Linköping Studies in Science and Technology
Licentiate Thesis No. 1797

Resource Allocation for Max-Min Fairness in Multi-Cell Massive MIMO

Trinh Van Chien

Division of Communication Systems
Department of Electrical Engineering (ISY)
Linköping University, 581 83 Linköping, Sweden
www.commsys.isy.liu.se

Linköping 2017

This is a Swedish Licentiate Thesis.
The Licentiate degree comprises 120 ECTS credits of postgraduate studies.

**Resource Allocation for Max-Min Fairness in Multi-Cell Massive
MIMO**

© 2017 Trinh Van Chien, unless otherwise noted.

ISBN 978-91-7685-387-0

ISSN 0280-7971

Printed in Sweden by LiU-Tryck, Linköping 2017

Abstract

Massive MIMO (multiple-input multiple-output) is considered as an heir of the multi-user MIMO technology and it has recently gained lots of attention from both academia and industry. By equipping base stations (BSs) with hundreds of antennas, this new technology can provide very large multiplexing gains by serving many users on the same time-frequency resources and thereby bring significant improvements in spectral efficiency (SE) and energy efficiency (EE) over the current wireless networks. The transmit power, pilot training, and spatial transmission resources need to be allocated properly to the users to achieve the highest possible performance. This is called resource allocation and can be formulated as design utility optimization problems. If the resource allocation in Massive MIMO is optimized, the technology can handle the exponential growth in both wireless data traffic and number of wireless devices, which cannot be done by the current cellular network technology.

In this thesis, we focus on two resource allocation aspects in Massive MIMO: The first part of the thesis studies if power control and advanced coordinated multipoint (CoMP) techniques are able to bring substantial gains to multi-cell Massive MIMO systems compared to the systems without using CoMP. More specifically, we consider a network topology with no cell boundary where the BSs can collaborate to serve the users in the considered coverage area. We focus on a downlink (DL) scenario in which each BS transmits different data signals to each user. This scenario does not require phase synchronization between BSs and therefore has the same backhaul requirements as conventional Massive MIMO systems, where each user is preassigned to only one BS. The scenario where all BSs are phase synchronized to send the same data is also included for comparison. We solve a total transmit power minimization problem in order to observe how much power Massive MIMO BSs consume to provide the requested quality of service (QoS) of each user. A max-min fairness optimization is also solved to provide every user with the same maximum QoS regardless of the propagation conditions.

The second part of the thesis considers a joint pilot design and uplink (UL)

power control problem in multi-cell Massive MIMO. The main motivation for this work is that the pilot assignment and pilot power allocation is momentous in Massive MIMO since the BSs are supposed to construct linear detection and precoding vectors from the channel estimates. Pilot contamination between pilot-sharing users leads to more interference during data transmission. The pilot design is more difficult if the pilot signals are reused frequently in space, as in Massive MIMO, which leads to greater pilot contamination effects. Related works have only studied either the pilot assignment or the pilot power control, but not the joint optimization. Furthermore, the pilot assignment is usually formulated as a combinatorial problem leading to prohibitive computational complexity. Therefore, in the second part of this thesis, a new pilot design is proposed to overcome such challenges by treating the pilot signals as continuous optimization variables. We use those pilot signals to solve different max-min fairness optimization problems with either ideal hardware or hardware impairments.

Populärvetenskaplig sammanfattning

Massiv MIMO betraktas som den nya generationen av flerantennteknik inom trådlös kommunikation och har fått stor uppmärksamhet från både akademien och industrin. Genom att utrusta basstationer med hundratala antenner kan Massiv MIMO ge höga datatakter och samtidigt använda mindre energi än nuvarande trådlösa nätverk. I Massiv MIMO är resursallokering ett viktigt verktyg för att ytterligare förbättra systemets prestanda. Genom resursallokering kan Massiv MIMO hantera den exponentiella tillväxten i både mängden trådlös datatrafik och antalet trådlösa enheter, vilket nuvarande system inte klarar av. I denna avhandling fokuserar vi på att optimera systemet så att prestandan för de minst privilegierade användarna maximeras. Först analyserar vi hur effekttreglering påverkar system där flera basstationer med ett stort antal antenner kan samarbeta för att betjäna användarna på ett optimalt sätt. Vi studerar även hur man kan gemensamt optimera pilotsekvenser och effekttreglering i upplänken.

Acknowledgments

I would like to send my gratitude to the main supervisor, Associate Professor Emil Björnson, for his valuable supervision and support. His advice, instruction, inspiration, and encouragement have been indispensable for my academic years. He is always dedicated to provide useful guidance whenever I need help. I would also like to send my sincere thanks to the co-supervisor, Professor Erik G. Larsson for giving me the great opportunity to pursue the Ph.D. degree in the Division of Communication Systems. He has been giving me insightful comments and suggestions to expand and complete my research perspectives. The fruitful results in this thesis would not been obtained without support from both supervisors.

I was lucky to have discussions with Dr. Hien Quoc Ngo during his working time at Linköping University. I learned a lot from his maturity and expertise in research. He was also willing to share and advise me in many things in my life from the beginning when I came to Linköping. Besides, the helpful and stimulating discussions with my colleagues vastly assisted me in my research during the last two years. I have indeed learned a lot from them. I am further very grateful to them for the warm and friendly work environment which makes me less lonely when away from home.

I would like to thank my family for their love and encouragement. They may not understand what I am working on, but the continuous support from them is what makes it possible to keep persistent activities in my research. Finally, the warmest thank should be sent to my dear friends for keeping in touch and being interested in my work.

Trinh Van Chien
Linköping, December 2017

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Preliminaries for Massive MIMO	5
1.2.1	Channel Hardening	5
1.2.2	Favorable Propagation	6
1.2.3	TDD and FDD Mode	7
1.3	Multi-Cell Massive MIMO Communications	8
1.3.1	Uplink Pilot Training Phase	8
1.3.2	Uplink Data Transmission	13
1.3.3	Downlink Data Transmission	14
1.4	Coordinated Multipoint (CoMP) Transmission	16
1.4.1	Non-Coherent Joint Transmission	16
1.4.2	Coherent Joint Transmission	20
1.4.3	Transmit Power Consumption at Base Stations	21
1.5	Optimization Preliminaries	22
1.5.1	Convex Optimization Problems	23
1.5.2	Linear Programming	24
1.5.3	Second-Order Cone Programming	24
1.5.4	Geometric Programming	25
1.5.5	Signomial Programming	25
1.5.6	Weighted Max-Min Fairness Optimization Problem	27
2	Contributions of the Thesis	29
2.1	Papers Included in the Thesis	29
2.2	Papers Not Included in the Thesis	31
	Bibliography	33

Included Papers

37

A Joint Power Allocation and User Association Optimization

39

1	Introduction	41
2	System Model and Achievable Performance	45
2.1	Uplink Channel Estimation	46
2.2	Downlink Data Transmission Model	48
2.3	Achievable Spectral Efficiency under Rayleigh Fading	50
3	Downlink Transmit Power Optimization for Massive MIMO Systems	53
4	Optimal Power Allocation and User Association by Linear Programming	55
4.1	Optimal Solution with Linear Programming	55
4.2	BS-User Association Principle	56
5	Max-min QoS Optimization	58
6	Numerical Results	60
7	Conclusion	66
8	Appendix	67
8.1	Proof of Proposition 1 and Theorem 1	67
8.2	Proof of Corollary 1	71
8.3	Proof of Corollary 2	72
8.4	Proof of Theorem 3	73
8.5	Proof of Corollary 3	74
8.6	Joint Power Allocation and User Association for Massive MIMO Systems with Coherent Joint Transmission	75
	References	77

B Joint Pilot Design and Uplink Power Allocation

83

1	Introduction	85
2	Pilot Designs for Massive MIMO Systems	88
2.1	Proposed Pilot Design	89
2.2	Other Pilot Designs	90
3	Uplink Massive MIMO Transmission	92
3.1	Channel Estimation with Arbitrary Pilots	92
3.2	Uplink Data Transmission	94
4	Max-min Fairness Optimization	96
4.1	Problem Formulation	97
4.2	Local Optimality Algorithm	100
5	Pilot Optimization for Cellular Massive MIMO Systems with Hardware Impairments	102

5.1	Channel Estimation under Hardware Impairments . .	103
5.2	UL Data Transmission and Max-min Fairness Opti- mization under Hardware Impairments	105
6	Generalization to Correlated Rayleigh fading	106
7	Numerical Results	108
8	Conclusion	115
9	Appendix	115
9.1	Proof of Lemma 5	115
9.2	Proof of Theorem 5	117
9.3	Proof of Theorem 6	118
	References	119

Chapter 1

Introduction

1.1 Motivation

The attenuations of the transmitted wireless signals due to, for example, scattering, shadowing by large obstacles, and long distances are fundamental challenges in radio propagation to provide reliable communications. Single-input single-output is the simplest form of communication systems, where the transmitter and receiver are equipped with only one antenna each. Hence, the receiver only observes one version of the transmitted signals at a given time instant and the transmitter cannot direct the signals towards the receiver, so it is only possible to achieve a high data throughput over short distances and even then, the system is affected by small-scale fading. In contrast, MIMO is a spatial multiplexing technology which utilizes multiple antennas at both the transmitter and receiver. Since a receiver observes many variants of the same transmitted signals, it can extract more efficiently the information to combat small-scale fading and enhance communication reliability. By having multiple antennas at the transmitter, directional beamforming can be used to steer the signal towards the receiver and achieve an amplification called the array gain. The transmitter can also simultaneously send multiple signals with different directional beamforming vectors, which increases the data rate and this is called the multiplexing gain. These are two fundamental improvements as compared to single-antenna scenarios. Academia and industry have investigated the MIMO technology for the last twenty years and recently it has been deployed in wireless standards, for instance, Wi-Fi (IEEE 802.11n, IEEE 802.11ac) and 4G (WiMAX, LTE) [1].

The number of wireless devices and the usage per device increase quickly, which has led to an exponential growth in the demand for data traffic [2]. This trend is expected to continue in the near future as recently reported

in [3]. Unfortunately, the current MIMO systems (e.g., Wi-Fi and 4G) cannot handle those demands due to limitations of only having a few antennas at the BSs. First, these systems can only provide a small array gain for the users and, second, the ability to serve multiple users on the given time-frequency resources is limited due to interference, which limits the multiplexing gain. Recently, Massive MIMO, the new generation of multi-user MIMO, has been considered as a potential technology for the next wireless network generations [4]. In Massive MIMO, the BSs are equipped with hundreds of antennas such that the impact of mutual interference, thermal noise, and small-scale fading can be almost eliminated by the array gain and related phenomena described below [5]. For a given time-frequency resource, a Massive MIMO BS is capable of serving tens of users simultaneously and therefore achieve high multiplexing gains that bring large enhancements in spectral efficiency (SE), measured in bit/s/Hz, and in energy efficiency (EE), measured in bit/J [6]. Massive MIMO systems provide higher data rates without the need for more bandwidth or deployment of more BSs.

