matrix model of communication via antenna arrays," *IEEE Transactions on information theory*, vol. 48, no. 9, pp. 2495–2506, Sep. 2002.

[46] A. G. Burr, "Capacity bounds and estimates for the finite scatterers MIMO wireless chanel," *IEEE Journal on Selected Areas in Communication*, vol. 21, no. 5, pp. 812–818, Jun. 2003.

[47] M. Sadeghi, E. Bjornson, E. G. Larsson, C. Yuen and T. Marzetta, "Joint unicast and multi-group multicast transmission in massive MIMO systems," *IEEE Transactions on Wireless Communications*, pp. 6375–6388, Jul. 2018.

[48] N. Q. Ngo, E. G. Larsson and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Transactions on Communications*, vol. 61, no. 6, Jun. 2013.

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

[50] M. Matthaiou, M. R. Mckay, P. J. Smith and J. A. Nossek, "On the condition number distribution of complex wishart matrices," *IEEE Transactions on Communications*, vol. 58, no. 6, pp. 1705–1717, Jun. 2010.

[51] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jul. 2012, pp. 1137–1141.

[52] L. Li, A. Ashikhmin and T. Marzetta, "Pilot contamination precoding for large scale antenna system: A max-min formulation," in *Proc. Allerton Conference on Communication, Control, and Computing*, Oct. 2013.

[53] A. Adhikary, A. Ashikhmin and T. L. Marzetta, "Uplink interference reduction in large-scale antenna systems," *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 2194–2206, May 2017.

[54] K. Appaiah, A. Ashikhmin and T. L. Marzetta, "Pilot contamination reduction in multi-user TDD systems," in *Proc. IEEE International Conference on Communications (ICC)*, May 2010, pp. 1–5.

[55] F. Fernandes, A. Ashikhmin and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[56] M. Filippou, D. Gesbert and H. Yin, "Decontaminating pilots in cognitive massive MIMO networks," in *Proc. International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2012, pp. 816–820.

[57] H. Yin, D. Gesbert, M. C. Filippou and Y. Liu, "Decontaminating pilots in massive MIMO systems," in *Proc. IEEE International Conference on Communications (ICC)*, Jun. 2013, pp. 3170–3175.

[58] N. Fatema, Y. Xiang and I. Natgunanathan, "Analysis of a semi blind pilot decontamination method in massive MIMO," in *Proc. International Telecommunication Networks and Applications Conference (ITNAC)*, Nov. 2017, pp. 1–6.

[59] 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)*, Apr. 2013, pp. 2890–2895.

[60] L. Dai, Z. Wang and Z. Yang, "Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 251–263, Feb. 2013.

[61] C. B. Peel, "On "dirty-paper coding"," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 112–113, May 2003.

[62] C. Li, Y. Iwanami and E. Okamoto, "A study on Dirty Paper Coding with a maximum beam in multiuser MIMO downlinks," in *Proc. IEEE Region 10 Conference*, Nov. 2010, pp. 2288–2293.

[63] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

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

[65] C. Windpassinger, R. F. H. Fischer and J. B. Huber, "Lattice-reduction-aided broadcast precoding," *IEEE Transactions on Communications*, vol. 52, no. 12, pp. 2057–2060, Dec. 2004.

[66] C. Lee, C. B. Chae, T. Kim, S. Choi and J. Lee, "Network massive MIMO for cell-boundary users: From a precoding normalization perspective," in *Proc. IEEE Globecom Workshops*, Dec. 2012, pp. 233–237.

[67] S. M. Kay, *Fundamental of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, NJ, USA: Prentice Hall, 1995.

[68] O. Saatlou, M. Omair Ahmad and M.N.S. Swamy, "Spectral efficiency of single cell massive MU-MIMO downlink TDD systems with optimal resource allocation," Under review for a journal publication.

[69] Q. Zhang, S. Jin, M. McKay, D. Morales-Jimenez and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5941–5955, Nov. 2015.

[70] T. M. Cover and J. A. Thomas, *Elements of information theory.* New York, USA: Wiley, 1991.

[71] Z. Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[72] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for non-convex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.

[73] M. Chiang, C. W. Tan, D. P. Palomar, D. O'neill and D. Julian, "Power control by geometric programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[74] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U.K.:Cambridge University Press, 2004.

