Linköping Studies in Science and Technology
Dissertation No. 2036

Spatial Resource Allocation in Massive MIMO Communication: *From Cellular to Cell-Free*

Trinh Van Chien

Linköping Studies in Science and Technology
Dissertations No. 2036

Spatial Resource Allocation in Massive MIMO Communications: *From Cellular to Cell-Free*

Trinh Van Chien

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

Linköping 2020

This is a Swedish Doctor of Philosophy thesis.
The Doctor of Philosophy degree comprises 240 ECTS credits of postgraduate studies.

**Spatial Resource Allocation in Massive MIMO Communications:
*From Cellular to Cell-Free***

© 2020 Trinh Van Chien, unless otherwise noted.

ISBN 978-91-7929-941-5

ISSN 0345-7524

Printed in Sweden by LiU-Tryck, Linköping 2020

Abstract

Massive MIMO (multiple-input multiple-output) is considered as an heir of the multi-user MIMO technology and it has gained lots of attention from both academia and industry since the last decade. By equipping base stations (BSs) with hundreds of antennas in a compact array or a distributed manner, 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 the five different resource allocation aspects in Massive MIMO communications: 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.

The third part of this thesis studies a two-layer decoding method that mitigates inter-cell interference in multi-cell Massive MIMO systems. In layer one, each BS estimates the channels to intra-cell users and uses the estimates for local decoding within the cell. This is followed by a second decoding layer where the BSs cooperate to mitigate inter-cell interference. An UL achievable SE expression is computed for arbitrary two-layer decoding schemes, while a closed-form expression is obtained for correlated Rayleigh fading channels, maximum-ratio combining (MRC), and large-scale fading decoding (LSFD) in the second layer. We formulate a sum SE maximization problem with both the data power and LSFD vectors as optimization variables. Since the problem is non-convex, we develop an algorithm based on the weighted minimum mean square error (MMSE) approach to obtain a stationary point with low computational complexity.

Motivated by recent successes of deep learning in predicting the solution to an optimization problem with low runtime, the fourth part of this thesis investigates the use of deep learning for power control optimization in Massive MIMO. We formulate the joint data and pilot power optimization for maximum sum SE in multi-cell Massive MIMO systems, which is a non-convex problem. We propose a new optimization algorithm, inspired by the weighted MMSE approach, to obtain a stationary point in polynomial time. We then use this algorithm together with deep learning to train a convolutional neural network to perform the joint data and pilot power control in sub-millisecond runtime. The solution is suitable for online optimization.

Finally, the fifth part of this thesis considers a large-scale distributed antenna system that serves the users by coherent joint transmission called Cell-free Massive MIMO. For a given user set, only a subset of the access points (APs) is likely needed to satisfy the users' performance demands. To find a flexible and energy-efficient implementation, we minimize the total power consumption at the APs in the DL, considering both the hardware-consumed and transmit powers, where APs can be turned off to reduce the former part. Even though this is a non-convex optimization problem, a globally optimal solution is obtained by solving a mixed-integer second-order cone program (SOCP). We also propose low-complexity algorithms that exploit group-sparsity or received power strength in the problem formulation.

Populärvetenskaplig sammanfattning

Massiv MIMO anses vara efterföljaren till MIMO-teknologin och har fått mycket uppmärksamhet både inom universitetsvärlden och i industrin. Denna teknologi använder hundratals antenner, antingen installerade i en kompakt gruppantenn eller fördelade över täckningsområdet, för att förbättra spektral-effektiviteten. Massiv MIMO kan ge goda multiplexvinster samt betydande förbättringar i spektraleffektivitet och energieffektivitet. Resursallokering spelar en viktig roll i förbättringen av prestandan i det trådlösa nätverket genom att det löser diverse funktionella optimeringsproblem. Användandet av resursallokering i massiv MIMO förser oss med en möjlig lösning för att hantera den exponentiella tillväxten både i trådlös datastrafik och i antalet trådlösa mottagare, vilken inte kan hanteras av nuvarande cellulär nätverksteknologi. I denna avhandling fokuserar vi på flera aspekter av resursallokering i flercelliga massiva MIMO-system:

- Vi försöker att minimera den totala transmissionseffekten för nerlänken i cellulära massiva MIMO-system när varje användare betjänas av en optimerad delmängd av basstationer. Vi härleder en undre gräns på den ergodiska spektraleffektiviteten när: *a)* kanalerna är Rayleighfäddande, *b)* antingen signalmaximerande förkodning eller inversförkodning används och *c)* basstationerna samarbetar genom att använda icke-koherent simultan transmission. Vi löser problemen med användarassociation och minimering av transmissionseffekten i nerlänken optimalt för ett givet spektraleffektivitetsvillkor. Dessutom löser vi ett rättvishetsproblem med användarspecifika vikter som maximerar den sämsta spektraleffektiviteten bland alla användare. Den optimala användarassociationen härleds, vilken skiljer sig klassisk från associationen som bygger på det största signal-till-brus-förhållandet.
- Vi optimerar pilotsignalerna och effekten som används för datasignalerna i

upplänken för att begränsa pilotkontaminering i massiva MIMO-system. Medan tidigare arbeten har behandlat pilotallokeringen som ett kombinatoriskt problem, så uppnår vi en enklare problemformulering genom att optimera direkt över pilotsekvenserna. För att göra det, så beräknar vi en undre gräns för spektraleffektiviteten i upplänken för Rayleighfäddande kanaler med signalmaximerande förkodning och godtyckliga pilotsekvenser. Vi maximerar sedan den sämsta spektraleffektiviteten bland alla användare med avseende på pilotsekvenserna och upplänks-effekten under en given effektbudget. Det resulterar i ett NP-svårt, signomialt problem, till vilket vi föreslår en effektiv algoritm med polynomisk komplexitet för att hitta ett lokalt optimum. Vi utökar också vårt optimerings ramverk till ett system vars prestanda är försämrat av hårvaruimperfektioner hos användarna.

- Vi studerar en tvådelad avkodningsmetod som begränsar intercell-interferensen i flercelliga massiva MIMO-system. I den första delen skattar varje basstation kanalen till användarna i sin egen cell och använder skattningen för att göra en lokal avkodning. Detta följs av en andra del med avkodning, där basstationerna samarbetar för att begränsa intercell-interferensen. En uppnåelig spektraleffektivitet beräknas för en godtycklig tvådelad avkodningsmetod, och ett slutet uttryck härleds för kanaler med korrelerad Rayleighfäddning, signalmaximerande förkodning och avkodning baserat på den storskaliga fäddningen i den andra delen. Vi formulerar ett icke-konvext optimeringsproblem för att maximera den totala spektraleffektiviteten, där både dataeffekterna och vektorerna som är baserade på den storskaliga fäddningen för avkodningen i den andra delen är optimeringsvariabler. För detta problem utvecklar vi en algoritm baserad på minimerandet av det viktade minsta kvadratfelet för att erhålla en stationär punkt med låg beräkningskomplexitet.
- Vi undersöker den simultana data- och piloteffektoptimeringen för att maximera den totala spektraleffektiviteten i ett flercelligt massivt MIMO-system, vilket inte är ett konvext problem. Vi föreslår först en ny optimeringsalgoritm, vilken är inspirerad av upplägget med att minimera det viktade minsta kvadratfelet, för att erhålla en stationär punkt i polynomisk tid. Sedan använder vi denna algoritm tillsammans med djupinlärning för att träna ett faltningsneuronnät till att utföra den simultana data- och piloteffektsregleringen på mindre än en millisekund, vilket gör den användbar för realtidsoptimering i ett verkligt flercelligt massivt MIMO-system.

- Vi undersöker effektiva lösningar på ett energieffektiviseringsproblem i cellfri massiv MIMO-kommunikation, en möjlig teknologi för 5G och framtida trådlösa nätverk. Vi minimerar den totala energiförbrukningen vid alla accesspunkter i nerlänkstransmissionen, när både hårdvaru- och transmissionseffekterna beaktas. Trots att detta inte är ett konvext problem, finner vi en globalt optimal lösning genom att lösa ett heltalskonprogram av andra ordningen med branch-and-bound-förfarandet. Vi föreslår också två algoritmer med låg komplexitet som ger en bra lokal lösning genom att utnyttja den glesa strukturen hos transmissions-effekterna för att endast aktivera en delmängd av accesspunkterna.

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 on many things in my life from the beginning when I came to Linköping. My special thanks should also send to Dr. Tuan Anh Le, Middlesex University London, Dr. Mohammad Reza Nakhai, King's College London, Dr. Tien Hoa Nguyen and Professor Van Duc Nguyen, Hanoi University of Science and Technology, for invaluable advice and productive collaborations. Besides, the helpful and stimulating discussions with my colleagues, such as Dr. Zheng Chen, Dr. Hei Victor Cheng, Dr. Marcus Karlsson, and Dr. Antonios Pitarokoilis, Dr. Özlem Tugfe Demir vastly assisted me in my research during the last five years. I have indeed learned a lot from them. During this journey, I would also send my sincere thanks to Dr. Kamil Senel, Dr. Meysam Sadeghi, Ema Becirovic, Ziya Gülgün, Daniel Verenzuela, Özgecan Özdoğan, and other doctoral students for fruitful discussions in technical seminars and coffee breaks. I am further very grateful to them all for the warm and friendly work environment, which makes me less lonely when away from home. Specifically, I am thankful to Amin Ghazanfari, my colleague, for all the moments we have shared in the office together with friendly discussions on professional and personal matters. Having wonderful time together, I

am also thankful to Giovanni Interdonato, Amin Ghazanfari, and other research fellows of 5Gwireless, a multi-partner European Training Network (ETN) project, within the framework of the H2020 Marie Skłodowska-Curie Innovative Training Networks (ITNs) and my gratitude should be sent to Professor Marco Di Renzo, the Project Coordinator. I further send my thanks to Dr. Christopher Mollén for helping me to prepare Populärvetenskaplig sammanfattning. He has also given me many other useful supports. I would like to thank Danyo Danev, Mikael Olofsson, and Håkan Johansson for their uncounted mentor support on my teaching activities in the last five years. While conducting research, I had an opportunity to visit OTEAcademy, Greece. I would like to extend my thanks to George Agapiou and other colleagues at OTE Academy who gave me a lot of kind support and assistance.

My research was financially supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 641985 (5Gwireless). It was also supported by ELLIIT and CENIIT. Moreover, Ericsson's Research Foundation supported me to participate and present my papers at the IEEE International Conference on Communications (ICC) in Shanghai, China. The support from these funders is always highly appreciated and was a determining factor for my success.

