

Linköping Studies in Science and Technology
Dissertations, No. 1642

Massive MIMO: Fundamentals and System Designs

Hien Quoc Ngo



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

Linköping 2015

Massive MIMO: Fundamentals and System Designs

© 2015 Hien Quoc Ngo, unless otherwise noted.

ISBN 978-91-7519-147-8

ISSN 0345-7524

Printed in Sweden by LiU-Tryck, Linköping 2015

Cảm ơn gia đình tôi, cảm ơn Em,

vì đã luôn bên cạnh tôi.

Abstract

The last ten years have seen a massive growth in the number of connected wireless devices. Billions of devices are connected and managed by wireless networks. At the same time, each device needs a high throughput to support applications such as voice, real-time video, movies, and games. Demands for wireless throughput and the number of wireless devices will always increase. In addition, there is a growing concern about energy consumption of wireless communication systems. Thus, future wireless systems have to satisfy three main requirements: i) having a high throughput; ii) simultaneously serving many users; and iii) having less energy consumption. Massive multiple-input multiple-output (MIMO) technology, where a base station (BS) equipped with very large number of antennas (collocated or distributed) serves many users in the same time-frequency resource, can meet the above requirements, and hence, it is a promising candidate technology for next generations of wireless systems. With massive antenna arrays at the BS, for most propagation environments, the channels become favorable, i.e., the channel vectors between the users and the BS are (nearly) pairwise orthogonal, and hence, linear processing is nearly optimal. A huge throughput and energy efficiency can be achieved due to the multiplexing gain and the array gain. In particular, with a simple power control scheme, Massive MIMO can offer uniformly good service for all users. In this dissertation, we focus on the performance of Massive MIMO. The dissertation consists of two main parts: fundamentals and system designs of Massive MIMO.

In the first part, we focus on fundamental limits of the system performance under practical constraints such as low complexity processing, limited length of each coherence interval, intercell interference, and finite-dimensional channels. We first study the potential for power savings of the Massive MIMO uplink with maximum-ratio combining (MRC), zero-forcing, and minimum mean-square error receivers, under perfect and imperfect channels. The energy and spectral efficiency tradeoff is investigated. Secondly, we consider a physical channel model where the angular domain is divided into a finite number of distinct directions. A lower bound on the capacity is derived, and the effect of pilot contamination in this finite-dimensional channel model is analyzed. Finally, some aspects of favorable propagation in Massive MIMO under Rayleigh fading and line-of-sight (LoS) channels are investigated. We show that both Rayleigh fading and LoS environments offer favorable propagation.

In the second part, based on the fundamental analysis in the first part, we propose some system designs for Massive MIMO. The acquisition of channel state information (CSI) is very important in Massive MIMO. Typically, the channels are estimated at the BS through uplink training. Owing to the limited length of the coherence interval, the system performance is limited by pilot contamination. To reduce the pilot contamination effect, we propose an eigenvalue-decomposition-based scheme to estimate the channel directly from the received data. The proposed scheme results in better performance compared with the conventional training schemes due to the reduced pilot contamination. Another important issue of CSI acquisition in Massive MIMO is how to acquire CSI at the users. To address this issue, we propose two channel estimation schemes at the users: i) a downlink “beamforming training” scheme, and ii) a method for blind estimation of the effective downlink channel gains. In both schemes, the channel estimation overhead is independent of the number of BS antennas. We also derive the optimal pilot and data powers as well as the training duration allocation to maximize the sum spectral efficiency of the Massive MIMO uplink with MRC receivers, for a given total energy budget spent in a coherence interval. Finally, applications of Massive MIMO in relay channels are proposed and analyzed. Specifically, we consider multi-pair relaying systems where many sources simultaneously communicate with many destinations in the same time-frequency resource with the help of a Massive MIMO relay. A Massive MIMO relay is equipped with many colocated or distributed antennas. We consider different duplexing modes (full-duplex and half-duplex) and different relaying protocols (amplify-and-forward, decode-and-forward, two-way relaying, and one-way relaying) at the relay. The potential benefits of massive MIMO technology in these relaying systems are explored in terms of spectral efficiency and power efficiency.

Populärvetenskaplig Sammanfattning

Det har skett en massiv tillväxt av antalet trådlöst kommunikerande enheter de senaste tio åren. Idag är miljarder av enheter anslutna och styrda över trådlösa nätverk. Samtidigt kräver varje enhet en hög dataakt för att stödja sina applikationer, som röstkommunikation, realtidsvideo, film och spel. Efterfrågan på trådlös dataakt och antalet trådlösa enheter kommer alltid att tillta. Samtidigt kan inte strömförbrukningen hos de trådlösa kommunikationssystemen tillåtas att öka. Således måste framtida trådlösa kommunikationssystem uppfylla tre huvudkrav: i) hög dataakt ii) kunna betjäna många användare samtidigt iii) lägre strömförbrukning.

Massiv MIMO (“multiple-input multiple output”), en teknik där basstationen är utrustad med ett stort antal antenner och samtidigt betjänar många användare över samma tid-frekvensresurs, kan uppfylla ovanstående krav. Följaktligen kan det betraktas som en lovande kandidat för nästa generations trådlösa system. För de flesta utbredningsmiljöer blir kanalen fördelaktig med en massiv antennuppställning (en uppställning av, låt säga, hundra antenner eller fler), det vill säga kanalvektorer mellan användare och basstation blir (nästan) parvis ortogonala, vilket gör linjär signalbehandling nästan optimal. Den höga dataakten och låga strömförbrukningen kan åstadkommas tack vare multiplexeringsvinsten och antennförstärkningen. I synnerhet kan massiv MIMO erbjuda en likformigt bra betjäning av alla användare med en enkel effektallokeringsmetod.