[75] M. Grant and S. Boyd. (2013), *CVX: Matlab software for disciplined convex programming.* Accessed: May. 2017. [online]. Available: http://cvxr.com/cvx.

[76] Z. Jiang, A. F. Molisch, G. Caire and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2868–2882, May 2015.

[77] J. Choi, Z. Chance, D. J. Love and U. Madhow, "Noncoherent trellis coded quantization: A practical limited feedback technique for massive MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 5016–5029, Dec. 2013.

[78] J. Choi, D. J. Love and T. Kim, "Trellis-extended codebooks and successive phase adjustment: A path from LTE-advanced to FDD massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2007–2016, Apr. 2015.

[79] J. Choi, D. J. Love and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 802–814, Oct. 2014.

[80] K. T. Truong and R. W. Heath, "The viability of distributed antennas for massive MIMO systems," in *Proc. Asilomar Conference on Signals, Systems and Computers (ACSSC)*, Nov. 2013, pp. 1318–1323.

[81] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2015, pp. 201–205.

[82] Y. G. Lim, C. B. Chae and G. Caire, "Performance analysis of massive MIMO for cell-boundary users," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6827–6842, Dec. 2015.

[83] H. Q. Ngo, L. N. Tran, T. Q. Duong, M. Matthaiou and E. G. Larsson, "Energy efficiency optimization for cell-free massive MIMO," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2017, pp. 1–5.

[84] S. Buzzi and C. D. Andrea, "Cell-free massive MIMO: User-centric approach," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706–709, Dec. 2017.

[85] J. Zhang, Y. Wei, E. Bjrnson, Y. Han and X. Li, "Spectral and energy efficiency of cell-free massive MIMO systems with hardware impairments," in *Proc. International Conference on Wireless Communications and Signal Processing (WCSP)*, Oct. 2017, pp. 1–6.

[86] T. X. Doan, H. Q. Ngo, T. Q. Duong and K. Tourki, "On the performance of multi-group multicast cell-free massive MIMO," *IEEE Communications Letters*, vol. 21, no. 12, pp. 2642–2645, Dec. 2017.

[87] Y. Li and G. A. Aruma Baduge, "NOMA-aided cell-free massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 950–953, Dec. 2018.

[88] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[89] A. M. Tulino and S. Verdu, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.

Appendix A

Proof of Proposition 3.1

Using the MRT precoder, we have $a_{ki} = \alpha_{MRT}\mathbf{h}_k^T\mathbf{h}_i^*$ [24]. Thus, $\mathbb{E}\{a_{ki}\}$ and $\mathrm{Var}(a_{ki})$ are given by [24]

$$\begin{cases} \mathbb{E}\{a_{ki}\} & = 0 & i \neq k \\ \mathbb{E}\{a_{kk}\} & = \frac{\sqrt{\tau_u p_u M}}{K(\tau_u p_u + 1)} & i = k \\ \mathrm{Var}(a_{ki}) & = \frac{1}{K} & i \ \forall \ k \end{cases} \tag{A.1}$$

Substituting (A.1) in (3.12), we can obtain $\hat{a}_{ki}$ and $\hat{a}_{kk}$ given by (3.18). To prove Proposition 3.1, we also need to calculate $\mathbb{E}\{|\epsilon_{ki}|^2\}$.

For $i = k$, using (3.12) and (3.18), $\mathbb{E}\{|\epsilon_{kk}|^2\}$ can be written as

$$\begin{aligned} \mathbb{E}\{|\epsilon_{kk}|^2\} = & \mathbb{E}\Bigg\{ \Bigg| \frac{K}{\tau_d p_p + K} a_{kk} - \frac{\sqrt{\tau_d p_p}}{\tau_d p_p + K} \tilde{n}_{p,kk} \\ & - \frac{K}{\tau_d p_p + K} \sqrt{\frac{\tau_u p_u M}{K(\tau_u p_u + 1)}} \Bigg|^2 \Bigg\} \\ = & \Big(\frac{K}{\tau_d p_p + K} \Big)^2 \mathbb{E}\{|a_{kk}|^2\} + \frac{\tau_d p_p}{(\tau_d p_p + K)^2}. \end{aligned} \tag{A.2}$$