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. I would like to thank Professor Nguyen Tien Son and his family and Dr Thanh Duc Nguyen and his family for their support in the last five years. Finally, the warmest thank should be sent to my all dear friends including Du Duc Ho, Minh Ha Le, Nastaran Meftahi, Dr. Thuong Nguyen Canh, Dr. Phuong Dung Trinh, and Dr. Lam Thanh Tu for keeping in touch, sharing social activities, and being interested in my work.

Trinh Van Chien
Linköping, January 2020

Contents

1	Introduction	3
1.1	Motivation	3
1.2	Papers Included in the Thesis	9
1.3	Papers Not Included in the Thesis	13
2	Preliminaries for Massive MIMO Communications	19
2.1	Channel Hardening	19
2.2	Favorable Propagation	20
2.3	TDD and FDD Mode	20
3	Multi-Cell Massive MIMO Communications	23
3.1	Uplink Pilot Training Phase	23
3.2	Uplink Data Transmission	28
3.3	Downlink Data Transmission	30
4	Coordinated Multipoint (CoMP) Transmission	33
4.1	Non-Coherent Joint Transmission	34
4.2	Coherent Joint Transmission	37
4.3	Pilot Decontamination by a Two-Layer Decoding System	38
4.4	Transmit Power Consumption at Base Stations	42
5	Optimization Preliminaries	45
5.1	Convex Optimization Problems	45
5.2	Linear Programming	46
5.3	Second-Order Cone Programming	47
5.4	Geometric Programming	47
5.5	Signomial Programming	48
5.6	Weighted Max-Min Fairness Optimization	49
5.7	Sum Spectral Efficiency Optimization	51
6	Artificial Neural Networks	53

6.1	Universal Approximation Theorem & Artificial Neural Networks	53
6.2	Artificial Intelligence based Wireless Model	56

Bibliography	59
---------------------	-----------

Included Papers	67
------------------------	-----------

A Joint Power Allocation and User Association Optimization	69
---	-----------

1	Introduction	71
2	System Model and Achievable Performance	75
2.1	Uplink Channel Estimation	76
2.2	Downlink Data Transmission Model	78
2.3	Achievable Spectral Efficiency under Rayleigh Fading	80
3	Downlink Transmit Power Optimization for Massive MIMO Systems	83
4	Optimal Power Allocation and User Association by Linear Programming	85
4.1	Optimal Solution with Linear Programming	85
4.2	BS-User Association Principle	86
5	Max-min QoS Optimization	88
6	Numerical Results	90
7	Conclusion	96
8	Appendix	97
8.1	Proof of Proposition 1 and Theorem 1	97
8.2	Proof of Corollary 1	101
8.3	Proof of Corollary 2	102
8.4	Proof of Theorem 3	103
8.5	Proof of Corollary 3	104
8.6	Joint Power Allocation and User Association for Massive MIMO Systems with Coherent Joint Transmission	105
	References	107

B Joint Pilot Design and Uplink Power Allocation	113
---	------------

1	Introduction	115
2	Pilot Designs for Massive MIMO Systems	118
2.1	Proposed Pilot Design	119
2.2	Other Pilot Designs	120
3	Uplink Massive MIMO Transmission	122
3.1	Channel Estimation with Arbitrary Pilots	122
3.2	Uplink Data Transmission	124

4	Max-min Fairness Optimization	127
4.1	Problem Formulation	127
4.2	Local Optimality Algorithm	129
5	Pilot Optimization for Cellular Massive MIMO Systems with Hardware Impairments	133
5.1	Channel Estimation under Hardware Impairments . .	134
5.2	UL Data Transmission and Max-min Fairness Opti- mization under Hardware Impairments	135
6	Generalization to Correlated Rayleigh fading	137
7	Numerical Results	138
8	Conclusion	146
9	Appendix	146
9.1	Proof of Lemma 2	146
9.2	Proof of Theorem 1	148
9.3	Proof of Theorem 2	149
	References	150

C Large-Scale-Fading Decoding in Cellular Massive MIMO Systems with Spatially Correlated Channels 155

1	Introduction	157
1.1	Main Contributions	160
2	System Model	161
2.1	Channel Estimation	161
2.2	Uplink Data Transmission	164
3	Uplink Performance Analysis	165
4	Data Power Control and LSFD Design for Sum SE Optimization	170
4.1	Problem Formulation	170
4.2	Iterative Algorithm	171
4.3	Sum SE Optimization Without Using LSFD	174
5	Numerical Results	176
5.1	Convergence	181
5.2	Impact of Spatial Correlation	182
5.3	Impact of Number of Antennas and Users	183
5.4	Performance of Regularized Zero-Forcing	184
6	Conclusion	184
7	Appendix	184
7.1	Useful Lemma and Definition	184
7.2	Proof of Theorem 1	185
7.3	Proof of Theorem 2 in case of MMSE	186
7.4	Proof of Theorem 2 in Case of EW-MMSE	188

7.5	Proof of Theorem 3	190
7.6	Proof of Theorem 4	192
References		196

D Power Control in Cellular Massive MIMO with Varying User Activity: A Deep Learning Solution 199

1	Introduction	201
2	Dynamic Massive MIMO System Model	205
2.1	Uplink Pilot Transmission Phase	205
2.2	Uplink Data Transmission Phase	207
3	Joint Pilot and Data Power Control for Sum Spectral Efficiency Optimization	208
3.1	Problem Formulation	208
3.2	Iterative Algorithm	210
4	A low-complexity solution with convolutional neural network	213
4.1	Conditions on Large-Scale Fading Model	214
4.2	Existence of a Single Neural Network for Joint Pilot and Data Power Control	217
4.3	Convolutional Neural Network Architecture with Residual Dense Connections	218
4.4	Dataset, Training, and Testing Phases	221
5	Numerical Results	222
5.1	Sum Spectral Efficiency & Power Consumption	223
5.2	Predicted Performance of PowerNet	224
5.3	Varying User Activity, Channel Models, and Data Label Effect	225
5.4	Runtime	227
6	Conclusion	228
7	Appendix	230
7.1	Proof of Theorem 1	230
7.2	Proof of Theorem 2	232
References		235

E Joint Power Allocation and Load Balancing Optimization for Energy-Efficient Cell-Free Massive MIMO Networks 241

1	Introduction	243
1.1	Main Contributions	245
2	System Model	246
2.1	Uplink Pilot Transmission	247
2.2	Downlink Performance Analysis	248

3	Total Power Minimization Problem	250
3.1	Problem Formulation	250
3.2	Globally Optimal Solution to the Total Power Minimization Problem	253
4	Two Suboptimal Algorithms With Lower Complexity	256
4.1	Utilizing Sparsity to Turn Off APs	257
4.2	Total Transmit Power Minimization and Turnoff APs	263
5	Numerical Results	264
6	Conclusion	270
7	Appendix	271
7.1	Proof of Lemma 4	271
7.2	Proof of Theorem 1	274
	References	275

List of Abbreviations

AP	access point
BS	base station
CoMP	coordinated multipoint
CSI	channel state information
DL	downlink
EE	energy efficiency
EW-MMSE	element-wise minimum mean square error
FDD	frequency division duplex
GPU	graphics processing unit
i.i.d.	independent and identically distributed
KKT	Karush-Kuhn-Tucker
LSFD	large-scale fading decoding
LS	least square
MIMO	multiple-input multiple-output
MR	maximum ratio
MRC	maximum-ratio combining
MMSE	minimum mean square error
NP	non-deterministic polynomial-time
PDF	probability density function
QoS	quality of service
SE	spectral efficiency
SINR	signal-to-interference-plus-noise ratio
SISO	single-input single-output
SNR	signal-to-noise ratio
SOC	second-order cone
SOCP	second-order cone program
TDD	time division duplex
UL	uplink
ZF	zero forcing

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 (SISO) 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. Therefore, 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, multiple-input multiple-output (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 the information more efficiently 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 data 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 base stations (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 many 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 [5,6]. 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 [7]. 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 uplink (UL) and downlink (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 [8]. 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 [5,9,10]. It is achieved by utilizing the fact that channel estimates obtained in the UL can be also used for the DL transmission, which is a physical characteristic called channel reciprocity. For this great benefit, Massive MIMO should therefore be deployed using TDD. This is further discussed in Section 2.3 of Chapter 2.
- *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 [11]. 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 Section 4.4, 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 [9] that improves the signal-to-noise ratio (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 [12]. 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 [13]. 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 nearly mutually orthogonal. This property is known as favorable propagation [14, 15]. 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 [16]. 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 [17] providing the initial framework of Massive MIMO with infinitely many antennas, numerous papers have analyzed various aspects of Massive MIMO systems in general [18, 19], and multi-cell Massive MIMO systems in particular [20]. The authors in [7] 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 quality of service (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 [21, 22]. Fortunately, power control algorithms are much easier to deploy in Massive MIMO since the SE expressions are only dependent on the large-scale fading coefficients thanks to the channel hardening property as demonstrated in [9, 23] 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, in Paper A [24], we establish a novel framework for joint user association and power allocation in the 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 [19, 25, 26] 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 [27, 28]. 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 [29]. 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 [11]. As a contribution of this thesis, in Paper B [30], 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.

Instead of combating pilot contamination, one can utilize decoding schemes where the BSs can cooperate with each other [31–33]. In the two-layer large-scale fading decoding (LSFD) framework, each BS applies an arbitrary local linear decoding method in the first layer, as in conventional Massive MIMO. The decoded signals are then gathered at a common central station that applies so-called LSFD vectors in a second-layer to combine the signals from multiple BSs to suppress pilot contamination and other inter-cell interference. The LSFD vectors are selected only based on the channel statistics (large-scale fading) and, therefore, there is no need for the BSs to share their local channel estimates. Previous works on LSFD have either considered uncorrelated Rayleigh fading channels [33] or the one-ring model for spatial correlation [34]. The latter paper optimizes the system with respect to network-wide max-min fairness, which is a criterion that gives all the users the same SE, but usually a very low such SE [35]. As presented in Paper C [36], this thesis generalizes the LSFD method to a scenario with arbitrary spatial correlation and also develop a method for joint power control and LSFD vector optimization in the system using the sum SE as the utility. We quantify the SE in a system with arbitrary processing in the two layers and then derive a closed-form expression for the case when maximum ratio combining (MRC) is used in the first layer. The LSFD vector that maximizes the SE follows in closed form. Additionally, an UL sum SE optimization problem with power constraints is formulated. Because it is a hard non-convex problem, we are not searching for the global optimum but develop an alternating low-complexity optimization algorithm that converges to a stationary point.