I denna avhandling börjar vi med att fokusera på grunderna av massiv MIMO. Speciellt kommer vi att studera de grundläggande begränsningarna av systemets prestanda i termer av spektral effektivitet och energieffektivitet när massiva antennuppställningar används. Detta kommer vi att göra med beaktande av praktiska begränsningar hos systemet, som lågkomplexitetsbehandling (till exempel linjär behandling av signaler), begränsad längd av varje koherensintervall, ofullständig kanalkännedom, intercell-interferens och ändlig-dimensionella kanaler. Dessutom undersöks några aspekter hos fördelaktig utbredning i massiv MIMO med rayleigh-fädningskanaler och kanaler med rakt sikt. Baserat på dessa grundläggande analyser föreslår vi sedan några systemkonstruktioner för massiv MIMO. Mer precist föreslår vi några

metoder för kanalskattning både för basstationen och för användarna, vilka ämnar minimera effekten av pilotkontaminering och kanalovisshet. Den optimala pilot- och dataeffekten så väl som valet av längden av träningsperioden studeras. Till slut föreslås och analyseras användandet av massiv MIMO i reläkanaler.

Acknowledgments

I would like to extend my sincere thanks to my supervisor, Prof. Erik G. Larsson, for his valuable support and supervision. His advice, guidance, encouragement, and inspiration have been invaluable over the years. Prof. Larsson always keeps an open mind in every academic discussion. I admire his critical eye for important research topics. I still remember when I began my doctoral studies, Prof. Larsson showed me the first paper on Massive MIMO and stimulated my interest for this topic. This thesis would not have been completed without his guidance and support.

I would like to thank Dr. Thomas L. Marzetta at Bell Laboratories, Alcatel-Lucent, USA, for his cooperative work, and for giving me a great opportunity to join his research group as a visiting scholar. It has been a great privilege to be a part of his research team. He gave me valuable help whenever I asked for assistance. I have learnt many useful things from him. I would also like to thank Dr. Alexei Ashikhmin and Dr. Hong Yang for making my visit at Bell Laboratories, Alcatel-Lucent in Murray Hill such a great experience.

I was lucky to meet many experts in the field. I am thankful to Dr. Michail Matthaiou at Queen's University Belfast, U.K., for his great cooperation. I have learnt a lot from his maturity and expertise. Many thanks to Dr. Trung Q. Duong at Queen's University Belfast, U.K., and Dr. Himal A. Suraweera at University of Peradeniya, Sri Lanka, for both technical and non-technical issues during the cooperative work. I would like to thank Dr. Le-Nam Tran at Maynooth University, Ireland, for his explanations and discussions on the optimization problems which helped me a lot. I am also thankful to all of my co-authors for the collaboration over these years: Dr. G. C. Alexandropoulos (France Research Center, Huawei Technologies Co. Ltd.), Prof. H-J. Zepernick (Blekinge Institute of Technology, Sweden), Dr. C. Yuen (Singapore University of Technology and Design, Singapore), Dr. A. K. Papazafeiropoulos (Imperial College, U.K.), Dr. H. Phan (University of Reading, U.K.), Dr. M. El Kashlan (Queen Mary University of London, U.K.), and Mr. L. Wang (Queen Mary University of London, U.K.).

The warmest thank to my colleagues at Communication Systems, ISY, Linköping University, for the stimulating discussions, and for providing the fun environment in which we have learnt and grown during the past 4+ years. Special thanks to my fellow PhD students: Chaitanya, Reza, Mirsad, Johannes, Antonios, Erik Axell, Victor, Christopher, and Marcus.

Finally, I would like to thank my family and friends, for their constant love, encouragement, and limitless support throughout my life.

Linköping, January 2015
Hien Quoc Ngo

Abbreviations

AF	Amplify-and-Forward
AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
BS	Base Station
CDF	Cumulative Distribution Function
CSI	Channel State Information
DF	Decode-and-Forward
DL	Downlink
DPC	Dirty Paper Coding
EVD	Eigenvalue Decomposition
FD	Full Duplex
FDD	Frequency Division Duplexing
HD	Half Duplex
i.i.d.	Independent and Identically Distributed
ILSP	Iterative Least-Square with Projection
LDPC	Low-Density Parity-Check
LTE	Long Term Evolution
LoS	Line-of-Sight
LS	Least-Squares
MAC	Multiple-Access Channel
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MMSE	Minimum Mean Square Error
MSE	Mean-Square Error
ML	Maximum Likelihood
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MU-MIMO	Multiuser MIMO

PDF	Probability Density Function
OFDM	Orthogonal Frequency Division Multiplexing
QAM	Quadrature Amplitude Modulation
RV	Random Variable
SEP	Symbol Error Probability
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
TDD	Time Division Duplexing
TWRC	Two-Way Relay Channel
UL	Uplink
ZF	Zero-Forcing

Contents

Abstract	v
Populärvetenskaplig Sammanfattning (in Swedish)	vii
Acknowledgments	ix
Abbreviations	xi
I Introduction	1
1 Motivation	3
2 Mutiuser MIMO Cellular Systems	7
2.1 System Models and Assumptions	7
2.2 Uplink Transmission	8
2.3 Downlink Transmission	9
2.4 Linear Processing	9
2.4.1 Linear Receivers (in the Uplink)	10
2.4.2 Linear Precoders (in the Downlink)	13
2.5 Channel Estimation	14
2.5.1 Channel Estimation in TDD Systems	14
2.5.2 Channel Estimation in FDD Systems	16
3 Massive MIMO	19
3.1 What is Massive MIMO?	19
3.2 How Massive MIMO Works	21
3.2.1 Channel Estimation	21
3.2.2 Uplink Data Transmission	21
3.2.3 Downlink Data Transmission	22
3.3 Why Massive MIMO	22
3.4 Challenges in Massive MIMO	23
3.4.1 Pilot Contamination	23
3.4.2 Unfavorable Propagation	24
3.4.3 New Standards and Designs are Required	24
4 Mathematical Preliminaries	25
4.1 Random Matrix Theory	25
4.2 Capacity Lower Bounds	26
5 Summary of Specific Contributions of the Dissertation	29