The main benefits of Massive MIMO systems are summarized as follows:

- *Higher scalability:* In small-scale MIMO systems, where each BS is only equipped with a few antennas, many different channel estimation methods can be used to achieve accurate channel estimates. However, when increasing the number of antennas, the transmission protocol must be properly designed to limit the channel estimation overhead or, more precisely, avoid that it grows proportionally to the number of antennas. There are two categories of protocols: frequency division duplex (FDD), i.e., the UL and DL transmissions operate at the same time and use different frequencies, and time division duplex (TDD), i.e., the UL and DL transmissions use the same frequency resource and operate in different time [7]. When using the TDD protocol, the channel estimation overhead is only proportional to the number of users and independent of the number of BS antennas [5,8,9]. This is achieved by utilizing the fact that channel estimates obtained in the UL can be also used for DL transmission, which is a physical characteristic called channel reciprocity. Massive MIMO should therefore be deployed using TDD. This is further discussed in Subsection 1.2.3.
- *Lower transmit power consumption:* When the number of BS antennas increases, meaning that the array gain (the power gain due to using multiple antennas) expands, the transmitted powers of the UL and DL can be significantly reduced while the desired SE for every user is maintained [10]. For the DL transmission, BSs transmit directional beams

into the directions where the users are located. Such beamforming focuses the signals on the individual users, which allows for reducing the transmit power and gaining higher EE. For the UL transmission, the total transmit power of all users can also be substantially reduced in many scenarios thanks to a large array gain obtained by the coherent combining of the received signals at the BS. The DL power consumption of each BS is briefly described in Subsection 1.4.3, while detailed analysis and simulation results are presented in Paper A.

- *Higher spectral efficiency:* In Massive MIMO systems, one of the key features to improve the SE is scaling up the number of BS antennas. This provides an array gain [8] that improves the SNR and SE of every user. In addition, the large number of antennas enables the BS to separate user signals in the spatial domain. Hence, with a large number of antennas, the BS can spatially multiplex a large number of users to improve the sum SE of the cell. Ideally, the sum SE can grow proportionally to the number of multiplexed users [11]. Unfortunately, classical MIMO systems only involve a few antennas, thus these gains are very limited and we expect substantially higher SE in Massive MIMO where BSs are equipped with hundreds of antennas.
- *Simpler signal processing:* In Massive MIMO, high SE can be obtained by using linear processing schemes, such as maximum ratio (MR) or zero forcing (ZF), which only cancel interference spatially and not by advanced coding and decoding schemes. This is in contrast to the optimal methods (dirty paper coding and successive interference cancellation) which are needed in conventional multi-user MIMO systems to achieve good performance [12]. Linear processing schemes work well when the system has a high ratio between the number of BS antennas and the number of served users, which leads to a set of user channels that are mutually nearly orthogonal. This property is known as favorable propagation [13, 14]. Apart from the transmission and reception processing, the signal processing needed for resource allocation can also be simplified in Massive MIMO. In particular, the channel hardening property makes the channel gains after MR or ZF processing more deterministic as the number of antennas increases [15]. This merit makes it possible to approximate the instantaneous gain with the average gain in resource allocation tasks and also alleviate the need for adapting the resource allocation to small-scale fading variations.

From the seminal paper [16] providing the initial framework of Massive MIMO with infinitely many antennas, numerous papers have analyzed various aspects

of Massive MIMO systems in general [17, 18], and multi-cell Massive MIMO systems in particular [19]. The authors in [6] studied the SE and EE for a finite number of antennas, but with fixed transmit power levels and ideal hardware. Power control is an important aspect of wireless communications in order to balance the effects of mutual interference and amplifying the power of the desired signals. Power control is challenging since the power allocated to increase the QoS for one user will contribute to interference at the other users. Power control in wireless networks has been studied for decades, but one big issue with existing algorithms is the complexity when deploying the algorithms in small-scale MIMO networks due to the fast variations of small-scale fading which require the power control to change very often [20]. Fortunately, power control algorithms are much easier to deploy in Massive MIMO since the SE expressions only dependent on the large-scale fading coefficients thanks to the channel hardening property as demonstrated in [8] and references therein. For simplicity, these prior works usually assumed that each BS serves a fixed set of equally many users. In practice, the user load is not uniformly distributed over the coverage area at any given time, and therefore some BSs may serve many more users than others. A good approach to deal with the BS-user association is letting all the BSs collaborate. For this reason, we established a novel framework for joint user association and power allocation in the downlink (DL) of Massive MIMO which allows a user to be served by a subset of the BSs. We consider both coherent and non-coherent joint transmission. In this work, we want to answer if advanced BS cooperation techniques can bring a significant reduction of transmit powers for Massive MIMO while the required SEs are maintained.

The pilot design is crucial in Massive MIMO systems [18] since every BS obtains instantaneous channel state information (CSI) from UL pilot signals, and then use them to construct the UL detection and DL precoding vectors. In prior works, the pilot design is divided into the two separate tasks: pilot assignment and pilot power control [21]. Pilot assignment consists of methods to assign each user with a pilot from an orthogonal pilot set to reduce interference in the pilot transmission, known as pilot contamination [22]. This is a challenging problem since different users are more or less susceptible to contamination. The best assignment solution is typically obtained by exhaustive search methods but such methods have exponential computational complexity. By utilizing imbalanced power allocation, pilot power control can give better channel estimation quality and reduce the coherent interference coming from the users utilizing the same pilot signals [10]. As a contribution of this thesis, we propose a new pilot design that can first

overcome the combinatorial problems. Furthermore, the new pilot design is a generalization of prior works and it performs both pilot assignment and pilot power control in a joint framework. We further compute a closed-form expression of the UL ergodic SE for Rayleigh fading channels when using the MR detection scheme. We use this closed-form expression to formulate a max-min fairness optimization that optimizes the weakest user SE. Numerical results demonstrate improvements of our proposal for multi-cell Massive MIMO systems over the prior works.

The rest of this chapter is organized as follows: Section 1.2 defines and explains basic terminologies which are used in Massive MIMO. Section 1.3 briefly presents the UL and DL transmission models in conventional multi-cell Massive MIMO systems. Meanwhile, the coordinated multipoint (CoMP) schemes comprising of coherent and non-coherent joint transmission for DL multi-cell Massive MIMO is described in Section 1.4. Some preliminaries of the classical optimization problems are introduced in Section 1.5.

1.2 Preliminaries for Massive MIMO

Massive MIMO is a multi-user system where each BS is equipped with hundreds of antennas. The system is designed to serve tens of users utilizing the same time-frequency resources. This section defines and explains the basic terminologies in Massive MIMO which are later used in this thesis.

1.2.1 Channel Hardening

Let $\mathbf{h} \in \mathbb{C}^M$ have random entries and stand for the channel response between an arbitrary BS and an arbitrary user, then the channel hardening property states that

$$\frac{\|\mathbf{h}\|^2}{\mathbb{E}\{\|\mathbf{h}\|^2\}} \rightarrow 1, \quad (1)$$

with almost sure convergence when $M \rightarrow \infty$. We stress that the channel hardening property is only satisfied under certain technical conditions on the correlation matrix $\mathbb{E}\{\mathbf{h}\mathbf{h}^H\}$ provided in [10], but these are typically satisfied by the channel models used in the communication field. Furthermore, (1) is interpreted as $\|\mathbf{h}\|^2$ being close to the expected value $\mathbb{E}\{\|\mathbf{h}\|^2\}$ if the BS is equipped with a sufficient large number of antennas. This important property demonstrates the disappearance of the small-scale fading effect and allows Massive MIMO systems to use the average channel gains, i.e., deterministic numbers, rather than the corresponding instantaneous values when computing the performance and making resource allocation decisions.

Both Paper A and Paper B utilize the channel hardening property to derive closed-form bounds on the UL/DL ergodic capacities which are independent of the small-scale fading realizations.

1.2.2 Favorable Propagation

Let $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{C}^M$ be random vectors which represent the channel responses between a BS and two different users. If these vectors are non-zero and orthogonal in the sense that

$$\mathbf{h}_1^H \mathbf{h}_2 = 0, \quad (2)$$

where $(\cdot)^H$ denotes the Hermitian transpose, then the BS can completely separate signals sent from the two users when it observes $\mathbf{y} = \mathbf{h}_1 s_1 + \mathbf{h}_2 s_2$. The signal sent from the first user is detected by simply computing the inner product between $\mathbf{y}$ and $\mathbf{h}_1$ as

$$\mathbf{h}_1^H \mathbf{y} = \mathbf{h}_1^H \mathbf{h}_1 s_1 + \mathbf{h}_1^H \mathbf{h}_2 s_2 = \|\mathbf{h}_1\|^2 s_1, \quad (3)$$

where the inner product between the two channel vectors disappears due to (2). The same approach can be applied for the second user: $\mathbf{h}_2^H \mathbf{y} = \|\mathbf{h}_2\|^2 s_2$. Here, we note that the BS needs perfect knowledge of $\mathbf{h}_1$ and $\mathbf{h}_2$ to compute these inner products. The channel orthogonal property in (2) is called favorable propagation, since the two users can communicate with the BS without affecting each other. In reality, the propagation channels may not offer favorable propagation due to the strict requirement in (2). However, an approximate form of the favorable propagation can be achieved, for example, in non-line-of-sight scenarios with rich scattering and in line-of-sight scenarios with distinct user angles [13]. For example, suppose the two channel vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ have independent random entries with zero mean, identical distribution, and bounded fourth-order moments, then

$$\frac{\mathbf{h}_1^H \mathbf{h}_2}{M} \rightarrow 0, \quad (4)$$

with almost sure convergence when $M \rightarrow \infty$ [5]. We refer to (4) as asymptotic favorable propagation, since if we divide all the terms in the second expression in (3) with M , the interference term will vanish asymptotically, while $\mathbf{h}_1^H \mathbf{h}_1/M$ goes to a non-zero constant.

When there are K users per cell, it is preferable to have $M \gg K$ if the interference from all users should be negligible.

Figure 1: Illustration of a basic TDD Massive MIMO transmission protocol, where the time-frequency resources are divided into the coherence intervals.