Deep learning has become a popular way to solve problems in a data-driven fashion, which has shown remarkable prediction accuracy in many applications, including wireless communications as well [37–39]. Deep learning is capable of providing efficient solutions, given that the complicated design and training phases have been successful. In Paper D [40], we first formulate a joint data and pilot power control for maximizing the ergodic sum SE. The

non-convexity of this problem is overcome by proposing a new algorithm, inspired by the weighted minimum mean square error (MMSE) approach, which finds a stationary point with polynomial complexity instead of seeking the global optimum with exponential computational complexity. We then explore the possibility of using deep learning to train a neural network to solve the power control problem with sub-millisecond runtime. To this end, we model the power optimization as a learning problem where the transmit powers should be predicted based on an input of large-scale fading coefficients. We train a deep convolutional neural network and show that it achieves very high accuracy in the power prediction.

Recently, a system with a massive number of access points (APs) distributed over a large area coherently serves a massive number of users, called Cell-free Massive MIMO, has attracted lots of interest [41–45]. In Paper E [46], we investigate how APs can be turned on and off in Cell-free Massive MIMO systems to save energy. This is useful in scenarios with varying user activity, user locations, and traffic requirements. We formulate an optimization problem where we minimize the total power consumption of all the APs in the DL transmission, considering both the hardware-consumed power and the transmit powers. Even though this is a non-convex optimization problem, we manage to obtain a globally optimal solution by solving a mixed-integer second-order cone (SOC) program with the branch and bound approach. We also propose two low complexity algorithms that yield a good local solution by exploiting the sparsity structure of the transmit powers (i.e., the zero power coefficients [47, 48]) to activate only a subset of the APs.

The rest of this thesis is organized as follows: The remaining parts of Chapter 1 give the list of included and excluded papers. Chapter 2 defines and explains basic terminologies which are used in Massive MIMO. Chapter 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 are described in Chapter 4. Some preliminaries of the classical optimization problems are introduced in Chapter 5. The fundamentals of artificial neural networks are presented in Chapter 6. Then the five included papers are provided.

1.2 Papers Included in the Thesis

This thesis focuses on 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. The third aspect is to exploit the large-scale-fading information to mitigate pilot contamination in a two-layer decoding Massive MIMO system. This is investigated in Paper C for a multi-cell Massive MIMO systems with either fixed transmit power or data power allocation. The fourth aspect is motivated from the use of deep learning for online power control. This is investigated in Paper D, which formulates and solves a sum SE optimization optimization for a multi-cell Massive MIMO systems with jointly optimizing the UL pilot and data powers. The solution to this algorithm is used to train a deep neural network, which utilizes only the large-scale-fading coefficients as the inputs to predict both the pilot and data powers in sub-milliseconds. Finally, the fifth aspect is to consider the joint power allocation and load balancing for a Cell-free Massive MIMO system, where many APs may be turned off for energy-efficient purposes. This is investigated in Paper E, which formulates and solves a mixed-integer SOC program.

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.

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 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

Published in: IEEE Transactions on Wireless Communications, volume 17, issue 3, pp. 2000 – 2015, March 2018.

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.

Paper C: Large-Scale-Fading Decoding in Cellular Massive MIMO Systems With Spatially Correlated Channels

Authored by: Trinh Van Chien, Christopher Mollén, and Emil Björnson

Published in: IEEE Transactions on Communications, volume 67, issue 4, pp. 2746 – 2762, April 2019.

Abstract: Massive MIMO systems can suffer from coherent intercell interference due to the phenomenon of pilot contamination. This paper investigates a two-layer decoding method that mitigates both coherent and non-coherent interference in multi-cell Massive MIMO. To this end, each BS first estimates the channels to intra-cell users using either MMSE or element-wise MMSE (EW-MMSE) estimation based on UL pilots. The estimates are used for local decoding on each BS followed by a second decoding layer where the BSs cooperate to mitigate inter-cell interference. An UL achievable SE expression is computed for arbitrary two-layer decoding schemes. A closed-form expression is then obtained for correlated Rayleigh fading, maximum-ratio combining, and the proposed LSFD in the second layer. We also formulate a sum SE maximization problem with both the data power and LSFD vectors as optimization variables. Since this is an NP-hard problem, we develop a low-complexity algorithm based on the weighted MMSE approach to obtain a local optimum. The numerical results show that both data power control and LSFD improve the sum SE performance over single-layer decoding multi-cell Massive MIMO systems.

Paper D: Sum Spectral Efficiency Maximization in Massive MIMO Systems: Benefits from Deep Learning

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

Submitted to: IEEE Transactions on Wireless Communications (Under Minor Revision).

Abstract: This paper considers the sum SE optimization problem in multi-cell Massive MIMO systems with a varying number of active users. This is formulated as a joint pilot and data power control problem. Since the problem is non-convex, we first derive a novel iterative algorithm that obtains a stationary point in polynomial time. To enable real-time implementation, we also develop a deep learning solution. The proposed neural network, PowerNet, only uses the large-scale fading information to predict both the pilot and data powers. The main novelty is that we exploit the problem structure to design a single neural network that can handle a dynamically varying number of active users; hence, PowerNet is simultaneously approximating many different

power control functions with varying number inputs and outputs. This is not the case in prior works and thus makes PowerNet an important step towards a practically useful solution. Numerical results demonstrate that PowerNet only loses 2% in sum SE, compared to the iterative algorithm, in a nine-cell system with up to 90 active users per in each coherence interval, and the runtime was only 0.03 ms on a graphics processing unit (GPU). When good data labels are selected for the training phase, PowerNet can yield better sum SE than by solving the optimization problem with one initial point.

Paper E: Joint Power Allocation and Load Balancing Optimization for Energy-Efficient Cell-Free Massive MIMO Networks

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

To be submitted to: IEEE Transactions on Wireless Communications.

Abstract: Cell-Free Massive MIMO is a potential technology for next-generation wireless networks to handle the growing number of users by allowing distributed access points (APs) jointly serving users in the coverage area. This paper investigates effective solutions to the energy-efficient concern in Cell-free Massive MIMO communications. In particular, we minimize the total power consumption at all APs with taking into account the hardware and transmit powers. This is a combinatorial optimization problem when seeking for the optimal subset of active APs, but we formulate it as a mixed-integer second-order cone (SOC) program, in which a globally optimal solution is able to obtain by the branch and bound approach. Motivated at work with large-scale networks, we also propose the two algorithms that yield local solutions by using the transmit powers and the sparsity structure as side information. We provide promising numerical results manifesting the superior improvements of the total power reduction of the new joint optimization framework compared with previous work only minimizing the total transmit power. More importantly, the proposed methods can handle the transmit power control and AP activation for large-scale networks with tens of APs and users.

1.3 Papers Not Included in the Thesis

This section contains the papers in the last 5 years contributed by the author, but these publications are not included in this thesis since they are either condense versions of the papers included in this thesis or outside the scope of this thesis.

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 detailed proofs of lemmas and theorems were omitted from this paper.

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.

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.

Proposals of Multipath Time-Variant Channel and Additive Coloured Noise Modelling for Underwater Acoustic OFDM-based Systems

Authored by: Do Viet Ha, Trinh Van Chien, and Van Duc Nguyen

Published in: International Journal of Wireless and Mobile Computing, volume 11, issue 4, pp. 329 – 338, February 2017.

Abstract: This paper presents the results of Underwater Acoustic (UWA) channel measurements, including the power delay profile and the Doppler power spectrum, and further uses them to analyse the Orthogonal Frequency Division Multiplexing (OFDM) system performance. The UWA channel simulation model is derived from the measurement data by applying a widely-used optimization algorithm called the ℓ_p -norm method. The close match of correlation functions of the measured and channel simulation model demonstrates the correctness of our proposed channel modeling. Moreover, we also propose to use the auto-regressive generation method for characterizing ambient noise in UWA systems. From the UWA channel simulation and coloured noise models, the performance of the OFDM system using different channel estimation techniques is investigated to analyze the impact of the acoustic medium as multipath, Doppler, and colored noise effects. The numerical results manifest superior improvements of the so-called sparse channel estimation over other traditional ones such as least squares or minimum mean square estimation.

Block Compressive Sensing of Image and Video with Nonlocal Lagrangian Multiplier and Patch-based Sparse Representation

Authored by: Trinh Van Chien, Khanh Quoc Dinh, Byeungwoo Jeon, and Martin Burger

Published in: Signal Processing: Image Communication, volume 55, pp. 93 – 106, March 2017.

Abstract: Although block compressive sensing (BCS) makes it tractable to sense large-sized images and video, its recovery performance has yet to be significantly improved because its recovered images or video usually suffer from blurred edges, loss of details, and high-frequency oscillatory artifacts, especially at a low subrate. This paper addresses these problems by designing a modified total variation technique that employs multi-block gradient processing, a denoised Lagrangian multiplier, and patch-based sparse representation. In the case of video, the proposed recovery method is able to exploit both spatial and temporal similarities. Simulation results confirm the improved performance of the proposed method for compressive sensing of images and video in terms of both objective and subjective qualities.

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.

Distributed Power Control in Downlink Cellular Massive MIMO Systems

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

Published in: IEEE International ITG Workshop on Smart Antennas (WSA), March 2018.

Abstract: This paper compares centralized and distributed methods to solve the power minimization problem with quality-of-service (QoS) constraints in the downlink (DL) of multi-cell Massive MIMO systems. In particular, we study the computational complexity, number of parameters that need to be exchanged between BSs, and the convergence of iterative implementations. Although a distributed implementation based on dual decomposition (which only requires statistical channel knowledge at each BS) typically converges to the global optimum after a few iterations, many parameters need to be exchanged to reach convergence.

A Successive Optimization Approach to Pilot Design for Multi-Cell Massive MIMO Systems

Authored by: Hayder Al-Salihi, Trinh Van Chien, Tuan Anh Le, and Mohammad Reza Nakhai

Published in: IEEE Communications Letters, volume 22, issue 5, pp. 1086 – 1089, May 2018.

Abstract: In this letter, we introduce a novel pilot design approach that minimizes the total mean square errors of the minimum mean square error estimators of all BSs subject to the transmit power constraints of individual users in the network, while tackling the pilot contamination in multi-cell Massive MIMO systems. First, we decompose the original non-convex problem into distributed optimization sub-problems at individual BSs, where each BS can optimize its own pilot signals given the knowledge of pilot signals from the remaining BSs. We then introduce a successive optimization approach to transform each optimization sub-problem into a linear matrix inequality form, which is convex and can be solved by available optimization packages. Simulation results confirm the fast convergence of the proposed approach and prevails a benchmark scheme in terms of providing higher accuracy.

A Power Efficient Pilot Design for Multi-Cell Massive MIMO Systems

Authored by: Tuan Anh Le, Trinh Van Chien, and Mohammad Reza Nakhai

Published in: IEEE Global Conference on Information and Signal Processing (GlobalSIP), November 2018.