5.1	Included Papers	29
5.2	Not Included Papers	35
6	Future Research Directions	37
II	Fundamentals of Massive MIMO	47
A	Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems	49
1	Introduction	52
2	System Model and Preliminaries	53
2.1	MU-MIMO System Model	53
2.2	Review of Some Results on Very Long Random Vectors	54
2.3	Favorable Propagation	55
3	Achievable Rate and Asymptotic ($M \rightarrow \infty$) Power Efficiency	56
3.1	Perfect Channel State Information	56
3.1.1	Maximum-Ratio Combining	58
3.1.2	Zero-Forcing Receiver	58
3.1.3	Minimum Mean-Squared Error Receiver	59
3.2	Imperfect Channel State Information	61
3.2.1	Maximum-Ratio Combining	63
3.2.2	ZF Receiver	64
3.2.3	MMSE Receiver	64
3.3	Power-Scaling Law for Multicell MU-MIMO Systems	66
3.3.1	Perfect CSI	67
3.3.2	Imperfect CSI	67
4	Energy-Efficiency versus Spectral-Efficiency Tradeoff	69
4.1	Single-Cell MU-MIMO Systems	69
4.1.1	Maximum-Ratio Combining	70
4.1.2	Zero-Forcing Receiver	71
4.2	Multicell MU-MIMO Systems	72
5	Numerical Results	73
5.1	Single-Cell MU-MIMO Systems	73
5.1.1	Power-Scaling Law	74
5.1.2	Energy Efficiency versus Spectral Efficiency Tradeoff	77
5.2	Multicell MU-MIMO Systems	78
6	Conclusion	79
A	Proof of Proposition 2	83
B	Proof of Proposition 3	84
B	The Multicell Multiuser MIMO Uplink with Very Large Antenna Arrays and a Finite-Dimensional Channel	87
1	Introduction	90
1.1	Contributions	91
1.2	Notation	92
2	System Model	92

2.1	Multi-cell Multi-user MIMO Model	92
2.2	Physical Channel Model	93
3	Channel Estimation	94
3.1	Uplink Training	94
3.2	Minimum Mean-Square Error Channel Estimation	95
4	Analysis of Uplink Data Transmission	96
4.1	The Pilot Contamination Effect	98
4.1.1	MRC Receiver	98
4.1.2	ZF Receiver	99
4.1.3	Uniform Linear Array	100
4.2	Achievable Uplink Rates	101
4.2.1	Maximum-Ratio Combining	102
4.2.2	Zero-Forcing Receiver	103
5	Numerical Results	104
5.1	Scenario I	105
5.2	Scenario II	107
6	Conclusions	110
A	Proof of Proposition 9	113
B	Proof of Theorem 1	114
C	Proof of Corollary 1	115
D	Proof of Corollary 2	116

C Aspects of Favorable Propagation in Massive MIMO 121

1	Introduction	124
2	Single-Cell System Model	124
3	Favorable Propagation	125
3.1	Favorable Propagation and Capacity	125
3.2	Measures of Favorable Propagation	126
3.2.1	Condition Number	126
3.2.2	Distance from Favorable Propagation	127
4	Favorable Propagation: Rayleigh Fading and Line-of-Sight Channels	127
4.1	Independent Rayleigh Fading	128
4.2	Uniform Random Line-of-Sight	129
4.3	Urns-and-Balls Model for UR-LoS	130
5	Examples and Discussions	132
6	Conclusion	133

III System Designs 137

D EVD-Based Channel Estimations for Multicell Multiuser MIMO with Very Large Antenna Arrays 139

1	Introduction	142
2	Multi-cell Multi-user MIMO Model	143
3	EVD-based Channel Estimation	144
3.1	Mathematical Preliminaries	144
3.2	Resolving the Multiplicative Factor Ambiguity	145
3.3	Implementation of the EVD-based Channel Estimation	146

4	Joint EVD-based Method and ILSP Algorithm	147
5	Numerical Results	148
6	Concluding Remarks	150
E	Massive MU-MIMO Downlink TDD Systems with Linear Pre-coding and Downlink Pilots	155
1	Introduction	158
2	System Model and Beamforming Training	159
2.1	Uplink Training	159
2.2	Downlink Transmission	160
2.3	Beamforming Training Scheme	161
3	Achievable Downlink Rate	162
3.1	Maximum-Ratio Transmission	163
3.2	Zero-Forcing	164
4	Numerical Results	164
5	Conclusion and Future Work	167
A	Proof of Proposition 10	169
B	Proof of Proposition 11	170
F	Blind Estimation of Effective Downlink Channel Gains in Massive MIMO	175
1	Introduction	178
2	System Model	179
3	Proposed Downlink Blind Channel Estimation Technique	180
3.1	Mathematical Preliminaries	181
3.2	Downlink Blind Channel Estimation Algorithm	182
3.3	Asymptotic Performance Analysis	182
4	Numerical Results	184
5	Concluding Remarks	185
G	Massive MIMO with Optimal Power and Training Duration Allocation	191
1	Introduction	194
2	Massive Multicell MIMO System Model	194
2.1	Uplink Training	195
2.2	Data Transmission	195
2.3	Sum Spectral Efficiency	196
3	Optimal Resource Allocation	197
4	Numerical Results	199
5	Conclusion	201
A	Proof of Proposition 13	203
H	Large-Scale Multipair Two-Way Relay Networks with Distributed AF Beamforming	207
1	Introduction	210
2	Multipair Two-Way Relay Channel Model	211
3	Distributed AF Transmission Scheme	211
3.1	Phase I	211