1.2.3 TDD and FDD Mode

The propagation channels vary over time and frequency. However, we divide the radio resources into coherence intervals in which the channels are static and frequency flat. We denote the number of symbols per coherence interval as τ_c .

There are two ways of implementing the DL and UL transmission over a given frequency band. In FDD mode, the bandwidth is split into two separate parts: one for the UL and one for the DL. Pilot signals are needed in both the DL and UL due to the frequency selective fading. If M and K are the number of BS antennas and users, respectively, then each pair of UL/DL coherence intervals need $M + K$ symbols dedicated to pilot training process and M symbols for feedback of the DL estimates.

There is an alternative TDD mode where the whole bandwidth is used for both DL and UL transmission but separated in time. If the system switches between DL and UL faster than the channels are changing, i.e., it takes place in the same coherence interval, then it is sufficient to learn the channels in only one of the directions. This leads to a pilot length of $\min(M, K)$ per coherence interval if we send pilots only in the most efficient direction. In the preferable Massive MIMO operating regime of $M \gg K$, where favorable propagation appears, we note that TDD systems should send pilots only in the UL and the pilot length becomes $\min(M, K) = K$.

In summary, FDD requires $M + K/2$ pilots and TDD requires K pilots per coherence interval. We conclude that TDD is the preferable mode since if the Massive MIMO systems work in the preferable operating regime, it not only requires shorter pilots than FDD but it is also highly scalable since the pilot length is independent of the number of BS antennas.

Figure 2: Uplink multi-cell Massive MIMO communications: the link between a user and its serving BS is considered as the desired link, while the links from this user to the other BSs are interference links.

1.3 Multi-Cell Massive MIMO Communications

In this section, we consider a cellular network with L cells each including a BS equipped with M antennas and serving K users. We assume that the system operates in TDD mode where the propagation channels vary over time and frequency. As described above, we divide the radio resources into coherence intervals of τ_c symbols in which the channels are static and frequency flat as shown in Figure 1. In each coherence interval, the UL training process utilizes τ_p symbols and the remaining symbols are dedicated to the UL and DL data transmissions. We also define the factors $\gamma^{\text{UL}}, \gamma^{\text{DL}} \in [0, 1]$ satisfying $\gamma^{\text{UL}} + \gamma^{\text{DL}} = 1$, such that $\gamma^{\text{UL}}(\tau_c - \tau_p)$ and $\gamma^{\text{DL}}(\tau_c - \tau_p)$ data symbols are assigned to the UL and DL transmissions, respectively in each coherence interval. We now review in more detail this transmission protocol to explain the main features and figure out the limitations of prior works which are seen as the motivations for the research presented in this thesis.

1.3.1 Uplink Pilot Training Phase

The UL transmission is schematically illustrated in Figure 2. The solid arrows are desired links each standing for the channel between a user and its serving BS. Meanwhile, the dashed links are the interfering links which come from a user in another cell and disturb the received signals at this BS.

The radio channels are unknown at the beginning of each coherence interval and the BSs estimate them in the pilot training phase. The estimates

will later be used to compute linear detection vectors in the UL and linear precoding vectors in the DL.

During the UL pilot transmission in an arbitrary coherence interval, the received baseband signal $\mathbf{Y}_l \in \mathbb{C}^{M \times \tau_p}$ at BS l is formulated as

$$\mathbf{Y}_l = \sum_{i=1}^L \sum_{t=1}^K \mathbf{h}_{i,t}^l \boldsymbol{\psi}_{i,t}^H + \mathbf{N}_l, \quad (5)$$

where $\mathbf{h}_{i,t}^l \in \mathbb{C}^M$ is the channel between user t in cell i and BS l and it comprises of both small-scale fading and large-scale fading. $\boldsymbol{\psi}_{i,t} \in \mathbb{C}^{\tau_p}$ denotes the deterministic pilot signal allocated to this user, while $\mathbf{N}_l \in \mathbb{C}^{M \times \tau_p}$ is Gaussian noise with the independent circularly symmetric complex Gaussian elements distributed as $\mathcal{CN}(0, \sigma_{\text{UL}}^2)$. The matrix

$$\boldsymbol{\Psi}_i = [\boldsymbol{\psi}_{i,1}, \dots, \boldsymbol{\psi}_{i,K}] \quad (6)$$

is the $\tau_p \times K$ matrix with the pilot signals used in cell i .

To estimate the channel from user k in cell l , the received signal $\mathbf{Y}_l$ in (5) is correlated with the pilot $\boldsymbol{\psi}_{l,k}$ of this user. We then obtain

$$\mathbf{y}_{l,k} = \mathbf{Y}_l \boldsymbol{\psi}_{l,k} = \sum_{i=1}^L \sum_{t=1}^K \mathbf{h}_{i,t}^l \boldsymbol{\psi}_{i,t}^H \boldsymbol{\psi}_{l,k} + \mathbf{N}_l \boldsymbol{\psi}_{l,k}. \quad (7)$$

Several channel estimation techniques can be applied to obtain an estimate of $\mathbf{h}_{l,k}^l$ from $\mathbf{y}_{l,k}$, for example, least square (LS) or minimum mean square error (MMSE) [23]. Here, we consider the MMSE estimate since this technique gives smaller estimation errors than LS. We assume that $\mathbf{h}_{i,t}^l \sim \mathcal{CN}(\mathbf{0}, \beta_{i,t}^l \mathbf{I}_M)$ where $\beta_{i,t}^l$ is the large-scale fading coefficient describing the path-loss and shadow fading. The channel model for $\mathbf{h}_{l,k}^l$ has zero mean, which is suitable for non-line-of-sight environments and is the scenario considered in this thesis. This model is known as uncorrelated Rayleigh fading. In Paper B, we also consider a correlated Rayleigh fading channel model. The MMSE channel estimate $\hat{\mathbf{h}}_{l,k}^l \in \mathbb{C}^M$ is formulated as

$$\begin{aligned} \hat{\mathbf{h}}_{l,k}^l &= \text{Cov}\{\mathbf{y}_{l,k}, \mathbf{h}_{l,k}^l\} (\text{Cov}\{\mathbf{y}_{l,k}, \mathbf{y}_{l,k}\})^{-1} \mathbf{y}_{l,k}, \\ &= \beta_{l,k}^l \|\boldsymbol{\psi}_{l,k}\|^2 \left(\sum_{i=1}^L \sum_{t=1}^K \beta_{i,t}^l |\boldsymbol{\psi}_{i,t}^H \boldsymbol{\psi}_{l,k}|^2 + \sigma_{\text{UL}}^2 \|\boldsymbol{\psi}_{l,k}\|^2 \right)^{-1} \mathbf{y}_{l,k}, \end{aligned} \quad (8)$$

where $\text{Cov}\{\cdot, \cdot\}$ denotes the covariance matrix of two random vectors. The estimation error $\mathbf{e}_{l,k}^l = \mathbf{h}_{l,k}^l - \hat{\mathbf{h}}_{l,k}^l$ with its covariance matrix

$$\left(\beta_{l,k}^l - (\beta_{l,k}^l)^2 \|\boldsymbol{\psi}_{l,k}\|^4 \left(\sum_{i=1}^L \sum_{t=1}^K \beta_{i,t}^l |\boldsymbol{\psi}_{i,t}^H \boldsymbol{\psi}_{l,k}|^2 + \sigma_{\text{UL}}^2 \|\boldsymbol{\psi}_{l,k}\|^2 \right)^{-1} \right) \mathbf{I}_M, \quad (9)$$

from which we see that the selection of pilot signals have a large impact on the estimation error. The pilot power $\|\boldsymbol{\psi}_{l,k}\|^2$ impacts how strong the desired pilot signal is and the squared inner products $|\boldsymbol{\psi}_{i,t}^H \boldsymbol{\psi}_{l,k}|^2$ determine how much interference that the users cause to each other during pilot transmission.

Hence, we conclude that the quality of the channel estimation depends on the pilot design. Many pilot signal structures have been proposed in the literature [11, 24–26], but they can be roughly classified into two main tasks: *pilot assignment* considers a set of well-designed pilot signals and aims at assigning these pilot signals to the users and *pilot power control* distributes a power budget to the pilot signals. We will give a brief review of such pilot designs by utilizing an orthonormal basis $\{\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{\tau_p}\}$ that spans all τ_p -length pilot signals, where $\boldsymbol{\phi}_k$ is the vector where the magnitude of the k th element equals 1 and the other elements equal 0. Two vectors of the basis satisfy

$$\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} = \begin{cases} 1 & \text{if } k = k', \\ 0 & \text{if } k \neq k'. \end{cases} \quad (10)$$

The basis matrix is then defined as $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{\tau_p}] \in \mathbb{C}^{\tau_p}$. One example of such a basis matrix is an identity matrix and an other example is a unitary matrix as in Figure 3. By using this orthonormal basis, the pilot signal designs in prior works can be categorized as follows and illustrated as in Figure 3:

- *Pilot assignment only*: This pilot design was proposed in [11, 24] and only focuses on the pilot assignment for a given set of orthogonal pilot signals. By using the permutation matrix $\boldsymbol{\Pi}_l \in \mathbb{R}_+^{\tau_p \times K}$ that has only one non-zero element in each row and at most one non-zero element in each column, which are denoted in a different color in Figure 3, the pilot signals $\boldsymbol{\Psi}_l$ in cell l are constructed from the basis matrix $\boldsymbol{\Phi}$ as $\boldsymbol{\Psi}_l = \sqrt{\hat{p}} \boldsymbol{\Phi} \boldsymbol{\Pi}_l$, where $\sqrt{\hat{p}}$ is the equal power level used by all users. This pilot design is shown in Figure 3a. Note that the k th column of the permutation matrix contains one non-zero element standing for the pilot signal index assigned to user k in cell l . We further note that there are many collections of permutation matrices $\boldsymbol{\Pi}_1, \dots, \boldsymbol{\Pi}_L$ that give the same result, since we can change the order of the basis vectors without affecting the performance. If we remove this ambiguity by fixing the assignment in the first cell, there are still $(K!)^{L-1}$ different combinations of $\boldsymbol{\Pi}_2, \dots, \boldsymbol{\Pi}_L$. Therefore, the computational complexity of the pilot assignment increases exponentially with the number of cells and number of users per cell. Many prior works on the pilot assignment topic only consider $\tau_p = K$ to limit the complexity, see for example [26].