Abstract: In this paper, we address the pilot contamination problem in multi-cell Massive MIMO systems. Particularly, we propose a pilot design scheme that simultaneously minimizes the channel estimation errors of all base BSs and the total pilot power consumption of all users subject to the transmit power constraint for every user in the network. We decompose the proposed non-convex problem into distributed optimization problems to be solved at each BS, assuming the knowledge of pilot signals of the other BSs. Then, we introduce a successive optimization approach to cast each distributed optimization problem into a convex linear matrix inequality form. Simulation results confirm that the proposed approach significantly reduces pilot power while maintaining the same level of channel estimation error as the recent work.

Two-Layer Decoding in Cellular Massive MIMO Systems with Spatial Channel Correlation

Authored by: Trinh Van Chien, Christopher Mollén, and Emil Björnson

Published in: IEEE International Conference on Communications (ICC), May 2019.

Summary: This paper contains selected parts of Paper C. We only consider the case of MMSE channel estimation. Because of space constraints, all the detailed proofs of lemmas and theorems were omitted from this paper.

Sum Spectral Efficiency Maximization in Massive MIMO Systems: Benefits from Deep Learning

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

Published in: IEEE International Conference on Communications (ICC), May 2019.

Summary: This paper contains selected parts of Paper C. We only consider the network with a fixed number of users per cell and all users are active. Because of space constraints, all the detailed proofs of lemmas and theorems were omitted from this paper.

Optimal Design of Energy-Efficient Cell-Free Massive MIMO: Joint Power Allocation and Load Balancing

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

Submitted to: IEEE International Conference on Acoustics, Speech, and
Signal Processing (ICASSP), May 2020.

Summary: This paper contains selected parts of Paper E. We only consider the mixed-integer SOC program and the first suboptimal algorithm which utilizes the sparsity structure of transmit power coefficients. Because of space constraints, the detailed proofs of lemmas and theorem were omitted.

Chapter 2

Preliminaries for Massive MIMO Communications

This chapter considers the multi-user Massive MIMO systems where each BS is equipped with hundreds of antennas. The system is designed to serve tens of users utilizing the same time-frequency resources. We define and explain the basic terminologies in Massive MIMO which are later used in this thesis.

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 $M \rightarrow \infty$ [9, 49]. 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 [11], 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. All the included papers 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.

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 [14]. 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, 11]. 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.

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

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

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.

Chapter 3

Multi-Cell Massive MIMO Communications

In this chapter, 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. 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.

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.

Figure 2: UL 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.

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 squares (LS) or MMSE [50]. 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 $\mathbf{h}_{l,k}^l \in \mathbb{C}^M$ 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 [12, 51–53], 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 $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{\tau_p}] \in \mathbb{C}^{\tau_p \times \tau_p}$. 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:

Figure 3: Illustration of the pilot designs in prior works: (a) The pilot design in [12, 51] which only focuses on the pilot assignment; (b) The pilot design in [52] which only focuses on the pilot power control; (c) A combined pilot design which involves both the pilot assignment and pilot power control.

- Pilot assignment only:* This pilot design was proposed in [12, 51] and only focuses on the pilot assignment for a given set of orthogonal pilot signals. By using the permutation matrix $\mathbf{\Pi}_l \in \mathbb{R}_+^{\tau_p \times K}$ 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 $\mathbf{\Psi}_l$ in cell l are constructed from the basis matrix $\mathbf{\Phi}$ as $\mathbf{\Psi}_l = \sqrt{\tilde{p}} \mathbf{\Phi} \mathbf{\Pi}_l$, where $\sqrt{\tilde{p}}$ 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 $\mathbf{\Pi}_1, \dots, \mathbf{\Pi}_L$ 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 $\mathbf{\Pi}_2, \dots, \mathbf{\Pi}_L$. 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 [53].
- Pilot power control only:* This pilot design was proposed in [52, 54] 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 $\mathbf{\Psi}_l = \mathbf{\Phi} \mathbf{P}_l$, 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 $\mathbf{\Psi}_l = \mathbf{\Phi} \mathbf{\Pi}_l \mathbf{P}_l$. 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 $\mathcal{A}_{l,k}$ 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

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

and it is distributed as

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

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

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

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 $\sum_{(i,t) \in \mathcal{P}_{l,k}} \beta_{i,t}^l \tilde{p}_{i,t}$. For a considered optimization problem such as the channel estimation error minimization or the SE maximization, selecting the best pilot reuse sets $\mathcal{P}_{l,k}$ 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 UL power control for multi-cell Massive MIMO systems is then investigated for either ideal hardware or hardware impairments.

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 [9]. 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 channels, 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 [9]. For example, by using MR detection with $\mathbf{v}_{l,k} = \hat{\mathbf{h}}_{l,k}^l$ and considering uncorrelated Rayleigh fading, (17) becomes

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

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

Figure 4: DL 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.

$\mathcal{P}_{l,k}$. 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.

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 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 [9], where the users only have access to the channel statistics since there are no DL pilots. The SINR expression in (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 Chapter 4.

Chapter 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 chapters. It may result in low SE for some users at the 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 [55]. CoMP can be roughly classified into three different categories [56, 57]: *coordinated scheduling/beamforming design* jointly designs the beamforming vectors and scheduling for a cluster/all users, *joint transmission* is interpreted as a simultaneous 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 chapter 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. In Paper E, we extend the framework to optimize the total power consumption problem (both hardware and transmit

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.

powers) for Cell-free Massive MIMO systems, where the BS antenna array is distributed in the coverage area.

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 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 $\mathbf{h}_{l,k} \in \mathbb{C}^M$. 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 [8]. In this section, 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

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

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 [9], 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 signal-to-interference-plus-noise ratio (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

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

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

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

where the SINR value is formulated as

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

- *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

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

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

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

where the SINR value is formulated as

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

- *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:

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

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.

4.2 Coherent Joint Transmission

With coherent joint transmission, it is assumed that all BSs transmit the same signals to a user,

$$\mathbf{x}_l = \sum_{t=1}^K \sqrt{\rho_{l,t}} \mathbf{w}_{l,t} s_t, \quad (37)$$

where s_t is the data signal from each BS to user t and $\mathbb{E}\{|s_t|^2\} = 1$. The received signal at user k is now formulated as

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

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

$$y_k = \underbrace{\sum_{i=1}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} s_k}_{\text{Desired signal}} + \underbrace{\sum_{i=1}^L \sum_{\substack{t=1 \\ t \neq k}}^K \sqrt{\rho_{i,t}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,t} s_t + n_k}_{\text{Interference + Noise}}. \quad (39)$$

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

$$R_k = \gamma^{\text{DL}} \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_k^{\text{DL}} \right) \right\}, \quad (40)$$

where the effective SINR value is

$$\text{SINR}_k^{\text{DL}} = \frac{\left| \sum_{i=1}^L \sqrt{\rho_{i,k}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,k} \right|^2}{\sum_{\substack{t=1 \\ t \neq k}}^K \left| \sum_{i=1}^L \sqrt{\rho_{i,t}} \mathbf{h}_{i,k}^H \mathbf{w}_{i,t} \right|^2 + \sigma_{\text{DL}}^2}. \quad (41)$$

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 derives 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. Meanwhile, Paper E investigates the power allocation and load balancing optimization for energy-efficient Cell-free Massive MIMO systems.

4.3 Pilot Decontamination by a Two-Layer Decoding System

A full collaboration of all BSs utilizing coherent joint transmission is problematic for large-scale systems, for example due to the delay, signaling, and

Figure 6: Desired signals are detected by the two-layer decoding technique.

synchronization issues that occur. Therefore, a collaboration of selected BSs may be preferred. Moreover, apart from working with the pilot design to combat pilot contamination, as mentioned in Section 3.1, we can utilize the pilot reuse pattern for more advanced decoding schemes. This approach is called LSFD, which allows the BSs in a cellular Massive MIMO system to collaborate with each other in mitigating pilot contamination effectively without sharing instantaneous channels, and therefore reducing the signaling. In a two-layer LSFD system, the decoding process to obtain the desired signals is implemented as follows:

- *In the first layer:* Each BS first applies an arbitrary local linear decoding such as MRC or ZF, preferably one that can suppress intra-cell interference effectively. It can be the same as in conventional Massive MIMO.
- *In the second layer:* The decoded signals from multiple BSs are then gathered at a common central station that applies so-called LSFD vectors to suppress pilot contamination and inter-cell interference.

We give a toy example to explain why LSFD vectors are useful to mitigate pilot contamination by considering a simple network with two BSs, each serving one user with the same index as their BS. The signal detection process by the two-layer decoding technique is schematically illustrated in Figure 6. For this network, there are four different channels $\{\mathbf{h}_{1,1}, \mathbf{h}_{1,2}, \mathbf{h}_{2,1}, \mathbf{h}_{2,2}\}$, in which $\mathbf{h}_{l,k} \sim \mathcal{CN}(\mathbf{0}, \beta_{l,k} \mathbf{I}_M)$ denotes the channel between BS l and user k , $l, k \in \{1, 2\}$.

In the pilot training phase, let us denote by $\psi \in \mathbb{C}$ the transmitted pilot signal shared by the two users with $|\psi|^2 = 1$, the received pilot signals at the

two BSs, $\mathbf{y}_1^p, \mathbf{y}_2^p \in \mathbb{C}^M$, are formulated as

$$\mathbf{y}_1^p = \mathbf{h}_{1,1}\psi + \mathbf{h}_{1,2}\psi + \mathbf{n}_1^p, \quad (42)$$

$$\mathbf{y}_2^p = \mathbf{h}_{2,1}\psi + \mathbf{h}_{2,2}\psi + \mathbf{n}_2^p, \quad (43)$$

where $\mathbf{n}_1^p, \mathbf{n}_2^p \sim \mathcal{CN}(\mathbf{0}, \sigma_{\text{UL}}^2 \mathbf{I}_M)$ are independent additive noise. The MMSE channel estimates of $\mathbf{h}_{1,1}$ and $\mathbf{h}_{2,2}$ are respectively formulated as

$$\hat{\mathbf{h}}_{1,1} = \frac{\beta_{1,1}}{\beta_{1,1} + \beta_{1,2} + \sigma_{\text{UL}}^2} \mathbf{y}_1^p = \gamma_{1,1} \mathbf{y}_1^p = \gamma_{1,1} \mathbf{h}_{1,1}\psi + \gamma_{1,1} \mathbf{h}_{1,2}\psi + \gamma_{1,1} \mathbf{n}_1^p, \quad (44)$$

$$\hat{\mathbf{h}}_{2,2} = \frac{\beta_{2,2}}{\beta_{2,1} + \beta_{2,2} + \sigma_{\text{UL}}^2} \mathbf{y}_2^p = \gamma_{2,2} \mathbf{y}_2^p = \gamma_{2,2} \mathbf{h}_{2,1}\psi + \gamma_{2,2} \mathbf{h}_{2,2}\psi + \gamma_{2,2} \mathbf{n}_2^p, \quad (45)$$

where the last equality in (44) and (45) are obtained by using (42) and (43), respectively. Moreover, $\gamma_{1,1}$ and $\gamma_{2,2}$ are given as

$$\gamma_{1,1} = \frac{\beta_{1,1}}{\beta_{1,1} + \beta_{1,2} + \sigma_{\text{UL}}^2}, \quad (46)$$

$$\gamma_{2,2} = \frac{\beta_{2,2}}{\beta_{2,1} + \beta_{2,2} + \sigma_{\text{UL}}^2}. \quad (47)$$