3.2	Phase II — Distributed AF Relaying	212
3.3	Asymptotic ($M \rightarrow \infty, K < \infty$) Performance	213
4	Achievable Rate for Finite M	214
4.1	Discussion of Results	215
4.1.1	Achievability of the Network Capacity	216
4.1.2	Power Scaling Laws	216
5	Numerical Results and Discussion	216
A	Derivation of (4)	219
B	Proof of Proposition 14	219

I Spectral Efficiency of the Multipair Two-Way Relay Channel with Massive Arrays 223

1	Introduction	226
2	System Models and Transmission Schemes	227
2.1	General Transmission Scheme	227
2.1.1	The First Phase — Training	227
2.1.2	The Second Phase — Multiple-Access Transmission of Payload Data	228
2.1.3	The Third Phase — Broadcast of Payload Data	229
2.1.4	Self-interference Reduction	229
2.2	Specific Transmission Schemes	230
2.2.1	Transmission Scheme I — Separate-Training ZF	230
2.2.2	Transmission Scheme II — Coupled-Training ZF	231
3	Asymptotic $M \rightarrow \infty$ Analysis	232
4	Lower Bound on the Capacity for Finite M	233
5	Numerical Results	234
6	Conclusion	236

J Multipair Full-Duplex Relaying with Massive Arrays and Linear Processing 241

1	Introduction	244
2	System Model	247
2.1	Channel Estimation	248
2.2	Data Transmission	249
2.2.1	Linear Receiver	249
2.2.2	Linear Precoding	250
2.3	ZF and MRC/MRT Processing	250
2.3.1	ZF Processing	250
2.3.2	MRC/MRT Processing	251
3	Loop Interference Cancellation with Large Antenna Arrays	252
3.1	Using a Large Receive Antenna Array ($N_{\text{rx}} \rightarrow \infty$)	252
3.2	Using a Large Transmit Antenna Array and Low Transmit Power ($p_{\text{R}} = E_{\text{R}}/N_{\text{tx}}$, where E_{R} is Fixed, and $N_{\text{tx}} \rightarrow \infty$)	253
4	Achievable Rate Analysis	254
5	Performance Evaluation	257
5.1	Power Efficiency	258
5.2	Comparison between Half-Duplex and Full-Duplex Modes	259
5.3	Power Allocation	260

6	Numerical Results	263
6.1	Validation of Achievable Rate Results	264
6.2	Power Efficiency	265
6.3	Full-Duplex Vs. Half-Duplex, Hybrid Relaying Mode	266
6.4	Power Allocation	268
7	Conclusion	269
A	Proof of Proposition 17	271
B	Proof of Theorem 3	273
B.1	Derive R_{SR_k}	273
B.2	Derive $R_{\text{RD},k}$	275
C	Proof of Theorem 4	275

Part I

Introduction

Chapter 1

Motivation

During the last years, data traffic (both mobile and fixed) has grown exponentially due to the dramatic growth of smartphones, tablets, laptops, and many other wireless data consuming devices. The demand for wireless data traffic will be even more in future [1–3]. Figures 1.1 shows the demand for mobile data traffic and the number of connected devices. Global mobile data traffic is expected to increase to 15.9 exabytes per month by 2018, which is about an 6-fold increase over 2014. In addition, the number of mobile devices and connections are expected to grow to 10.2 billion by 2018. New technologies are required to meet this demand. Related to wireless data traffic, the key parameter to consider is wireless throughput (bits/s) which is defined as:

$$\text{Throughput} = \text{Bandwidth (Hz)} \times \text{Spectral efficiency (bits/s/Hz)}.$$

Clearly, to improve the throughput, some new technologies which can increase the bandwidth or the spectral efficiency or both should be exploited. In this thesis, we focus on techniques which improve the spectral efficiency. A well-known way to increase the spectral efficiency is using multiple antennas at the transceivers.

In wireless communication, the transmitted signals are being attenuated by fading due to multipath propagation and by shadowing due to large obstacles between the transmitter and the receiver, yielding a fundamental challenge for reliable communication. Transmission with multiple-input multiple-output (MIMO) antennas is a well-known diversity technique to enhance the reliability of the communication. Furthermore, with multiple antennas, multiple streams can be sent out and hence, we can obtain a multiplexing gain which significantly improves the communication capacity. MIMO systems have gained significant attention for the past decades, and are now being incorporated into several new generation wireless standards (e.g., LTE-Advanced, 802.16m).

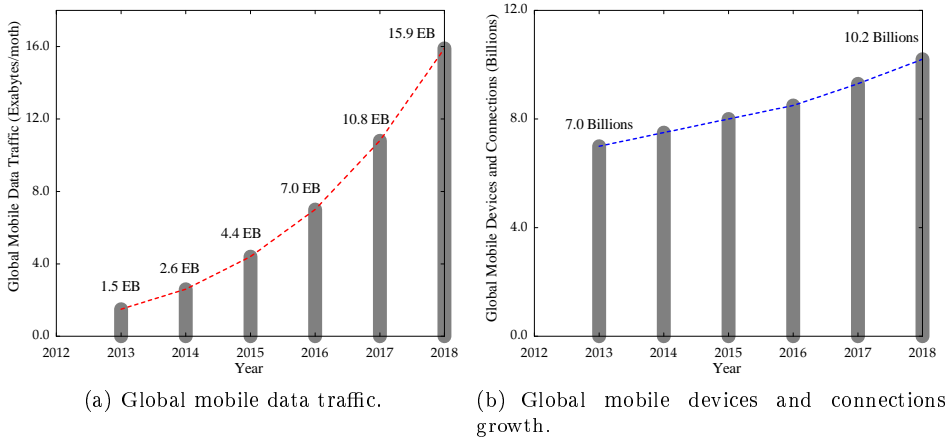


Figure 1.1: Demand for mobile data traffic and number of connected devices. (Source: Cisco [3])