- *Pilot power control only:* This pilot design was proposed in [25, 27] and concentrates on the pilot power control, while assigning $\tau_p = K$ pilots in a predefined manner in each cell (k th pilot to the k th user). The pilot signals assigned to the K users in cell l are $\Psi_l = \Phi \mathbf{P}_l$, where $\mathbf{P}_l = \text{diag}(\sqrt{\tilde{p}_{l,1}}, \dots, \sqrt{\tilde{p}_{l,K}})$ consists of the square roots of the powers for all users in cell l . Here $\text{diag}(\mathbf{x})$ denotes the diagonal matrix with the vector $\mathbf{x}$ on the diagonal. This is illustrated in Figure 3b. $\tilde{p}_{l,k}$ is the pilot power of user k in cell l .
- *A combination of pilot assignment and pilot power control:* This is basically a combination of the two pilot designs above with $\Psi_l = \Phi \Pi_l \mathbf{P}_l$. It is more costly than the previous designs and is visualized in Figure 3c since for a given power set, we need to find the best pilot reuse set for each user using a utility function, for example, mean squared error. This is considered in Paper B as benchmark.

In order to observe more clearly of how a pilot design effects to the channel estimation quality, we now consider the combination of pilot assignment and pilot power control design. If user k in cell l transmits its pilot signal using the power $\tilde{p}_{l,t}$ and $\mathcal{S}_{l,k}$ is the set of indices of all users that use the same pilot as user k in cell l (including the user itself), the MMSE estimator in (8) gives

$$\hat{\mathbf{h}}_{l,k}^l = \frac{\beta_{l,k}^l \tilde{p}_{l,k}}{\sum_{(i,t) \in \mathcal{S}_{l,k}} \beta_{i,t}^l \tilde{p}_{i,t} + \sigma_{\text{UL}}^2} \mathbf{y}_{l,k} \quad (11)$$

and it is distributed as

$$\hat{\mathbf{h}}_{l,k}^l \sim \mathcal{CN} \left(\mathbf{0}, \frac{(\beta_{l,k}^l)^2 \tilde{p}_{l,k}}{\sum_{(i,t) \in \mathcal{S}_{l,k}} \beta_{i,t}^l \tilde{p}_{i,t} + \sigma_{\text{UL}}^2} \mathbf{I}_M \right). \quad (12)$$

The channel estimation error $\mathbf{e}_{l,k}^l = \mathbf{h}_{l,k}^l - \hat{\mathbf{h}}_{l,k}^l$ is distributed as

$$\mathbf{e}_{l,k}^l \sim \mathcal{CN} \left(\mathbf{0}, \left(\beta_{l,k}^l - \frac{(\beta_{l,k}^l)^2 \tilde{p}_{l,k}}{\sum_{(i,t) \in \mathcal{S}_{l,k}} \beta_{i,t}^l \tilde{p}_{i,t} + \sigma_{\text{UL}}^2} \right) \mathbf{I}_M \right). \quad (13)$$

Based on (13), the channel estimation quality is affected negatively from the users utilizing the same pilot signal as user k in cell l and it shows up in the term $\sum_{(i,t) \in \mathcal{S}_{l,k}} \beta_{i,t}^l \tilde{p}_{i,t}$. For a considered optimization problem such as the channel estimation error minimization or the SE maximization, selecting

Figure 3: Illustration of the pilot designs in prior works: (a) The pilot design in [11, 24] which only focuses on the pilot assignment; (b) The pilot design in [25] which only focuses on the pilot power control; (c) A combined pilot design which involves both the pilot assignment and pilot power control.

the best pilot reuse sets $\mathcal{P}_{i,k}$ for all users in the L cells is a combinatorial assignment problem and it is intractable for large-scale networks due to a huge amount of possible combinations as aforementioned. In Paper B, we overcome the combinatorial pilot assignment problem by proposing a new pilot design which treats pilot signals as continuous optimization variables instead of vectors from an orthogonal basis. The performance of joint pilot sequence design and uplink power control for multi-cell Massive MIMO systems is then investigated for either ideal hardware or hardware impairments.

1.3.2 Uplink Data Transmission

In the UL transmission, the K users in a cell are sending data signals to the serving BS. All users in the network cause mutual interference to each other, i.e., intra-cell interference and inter-cell interference. We assume that an arbitrary user t in cell i transmits the data signal $x_{i,t} \sim \mathcal{CN}(0, 1)$. At BS l , the $M \times 1$ received signal vector is the superposition of all transmitted signals and formulated as

$$\mathbf{y}_l = \sum_{i=1}^L \sum_{t=1}^K \sqrt{p_{i,t}} \mathbf{h}_{i,t}^l x_{i,t} + \mathbf{n}_l, \quad (14)$$

where $p_{i,t}$ is the transmit power that the user allocates to the signal $x_{i,t}$ and the additive noise follows a complex Gaussian distribution, $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{0}, \sigma_{\text{UL}}^2 \mathbf{I}_M)$. BS l then selects a detection vector $\mathbf{v}_{l,k} \in \mathbb{C}^M$ to detect the transmitted signal by applying it to the received signal in (14) as

$$\mathbf{v}_{l,k}^H \mathbf{y}_l = \sum_{i=1}^L \sum_{t=1}^K \sqrt{p_{i,t}} \mathbf{v}_{l,k}^H \mathbf{h}_{i,t}^l x_{i,t} + \mathbf{v}_{l,k}^H \mathbf{n}_l. \quad (15)$$

Since the exact ergodic channel capacity for the case of imperfect channels is unknown, we need to use an alternative metric for the communication performance. In this thesis, we consider a lower bound on the UL ergodic capacity of the channel to user k in cell l which is

$$R_{l,k}^{\text{UL}} = \gamma^{\text{UL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left(1 + \text{SINR}_{l,k}^{\text{UL}} \right), \quad (16)$$

where the effective SINR value, denoted by $\text{SINR}_{l,k}^{\text{UL}}$, is

$$\text{SINR}_{l,k}^{\text{UL}} = \frac{p_{l,k} |\mathbb{E}\{\mathbf{v}_{l,k}^H \mathbf{h}_{l,k}^l\}|^2}{\sum_{i=1}^L \sum_{t=1}^K p_{i,t} \mathbb{E}\{|\mathbf{v}_{l,k}^H \mathbf{h}_{i,t}^l|^2\} - p_{l,k} |\mathbb{E}\{\mathbf{v}_{l,k}^H \mathbf{h}_{l,k}^l\}|^2 + \sigma_{\text{UL}}^2 \mathbb{E}\{\|\mathbf{v}_{l,k}\|^2\}}. \quad (17)$$

We note that the lower bound in (16) on the capacity is obtained by using the use-and-then-forget bounding technique from [8]. Here, the effective SINR value means that the lower bound in (16) is equivalent to the capacity of a Gaussian channel that has the SNR equal to the SINR in (17). The capacity bound (16) is applicable for any channel distributions and detection vectors. The lower bound in (16) is measured in bit/s/Hz and will be called an SE.

For the special case of uncorrelated Rayleigh fading, i.e., $\mathbf{h}_{l,k}^l \sim \mathcal{CN}(\mathbf{0}, \beta_{l,k}^l \mathbf{I}_M)$, closed-form expressions of the UL ergodic capacity bound can be obtained for some linear detection schemes by computing the moments of Gaussian distributions [8]. For example, by using MR detection with $\mathbf{v}_{l,k} = \hat{\mathbf{h}}_{l,k}^l$ and considering uncorrelated Rayleigh fading, (17) becomes

$$\text{SINR}_{l,k}^{\text{UL,MR}} = \frac{M p_{l,k} \frac{\bar{p}_{l,k} (\beta_{l,k}^l)^2}{\sum_{(i,t) \in \mathcal{P}_{l,k}} \bar{p}_{i,t} \beta_{i,t}^l + \sigma_{\text{UL}}^2}}{M \sum_{(i,t) \in \mathcal{P}_{l,k} \setminus \{(l,k)\}} p_{i,t} \frac{\bar{p}_{i,t} (\beta_{i,t}^l)^2}{\sum_{(i',t') \in \mathcal{P}_{l,k}} \bar{p}_{i',t'} \beta_{i',t'}^l + \sigma_{\text{UL}}^2} + \sum_{i=1}^L \sum_{t=1}^K p_{i,t} \beta_{i,t}^l + \sigma_{\text{UL}}^2}}. \quad (18)$$

From (18), we observe that the SE depends on the data and pilot power allocation. The SE of each user also depends on the pilot assignment set $\mathcal{P}_{l,k}$. Therefore, as a contribution of this thesis, we would like to answer the question: how much can a multi-cell Massive MIMO system improve the SE by jointly optimizing the UL transmit powers and pilot sequence design? To answer this question, we compare a new optimized pilot design with the designs in related works. The channel estimates obtained with these schemes are used to compute (17). The pilot and data transmit powers are then optimized using a max-min fairness optimization problem. This work is presented in detail in Paper B.

1.3.3 Downlink Data Transmission

We now consider the DL transmission of a multi-cell Massive MIMO network, where the BSs are transmitting signals to their users as shown in Figure 4. For an arbitrary BS l , we let $\mathbf{x}_l \in \mathbb{C}^M$ denote the transmit signals intended for its K users. By applying linear precoding, this transmit signal vector is computed as

$$\mathbf{x}_l = \sum_{t=1}^K \sqrt{\rho_{l,t}} \mathbf{w}_{l,t} s_{l,t}, \quad (19)$$

where the intended payload symbol $s_{l,t}$ for user t in cell l has unit transmit power $\mathbb{E}\{|s_{l,t}|^2\} = 1$ and $\rho_{l,t}$ denotes the transmit power allocated to this particular user. Moreover, $\mathbf{w}_{l,t} \in \mathbb{C}^M$, for $t = 1, \dots, K$, are the corresponding

Figure 4: Downlink multi-cell Massive MIMO communication: The links from a BS to users in its coverage area are considered as the desired links, while the other links are the interfering links.

linear precoding vectors that determine the spatial directivity of the signal sent to each user. The received signal $y_{l,k} \in \mathbb{C}$ at user k in cell l is modeled as

$$y_{l,k} = \sum_{i=1}^L (\mathbf{h}_{l,k}^i)^H \mathbf{x}_i + n_{l,k}, \quad (20)$$

where $n_{l,k} \sim \mathcal{CN}(0, \sigma_{\text{DL}}^2)$ is the additive Gaussian noise. In the DL, a lower bound on the ergodic capacity of an arbitrary user k in cell l is

$$R_{l,k}^{\text{DL}} = \gamma^{\text{DL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left(1 + \text{SINR}_{l,k}^{\text{DL}} \right), \quad (21)$$

where the effective SINR value, $\text{SINR}_{l,k}^{\text{DL}}$, is computed as

$$\text{SINR}_{l,k}^{\text{DL}} = \frac{\rho_{l,k} |\mathbb{E}\{(\mathbf{h}_{l,k}^l)^H \mathbf{w}_{l,k}\}|^2}{\sum_{i=1}^L \sum_{t=1}^K \rho_{i,t} \mathbb{E}\{|\mathbf{h}_{l,k}^i|^2\} - \rho_{l,k} |\mathbb{E}\{(\mathbf{h}_{l,k}^l)^H \mathbf{w}_{l,k}\}|^2 + \sigma_{\text{DL}}^2}. \quad (22)$$

This bound follows from a standard capacity bounding technique from [8], where the users only have access to the channel statistics since there are no DL pilots. (22) is derived by assuming that every BS serves K users and every user is preassigned to one BS only. This may result in low SE for cell-edge users who are far away from the serving BS and contaminated strongly by mutual interference from neighbor cells. An alternative is that some users are

connected to multiple BSs, which jointly transmit to the user. To understand how much a multi-cell Massive MIMO system can gain in SE from using joint transmission, it is necessary to model, optimize, and compare different forms of joint transmission schemes. This study is provided in Paper A and the two different joint transmission schemes are briefly summarized in the next sections.