In the UL data transmission, let us denote the transmitted signal from user k in cell k by s_k , then the received signals at the two BSs, $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{C}^M$ are formulated as

$$\mathbf{y}_1 = \mathbf{h}_{1,1}s_1 + \mathbf{h}_{1,2}s_2 + \mathbf{n}_1, \quad (48)$$

$$\mathbf{y}_2 = \mathbf{h}_{2,1}s_1 + \mathbf{h}_{2,2}s_2 + \mathbf{n}_2, \quad (49)$$

where $\mathbf{n}_1, \mathbf{n}_2 \sim \mathcal{CN}(\mathbf{0}, \sigma_{\text{UL}}^2 \mathbf{I}_M)$ are independent additive noise. In the first layer, each BS uses an MRC vector, which is defined as

$$\mathbf{v}_1 = \frac{\hat{\mathbf{h}}_{1,1}}{M\psi^*\gamma_{1,1}} = \frac{\mathbf{h}_{1,1}}{M} + \frac{\mathbf{h}_{1,2}}{M} + \frac{\mathbf{n}_1^p}{M\psi^*}, \quad (50)$$

$$\mathbf{v}_2 = \frac{\hat{\mathbf{h}}_{2,2}}{M\psi^*\gamma_{2,2}} = \frac{\mathbf{h}_{2,1}}{M} + \frac{\mathbf{h}_{2,2}}{M} + \frac{\mathbf{n}_2^p}{M\psi^*}, \quad (51)$$

where ψ^* is the complex conjugate of ψ . Substituting (50) and (51) into (48) and (49), the signal estimates $\hat{s}_1, \hat{s}_2 \in \mathbb{C}$ are locally obtained at the two BSs as

$$\hat{s}_1 = \mathbf{v}_1^H \mathbf{y}_1 = \left(\frac{\mathbf{h}_{1,1}^H}{M} + \frac{\mathbf{h}_{1,2}^H}{M} + \frac{(\mathbf{n}_1^p)^H}{M\psi} \right) (\mathbf{h}_{1,1}s_1 + \mathbf{h}_{1,2}s_2 + \mathbf{n}_1), \quad (52)$$

$$\hat{s}_2 = \mathbf{v}_2^H \mathbf{y}_2 = \left(\frac{\mathbf{h}_{2,1}^H}{M} + \frac{\mathbf{h}_{2,2}^H}{M} + \frac{(\mathbf{n}_2^p)^H}{M\psi^*} \right) (\mathbf{h}_{2,1}s_1 + \mathbf{h}_{2,2}s_2 + \mathbf{n}_2). \quad (53)$$

When the number of antennas goes to infinity, the incoherent interference and noise vanish thanks to the favorable propagation property described in Section 2.2. It means that each of the following inner product converges to zero as $M \rightarrow \infty$:

$$\frac{\mathbf{h}_{1,1}^H \mathbf{h}_{1,2}}{M}, \frac{\mathbf{h}_{1,1}^H \mathbf{n}_1}{M}, \frac{\mathbf{h}_{1,2}^H \mathbf{h}_{1,1}}{M}, \frac{\mathbf{h}_{1,2}^H \mathbf{n}_1}{M}, \frac{(\mathbf{n}_1^p)^H \mathbf{h}_{1,1}}{M\psi}, \frac{(\mathbf{n}_1^p)^H \mathbf{h}_{1,2}}{M\psi}, \frac{(\mathbf{n}_1^p)^H \mathbf{n}_1}{M\psi} \rightarrow 0, \quad (54)$$

$$\frac{\mathbf{h}_{2,1}^H \mathbf{h}_{2,2}}{M}, \frac{\mathbf{h}_{2,1}^H \mathbf{n}_2}{M}, \frac{\mathbf{h}_{2,2}^H \mathbf{h}_{2,1}}{M}, \frac{\mathbf{h}_{2,2}^H \mathbf{n}_2}{M}, \frac{(\mathbf{n}_2^p)^H \mathbf{h}_{2,1}}{M\psi}, \frac{(\mathbf{n}_2^p)^H \mathbf{h}_{2,2}}{M\psi}, \frac{(\mathbf{n}_2^p)^H \mathbf{n}_2}{M\psi} \rightarrow 0, \quad (55)$$

whilst the desired signals and pilot contamination remain as a consequence of the channel hardening property described in Section 2.1:

$$\frac{\mathbf{h}_{1,1}^H \mathbf{h}_{1,1}}{M} \rightarrow \beta_{1,1}, \frac{\mathbf{h}_{1,2}^H \mathbf{h}_{1,2}}{M} \rightarrow \beta_{1,2}, \frac{\mathbf{h}_{2,1}^H \mathbf{h}_{2,1}}{M} \rightarrow \beta_{2,1}, \frac{\mathbf{h}_{2,2}^H \mathbf{h}_{2,2}}{M} \rightarrow \beta_{2,2}. \quad (56)$$

Therefore, the decoded signals in (52) and (53) are given as

$$\begin{bmatrix} \hat{s}_1 \\ \hat{s}_2 \end{bmatrix} = \begin{bmatrix} \beta_{1,1}s_1 + \beta_{1,2}s_2 \\ \beta_{2,1}s_1 + \beta_{2,2}s_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{bmatrix}}_{=\mathbf{B}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}. \quad (57)$$

The asymptotic SE obtained by linear decoding in the first layer is limited by mutual interference since each BS observes a linear combination of the two signals. In the second layer, a central station will process $\hat{s}_1, \hat{s}_2$ to mitigate the remaining interference as

$$\mathbf{B}^{-1} \begin{bmatrix} \hat{s}_1 \\ \hat{s}_2 \end{bmatrix} = \mathbf{B}^{-1} \mathbf{B} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}. \quad (58)$$

In (58), $\mathbf{B}^{-1}$ exists only when the β -terms give $\det(\mathbf{B}) \neq 0$, which happens when $\beta_{1,1}\beta_{2,2} \neq \beta_{1,2}\beta_{2,1}$ and is typically the case in practice. The rows of the inverse matrix $\mathbf{B}^{-1}$ are denoted as the LSFD vectors since they only depend on the large scale fading coefficients. Therefore, the central station does not need to know the instantaneous channels. The network can achieve an unbounded SE when $M \rightarrow \infty$ since the decoded signals in (58) are free from noise and interference.

In Paper C, we study the two-layer decoding method to mitigate inter-cell interference in multi-cell Massive MIMO systems with correlated Rayleigh fading channels. An UL achievable SE expression is computed for arbitrary two-layer decoding schemes, while a closed-form expression is obtained for MR detection in the first layer and LSFD in the second layer. We formulate

Figure 7: The block diagram of a base station.

a non-convex sum SE maximization problem with both the data powers and LSFD vectors as optimization variables and develop an algorithm based on the weighted MMSE approach to obtain a stationary point with low computational complexity.

4.4 Transmit Power Consumption at Base Stations

Each Massive MIMO BS may be equipped with hundreds of antennas and simultaneously serve multiple users. Therefore an accurate model of the power consumption at the BSs is essential for power optimization problems. In this thesis, we use a simplified block diagram of a BS [58] as shown in Figure 7 when measuring the power consumption. As for classical MIMO BSs, the total power consumption at each Massive MIMO BS includes a dissipation part, which is determined by the hardware technology, and a transmit part, which is a function of the transmitted signals. More precisely, the power consumption at each BS spends power on four basic modules: Power amplifier, analog processing, digital processing, and fronthaul signaling. For BS l , it can be mathematically modeled as [59]

$$P_l = \begin{cases} P_{\text{active},l} + \Delta_l P_{\text{trans},l} & \text{if } P_{\text{trans},l} \neq 0, \\ P_{\text{sleep},l} & \text{if } P_{\text{trans},l} = 0, \end{cases} \quad (59)$$

where $P_{\text{sleep},l}$ is the power consumption at BS l when it is in sleep mode (on stand by). $P_{\text{active},l}$ denotes the hardware dissipation power to process the signals in both analog and digital domains, which is further divided as

$$P_{\text{active},l} = P_{\text{tc},l} + BP_{\text{bt},l} \sum_{k=1}^K R_{lk}, \quad (60)$$

where $P_{\text{tc},l}$ models the static power consumption of the transceiver chains connected to active antennas and the traffic-independent power of the analog

processing, digital processing, and fronthaul connections (please refer to Figure 7), which is defined as

$$P_{\text{tc},l} = M P_{\text{tc},l0} + P_{l0}. \quad (61)$$

Here, $P_{\text{tc},l0}$ is the internal power that each antenna consumes to run the circuit components such as converters, filters, and mixers. P_{l0} is the power consumption of other things that are not proportional to the number of antennas such as the fronthaul connections from the center processing unit to the BS and an local oscillator. In (60), $P_{\text{bt},l}$ (measured in Watt per bit/s) is the traffic-varying power consumption (of the fronthaul and baseband processing, including channel coding) that is proportional to the SE and the system bandwidth B Hz. In (59), the scaling factor $\Delta_l \geq 1$ denotes the inefficiency factor of the power amplifier.

From (23), the transmit power $P_{\text{trans},l}$ at BS l is obtained for the case of non-coherent joint transmission as

$$P_{\text{trans},l} = \sum_{k=1}^K \|\mathbf{w}_{l,k}\|^2 \mathbb{E}\{|s_{l,k}|^2\} = \sum_{k=1}^K \rho_{l,k} \|\mathbf{w}_{l,k}\|^2, \quad (62)$$

and in the case of coherent joint transmission, the transmit power is formulated based on (37) as

$$P_{\text{trans},l} = \sum_{k=1}^K \|\mathbf{w}_{l,k}\|^2 \mathbb{E}\{|s_k|^2\} = \sum_{k=1}^K \rho_{l,k} \|\mathbf{w}_{l,k}\|^2. \quad (63)$$

From (62) and (63), 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 this thesis, Paper A considers a transmit power optimization problem of the form

$$\begin{aligned} & \underset{\{\rho_{l,k} \geq 0\}}{\text{minimize}} && \sum_{i=1}^L P_{\text{trans},i} \\ & \text{subject to} && R_k \geq \xi_k, \forall k = 1, \dots, K, \\ & && P_{\text{trans},l} \leq P_{\text{max},l}, \forall l = 1, \dots, L, \end{aligned} \quad (64)$$

where ξ_k is the required SE of user k which is a fixed parameter measured in bit/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 (64)

to investigate the joint power allocation and user association problems in Massive MIMO in Paper A.