We also know that $\mathbb{E}\{|a_{kk}|^2\} = \alpha^2_{MRT}\big(\frac{\tau_u p_u}{\tau_u p_u+1}\big)^2 M(M+1) + \alpha^2_{MRT}\frac{\tau_u p_u}{(\tau_u p_u+1)^2}M$ [24]. Substituting $\mathbb{E}\{|a_{kk}|^2\}$ into (A.2), we have

$$\mathbb{E}\{|\epsilon_{kk}|^2\} = \frac{1}{\tau_d p_p + K}. \tag{A.3}$$

For $i \neq k$, using (3.13) and (3.18), $\mathbb{E}\{|\epsilon_{ki}|^2\}$ can be written as

$$\begin{aligned}\mathbb{E}\{|\epsilon_{ki}|^2\} &= \mathbb{E}\{|\frac{K}{\tau_d p_p + K}a_{ki} - \frac{\sqrt{\tau_d p_p}}{\tau_d p_p + K}\tilde{n}_{p,ki}|^2\} \\ &= \big(\frac{K}{\tau_d p_p + K}\big)^2\mathbb{E}\{|a_{ki}|^2\} + \frac{\tau_d p_p}{(\tau_d p_p + K)^2}.\end{aligned} \tag{A.4}$$

We also know that $\mathbb{E}\{|a_{ki}|^2\} = \frac{1}{K}$ [24]. Substituting $\mathbb{E}\{|a_{ki}|^2\}$ into (A.4), we have

$$\mathbb{E}\{|\epsilon_{ki}|^2\} = \frac{1}{\tau_d p_p + K}. \tag{A.5}$$

Substituting (A.1) in (3.13) and having (A.3) and (A.5), we conclude the proof of Proposition 1.

Appendix B

Proof of Proposition 3.2

To prove *Proposition 3.2*, first, we calculate $\mathbb{E}\{|\hat{a}_{ki}|^2\}$ for MRT precoding.

$\mathbb{E}\{|\hat{a}_{ki}|^2\}$:

For $i=k$, from (3.18) and using the fact that $\mathbb{E}\{|a_{kk}|^2\} = \alpha^2_{MRT}\big(\frac{\tau_u p_u}{\tau_u p_u+1}\big)^2 M(M+1) + \alpha^2_{MRT}\frac{\tau_u p_u}{(\tau_u p_u+1)^2}M$ and $\mathbb{E}\{a_{kk}\} = \sqrt{\frac{\tau_u p_u M}{K(\tau_u p_u+1)}}$ [24], we have

$$\begin{aligned}
\mathbb{E}\{|\hat{a}_{kk}|^2\} &= \frac{1}{(\tau_d p_p+K)^2}\mathbb{E}\left[\left|\tau_d p_p a_{kk}+\sqrt{\tau_d p_p}\tilde{n}^T_{pkk}+K\mathbb{E}\{a_{kk}\}\right|^2\right] \\
&=\frac{1}{(\tau_d p_p+K)^2}\Bigg[\mathbb{E}\{|a_{kk}|^2\}(\tau_d p_p)^2+2K\big(\mathbb{E}\{a_{kk}\}\big)^2(\tau_d p_p) \\
&+(K\mathbb{E}\{a_{kk}\})^2\Bigg]=\frac{1}{(\tau_d p_p+K)^2}\Bigg[\Bigg(\alpha^2_{MRT}\frac{\tau_u p_u}{(\tau_u p_u+1)^2}M \\
&+\alpha^2_{MRT}\big(\frac{\tau_u p_u}{\tau_u p_u+1}\big)^2 M(M+1)\Bigg)(\tau_d p_p)^2 + \\
&+\big(\frac{2\tau_u p_u M}{\tau_u p_u+1}+1\big)\tau_d p_p + \frac{\tau_u p_u MK}{(\tau_u p_u+1)}\Bigg]
\end{aligned} \tag{B.1}$$

For $i\neq k$, from (3.18) and using the fact that $\mathbb{E}\{|a_{ki}|^2\}=\frac{1}{K}$ and $\mathbb{E}\{a_{ki}\}=0$ [24], we have

$$\mathbb{E}\{|\hat{a}_{ki}|^2\}=\frac{\tau_d p_p}{(\tau_d p_p+K)^2}\ \big[\frac{\tau_d p_p}{K}+1\big] \tag{B.2}$$

Then, by substituting (B.1) and (B.2) into (3.23), the proof of *Proposition 3* is completed.