1.4 Coordinated Multipoint (CoMP) Transmission

In a classical cellular Massive MIMO system, each BS serves K users that are exclusively assigned to that BS, which creates disjoint cells as demonstrated in the previous sections. It may result in low SE for some users at cell edge due to the weak received signal from the home BS and strong interference from neighboring cells. CoMP is one potential method to deal with this issue. In CoMP, multiple BSs collaborate to serve a user. This method can increase the sum SE for the entire system as well as the cell-edge users [28]. CoMP can be roughly classified into three different categories [29]: *coordinated scheduling/ beamforming design* jointly designs the beamforming vectors and scheduling for a cluster/all users, *joint transmission* is interpreted as a simultaneously transmission of data signals to a user from multiple BSs, and *transmission point selection* selects the best BS to serve a user.

In this thesis, we focus on the joint transmission which is collaboration among BSs in the data transmission phase. This is the most advanced form of CoMP and therefore serves as an upper bound on the achievable performance. Figure 5 demonstrates that in the coordination area, there is no cell boundary and all users are potentially served by multiple BSs. If we design the CoMP system properly, the sum SE and per-user SE is higher than in conventional multi-user MIMO systems without CoMP. The main scope of this section is to outline the two main CoMP joint transmission schemes considered in this thesis: coherent and non-coherent joint transmission. In Paper A, we use them to formulate a total transmit power optimization problem for Massive MIMO system under limited power budgets together with the QoS requirements for all users.

1.4.1 Non-Coherent Joint Transmission

In non-coherent joint transmission, multiple BSs can send simultaneous signals to a user, but each data signal is independent from the other ones. This does not require phase-coherence between BSs, therefore it is called non-coherent transmission. However, it will require successive decoding at

Figure 5: Illustration of a Massive MIMO system utilizing CoMP with no cell boundary in the coordination area. All links from the BSs to a user can be considered to as desired links.

the user side.

We consider a network comprising L BSs each equipped with M antennas and able to serve K users. Since there are no cell boundaries in this network, the channel between user k , $k = 1, \dots, K$, and BS l , $l = 1, \dots, L$, is now denoted as $\mathbf{h}_{l,k} \in \mathbf{C}^M$. The transmitted signal at BS l is formulated as

$$\mathbf{x}_l = \sum_{t=1}^K \sqrt{\rho_{l,t}} \mathbf{w}_{l,t} s_{l,t}, \quad (23)$$

where $s_{l,t}$ is the independent data signal from BS l to user t and $\mathbb{E}\{|s_{l,t}|^2\} = 1$, while $\rho_{l,t}$ is the power that BS l allocates to the signal $s_{l,t}$. The corresponding precoding vector used for this user is denoted as $\mathbf{w}_{l,t}$. The received baseband signal at user k is formulated as

$$y_k = \sum_{l=1}^L \mathbf{h}_{l,k}^H \mathbf{x}_l + n_k, \quad (24)$$

where $n_k \sim \mathcal{CN}(0, \sigma_{\text{DL}}^2)$ denotes complex Gaussian noise. Plugging (23) into (24), we obtain

$$y_k = \underbrace{\sum_{l=1}^L \sqrt{\rho_{l,k}} \mathbf{h}_{l,k}^H \mathbf{w}_{l,k} s_{l,k}}_{\text{Desired signals}} + \underbrace{\sum_{l=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \sqrt{\rho_{l,t}} \mathbf{h}_{l,t}^H \mathbf{w}_{l,t} s_{l,t}}_{\text{Interference + Noise}} + n_k. \quad (25)$$

The above equation indicates that the desired signals for user k may come from all L BSs and contribute to increasing the achievable SE. The different desired signals from the L BSs are decoded by using successive interference cancellation [7]. In this subsection, we assume that all users have perfect CSI while the case in which the users have no CSI is considered in Paper A. In more detail, user k decodes the potentially desired signals in L stages as follows:

- *In the first stage*, user k will decode the transmitted signal from BS 1. The received signal in (25) is now reformulated as

$$y_{1,k} = y_k = \underbrace{\sqrt{\rho_{1,k}} \mathbf{h}_{1,k}^H \mathbf{w}_{1,k} s_{1,k}}_{\text{Desired signal}} + \underbrace{\sum_{i=2}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} s_{i,k}}_{\text{Interference}} + \underbrace{\sum_{i=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \sqrt{\rho_{i,t}} \mathbf{h}_{i,t}^H \mathbf{w}_{i,t} s_{i,t}}_{\text{Noise}} + n_k. \quad (26)$$

In the above equation, the first term denotes the desired signal from BS 1, while the second term involves mutual interference and noise. Using a capacity bounding technique in the case of perfect CSI [8], we obtain a lower bound on the ergodic capacity of user k and BS 1 as

$$R_{1,k} = \gamma^{\text{DL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_{1,k}^{\text{DL}} \right) \right\}, \quad (27)$$

where the SINR value, $\text{SINR}_{1,k}^{\text{DL}}$, is computed as

$$\text{SINR}_{1,k}^{\text{DL}} = \frac{\rho_{1,k} |\mathbf{h}_{1,k}^H \mathbf{w}_{1,k}|^2}{\sum_{i=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \rho_{i,t} |\mathbf{h}_{i,k}^H \mathbf{w}_{i,t}|^2 + \sum_{i=2}^L \rho_{i,k} |\mathbf{h}_{i,k}^H \mathbf{w}_{i,k}|^2 + \sigma_{\text{DL}}^2}. \quad (28)$$

- *In the second stage*, After decoding successfully the desired signals from BS 1, user k subtracts the decoded signal from BS 1, and then recover the transmitted signal from BS 2 by

$$y_{2,k} = y_k - \sqrt{\rho_{1,k}} \mathbf{h}_{1,k}^H \mathbf{w}_{1,k} s_{1,k} = \underbrace{\sqrt{\rho_{2,k}} \mathbf{h}_{2,k}^H \mathbf{w}_{2,k} s_{2,k}}_{\text{Desired signal}} + \underbrace{\sum_{i=3}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} s_{i,k}}_{\text{Interference}} + \underbrace{\sum_{i=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \sqrt{\rho_{i,t}} \mathbf{h}_{i,t}^H \mathbf{w}_{i,t} s_{i,t}}_{\text{Noise}} + n_k. \quad (29)$$

Using the same capacity bounding technique, the lower bound on the ergodic capacity of BS 2 and user k is

$$R_{2,k} = \gamma^{\text{DL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_{2,k}^{\text{DL}} \right) \right\}, \quad (30)$$

where the SINR value is formulated as

$$\text{SINR}_{2,k}^{\text{DL}} = \frac{\rho_{2,k} |\mathbf{h}_{2,k}^H \mathbf{w}_{2,k}|^2}{\sum_{\substack{L \\ i=1}} \sum_{\substack{K \\ t=1}} \rho_{i,t} |\mathbf{h}_{i,t}^H \mathbf{w}_{i,t}|^2 + \sum_{\substack{L \\ i=3}} \rho_{i,k} |\mathbf{h}_{i,k}^H \mathbf{w}_{i,k}|^2 + \sigma_{\text{DL}}^2}. \quad (31)$$

- *In the l th stage*, by processing in the same way, user k recovers the transmitted signal from BS l by subtracting the first $l-1$ recovered signals as

$$\begin{aligned} y_{l,k} &= y_k - \sum_{i=1}^{l-1} \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} s_{i,k} \\ &= \underbrace{\sqrt{\rho_{l,k}} \mathbf{h}_{l,k}^H \mathbf{w}_{l,k} s_{l,k}}_{\text{Desired signal}} + \underbrace{\sum_{i=l+1}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} s_{i,k} + \sum_{\substack{L \\ i=1}} \sum_{\substack{K \\ t=1}} \sqrt{\rho_{i,t}} \mathbf{h}_{i,t}^H \mathbf{w}_{i,t} s_{i,t} + n_k}_{\text{Interference + Noise}}. \end{aligned} \quad (32)$$

The lower bound on the ergodic capacity between user k and BS l is computed as

$$R_{l,k} = \gamma^{\text{DL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_{l,k}^{\text{DL}} \right) \right\}, \quad (33)$$

where the SINR value is formulated as

$$\text{SINR}_{l,k}^{\text{DL}} = \frac{\rho_{l,k} |\mathbf{h}_{l,k}^H \mathbf{w}_{l,k}|^2}{\sum_{\substack{L \\ i=1}} \sum_{\substack{K \\ t=1}} \rho_{i,t} |\mathbf{h}_{i,t}^H \mathbf{w}_{i,t}|^2 + \sum_{\substack{L \\ i=l+1}} \rho_{i,k} |\mathbf{h}_{i,k}^H \mathbf{w}_{i,k}|^2 + \sigma_{\text{DL}}^2}. \quad (34)$$

- *Finally*, after successfully decoding the desired signals from all BSs, user k obtains the total SE which is summation of the SE from all L BSs:

$$R_k = \sum_{i=1}^L R_{i,k} = \gamma^{\text{DL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_k^{\text{DL}} \right) \right\} [\text{b/s/Hz}], \quad (35)$$

where the SINR value, denoted by $\text{SINR}_k^{\text{DL}}$, is

$$\text{SINR}_k^{\text{DL}} = \frac{\sum_{i=1}^L \rho_{i,k} |\mathbf{h}_{i,k}^H \mathbf{w}_{i,k}|^2}{\sum_{i=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \rho_{i,t} |\mathbf{h}_{i,t}^H \mathbf{w}_{i,t}|^2 + \sigma_{\text{DL}}^2}. \quad (36)$$

From the SE expression in (35), we observe that non-coherent joint transmission leads an effective SINR expression where the numerator of (36) is a superposition of the desired signals from all BSs. However, the mutual interference term in the denominator also contains more terms than in a conventional Massive MIMO system. Hence, we need to carefully optimize the SEs to see clear gains from non-coherent transmission. In Paper A, we provide a detailed theoretical analysis and simulation results for this CoMP technique, using estimated channels instead of perfect CSI.