The effort to exploit the spatial multiplexing gain has been shifted from MIMO to multiuser MIMO (MU-MIMO), where several users are simultaneously served by a multiple-antenna base station (BS). With MU-MIMO setups, a spatial multiplexing gain can be achieved even if each user has a single antenna [4]. This is important since users cannot support many antennas due to the small physical size and low-cost requirements of the terminals, whereas the BS can support many antennas. MU-MIMO does not only reap all benefits of MIMO systems, but also overcomes most of propagation limitations in MIMO such as ill-behaved channels. Specifically, by using scheduling schemes, we can reduce the limitations of ill-behaved channels. Line-of-sight propagation, which causes significant reduction of the performance of MIMO systems, is no longer a problem in MU-MIMO systems. Thus, MU-MIMO has attracted substantial interest [4–9].

There always exists a tradeoff between the system performance and the implementation complexity. The advantages of MU-MIMO come at a price:

- Multiuser interference: the performance of a given user may significantly degrade due to the interference from other users. To tackle this problem, interference reduction or cancellation techniques, such as maximum likelihood multiuser detection for the uplink [10], dirty paper coding (DPC) techniques for the downlink [11], or interference alignment [12], should be used. These techniques are complicated and have high computational complexity.
- Acquisition of channel state information: in order to achieve a high spatial multiplexing gain, the BS needs to process the received signals coherently. This requires accurate and timely acquisition of channel state information (CSI). This can be challenging, especially in high mobility scenarios.

- User scheduling: since several users are served on the same time-frequency resource, scheduling schemes which optimally select the group of users depending on the precoding/detection schemes, CSI knowledge etc., should be considered. This increases the cost of the system implementation.

The more antennas the BS is equipped with, the more degrees of freedom are offered and hence, more users can simultaneously communicate in the same time-frequency resource. As a result, a huge sum throughput can be obtained. With large antenna arrays, conventional signal processing techniques (e.g. maximum likelihood detection) become prohibitively complex due to the high signal dimensions. The main question is whether we can obtain the huge multiplexing gain with low-complexity signal processing and low-cost hardware implementation.

In [13], Marzetta showed that the use of an excessive number of BS antennas compared with the number of active users makes simple linear processing nearly optimal. More precisely, even with simple maximum-ratio combining (MRC) in the uplink or maximum-ratio transmission (MRT) in the downlink, the effects of fast fading, intracell interference, and uncorrelated noise tend to disappear as the number of BS station antennas grows large. MU-MIMO systems, where a BS with a hundred or more antennas simultaneously serves tens (or more) of users in the same time-frequency resource, are known as *Massive MIMO* systems (also called very large MU-MIMO, hyper-MIMO, or full-dimension MIMO systems). In Massive MIMO, it is expected that each antenna would be contained in an inexpensive module with simple processing and a low-power amplifier. The main benefits of Massive MIMO systems are:

- (1) *Huge spectral efficiency and high communication reliability*: Massive MIMO inherits all gains from conventional MU-MIMO, i.e., with M -antenna BS and K single-antenna users, we can achieve a diversity of order M and a multiplexing gain of $\min(M, K)$. By increasing both M and K , we can obtain a huge spectral efficiency and very high communication reliability.
- (2) *High energy efficiency*: In the uplink Massive MIMO, coherent combining can achieve a very high array gain which allows for substantial reduction in the transmit power of each user. In the downlink, the BS can focus the energy into the spatial directions where the terminals are located. As a result, with massive antenna arrays, the radiated power can be reduced by an order of magnitude, or more, and hence, we can obtain high energy efficiency. For a fixed number of users, by doubling the number of BS antennas, while reducing the transmit power by two, we can maintain the original the spectral efficiency, and hence, the radiated energy efficiency is doubled.
- (3) *Simple signal processing*: For most propagation environments, the use of an excessive number of BS antennas over the number of users yields favorable propagation where the channel vectors between the users and the BS are

pairwisely (nearly) orthogonal. Under favorable propagation, the effect of interuser interference and noise can be eliminated with simple linear signal processing (linear precoding in the downlink and linear decoding in the uplink). As a result, simple linear processing schemes are nearly optimal. Another key property of Massive MIMO is channel hardening. Under some conditions, when the number of BS antennas is large, the channel becomes (nearly) deterministic, and hence, the effect of small-scale fading is averaged out. The system scheduling, power control, etc., can be done over the large-scale fading time scale instead of over the small-scale fading time scale. This simplifies the signal processing significantly.

Massive MIMO is a promising candidate technology for next-generation wireless systems. Recently, there has been a great deal of interest in this technology [14–18]. Although there is much research work on this topic, a number of issues still need to be tackled before reducing Massive MIMO to practice [19–26].

Inspired by the above discussion, in this dissertation, we study the fundamentals of Massive MIMO including favorable propagation aspects, spectral and energy efficiency, and effects of finite-dimensional channel models. Capacity bounds are derived and analysed under practical constraints such as low-complexity processing, imperfect CSI, and intercell interference. Based on the fundamental analysis of Massive MIMO, resource allocation as well as system designs are also proposed.

In the following, brief introductions to multiuser MIMO and Massive MIMO are given in Chapter 2 and Chapter 3, respectively. In Chapter 4, we provide some mathematical preliminaries which will be used throughout the thesis. In Chapter 5, we list the specific contributions of the thesis together with a short description of the included papers. Finally, future research directions are discussed in Chapter 6.

Chapter 2

Mutiuser MIMO Cellular Systems

Massive MIMO is a MU-MIMO cellular system where the number of BS antennas and the number users are large. In this section, we will provide the basic background of MU-MIMO cellular systems in terms of communication schemes and signal detection, for both the uplink and downlink. For the sake of simplicity, we limit our discussions to the single-cell systems.