Appendix C

Proof of Proposition 4.1

Before proving *Proposition 4.1*, we first calculate $\hat{a}_{ki}$. Using the definition a_{ki}, we have

$$a_{ki} = \alpha \mathbf{g}_k^T \hat{\mathbf{g}}_i^* = \alpha(\hat{\mathbf{g}}_k^T \hat{\mathbf{g}}_i^* + \epsilon_k^T \hat{\mathbf{g}}_i^*). \tag{C.1}$$

Using this definition and (4.11), for $i \neq k$, we have $\mathbb{E}\{a_{ki}\} = 0$ and $\mathbb{E}\{\tilde{y}_{p,k,i}\} = 0$. For $i = k$, we also have

$$\begin{aligned} \mathbb{E}\{a_{kk}\} &= \frac{\alpha\ M\tau_u p_u \beta_k^2}{1 + \tau_u p_u \beta_k}, \\ \mathbb{E}\{\tilde{y}_{p,kk}\} &= \frac{\sqrt{\tau_d p_{p_k}}\alpha\ M\tau_u p_u \beta_k^2}{1 + \tau_u p_u \beta_k}. \end{aligned} \tag{C.2}$$

Using (4.11) and (C.2), for $i = k$ and $i \neq k$, we respectively have

$$\begin{aligned} \mathbb{E}\{a_{kk}\ \tilde{y}_{p,kk}\} &= \Big(\frac{\alpha\ \tau_u p_u \beta_k^2}{1 + \tau_u p_u \beta_k}\Big)^2 (M^2+M) + \frac{\alpha^2 \tau_u p_u M \beta_k^3}{(1 + \tau_u p_u \beta_k)^2}, \\ \mathbb{E}\{a_{ki}\ \tilde{y}_{p,ki}\} &= \frac{M\alpha^2 \tau_u p_u \beta_k \beta_i^2}{1 + \tau_u p_u \beta_i}. \end{aligned} \tag{C.3}$$

The first equation is obtained by employing the fact that $\alpha^2 \mathbb{E}\{\| \hat{g}_k \|^4\} = \alpha^2 \big(\frac{\tau_u p_u \beta_k^2}{1+\tau_u p_u \beta_k}\big)^2 M(M+1)$ (please see Lemma 2.9 in [89]). Knowing the fact that $\text{cov}(a_{ki}, \tilde{y}_{p,ki}) = \mathbb{E}\{a_{ki}\tilde{y}_{p,ki}\} -$

$\mathbb{E}\{a_{ki}\}\mathbb{E}\{\tilde{y}_{p,ki}\}$ and employing (C.2), for $i = k$ and $i \neq k$, we have

$$\text{cov}\{a_{ki}\ \tilde{y}_{p,ki}\} = M\tau_u p_u \sqrt{\tau_d p_{p_k}}\ \gamma_{ki}, \tag{C.4}$$

Moreover, we know that $\text{cov}(\tilde{y}_{p,ki}, \tilde{y}_{p,ki}) = \mathbb{E}\{\tilde{y}_{p,ki}\tilde{y}_{p,ki}\} - \mathbb{E}\{\tilde{y}_{p,ki}\}\mathbb{E}\{\tilde{y}_{p,ki}\}$ or equivalently, for $i \neq k$ and $i = k$, we have

$$\begin{aligned}\text{cov}\{\tilde{y}_{p,ki}\ \tilde{y}_{p,ki}\} &= \mathbb{E}\{|\sqrt{\tau_d p_{p_k}} a_{ki} + \tilde{n}_{p,ki}|^2\} - |\mathbb{E}\{\tilde{y}_{p,ki}\}|^2 \\ &= M\tau_u p_u \sqrt{\tau_d p_{p_k}}\ \gamma_{ki} + 1 \end{aligned} \tag{C.5}$$