1.4.2 Coherent Joint Transmission

With coherent joint transmission, it is assumed that all BSs transmit the same signals to a user, and therefore the received signal at user k is now formulated as

$$y_k = \sum_{i=1}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} s_k + \sum_{i=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \sqrt{\rho_{i,t}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,t} s_t + n_k. \quad (37)$$

The main disadvantages of coherent joint transmission is the stricter synchronization requirement among all BSs in the coordination area to transmit the same signals coherently. In addition, the same data signals need to be conveyed to multiple BSs, which increases the backhaul signaling. However, at the receiver side, users decode the desired signals as usual and therefore the computational decoding complexity is reduced compared to non-coherent joint transmission. Similar to the previous section, we assume the users have perfect channel knowledge and therefore the decoding process is formulated as

$$y_k = \underbrace{\sum_{i=1}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} s_k}_{\text{Desired signal}} + \underbrace{\sum_{i=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \sqrt{\rho_{i,t}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,t} s_t + n_k}_{\text{Interference + Noise}}. \quad (38)$$

Using the same capacity bounding technique as before, a lower bound on the ergodic capacity of user k is

$$R_k = \gamma^{\text{DL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 (1 + \text{SINR}_k^{\text{DL}}) \right\}, \quad (39)$$

where the effective SINR value is

$$\text{SINR}_k^{\text{DL}} = \frac{\left| \sum_{i=1}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} \right|^2}{\sum_{\substack{i=1 \\ i \neq k}}^K \left| \sum_{t=1}^L \sqrt{\rho_{i,t}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,t} \right|^2 + \sigma_{\text{DL}}^2}. \quad (40)$$

The SE of user k in the case of BSs using coherent joint transmission is expected to be better than both non-coherent joint transmission and classical cellular networks. Paper A gives detailed numerical results and derive closed-form lower bounds on the ergodic capacity for this CoMP technique but for the practical case where no CSI is available at the users.

1.4.3 Transmit Power Consumption at Base Stations

Each Massive MIMO BS may be equipped with hundreds of antennas and simultaneously serve multiple users. Therefore a model of the power consumption at the BSs in Massive MIMO networks is necessary. As for classical MIMO BSs, the total power consumption at each Massive MIMO BS includes a static part, which is determined by the hardware technology, and a dynamic part, which is a function of the transmitted signals. For BS l , it can be expressed as [30]

$$P_l = \begin{cases} P_{\text{active},l} + \Delta_l P_{\text{trans},l} & \text{if } P_{\text{trans},l} \neq 0, \\ P_{\text{sleep},l} & \text{if } P_{\text{trans},l} = 0, \end{cases} \quad (41)$$

where $P_{\text{sleep},l}$ is the sleep mode power consumption at BS l . The scaling factor $\Delta_l \geq 1$ denotes the amplifier inefficiency factor of the power amplifier. The transmit power $P_{\text{trans},l}$ at BS l is obtained for the case of non-coherent joint transmission as

$$P_{\text{trans},l} = \sum_{k=1}^K \|\mathbf{w}_{l,k}\|^2 \mathbb{E}\{|s_{l,k}|^2\} = \sum_{k=1}^K \rho_{l,k} \|\mathbf{w}_{l,k}\|^2, \quad (42)$$

and in the case of coherent joint transmission the transmit power is

$$P_{\text{trans},l} = \sum_{k=1}^K \|\mathbf{w}_{l,k}\|^2 \mathbb{E}\{|s_k|^2\} = \sum_{k=1}^K \rho_{l,k} \|\mathbf{w}_{l,k}\|^2. \quad (43)$$

From (42) and (43), the same expression of transmit power is observed for coherent and non-coherent joint transmission. Furthermore, these equations apply for an arbitrary precoding scheme. The important goal of transmit power control is to optimize the transmit power for every active user. In contrast, the sleep mode power depends on the hardware technology and needs to be optimized when the circuits are designed and manufactured. In this thesis, we consider a transmit power optimization problem of the form

$$\begin{aligned} & \underset{\{\rho_{l,k} \geq 0\}}{\text{minimize}} && \sum_{i=1}^L P_{\text{trans},i} \\ & \text{subject to} && R_k \geq \xi_k, \forall k, \\ & && P_{\text{trans},l} \leq P_{\text{max},l}, \end{aligned} \tag{44}$$

where ξ_k is the required SE of user k which is a fixed parameter measured in b/s/Hz. $P_{\text{max},l}$ is the maximum transmit power that BS l can supply. This optimization problem minimizes the transmit power of all BSs with required SE for each user and a limited power budget at every BS. We will use (44) to investigate the joint power allocation and user association problems in Massive MIMO as shown in Paper A.

1.5 Optimization Preliminaries

This section presents preliminaries of optimization theory, including some basic optimization classes and properties. In optimization theory, an optimization problem on standard form is formulated as

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && f_0(\mathbf{x}) \\ & \text{subject to} && f_i(\mathbf{x}) \leq b_i, \quad i = 1, \dots, m, \end{aligned} \tag{45}$$

where the vector $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$ denotes the optimization variable which originates from a domain $\mathcal{X} \subseteq \mathbb{R}^n$. The function $f_0(\mathbf{x})$ is the objective function, while the functions $f_i(\mathbf{x}), \forall i = 1, \dots, m$, are the inequality constraint functions. The constants $b_i \in \mathbb{R}, \forall i = 1, \dots, m$, are the bounds of the inequality constraints. If $\mathbf{x}$ makes a constraint function satisfied, then it is called a feasible point of that constraint. The feasible domain for a constraint is the set of all feasible points. The intersection of all the feasible domains is defined as the feasible region of the optimization problem.

A locally optimal solution $\mathbf{x}_0$ produces the smallest objective function $f_0(\mathbf{x})$ of the problem (45) among the $\mathbf{x} \in \mathcal{X}$ in the vicinity of $\mathbf{x}_0$, but this condition may not lead to the smallest objective function when considering

the entire feasible domain. In contrast, the globally optimal solution yields the smallest objective function among all the feasible points.

1.5.1 Convex Optimization Problems

We now focus on convex optimization problems for which one can show that every local optimal solution is also a global optimal solution, which makes these problems relatively easy to solve.

We first introduce the definition of a convex set. In particular, $\mathcal{X}$ is a convex set if for any $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathcal{X}$ and $a_1, \dots, a_m$ with $a_1 + \dots + a_m = 1$, we have

$$a_1 \mathbf{x}_1 + \dots + a_m \mathbf{x}_m \in \mathcal{X}. \quad (46)$$

We introduce the definition of convex functions: For all $\mathbf{x}, \tilde{\mathbf{x}} \in \mathcal{X}$ and $\alpha_1, \alpha_2 \in \mathbb{R}_+$ with $\alpha_1 + \alpha_2 = 1$ and $a_1 \mathbf{x} + a_2 \tilde{\mathbf{x}} \in \mathcal{X}$, the functions $f_i, \forall i = 0, \dots, m$, satisfy

$$f_i(\alpha_1 \mathbf{x} + \alpha_2 \tilde{\mathbf{x}}) \leq \alpha_1 f_i(\mathbf{x}) + \alpha_2 f_i(\tilde{\mathbf{x}}), \quad (47)$$

then (45) is a convex optimization problem. A vector $\mathbf{x}^*$ is the optimal solution to (45) if it yields the smallest objective value among all feasible values $\mathbf{x} \in \mathcal{X}$ that satisfies all the constraints. There are several important properties of convex optimization problems:

- Since the feasible domains of the objective and constraint functions are convex sets, the feasible region of the optimization problem is also a convex set. It ensures that an infeasible solution is not generated when solving the optimization problem.
- The convex objective function of a convex problem guarantees that all local optimums are also the global optimum. Therefore, if we can find a local solution, using any search algorithm, then this local optimum is the global optimum.

In general, solving a convex optimization problem requires a computational complexity of the order of $\mathcal{O}(\max\{n^3, n^2 m, F\})$, where F is the cost of evaluating the first and second derivative of the objective and constraint functions [31]. However, the exact computational complexity depends on the methods involved to solve the optimization problems.

An optimization problem that does not satisfy (47) is non-convex. Attaining the global optimum for a non-convex problems generally require algorithms that explicitly searches for the global optimum. In many cases, these algorithms have exponential computational complexity, thus a local optimum is usually preferred when dealing with such problems in practice [32].

1.5.2 Linear Programming

Linear optimization is an important special case of convex optimization problems that is formulated as

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i = 1, \dots, m, \end{aligned} \quad (48)$$

where $\mathbf{c}$ and $\mathbf{a}_i, \forall i, \in \mathbb{R}^n$ are vectors. In comparison to (45), the objective function $f_0(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$ is now a linear function of $\mathbf{x}$. Mapping to (45), the i th constraint function is formulated as $f_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{x}$ and it is a linear function of variable $\mathbf{x}$. There is no simple analytical formula for the solution to a linear program in general [31], but it is a convex problem. Consequently, the globally optimal solution can be obtained in polynomial time by using a general purpose optimization toolboxes, as CVX [33]. Linear programming has quite low computational complexity since there is no need to compute the second derivative of the objective and constraint functions. In general, the computational complexity is, for instance, of the order of $\mathcal{O}(n^2 m)$ if $m \geq n$ [31].

In Paper A, we will prove that the total transmit power minimization problem in the case of non-coherent joint transmission with Rayleigh fading and MRT or ZF precoding belong to the linear programming class.

1.5.3 Second-Order Cone Programming

We now consider another popular optimization class called second order cone programs (SOCP), which is also a special case of convex programs. The standard form is defined as

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T \mathbf{x} + d_i, \quad i = 1, \dots, m, \end{aligned} \quad (49)$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A}_i \in \mathbb{R}^{n_i \times n}$, $\mathbf{c}_i \in \mathbb{R}^{n_i}$, and $d_i \in \mathbb{R}$ are constant parameters. The constraints in (49) are called second-order cone constraints. A SOCP is convex and, therefore, the globally optimal solution is obtained in polynomial time by using a general purpose optimization toolbox such as CVX [33]. The SOCP problems have higher complexity than linear programs since they require to evaluate the second derivative of the constraints [34]. The total transmit power minimization problem in the case of coherent joint transmission is a SOCP as demonstrated in Paper A.