Besides, for a Cell-Free Massive MIMO system with L APs serving K users by coherent joint transmission, Paper E considers a total power consumption optimization problem of the form

$$\begin{aligned}
& \underset{\{\rho_{l,k} \geq 0\}, \mathcal{A}}{\text{minimize}} && \sum_{l \in \mathcal{A}} (P_{\text{active},l} + \Delta_l P_{\text{trans},l}) + \sum_{l \notin \mathcal{A}} P_{\text{sleep},l} \\
& \text{subject to} && R_k \geq \xi_k, \forall k = 1, \dots, K, \\
& && P_{\text{trans},l} \leq P_{\text{max},l}, \forall l \in \mathcal{A},
\end{aligned} \tag{65}$$

where $\mathcal{A} \in \{1, \dots, L\}$ denotes the set of APs which are in active mode. We observe that the objective function to problem (65) can be reformulated as

$$\sum_{l \in \mathcal{A}} (P_{\text{active},l} - P_{\text{sleep},l} + \Delta_l P_{\text{trans},l}) + \sum_{l=1}^L P_{\text{sleep},l}. \tag{66}$$

In (66), the second term is independent of the optimization variables, so we can remove it from the optimization problem without effecting the optimal solution. By letting $\tilde{P}_{\text{active},l} = P_{\text{active},l} - P_{\text{sleep},l}$, problem (65) has the same optimal solution as the following optimization problem

$$\begin{aligned}
& \underset{\{\rho_{l,k} \geq 0\}, \mathcal{A}}{\text{minimize}} && \sum_{l \in \mathcal{A}} \tilde{P}_{\text{active},l} + \Delta_l P_{\text{trans},l} \\
& \text{subject to} && R_k \geq \xi_k, \forall k = 1, \dots, K, \\
& && P_{\text{trans},l} \leq P_{\text{max},l}, \forall l \in \mathcal{A}.
\end{aligned} \tag{67}$$

Solving problem (67) is still challenging since it contains both the discrete and continuous optimization variables. We will propose the solutions to this problem in Paper E. In this paper, we actually set $P_{\text{sleep},l} = 0, \forall l$, but including the sleep power consumption does not change to the optimization structure, and therefore the same proposed algorithms can be applied.

Chapter 5

Optimization Preliminaries

This chapter presents preliminaries of optimization theory, including some basic optimization classes and properties, which are used in the included papers. In optimization theory, an optimization problem on standard form is formulated as

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && f_0(\mathbf{x}) \\ & \text{subject to} && f_i(\mathbf{x}) \leq b_i, \quad i = 1, \dots, m, \end{aligned} \tag{68}$$

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 (68) 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.

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

$$a_1 \mathbf{x}_1 + \dots + a_m \mathbf{x}_m \in \mathcal{X}. \quad (69)$$

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 $\alpha_1 \mathbf{x} + \alpha_2 \tilde{\mathbf{x}} \in \mathcal{X}$, the functions $f_i, \forall i = 0, \dots, m$, satisfy

$$f_i(\alpha_1 \mathbf{x} + \alpha_2 \tilde{\mathbf{x}}) \leq \alpha_1 f_i(\mathbf{x}) + \alpha_2 f_i(\tilde{\mathbf{x}}), \quad (70)$$

then (68) is a convex optimization problem. A vector $\mathbf{x}^*$ is the optimal solution to (68) 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 $\mathcal{O}(\max\{n^3, n^2m, F\})$, where F is the cost of evaluating the first and second derivative of the objective and constraint functions [60]. However, the exact computational complexity depends on the methods involved to solve the optimization problems.

An optimization problem that does not satisfy (70) 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 [61].

5.2 Linear Programming

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

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i = 1, \dots, m, \end{aligned} \quad (71)$$

where $\mathbf{c}$ and $\mathbf{a}_i, \forall i, \in \mathbb{R}^n$ are vectors. In comparison to (68), the objective function $f_0(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$ is now a linear function of $\mathbf{x}$. Mapping to (68), 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 [60], 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 [62]. 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 $m \geq n$ [60].

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.

5.3 Second-Order Cone Programming

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

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T \mathbf{x} + d_i, \quad i = 1, \dots, m, \end{aligned} \quad (72)$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A}_i \in \mathbb{R}^{n_i \times n}$, $\mathbf{c}_i \in \mathbb{R}^n$, and $d_i \in \mathbb{R}$ are constant parameters. The constraints in (72) are called SOC 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 [62]. The SOCP problems have higher complexity than linear programs since they require to evaluate the second derivative of the constraints [63, 64]. The total transmit power minimization problem in the case of coherent joint transmission is a SOC program as demonstrated in Paper A.

5.4 Geometric Programming

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

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} && \sum_{m=1}^M c_{m,0} \prod_{n=1}^N x_n^{a_{n,m,0}} \\ & \text{subject to} && \sum_{m=1}^M c_{m,i} \prod_{n=1}^N x_n^{a_{n,m,i}} \leq 1, \quad i = 1, \dots, m, \end{aligned} \quad (73)$$

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 (73) becomes

$$\begin{aligned} \underset{\{y_n\}}{\text{minimize}} \quad & \ln \left(\sum_{m=1}^M c_{m,0} e^{\sum_{n=1}^N y_n a_{n,m,0}} \right) \\ \text{subject to} \quad & \ln \left(\sum_{m=1}^M c_{m,i} e^{\sum_{n=1}^N y_n a_{n,m,i}} \right) \leq 0, i = 1, \dots, m. \end{aligned} \quad (74)$$

Since the weighted log-sum-exponentials functions are convex, (74) is a convex problem. Therefore, we can obtain the globally optimal solution to (74) in tractable time by using interior-point methods [60]. 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 [60, 65, 66]. This optimization class will be utilized in Paper B when we work with joint pilot design and UL power control for multi-cell Massive MIMO.

5.5 Signomial Programming

A signomial program has the same structure as that of a geometric program in (73), 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 [67]. 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 [68]:

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.

5.6 Weighted Max-Min Fairness Optimization

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

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \min_{i \in \{1, \dots, m\}} \quad \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})w_i}, \quad (75)$$

where $w_i > 0$ is the weight value of the function $f_i(\mathbf{x})/g_i(\mathbf{x})$. We solve the problem (75) by converting it to the epi-graph form as

$$\begin{aligned} & \underset{\xi, \mathbf{x} \in \mathcal{X}}{\text{maximize}} && \xi \\ & \text{subject to} && f_i(\mathbf{x}) - g_i(\mathbf{x})w_i\xi \geq 0, \quad i = 1, \dots, m, \end{aligned} \quad (76)$$

where the auxiliary variable ξ moves the objective function of problem (75) to the constraints in (76). We stress that it is possible to obtain the globally optimal solution to (76) 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 (75) can be obtained by using a general-purpose optimization toolbox such as CVX [62]. Alternatively, if (76) is a convex problem for a given ξ , i.e.,

$$\begin{aligned} & \underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} && 0 \\ & \text{subject to} && f_i(\mathbf{x}) - g_i(\mathbf{x})w_i\xi \geq 0, \quad i = 1, \dots, m, \end{aligned} \quad (77)$$

is convex, then the optimal solution to (75) is obtained by using the bisection search for the largest feasible values of ξ . The details of optimizing (76) via utilizing bisection search is given in Algorithm 1.

Algorithm 1 Weighted max-min fairness optimization with bisection

Result: Solve optimization problem (75).