Substituting (C.2), (C.4) and (C.5) into (4.12), for $i \neq k$ and $i = k$, we respectively have

$$\begin{aligned}\hat{a}_{ki} &= \frac{M\tau_u p_u \sqrt{\tau_d p_{p_k}}\ \gamma_{ki}}{M\tau_u p_u \tau_d p_{p_k}\ \gamma_{ki} + 1}\ (\tilde{y}_{p,k,i}), \\ \hat{a}_{kk} &= \frac{\alpha M\tau_u p_u \beta_k^2}{1 + \tau_u p_u \beta_k} + \frac{M\tau_u p_u \sqrt{\tau_d p_{p_k}}\ \gamma_{kk}}{M\tau_u p_u \tau_d p_{p_k}\ \gamma_{kk} + 1}\ \Big(\tilde{y}_{p,k,k} \\ &\quad - \frac{M\alpha^2 \tau_u p_u \sqrt{\tau_d p_{p_k}} \beta_k^2}{1 + \tau_u p_u \beta_i}\Big). \end{aligned} \tag{C.6}$$

To prove *Proposition 4.1*, we calculate $\mathbb{E}\{|\hat{a}_{ki}|^2\}$ and $\mathbb{E}\{|\epsilon_{ki}|^2\}$. Using (C.6), for $i \neq k$ and $i = k$, we respectively have

$$\begin{aligned}\mathbb{E}\{|\hat{a}_{ki}|^2\} &= \frac{M\tau_u p_u\ \gamma_{ki}^2}{\gamma_{ki} + \frac{1}{M\tau_u p_u \tau_d p_{p_i}}}\ , \\ \mathbb{E}\{|\hat{a}_{kk}|^2\} &= \frac{M^2\alpha^2(\tau_u p_u)^2 \beta_k^4}{(1 + \tau_u p_u \beta_k)^2} + \frac{M\tau_u p_u\ \gamma_{kk}^2}{\gamma_{kk} + \frac{1}{M\tau_u p_u \tau_d p_{p_k}}}. \end{aligned} \tag{C.7}$$

Knowing the fact that $\varepsilon_{ki} = a_{ki} - \hat{a}_{ki}$ and using (C.6), for $i \neq k$ and $i = k$, we have

$$\begin{aligned}\mathbb{E}\{|\varepsilon_{ki}|^2\} &= \mathbb{E}\{|a_{ki} - \hat{a}_{ki}|^2\} \\ &= \frac{\gamma_{ki}}{\frac{1}{M\tau_u p_u} + \tau_d p_{p_i} \gamma_{ki}}\ . \end{aligned} \tag{C.8}$$

Substituting (C.7) and (C.8) into (4.15), we conclude the proof of *Proposition 4.1.*

Appendix D

Proof of Lemma 5.1

Using ZF detector, i.e., $\mathbf{V} = \mathbf{T}(\mathbf{T}^{\dagger}\mathbf{T})^{-1}$, (5.14) becomes

$$R_{p,k} = \mathbb{E}\Bigg\{ \log_2 \left(1 + \frac{p_k}{\left[(\mathbf{T}^{\dagger}\mathbf{T})^{-1}\right]_{kk}} \right) \Bigg\}. \tag{D.1}$$

Since $\log_2(1+\frac{1}{x})$ is a convex function, by using the Jensen's inequality $Ef(x) \leq f(E(x))$, we obtain the following lower bound

$$\tilde{R}_{p,k} = \log_2 \left(1 + \frac{p_k}{\mathbb{E}\Bigg\{ \left[(\mathbf{T}^{\dagger}\mathbf{T})^{-1}\right]_{kk} \Bigg\}} \right). \tag{D.2}$$

By using (5.2) and (5.9), we have

$$\begin{aligned} \mathbb{E}\Bigg\{ \left[(\mathbf{T}^{\dagger}\mathbf{T})^{-1}\right]_{kk} \Bigg\} = \frac{L}{M}\mathbb{E}\Bigg\{ \left[(\mathbf{G}^{\dagger}\ \mathbf{G})^{-1}\right]_{kk} \Bigg\} = \\ \frac{L}{M\ \beta_k}\mathbb{E}\Bigg\{ \left[(\mathbf{H}^{\dagger}\mathbf{H})^{-1}\right]_{kk} \Bigg\} = \frac{L}{M\beta_k K}\mathbb{E}\Bigg\{ \mathrm{tr}\left[(\mathbf{H}^{\dagger}\mathbf{H})^{-1}\right] \Bigg\}, \end{aligned} \tag{D.3}$$

since the elements of $\mathbf{H}$ are i.i.d with zero mean and unit variance gaussian distribution. In this case, $\mathbf{H}^{\dagger}\mathbf{H}$ is a Wishart matrix with L degrees of freedom. To calculate