2.1 System Models and Assumptions

We consider a MU-MIMO system which consists of one BS and K active users. The BS is equipped with M antennas, while each user has a single-antenna. In general, each user can be equipped with multiple antennas. However, for simplicity of the analysis, we limit ourselves to systems with single-antenna users. See Figure 2.1. We assume that all K users share the same time-frequency resource. Furthermore, we assume that the BS and the users have perfect CSI. The channels are acquired at the BS and the users during the training phase. The specific training schemes depend on the system protocols (frequency-division duplex (FDD) or time-division duplex (TDD)), and will be discussed in detail in Section 2.5.

Let $\mathbf{H} \in \mathbb{C}^{M \times K}$ be the channel matrix between the K users and the BS antenna array, where the k th column of \mathbf{H} , denoted by \mathbf{h}_k , represents the $M \times 1$ channel vector between the k th user and the BS. In general, the propagation channel is modeled via large-scale fading and small-scale fading. But in this chapter, we ignore large-scale fading, and further assume that the elements of \mathbf{H} are i.i.d. Gaussian distributed with zero mean and unit variance.

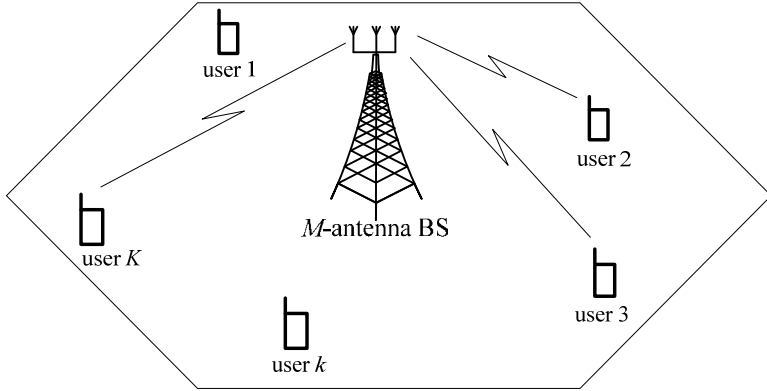


Figure 2.1: Multiuser MIMO Systems. Here, K single-antenna users are served by the M -antenna BS in the same time-frequency resource.

2.2 Uplink Transmission

Uplink (or reverse link) transmission is the scenario where the K users transmit signals to the BS. Let s_k , where $\mathbb{E}\{|s_k|^2\} = 1$, be the signal transmitted from the k th user. Since K users share the same time-frequency resource, the $M \times 1$ received signal vector at the BS is the combination of all signals transmitted from all K users:

$$\mathbf{y}_{\text{ul}} = \sqrt{p_{\text{u}}} \sum_{k=1}^K \mathbf{h}_k s_k + \mathbf{n} \quad (2.1)$$

$$= \sqrt{p_{\text{u}}} \mathbf{H} \mathbf{s} + \mathbf{n}, \quad (2.2)$$

where p_{u} is the average signal-to-noise ratio (SNR), $\mathbf{n} \in \mathbb{C}^{M \times 1}$ is the additive noise vector, and $\mathbf{s} \triangleq [s_1 \dots s_K]^T$. We assume that the elements of \mathbf{n} are i.i.d. Gaussian random variables (RVs) with zero mean and unit variance, and independent of \mathbf{H} .

From the received signal vector \mathbf{y}_{ul} together with knowledge of the CSI, the BS will coherently detect the signals transmitted from the K users. The channel model (2.2) is the multiple-access channel which has the sum-capacity [27]

$$C_{\text{ul,sum}} = \log_2 \det \left(\mathbf{I}_K + p_{\text{u}} \mathbf{H}^H \mathbf{H} \right). \quad (2.3)$$

The aforementioned sum-capacity can be achieved by using the successive interference cancellation (SIC) technique [28]. With SIC, after one user is detected, its signal is subtracted from the received signal before the next user is detected.

2.3 Downlink Transmission

Downlink (or forward link) is the scenario where the BS transmits signals to all K users. Let $\mathbf{x} \in \mathbb{C}^{M \times 1}$, where $\mathbb{E}\{\|\mathbf{x}\|^2\} = 1$, be the signal vector transmitted from the BS antenna array. Then, the received signal at the k th user is given by

$$y_{\text{dl},k} = \sqrt{p_d} \mathbf{h}_k^T \mathbf{x} + z_k, \quad (2.4)$$

where p_d is the average SNR and z_k is the additive noise at the k th user. We assume that z_k is Gaussian distributed with zero mean and unit variance. Collectively, the received signal vector of the K users can be written as

$$\mathbf{y}_{\text{dl}} = \sqrt{p_d} \mathbf{H}^T \mathbf{x} + \mathbf{z}, \quad (2.5)$$

where $\mathbf{y}_{\text{dl}} \triangleq [y_{\text{dl},1} \ y_{\text{dl},2} \ \dots \ y_{\text{dl},K}]^T$ and $\mathbf{z} \triangleq [z_1 \ z_2 \ \dots \ z_K]^T$. The channel model (2.5) is the broadcast channel whose sum-capacity is known to be

$$C_{\text{sum}} = \max_{\substack{\{q_k\} \\ q_k \geq 0, \sum_{k=1}^K q_k \leq 1}} \log_2 \det \left(\mathbf{I}_M + p_d \mathbf{H}^* \mathbf{D}_q \mathbf{H}^T \right), \quad (2.6)$$

where \mathbf{D}_q is the diagonal matrix whose k th diagonal element is q_k . The sum-capacity (2.6) can be achieved by using the dirty-paper coding (DPC) technique.