1.5.4 Geometric Programming

We now study geometric optimization problems, which are of the form

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && \sum_{m=1}^M c_{m,0} \prod_{n=1}^N x_n^{a_{n,m,0}} \\ & \text{subject to} && \sum_{m=1}^M c_{m,i} \prod_{n=1}^N x_n^{a_{n,m,i}} \leq 1, i = 1, \dots, m, \end{aligned} \quad (50)$$

where all coefficients $c_{m,i}, i = 0, \dots, m$, are nonnegative and the exponents $a_{n,m,i}$ are real numbers. Geometric programs may be convex in some particular scenarios, but they are generally non-convex. However, by exploiting a hidden convex structure, geometric problems can be converted to convex problems. Let us make the change of variable $x_n = e^{y_n}, \forall n$, and then taking the natural logarithm of the objective and constraint functions, the optimization problem (50) becomes

$$\begin{aligned} & \underset{\{y_n\}}{\text{minimize}} && \ln \left(\sum_{m=1}^M c_{m,0} e^{\sum_{n=1}^N y_n a_{n,m,0}} \right) \\ & \text{subject to} && \ln \left(\sum_{m=1}^M c_{m,i} e^{\sum_{n=1}^N y_n a_{n,m,i}} \right) \leq 0, i = 1, \dots, m. \end{aligned} \quad (51)$$

Since the weighted log-sum-exponentials functions are convex, (51) is a convex problem. Therefore, we can obtain the globally optimal solution to (51) in tractable time by using interior-point methods. The computational cost is higher than the linear or SOCP problems since the cost of evaluating the first and second derivatives of the objective and constraint functions is complicated in many applications [31]. This optimization class will be utilized in Paper B when we work with joint pilot design and uplink power control for multi-cell Massive MIMO.

1.5.5 Signomial Programming

A signomial program has the same structure as that of a geometric program in (50), but at least one of the coefficients $c_{m,i}$ has a negative value. We note that a signomial program is non-convex, so finding the globally optimal solution is attained with the extremely high computational complexity [35]. Nonetheless, we may find a local solution by a successive approximation approach if the signomial optimization problem is bounded by a convex problem with the approximated convex constraints. In general, if $f_i(\mathbf{x}) =$

$\sum_{m=1}^M c_{m,i} \prod_{n=1}^N x_n^{a_{n,m,i}}$ is a signomial function, we can upper bound it by a convex function $\hat{f}_i(\mathbf{x})$, i.e., $f_i(\mathbf{x}) \leq \hat{f}_i(\mathbf{x})$. The solution of the successive approximation approach converges to a stationary point of the original signomial problem if at the n th iteration the following conditions are satisfied [36]:

1. $f_i(\mathbf{x}^{(n)}) \leq \hat{f}_i(\mathbf{x}^{(n)}), \forall \mathbf{x}^{(n)} \in \mathcal{X}$.
2. $f_i(\mathbf{x}^{*,(n-1)}) = \hat{f}_i(\mathbf{x}^{*,(n-1)})$, where $\mathbf{x}^{*,(n-1)}$ is the optimal solution of the approximated optimization in the $(n-1)$ th iteration.
3. $\nabla f_i(\mathbf{x}^{*,(n-1)}) = \nabla \hat{f}_i(\mathbf{x}^{*,(n-1)})$, where ∇ is the first-order derivative operator.

The first condition ensures that the globally optimal solution to the approximated optimization problem is also feasible to the original signomial problem. The second condition guarantees that the solution of each iteration decreases the objective function monotonically. Finally, the third condition makes sure that the Karush-Kuhn-Tucker (KKT) conditions of the original signomial problem and the approximated problem coincide after a number of iterations. The main steps to find that local optimum are summarized as follows:

1. Set up the initial values of the optimization variables and then compute the required parameters of the approximated functions.
2. Solve the approximated convex problem to obtain the optimal solution with the given required parameters.
3. Update the required parameter of the approximated functions from the optimal solution obtained in Step 2.
4. Repeat Steps 2 and 3 until the algorithm converges.

The computational complexity for finding a local optimum to the signomial optimization problem is directly proportional to the computational complexity of the convex problem that is solved in each iteration. In Paper B, we first observe that our max-min fairness optimization with the proposed pilot structure is a signomial program, then we apply the above four steps to obtain a local optimum. Furthermore, the special properties of the approximated functions which are utilized in Paper B allow us to analytically prove the convergence of the proposed successive approximation approach.

1.5.6 Weighted Max-Min Fairness Optimization Problem

One target of Massive MIMO systems is to provide a uniformly good service for all users in the network. This can be achieved by a weighted max-min fairness problem. In other words, we will maximize the lowest SE over all the users, possibly with some user specific weighting. Mathematically, a max-min fairness optimization problem can be formulated as

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \min_{i \in \{1, \dots, m\}} \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})w_i}, \quad (52)$$

where $w_i > 0$ is the weight value of the function $f_i(\mathbf{x})/g_i(\mathbf{x})$. We solve the problem (52) by converting it to the epi-graph representation as

$$\begin{aligned} & \underset{\xi, \mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \xi \\ & \text{subject to} \quad f_i(\mathbf{x}) - g_i(\mathbf{x})w_i\xi \geq 0, \quad i = 1, \dots, m. \end{aligned} \quad (53)$$

We stress that it is possible to obtain the globally optimal solution to (52) if all functions $f_i(\mathbf{x}) - g_i(\mathbf{x})w_i\xi, \forall i$, are concave. In this sense, the global optimum to (52) can be obtained by using a general-purpose toolbox as CVX [33]. Alternatively, if (53) is a convex problem for given ξ , i.e.,

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad 0 \\ & \text{subject to} \quad f_i(\mathbf{x}) - g_i(\mathbf{x})w_i\xi \geq 0, \quad i = 1, \dots, m, \end{aligned} \quad (54)$$

is convex, then the optimal solution to (52) is obtained by using the bisection search over possible values of ξ . The detail of optimizing (53) via utilizing bisection search is given in Algorithm 1.

In Paper A, we investigate the weighted max-min fairness optimization for the CoMP frameworks where multiple BSs can collaborate to serve all users. Meanwhile, the application of the weighted max-min fairness optimization to joint pilot design and UL power optimization is studied in Paper B.

Algorithm 1 Weighted max-min fairness optimization with bisection

Result: Solve optimization in (52).

Input: Initial upper bound ξ_0^{upper} , initiate $\mathbf{x}^{\text{lower}}$ and line-search accuracy δ ;

Set $\xi^{\text{lower}} = 0$; $\xi^{\text{upper}} = \xi_0^{\text{upper}}$;

while $\xi^{\text{upper}} - \xi^{\text{lower}} > \delta$ **do**

Set $\xi^{\text{candidate}} = (\xi^{\text{upper}} + \xi^{\text{lower}})/2$;

if (54) is infeasible for $\xi^{\text{candidate}}$, **then**

Set $\xi^{\text{upper}} = \xi^{\text{candidate}}$;

else

Set $\{\mathbf{x}^{\text{lower}}\}$ as the solution to (54);

Set $\xi^{\text{lower}} = \xi^{\text{candidate}}$;

end if

end while

Set $\xi_{\text{final}}^{\text{lower}} = \xi^{\text{lower}}$ and $\xi_{\text{final}}^{\text{upper}} = \xi^{\text{upper}}$;

Output: Final interval $[\xi_{\text{final}}^{\text{lower}}, \xi_{\text{final}}^{\text{upper}}]$ and the optimal solution $\mathbf{x}^* = \mathbf{x}^{\text{lower}}$;

Chapter 2

Contributions of the Thesis

This thesis focuses on two important aspects of resource allocation which improves the performance of multi-cell Massive MIMO systems. The first aspect is to jointly optimize the power allocation and user association for DL Massive MIMO systems. We study this with either coherent or non-coherent joint transmission. This work is presented in Paper A. The second aspect is the pilot design and UL power control. This is investigated in Paper B for systems with either ideal hardware or hardware impairments. In this section, we provide the publication information for these papers and further list other publications that are not included since they are preliminary versions of the included papers or not within the main scope of this thesis.

2.1 Papers Included in the Thesis

Paper A: Joint Power Allocation and User Association Optimization for Massive MIMO Systems

Authored by: Trinh Van Chien, Emil Björnson, and Erik G. Larsson

Published in: Transactions on Wireless Communications, volume 15, issue 9, pp. 6384 – 6399, September 2016.

Abstract: This work investigates the joint power allocation and user association problem in multi-cell Massive MIMO DL systems. The target is to minimize the total transmit power consumption when each user is served by an optimized subset of the BSs, using either non-coherent joint transmission or coherent joint transmission. We first derive a lower bound on the ergodic spectral efficiency (SE), which is applicable for any channel distribution and precoding scheme. Closed-form expressions are obtained for Rayleigh fading

channels with either MRT or ZF precoding. From these bounds, we further formulate the DL power minimization problems with fixed SE constraints for the users. These problems are proved to be solvable as convex optimization problems, i.e., linear programs for non-coherent joint transmission or as second order cone programs for coherent joint transmission. Hence, the optimization problems give the optimal power allocation and BS user association with low complexity. Furthermore, we formulate and solve a max-min fairness problem which maximizes the worst SE among the users, and we show that it can be efficiently solved to obtain the optimal solutions. Simulations manifest that the proposed methods provide good SE for the users using less transmit power than in small-scale systems and the optimal user association can effectively balance the load between BSs when needed. Even though our framework allows the joint transmission from multiple BSs, there is an overwhelming probability that only one BS is associated with each user at the optimal solution.

Paper B: Joint Pilot Design and Uplink Power Allocation in Multi-Cell Massive MIMO Systems

Authored by: Trinh Van Chien, Emil Björnson, and Erik G. Larsson

Submitted to: IEEE Transactions on Wireless Communications, March 2017.
A minor revision of the paper is currently under review.

Abstract: This work considers pilot design to mitigate pilot contamination and providing good service for everyone in multi-cell Massive MIMO systems. Different from prior works which model the pilot design as a combinatorial assignment problem, we treat the pilot signals as continuous optimization variables. We compute a lower bound on the UL capacity for Rayleigh fading channels with MR detection that can be applied with arbitrary pilot signals. We further formulate the max-min fairness problem under power budget constraints, with the pilot signals and data powers as optimization variables. Although this optimization problem is NP-hard due to signomial constraints, we demonstrate how to obtain the globally optimal solution. Hence we propose an efficient algorithm to obtain a local optimum with polynomial complexity. Our framework serves as a benchmark for pilot design in scenarios with either ideal or non-ideal hardware. Numerical results manifest that the proposed optimization algorithms are nearly optimal and the new pilot structure and optimization bring large gains over the state-of-the-art suboptimal pilot designs.