$\mathbb{E}\Big\{\mathrm{tr}\big[(\mathbf{H}^{\dagger}\mathbf{H})^{-1}\big]\Big\}$, we use the properties of Wishart matrices [89] as

$$\mathbb{E}\Big\{\mathrm{tr}(\mathbf{W}^{-1})\Big\} = m/(n-m), \tag{D.4}$$

where $\mathbf{W} \sim \mathrm{w}_m(n, \mathbf{I}_n)$ is a central complex Wishart matrix with $n(n > m)$ degrees of freedom. In view of this, we obtain

$$\mathbb{E}\Big\{\mathrm{tr}(\mathbf{H}^{\dagger}\mathbf{H})^{-1}\Big\} = K/(L-K), \tag{D.5}$$

with $L \geq K+1$. By substituting (D.5) into (D.3) and then, (D.3) into (D.2), the proof is completed.

Appendix E

Proof of Lemma 5.2

When BS employes MRC detector, i.e., $\mathbf{V} = \mathbf{T}$, then (5.14) becomes

$$R_{p,k} = \mathbb{E}\Bigg\{\log_2\Bigg(1+\frac{p_k \parallel \mathbf{t}_k \parallel^4}{\sum_{i=1,i\neq k} p_i \mid \mathbf{t}_k^\dagger \mathbf{t}_i \mid^2 + \parallel \mathbf{t}_k \parallel^2}\Bigg)\Bigg\}. \tag{E.1}$$

Using the Jensen's inequality $Ef(x) \leq f(E(x))$, we have

$$\tilde{R}_{p,k}^{mrc} = \log_2\Bigg(1+\Bigg(\mathbb{E}\Bigg\{\frac{\sum_{i=1,i\neq k} p_i \mid \tilde{\mathbf{t}}_i \mid^2 +1}{p_k \parallel \mathbf{t}_k \parallel^2}\Bigg\}\Bigg)^{-1}\Bigg). \tag{E.2}$$

where $\tilde{\mathbf{t}}_i \triangleq \frac{\mathbf{t}_k^\dagger \mathbf{t}_i}{\|\mathbf{t}_k\|}$. Conditioned on $\mathbf{t}_k$, $\tilde{\mathbf{t}}_i$ is a Gaussian random variable (RV) with zero mean (see Appendix of [11]). Using (5.2) and (5.9), the variance of $\tilde{\mathbf{t}}_i$ is obtained as $\mathbb{E}\{\tilde{\mathbf{t}}_i^\dagger \tilde{\mathbf{t}}_i\} = \frac{M}{L}\beta_i$ which is independent of $\mathbf{t}_k$. Hence, $\tilde{\mathbf{t}}_i$ is Gaussian distributed and does not depend on $\mathbf{t}_k$, i.e., $\tilde{\mathbf{t}}_i \sim \mathcal{CN}(0, \frac{M}{L}\beta_i)$. As a result, we have

$$\begin{aligned} &\mathbb{E}\Bigg\{\frac{\sum_{i=1,i\neq k} p_i|\tilde{\mathbf{t}}_i|^2+1}{p_k\|\mathbf{t}_k\|^2}\Bigg\} \\ &\quad = \quad \Bigg(\sum_{i=1,i\neq k} p_i\mathbb{E}\{\mid \tilde{\mathbf{t}}_i \mid^2\} + 1\Bigg)\mathbb{E}\Big\{\frac{1}{p_k \parallel \mathbf{t}_k \parallel^2}\Big\} \\ &\quad = \quad \Bigg(\sum_{i=1,i\neq k} \frac{M}{L} p_i\beta_i + 1\Bigg)\mathbb{E}\Big\{\frac{1}{p_k \parallel \mathbf{t}_k \parallel^2}\Big\}. \end{aligned} \tag{E.3}$$

Using (D.4), we have

$$\mathbb{E}\Big\{\frac{1}{p_k \parallel \mathbf{t}_k \parallel^2}\Big\} = \frac{L}{p_k M \beta_k (L-1)}. \tag{E.4}$$

By substituting (E.4) into (E.3) and then, (E.3) into (E.2), the proof is completed.