2.4 Linear Processing

To obtain optimal performance, complex signal processing techniques must be implemented. For example, in the uplink, the maximum-likelihood (ML) multiuser detection can be used. With ML multiuser detection, the BS has to search all possible transmitted signal vectors \mathbf{s} , and choose the best one as follows:

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathcal{S}^K} \|\mathbf{y}_{\text{ul}} - \sqrt{p_u} \mathbf{H} \mathbf{s}\|^2 \quad (2.7)$$

where \mathcal{S} is the finite alphabet of s_k , $k = 1, 2, \dots, K$. The problem (2.7) is a least-squares (LS) problem with a finite-alphabet constraint. The BS has to search over $|\mathcal{S}|^K$ vectors, where $|\mathcal{S}|$ denotes the cardinality of the set \mathcal{S} . Therefore, ML has a complexity which is exponential in the number of users.

The BS can use linear processing schemes (linear receivers in the uplink and linear precoders in the downlink) to reduce the signal processing complexity. These schemes are not optimal. However, when the number of BS antennas is large, it is shown in [13, 14] that linear processing is nearly-optimal. Therefore, in this thesis, we will consider linear processing. The details of linear processing techniques are presented in the following sections.

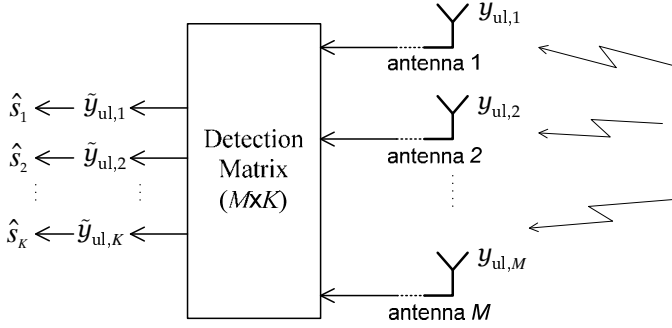


Figure 2.2: Block diagram of linear detection at the BS.

2.4.1 Linear Receivers (in the Uplink)

With linear detection schemes at the BS, the received signal \mathbf{y}_{ul} is separated into K streams by multiplying it with an $M \times K$ linear detection matrix, \mathbf{A} :

$$\tilde{\mathbf{y}}_{\text{ul}} = \mathbf{A}^H \mathbf{y}_{\text{ul}} = \sqrt{p_{\text{u}}} \mathbf{A}^H \mathbf{H} \mathbf{s} + \mathbf{A}^H \mathbf{n}. \quad (2.8)$$

Each stream is then decoded independently. See Figure 2.2. The complexity is on the order of $K|\mathcal{S}|$. From (2.8), the k th stream (element) of $\tilde{\mathbf{y}}_{\text{ul}}$, which is used to decode s_k , is given by

$$\tilde{y}_{\text{ul},k} = \underbrace{\sqrt{p_{\text{u}}} \mathbf{a}_k^H \mathbf{h}_k s_k}_{\text{desired signal}} + \underbrace{\sqrt{p_{\text{u}}} \sum_{k' \neq k}^K \mathbf{a}_k^H \mathbf{h}_{k'} s_{k'}}_{\text{interuser interference}} + \underbrace{\mathbf{a}_k^H \mathbf{n}}_{\text{noise}}, \quad (2.9)$$

where \mathbf{a}_k denotes the k th column of \mathbf{A} . The interference plus noise is treated as effective noise, and hence, the received signal-to-interference-plus-noise ratio (SINR) of the k th stream is given by

$$\text{SINR}_k = \frac{p_{\text{u}} |\mathbf{a}_k^H \mathbf{h}_k|^2}{p_{\text{u}} \sum_{k' \neq k}^K |\mathbf{a}_k^H \mathbf{h}_{k'}|^2 + \|\mathbf{a}_k\|^2}. \quad (2.10)$$

We now review some conventional linear multiuser receivers.

a) *Maximum-Ratio Combining receiver:*

With MRC, the BS aims to maximize the received signal-to-noise ratio (SNR) of each stream, ignoring the effect of multiuser interference. From (2.9), the

k th column of the MRC receiver matrix \mathbf{A} is:

$$\begin{aligned}\mathbf{a}_{\text{mrc},k} &= \underset{\mathbf{a}_k \in \mathbb{C}^{M \times 1}}{\operatorname{argmax}} \frac{\text{power}(\text{desired signal})}{\text{power}(\text{noise})} \\ &= \underset{\mathbf{a}_k \in \mathbb{C}^{M \times 1}}{\operatorname{argmax}} \frac{p_u |\mathbf{a}_k^H \mathbf{h}_k|^2}{\|\mathbf{a}_k\|^2}.\end{aligned}\quad (2.11)$$

Since

$$\frac{p_u |\mathbf{a}_k^H \mathbf{h}_k|^2}{\|\mathbf{a}_k\|^2} \leq \frac{p_u \|\mathbf{a}_k\|^2 \|\mathbf{h}_k\|^2}{\|\mathbf{a}_k\|^2} = p_u \|\mathbf{h}_k\|^2,$$

and equality holds when $\mathbf{a}_k = \text{const} \cdot \mathbf{h}_k$, the MRC receiver is: $\mathbf{a}_{\text{mrc},k} = \text{const} \cdot \mathbf{h}_k$. Plugging $\mathbf{a}_{\text{mrc},k}$ into (2.10), the received SINR of the k th stream for MRC is given by

$$\text{SINR}_{\text{mrc},k} = \frac{p_u \|\mathbf{h}_k\|^4}{p_u \sum_{k' \neq k}^K |\mathbf{h}_k^H \mathbf{h}_{k'}|^2 + \|\mathbf{h}_k\|^2} \quad (2.12)$$

$$\rightarrow \frac{\|\mathbf{h}_k\|^4}{\sum_{k' \neq k}^K |\mathbf{h}_k^H \mathbf{h}_{k'}|^2}, \text{ as } p_u \rightarrow \infty. \quad (2.13)$$