2.2 Papers Not Included in the Thesis

Paper C: Downlink Power Control for Massive MIMO Cellular Systems with Optimal User Association

Authored by: Trinh Van Chien, Emil Björnson, and Erik G. Larsson

Published in: IEEE International Conference on Communications (ICC), May 2016.

Summary: This paper contains selected parts of Paper A. We only consider the case of non-coherent joint transmission and assume that there are orthogonal pilot signals for all users. Because of space constraints, all proof details of lemmas and theorems were omitted from this paper.

Paper D: Multi-Cell Massive MIMO Performance with Double Scattering Channels

Authored by: Trinh Van Chien, Emil Björnson, and Erik G. Larsson

Published in: IEEE International Workshop on Computer-Aided Modeling Analysis and Design of Communication Links and Networks (CAMAD), October 2016.

Abstract: This paper investigates the SE of multi-cell MIMO using different channel models. Prior works have derived closed-form SE bounds and approximations for Gaussian distributed channels, while we consider the *double scattering model*—a prime example of a non-Gaussian channel for which it is intractable to obtain closed form SE expressions. The channels are estimated using limited resources, which gives rise to pilot contamination, and the estimates are used for linear detection and to compute the SE numerically. Analytical and numerical examples are used to describe the key behaviors of the double scattering models, which differ from conventional Massive MIMO models. Finally, we provide multi-cell simulation results that compare the double scattering model with uncorrelated Rayleigh fading and explain under what conditions we can expect to achieve similar SEs.

Paper E: Massive MIMO Communications

Authored by: Trinh Van Chien and Emil Björnson

Published in: in 5G Mobile Communications, W. Xiang et al. (eds.), pp. 77-116, Springer, January 2017.

Abstract: Every new network generation needs to make a leap in area data throughput, to manage the growing wireless data traffic. The Massive MIMO technology can bring at least ten-fold improvements in area throughput by increasing the spectral efficiency (bit/s/Hz/cell), while using the same bandwidth and density of base stations as in current networks. These extraordinary gains are achieved by equipping the base stations with arrays of a hundred antennas to enable spatial multiplexing of tens of user terminals. This chapter overviews and explains the basic motivations and communication theory behind the Massive MIMO technology, and provides implementation-related design guidelines.

Paper F: Joint Pilot Sequence Design and Power Control for Max-Min Fairness in Uplink Massive MIMO

Authored by: Trinh Van Chien, Emil Björnson, and Erik G. Larsson

Published in: IEEE International Conference on Communications (ICC), May 2017.

Summary: This paper includes selected parts of Paper B. In this paper, we only optimize the pilot assignment and pilot transmit powers to mitigate pilot contamination in Massive MIMO systems with the fixed transmit power for the data. The analysis and simulation results are only demonstrated for uncorrelated Rayleigh fading channels and ideal hardware. Moreover, the detailed proofs of lemmas and theorems were omitted due to space limitations.

Bibliography

- [1] Q. Li, G. Li, W. Lee, M.-I. Lee, D. Mazzaresse, B. Clerckx, and Z. Li, “MIMO techniques in WiMAX and LTE: A feature overview,” *IEEE Commun. Mag.*, vol. 48, pp. 86–92, 2010.
- [2] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, “What will 5G be?” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [3] Cisco, “Visual networking index: Global mobile data traffic forecast update, 2016-2021,” Tech. Rep., Feb. 2017.
- [4] T. E. Bogale and L. B. Le, “Massive MIMO and mmwave for 5G wireless HetNet: Potential benefits and challenges,” *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, 2016.
- [5] E. Björnson, E. G. Larsson, and T. L. Marzetta, “Massive MIMO: Ten myths and one critical question,” *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114 – 123, 2016.
- [6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, 2013.
- [7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [8] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [9] J. Flordelis, F. Rusek, F. Tufvesson, E. G. Larsson, and O. Edfors, “Massive MIMO performance-TDD versus FDD: What do measurements say?” *arXiv preprint arXiv:1704.00623*, 2017.

- [10] E. Björnson, J. Hoydis, and L. Sanguinetti, “Massive MIMO networks: Spectral, energy, and hardware efficiency,” *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154 – 655, 2017.
- [11] E. Björnson, E. G. Larsson, and M. Debbah, “Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?” *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293–1308, 2016.
- [12] L. Sanguinetti and H. Poor, “Fundamentals of multi-user MIMO communications,” in *New Directions in Wireless Communications Research*, V. Tarokh, Ed. Springer US, 2009, pp. 139–173.
- [13] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Aspects of favorable propagation in massive MIMO,” in *Proc. IEEE EUSIPCO*, 2014.
- [14] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, “Massive MIMO performance evaluation based on measured propagation data,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, 2015.
- [15] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.
- [16] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.
- [17] F. Rusek, D. Persson, B. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, “Scaling up MIMO: Opportunities and challenges with very large arrays,” *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, 2013.
- [18] G. Fodor, N. Rajatheva, W. Zirwas, L. Thiele, M. Kurras, K. Guo, A. Tolli, J. H. Sorensen, and E. de Carvalho, “An overview of Massive MIMO technology components in METIS,” *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 155–161, 2017.
- [19] H. H. Yang, G. Geraci, T. Q. Quek, and J. G. Andrews, “Cell-edge-aware precoding for downlink Massive MIMO cellular networks,” *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3344–3358, 2017.
- [20] E. Björnson and E. Jorswieck, “Optimal resource allocation in coordinated multi-cell systems,” *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, 2013.

-
- [21] S. Jin, M. Li, Y. Huang, Y. Du, and X. Gao, "Pilot scheduling schemes for multi-cell massive multiple-input-multiple-output transmission," *IET Communications*, vol. 9, no. 5, pp. 689–700, 2015.
- [22] V. Saxena, G. Fodor, and E. Karipidis, "Mitigating pilot contamination by pilot reuse and power control schemes for Massive MIMO systems," in *VTC Spring*, 2015, pp. 1–6.
- [23] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [24] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Commun.*, vol. 10, no. 8, pp. 2640–2651, 2011.
- [25] H. V. Cheng, E. Björnson, and E. G. Larsson, "Optimal pilot and payload power control in single-cell Massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2363 – 2378, 2017.
- [26] X. Zhu, Z. Wang, L. Dai, and C. Qian, "Smart pilot assignment for Massive MIMO," *IEEE Commun. Letters*, vol. 19, no. 9, pp. 1644 – 1647, 2015.
- [27] K. Guo, Y. Guo, G. Fodor, and G. Ascheid, "Uplink power control with MMSE receiver in multi-cell MU-Massive-MIMO systems," in *Proc. IEEE ICC*, 2014, pp. 5184–5190.
- [28] M. Boldi, A. Tölö, M. Olsson, E. Hardouin, T. Svensson, F. Boccardi, L. Thiele, and V. Jungnickel, "Coordinated multipoint (CoMP) systems," in *Mobile and Wireless Communications for IMT-Advanced and Beyond*, A. Osseiran, J. Monserrat, and W. Mohr, Eds. Wiley, 2011, pp. 121–155.
- [29] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: Deployment scenarios and operational challenges," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 44–50, 2012.
- [30] G. Auer, O. Blume, V. Giannini, I. Godor, M. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, and W. Wajda, *D2.3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown*. INFOS-ICT-247733 EARTH, ver. 2.0, 2012. [Online]. Available: <http://www.ict-earth.eu/>

- [31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [32] Y. Zhao, T. Larsson, D. Yuan, E. Rönnerberg, and L. Lei, “Power efficient uplink scheduling in SC-FDMA: benchmarking by column generation,” *Optimization and Engineering*, vol. 17, no. 4, pp. 695–725, 2016.
- [33] CVX Research Inc., “CVX: Matlab software for disciplined convex programming, academic users,” <http://cvxr.com/cvx>, 2015.
- [34] F. Alizadeh and D. Goldfarb, “Second-order cone programming,” *Mathematical programming*, vol. 95, no. 1, pp. 3–51, 2003.
- [35] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, “MM algorithms for geometric and signomial programming,” *Mathematical Programming*, vol. 143, no. 1, pp. 339 – 356, 2014.
- [36] M. Chiang, C. W. Tan, D. P. Palomar, D. O. Neill, and D. Julian, “Power control by geometric programming,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, 2007.

Included Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-144221>

Other Recently Published Theses From
The Division of Communication Systems
Department of Electrical Engineering (ISY)
Linköping University, Sweden

Hei Victor Cheng, *Aspects of Power Allocation in Massive MIMO*, Linköping Studies in Science and Technology. Licentiate Thesis, No. 1767, 2016.

Marcus Karlsson, *Aspects of Massive MIMO*, Linköping Studies in Science and Technology. Licentiate Thesis, No. 1764, 2016.

Christopher Mollén, *On Massive MIMO Base Stations with Low-End Hardware*, Linköping Studies in Science and Technology. Licentiate Thesis, No. 1756, 2016.

Antonios Pitarokoilis, *Phase Noise and Wideband Transmission in Massive MIMO*, Linköping Studies in Science and Technology. Dissertations, No. 1756, 2016.

Anu Kalidas M. Pillai, *Signal Reconstruction Algorithms for Time-Interleaved ADCs*, Linköping Studies in Science and Technology. Dissertations, No. 1672, 2015.

Ngô Quốc Hiên, *Massive MIMO: Fundamentals and System Designs*, Linköping Studies in Science and Technology. Dissertations, No. 1642, 2015.

Mirsad Ćirkić, *Efficient MIMO Detection Methods*, Linköping Studies in Science and Technology. Dissertations, No. 1570, 2014.

Reza Moosavi, *Improving the Efficiency of Control Signaling in Wireless Multiple Access Systems*, Linköping Studies in Science and Technology. Dissertations, No. 1556, 2014.

Johannes Lindblom, *The MISO Interference Channel as a Model for Non-Orthogonal Spectrum Sharing*, Linköping Studies in Science and Technology. Dissertations, No. 1555, 2014.

Antonios Pitarokoilis, *On the Performance of Massive MIMO Systems with Single Carrier Transmission and Phase Noise*, Linköping Studies in Science and Technology. Licentiate Thesis, No. 1618, 2013.

Tumula V. K. Chaitanya, *HARQ Systems: Resource Allocation, Feedback Error Protection, and Bits-to-Symbol Mappings*, Linköping Studies in Science and Technology. Dissertations, No. 1555, 2013.