Input: Initial upper bound ξ_0^{upper} , initiate $\mathbf{x}^{\text{lower}} = \mathbf{0}$ and line-search accuracy $\delta > 0$;

```

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 (77) is infeasible for  $\xi = \xi^{\text{candidate}}$ , then
        Set  $\xi^{\text{upper}} = \xi^{\text{candidate}}$ ;
    else
        Set  $\{\mathbf{x}^{\text{lower}}\}$  as the solution to (77);
        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 $\mathbf{x}^* = \mathbf{x}^{\text{lower}}$;

In Paper A, we investigate the weighted max-min fairness optimization problem for the CoMP frameworks where multiple BSs can collaborate to serve all users, using either coherent or non-coherent joint transmission.

Meanwhile, the application of weighted max-min fairness optimization to joint pilot design and UL power optimization is studied in Paper B.

5.7 Sum Spectral Efficiency Optimization

In order to improve the minimum SE, the solution to the max-min optimization problem may need to sacrifice the SE of other users significantly, especially users with the average SE. An alternative utility metric is to maximize the sum SE in the system, which is of the form

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^m \log_2 \left(1 + \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} \right), \quad (78)$$

where $f_i(\mathbf{x}) \geq 0, g_i(\mathbf{x}) > 0$ and the fraction $f_i(\mathbf{x})/g_i(\mathbf{x})$ stands for a general form of the SINR expression. Problem (78) is generally non-convex due to the non-convexity of the objective function by the SINR structures. Therefore, a global optimum to (78) is difficult to attain in any non-trivial setup in many scenarios [69]. For further processing, we will convert the sum SE optimization problem into the weighted MMSE optimization problem [70]. Specifically, for the user associated with the SINR value defined by the fraction $f_i(\mathbf{x})/g_i(\mathbf{x})$ we can formulate an equivalent SISO transmission scheme having the same SE, where the received signal is

$$y_i = \sqrt{f_i(\mathbf{x})} s_i + \sqrt{g_i(\mathbf{x})} n_i, \quad (79)$$

where s_i is data symbol sent by user i with $\mathbb{E}\{s_i^2\} = 1$. The parameter $n_i \sim \mathcal{N}(0, 1)$ denotes the additive noise and y_i is the received signal at the receiver. A beamforming coefficient $u_i \in \mathbb{R}$ is used to decode the desired signal as

$$\hat{s}_i = u_i y_i = u_i \sqrt{f_i(\mathbf{x})} s_i + u_i \sqrt{g_i(\mathbf{x})} n_i. \quad (80)$$

From (80), we therefore can compute the mean square error of the decoding process as

$$\begin{aligned} e_i &= \mathbb{E} \{ (\hat{s}_i - s_i)^2 \} = \left(u_i \sqrt{f_i(\mathbf{x})} - 1 \right)^2 + u_i^2 g_i(\mathbf{x}) \\ &= u_i^2 (f_i(\mathbf{x}) + g_i(\mathbf{x})) - 2u_i \sqrt{f_i(\mathbf{x})} + 1. \end{aligned} \quad (81)$$

Mathematically, problem (78) has the same globally optimal value of the variable $\mathbf{x}$ as the following problem

$$\underset{\mathbf{x} \in \mathcal{X}, \{u_i\}, \{w_i \geq 0\}}{\text{minimize}} \quad \sum_{i=1}^m w_i e_i - \ln(w_i). \quad (82)$$

More precisely, to prove the equivalence of problems (78) and (82), we first compute the optimal solution to u_i from the quadratic function (81) with respect to u_i as

$$u_i^* = \frac{\sqrt{f_i(\mathbf{x})}}{f_i(\mathbf{x}) + g_i(\mathbf{x})}. \quad (83)$$

Plugging (83) into (81), we obtain a new expression for e_i as

$$e_i = \frac{g_i(\mathbf{x})}{f_i(\mathbf{x}) + g_i(\mathbf{x})}. \quad (84)$$

From (82), the optimal solution to w_i is $w_i^* = e_i^{-1}$. Consequently, together with (84), problem (82) is equivalent (having the same optimal solution to $\mathbf{x}$) to solving

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} \quad m - \sum_{i=1}^m \ln \left(1 + \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} \right). \quad (85)$$

Since (85) is easily converted to (78), the proof of the equivalence between problems (78) and (85) is completed. We stress that (82) is still a non-convex problem, but it provides an element-wise convex structure: Problem (82) is convex with respect to one of the set of variables when the other sets are fixed. Paper C uses the equivalence of problems (78) and (82) to propose a solution to a joint LSFD design and data power control problem. Paper D also utilizes this equivalence to formulate and solve a joint pilot and data power control problem for dynamic cellular Massive MIMO systems, where each BS serves a different number of users.

Chapter 6

Artificial Neural Networks

Wireless systems have observed a dramatically increasing computational complexity when serving many users on the same radio resource using many antennas, which challenges the current approaches to optimize the system design and operation. Artificial neural networks and deep learning are potentially emerging techniques to be utilized in future wireless communications to solve the computational complexity issues. This chapter provides a brief discussion of artificial neural networks which aims at giving the methodologies and motivations of this essential tool for designing and operating future wireless systems.

6.1 Universal Approximation Theorem & Artificial Neural Networks

Each aspect of a wireless system can be optimized using mathematical models, functions, and algorithms, but the complexity is often too high for real-time usage. To cope with this issue, artificial neural networks, inspired by animal brains, are computing systems that can “learn” the features of functions from the input data. A seminal result is the existence of a single hidden layer neural network with a finite number of neurons (please see Figure 8 for an example) that can approximate any continuous function $f(\mathbf{x})$ on a compact domain [71, 72]. For a given accuracy $\delta > 0$, there exists an integer S and a feedforward neural network $\text{Net}_S(\mathbf{x}, \Theta)$ with S hidden neurons for which we can learn the characteristics of the continuous function $f(\mathbf{x})$ in the sense that

$$\sup_{\mathbf{x} \in \mathcal{X}} \|\text{Net}_S(\mathbf{x}, \Theta) - f(\mathbf{x})\| \leq \delta, \quad (86)$$

Figure 8: A single hidden layer feedforward neural network.

where Θ includes all weights and biases which must be optimized to approximate the function $f(\mathbf{x})$.¹ The optimized weight and bias values in Θ are generally unknown, but we can obtain a good solution to Θ by using either supervised learning or unsupervised learning [39, 73]. This thesis focuses on the former method, in which one collects the training data comprising of the different realizations of the input $\mathbf{x}$ and the corresponding output $f(\mathbf{x})$. The training data set contains both input and out output data, which are labeled in pairs. This method is a preferable option to train a deep neural network to behave as a given function when a subset of inputs together with the corresponding desired outputs (the exact value of $f(\mathbf{x})$ for a given $\mathbf{x}$) are available. We stress that if computing the exact value of $f(\mathbf{x})$ is too costly and an algorithm only can yield an approximation, this learning technique can still be applied with an acceptable level of performance.

Figure 8 gives an example of an artificial neural network with the two inputs: x_1, x_2 , the two outputs: y_1, y_2 , and one hidden layer comprising of three neurons: h_1, h_2 , and h_3 . By collecting a sufficiently large data set for training, the activities to train this neural network using the supervised learning are implemented using the forward and back propagation methods.

In the forward propagation, from the two input values x_1 and x_2 , this neural network will learn/predict the two output values y_1 and y_2 . In the hidden layer, the intermediate value stored at each neuron is respectively

¹In general, the inputs of a neural network can include other information instead of only $\mathbf{x}$ and Θ .

computed as

$$h_1 = g(x_1 w_{11}^1 + x_2 w_{21}^1 + b_1^1), \quad (87)$$

$$h_2 = g(x_1 w_{12}^1 + x_2 w_{22}^1 + b_2^1), \quad (88)$$

$$h_3 = g(x_1 w_{11}^1 + x_2 w_{21}^1 + b_3^1), \quad (89)$$

where $g(\cdot)$ is the activation function, e.g., the rectifier gives $f(x) = \max(0, x)$ or the sigmoid function gives $f(x) = 1/(1 + e^{-x})$, to handle the nonlinear properties in the data set. $w_{ij}^1, i, j \in \{1, 2\}$, denote the weights used to transfer the input values to the hidden layer. $b_i^1, i \in \{1, 2, 3\}$, are the biases used by the hidden layer. The biases values are added to the results for better fitting the data. This is because the biases allow the activation function to be shifted to the left or right. The hidden layer propagates information to the output layer as

$$y_1 = g(h_1 w_{11}^2 + h_2 w_{21}^2 + h_3 w_{31}^2 + b_1^2), \quad (90)$$

$$y_2 = g(h_1 w_{12}^2 + h_2 w_{22}^2 + h_3 w_{32}^2 + b_2^2), \quad (91)$$

where $w_{ij}^2, i \in \{1, 2, 3\}, j \in \{1, 2\}$, denote the weights from the hidden layer to the output layer. $b_i^2, i \in \{1, 2\}$, are the biases used by the output layer.

In the back propagation, a loss function as $l(\{w_{i,j}^1, w_{i,j}^2, b_i^1, b_i^2\})$ is needed to determine how well the neural network approximates the desired function. It can, for example, be selected as

$$l(\{w_{i,j}^1, w_{i,j}^2, b_i^1, b_i^2\}) = |y_1 - y_1^*|^2 + |y_2 - y_2^*|^2, \quad (92)$$

where y_1^* and y_2^* are the output data used as reference for the training process. The back propagation updates the weights and biases by using, for example the gradient descent as

$$w_{i,j}^1 \leftarrow w_{i,j}^1 - \alpha \partial l(\{w_{i,j}^1, w_{i,j}^2, b_i^1, b_i^2\}) / \partial w_{i,j}^1, \forall i, j, \quad (93)$$

$$w_{i,j}^2 \leftarrow w_{i,j}^2 - \alpha \partial l(\{w_{i,j}^1, w_{i,j}^2, b_i^1, b_i^2\}) / \partial w_{i,j}^2, \forall i, j, \quad (94)$$

$$b_i^1 \leftarrow b_i^1 - \alpha \partial l(\{w_{i,j}^1, w_{i,j}^2, b_i^1, b_i^2\}) / \partial b_i^1, \forall i, \quad (95)$$

$$b_i^2 \leftarrow b_i^2 - \alpha \partial l(\{w_{i,j}^1, w_{i,j}^2, b_i^1, b_i^2\}) / \partial b_i^2, \forall i, \quad (96)$$

where the hyperparameter $0 < \alpha < 1$ is the learning rate, whose value may be set before the training process begins or can be adapted during the training phase. The learning rate controls how fast a neural network learns the

input-output relation. In general, a large learning rate may lead to unstable training and a tiny value leads to a failure to train the neural network. A proper selection of the learning rate is crucial and heavily dependent on each particular problem. The forward and back propagation are repeated until the desired accuracy is obtained. While the forward propagation is used for both training and testing phases, the back propagation is only applied for the training phases.

We stress that the universal approximation theorem proves the existence of a basic wide neural network with one hidden layer and a sufficiently large number of neurons to approximate any given function. This obtained single hidden layer neural network is somewhat flexible in the sense that we can vary the number of inputs, outputs, and neurons upon concrete applications. However, the universal approximation theorem still has the following main limitations [74]:

- The number of neurons required to obtain a desired prediction accuracy is not given. Therefore, we must carefully design the set of neurons for every specific problems to achieve the desired prediction accuracy.
- It is limited by a single hidden layer with many neurons, while researchers have recently shown that it is convenient to use multiple hidden layers to improve the prediction accuracy for a given number of neurons and to reduce the number of required neurons [75].
- It does not show how to configure an artificial neural network, for instance, the type of activation functions and the number of hidden layers to achieve a given prediction accuracy.

Recently, artificial neural networks with multiple hidden layers (deep learning) have become a popular trend to solve problems in a data-driven fashion and have shown great performance in various applications such as image restoration [76], pattern recognition [39], and power allocation [74]. For a single-hidden layer neural network, established by the universal approximation theorem, an equivalent narrow neural network with multiple hidden layers (or usually called a deep neural network) can be constructed as in [75].

6.2 Artificial Intelligence based Wireless Model

Classical optimization algorithms solving resource allocation problems in wireless communications are usually implemented in an iterative manner. However, these algorithms require many iterations and high computational

Figure 9: A learning-to-optimize framework applied for resource allocation problems in Massive MIMO communications.

complexity in many applications, especially far away from a real-time implementation [21]. As a potential solution to this issue, one way of exploiting artificial neural networks for resource allocation in Massive MIMO communications is given in Figure 9a. This indicates that deep neural networks should not entirely replace the classic optimization algorithms, but be used in the training phase. In more detail, an optimization algorithm will be treated as a black box and its properties are learned by utilizing an artificial neural network as shown in Figure 9a. The learning is achieved by exploiting the relation between the inputs and outputs of an optimization algorithm.

As an example in Massive MIMO communication, the inputs may be large-scale fading coefficients, while the outputs may be transmit powers that the BSs allocate for users. This is the same inputs and outputs as for a power optimization algorithm. In the training phase, the optimization algorithm is used to attain optimized solutions, either optimal or suboptimal, for many instances of large-scale fading coefficients, which are used to form a training data set for this artificial neural network. The network is trained by adapting weights and biases to learn the input-output relation, in the sense that a set of transmit powers will be designed for a given set of large-scale fading coefficients. In the testing phase in Figure 9b, the trained artificial neural network is used to predict the optimized solution whenever new input information is available. We again notice that the testing phase is of low computational complexity since the artificial neural networks only use the forward propagation once and the optimization algorithm is removed.

When applying an artificial neural network for resource allocation by utilizing supervised learning, we should note the following [38]:

- An essential condition to obtain a high prediction accuracy is that the training phase of an artificial neural network should utilize a sufficiently large training data set. This requires to design efficient optimization algorithms, which are able to generate such a data set in polynomial time. However, we notice that the training phase can be implemented offline, so optimization algorithms with reasonably high complexity are affordable. We stress that real-time constraints in the training phase are not applied to our work.
- The learning-to-optimize framework in Figure 9 can be exploited to reduce the runtime by utilizing modern computing capacity to execute complex optimization algorithms and to implement the training and testing phases of artificial neural networks. A multi-core computer with graphics processing units can perform the learning-to-optimize framework at a much faster speed than a traditional computer with only a central processing unit.

In Paper D, we consider the application of neural networks for power control in cellular Massive MIMO systems, where each BS serves a different number of users. We formulate a new sum SE optimization problem with the pilot and data powers as optimization variables. Since it is a non-convex problem, a globally optimal solution to pilot and data powers is not available in polynomial time. Therefore, an iterative algorithm is proposed to obtain a stationary point to this optimization problem, but its complexity is too high for real-time use. However the stationary points can be used to train a neural network. In particular, by exploiting the framework supervised learning in Figure 9, we further construct and train a neural network to predict the outputs of both pilot and data powers from the inputs of large-scale fading coefficients. In order to achieve highly accurate prediction performance of the pilot and data powers from the large-scale fading coefficients, we base our contribution on the deep architectures of a multiple hidden layer structure.

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] T. Van Chien and E. Björnson, “Massive MIMO communications,” in *5G Mobile Communications*. Springer, 2017, pp. 77–116.
- [7] 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.
- [8] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [9] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [10] J. Flordelis, F. Rusek, F. Tufvesson, E. G. Larsson, and O. Edfors, “Massive MIMO performance-TDD versus FDD: What do measurements say?” *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, 2018.

- [11] 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.
- [12] 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.
- [13] 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.
- [14] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Aspects of favorable propagation in Massive MIMO,” in *Proc. IEEE EUSIPCO*, 2014.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.

- [22] T. V. Chien, E. Björnson, E. G. Larsson, and T. A. Le, “Distributed power control in downlink cellular Massive MIMO systems,” in *Proc. ITG Workshop on Smart Antennas (WSA)*, 2018.
- [23] T. V. Chien, E. Björnson, and E. G. Larsson, “Downlink power control for Massive MIMO cellular systems with optimal user association,” in *Proc. IEEE ICC*, 2016.
- [24] —, “Joint power allocation and user association optimization for Massive MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6384 – 6399, 2016.
- [25] —, “Joint pilot sequence design and power control for max-min fairness in uplink Massive MIMO,” in *Proc. IEEE ICC*, 2017.
- [26] H. Al-Salihi, T. Van Chien, T. A. Le, and M. R. Nakhai, “A successive optimization approach to pilot design for multi-cell Massive systems,” *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 1086–1089, 2018.
- [27] 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.
- [28] T. A. Le, T. Van Chien, and M. R. Nakhai, “A power efficient pilot design for multi-cell Massive MIMO systems,” in *Proc. IEEE GlobalSIP*. IEEE, 2018, pp. 823–827.
- [29] 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.
- [30] T. Van Chien, E. Björnson, and E. G. Larsson, “Joint pilot design and uplink power allocation in multi-cell Massive MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2000–2015, 2018.
- [31] T. Van Chien, C. Mollén, and E. Björnson, “Two-layer decoding in cellular Massive MIMO systems with spatial channel correlation,” in *Proc. IEEE ICC*, 2019.
- [32] A. Ashikhmin, L. Li, and T. L. Marzetta, “Interference reduction in multi-cell Massive MIMO systems with large-scale fading precoding,” *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6340 – 6361, 2018.

- [33] A. Adhikary, A. Ashikhmin, and T. L. Marzetta, "Uplink interference reduction in large-scale antenna systems," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2194–2206, 2017.
- [34] A. Adhikary and A. Ashikhmin, "Uplink Massive MIMO for channels with spatial correlation," in *Proc. IEEE GLOBECOM*, 2018.
- [35] 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. [Online]. Available: <http://dx.doi.org/10.1561/20000000093>
- [36] T. Van Chien, C. Mollén, and E. Björnson, "Large-scale-fading decoding in cellular Massive MIMO systems with spatially correlated channels," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 2746 – 2762, 2019.
- [37] T. Van Chien, E. Björnson, and E. G. Larsson, "Sum spectral efficiency maximization in Massive MIMO systems: Benefits from deep learning," in *Proc. IEEE ICC*, 2019.
- [38] A. Zappone, M. Di Renzo, M. Debbah, T. T. Lam, and X. Qian, "Model-aided wireless artificial intelligence: Embedding expert knowledge in deep neural networks towards wireless systems optimization," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 60–69, 2019.
- [39] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [40] T. Van Chien, T. N. Canh, E. Björnson, and E. G. Larsson, "Power control in cellular Massive MIMO with varying user activity: A deep learning solution," *IEEE Trans. Wireless Commun.*, 2019. [Online]. Available: <http://arXiv:1901.03620>
- [41] H. Q. Ngo, A. E. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free Massive MIMO: Uniformly great service for everyone," in *Proc. IEEE SPAWC*, 2015.
- [42] E. Björnson and L. Sanguinetti, "Scalable cell-free Massive MIMO systems," *arXiv*, Aug 2019. [Online]. Available: <https://arxiv.org/abs/1908.03119>
- [43] T. Van Chien, E. Björnson, and E. G. Larsson, "Optimal design of energy-efficient Cell-free Massive MIMO: Joint power allocation and load balancing," in *Proc. IEEE ICASSP*, 2019.

-
- [44] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free Massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, 2017.
 - [45] T. H. Nguyen, T. K. Nguyen, H. D. Han, and V. D. Nguyen, "Optimal power control and load balancing for uplink cell-free multi-user Massive MIMO," *IEEE Access*, vol. 6, pp. 14 462–14 473, 2018.
 - [46] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and load balancing optimization for energy-efficient Cell-free Massive MIMO networks," *IEEE Trans. Wireless Commun.*, 2019, to be submitted.
 - [47] T. Van Chien, K. Q. Dinh, B. Jeon, and M. Burger, "Block compressive sensing of image and video with nonlocal lagrangian multiplier and patch-based sparse representation," *Signal Processing: Image Communication*, vol. 54, pp. 93–106, 2017.
 - [48] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.
 - [49] T. Van Chien, E. Björnson, and E. G. Larsson, "Multi-cell Massive MIMO performance with double scattering channels," in *Proc. IEEE CAMAD*. IEEE, 2016, pp. 231–236.
 - [50] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
 - [51] 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.
 - [52] 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.
 - [53] 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.
 - [54] 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.

- [55] M. Boldi, A. Tölili, 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.
- [56] 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.
- [57] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, “Multi-cell MIMO cooperative networks: A new look at interference,” *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, 2010.
- [58] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, O. Blume, and A. Fehske, “How much energy is needed to run a wireless network?” *IEEE Wireless Commun. Mag.*, vol. 18, no. 5, pp. 40–49, 2011.
- [59] 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*. INFSO-ICT-247733 EARTH, ver. 2.0, 2012. [Online]. Available: <http://www.ict-earth.eu/>
- [60] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [61] 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.
- [62] CVX Research Inc., “CVX: Matlab software for disciplined convex programming, academic users,” <http://cvxr.com/cvx>, 2015.
- [63] F. Alizadeh and D. Goldfarb, “Second-order cone programming,” *Mathematical programming*, vol. 95, no. 1, pp. 3–51, 2003.
- [64] A. Ghazanfari, A. Tölili, and H. Pennanen, “Sum power minimization for cellular systems with underlay D2D communications,” in *Proc. CROWN-COM*, 2014, pp. 45–50.
- [65] A. Ghazanfari, E. Björnson, and E. G. Larsson, “Optimized power control for Massive MIMO with underlaid D2D communications,” *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 2763–2778, 2019.

- [66] A. Ghazanfari, H. V. Cheng, E. Björnson, and E. G. Larsson, “A fair and scalable power control scheme in multi-cell massive mimo,” in *Proc. ICASSP*. IEEE, 2019, pp. 4499–4503.
- [67] 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.
- [68] 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.
- [69] V. Annapureddy and V. Veeravalli, “Sum capacity of MIMO interference channels in the low interference regime,” *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2565–2581, 2011.
- [70] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, “Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, 2008.
- [71] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [72] T. O’Shea and J. Hoydis, “An introduction to deep learning for the physical layer,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.
- [73] W. Lee, M. Kim, and D.-H. Cho, “Deep power control: Transmit power control scheme based on convolutional neural network,” *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1276–1279, 2018.
- [74] A. Zappone, M. Di Renzo, and M. Debbah, “Wireless networks design in the era of deep learning: Model-based, ai-based, or both?” *arXiv preprint arXiv:1902.02647*, 2019.
- [75] Z. Lu, H. Pu, F. Wang, Z. Hu, and L. Wang, “The expressive power of neural networks: A view from the width,” in *Advances in Neural Information Processing Systems*, 2017, pp. 6231–6239.
- [76] T. N. Canh and T. V. Chien, “Error resilient deep compressive sensing,” in *Proc. IEEE ICCE*, 2020, to be submitted.

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-162582>

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

Amin Ghazanfari, *Power Control for Multi-Cell Massive MIMO*, Linköping Studies in Science and Technology. Licentiate Thesis, No. 1852, 2019.

Giovanni Interdonato, *Signal Processing Aspects of Cell-Free Massive MIMO*, Linköping Studies in Science and Technology. Licentiate Thesis, No. 1817, 2018.

Marcus Karlsson, *Blind Massive MIMO Base Stations: Downlink Transmission and Jamming*, Linköping Studies in Science and Technology. Dissertations, No. 1950, 2018.

Daniel Verenzuela, *Aspects of Alternative Massive MIMO Designs: Superimposed Pilots and Mixed-ADCs*, Linköping Studies in Science and Technology. Licentiate Thesis, No. 1803, 2018.

Hei Victor Cheng, *Optimizing Massive MIMO: Precoder Design and Power Allocation*, Linköping Studies in Science and Technology. Dissertations, No. 1929, 2018.

Christopher Mollén, *High-End Performance with Low-End Hardware: Analysis of Massive MIMO Base Station Transceivers*, Linköping Studies in Science and Technology. Dissertations, No. 1896, 2017.

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.

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.

FACULTY OF SCIENCE AND ENGINEERING

Linköping Studies in Science and Technology,
Dissertation No. 2036, 2020

Division of Communication Systems,
Department of Electrical Engineering (ISY)

Linköping University
SE-581 83 Linköping, Sweden

www.liu.se