- Advantage: the signal processing is very simple since the BS just multiplies the received vector with the conjugate-transpose of the channel matrix \mathbf{H} , and then detects each stream separately. **More importantly, MRC can be implemented in a distributed manner.** Furthermore, at low p_u , $\text{SINR}_{\text{mrc},k} \approx p_u \|\mathbf{h}_k\|^2$. This implies that at low SNR, MRC can achieve the same array gain as in the case of a single-user system.
- Disadvantage: as discussed above, since MRC neglects the effect of multiuser interference, it performs poorly in **interference-limited** scenarios. This can be seen in (2.13), where the SINR is upper bounded by a constant (with respect to p_u) when p_u is large.

b) *Zero-Forcing Receiver:*

By contrast to MRC, zero-forcing (ZF) receivers take the interuser interference into account, but neglect the effect of noise. With ZF, the multiuser interference is completely nulled out by projecting each stream onto the orthogonal complement of the interuser interference. More precisely, the k th column of the ZF receiver matrix satisfies:

$$\begin{cases} \mathbf{a}_{\text{zf},k}^H \mathbf{h}_k \neq 0 \\ \mathbf{a}_{\text{zf},k}^H \mathbf{h}_{k'} = 0, \quad \forall k' \neq k. \end{cases} \quad (2.14)$$

The ZF receiver matrix, which satisfies (2.14) for all k , is the pseudo-inverse of the channel matrix \mathbf{H} . With ZF, we have

$$\tilde{\mathbf{y}}_{\text{ul}} = \left(\mathbf{H}^H \mathbf{H} \right)^{-1} \mathbf{H}^H \mathbf{y}_{\text{ul}} = \sqrt{p_u} \mathbf{s} + \left(\mathbf{H}^H \mathbf{H} \right)^{-1} \mathbf{H}^H \mathbf{n}. \quad (2.15)$$

This scheme requires that $M \geq K$ (so that the matrix $\mathbf{H}^H \mathbf{H}$ is invertible). We can see that each stream (element) of $\tilde{\mathbf{y}}_{\text{ul}}$ in (2.15) is free of multiuser interference. The k th stream of $\tilde{\mathbf{y}}_{\text{ul}}$ is used to detect s_k :

$$\tilde{y}_{\text{ul},k} = \sqrt{p_u} s_k + \tilde{n}_k, \quad (2.16)$$

where \tilde{n}_k denotes the k th element of $\left(\mathbf{H}^H \mathbf{H}\right)^{-1} \mathbf{H}^H \mathbf{n}$. Thus, the received SINR of the k th stream is given by

$$\text{SINR}_{\text{zf},k} = \frac{p_u}{\left[\left(\mathbf{H}^H \mathbf{H}\right)^{-1}\right]_{kk}}. \quad (2.17)$$

- Advantage: the signal processing is simple and ZF works well in interference-limited scenarios. The SINR can be made as high as desired by increasing the transmit power.
- Disadvantage: since ZF neglects the effect of noise, it works poorly under noise-limited scenarios. Furthermore, if the channel is not well-conditioned then the pseudo-inverse amplifies the noise significantly, and hence, the performance is very poor. Compared with MRC, ZF has a higher implementation complexity due to the computation of the pseudo-inverse of the channel gain matrix.

c) *Minimum Mean-Square Error Receiver:*

The linear minimum mean-square error (MMSE) receiver aims to minimize the mean-square error between the estimate $\mathbf{A}^H \mathbf{y}_{\text{ul}}$ and the transmitted signal \mathbf{s} . More precisely,

$$\mathbf{A}_{\text{mmse}} = \arg \min_{\mathbf{A} \in \mathbb{C}^{M \times K}} \mathbb{E} \left\{ \left\| \mathbf{A}^H \mathbf{y}_{\text{ul}} - \mathbf{s} \right\|^2 \right\} \quad (2.18)$$

$$= \arg \min_{\mathbf{A} \in \mathbb{C}^{M \times K}} \sum_{k=1}^K \mathbb{E} \left\{ |\mathbf{a}_k^H \mathbf{y}_{\text{ul}} - s_k|^2 \right\}. \quad (2.19)$$

where \mathbf{a}_k is the k th column of \mathbf{A} . Therefore, the k th column of the MMSE receiver matrix is [47]

$$\mathbf{a}_{\text{mmse},k} = \arg \min_{\mathbf{a}_k \in \mathbb{C}^{M \times 1}} \mathbb{E} \left\{ |\mathbf{a}_k^H \mathbf{y}_{\text{ul}} - s_k|^2 \right\} \quad (2.20)$$

$$= \text{cov}(\mathbf{y}_{\text{ul}}, \mathbf{y}_{\text{ul}})^{-1} \text{cov}(s_k, \mathbf{y}_{\text{ul}})^H \quad (2.21)$$

$$= \sqrt{p_u} \left(p_u \mathbf{H} \mathbf{H}^H + \mathbf{I}_M \right)^{-1} \mathbf{h}_k, \quad (2.22)$$

where $\text{cov}(\mathbf{v}_1, \mathbf{v}_2) \triangleq \mathbb{E} \{ \mathbf{v}_1 \mathbf{v}_2^H \}$, where \mathbf{v}_1 and \mathbf{v}_2 are two random column vectors with zero-mean elements.

It is known that the MMSE receiver maximizes the received SINR. Therefore, among the MMSE, ZF, and MRC receivers, MMSE is the best. We can see

from (2.22) that, at high SNR (high p_u), ZF approaches MMSE, while at low SNR, MRC performs as well as MMSE. Furthermore, substituting (2.22) into (2.10), the received SINR for the MMSE receiver is given by

$$\text{SINR}_{\text{mmse},k} = p_u \mathbf{h}_k^H \left(p_u \sum_{i \neq k}^K \mathbf{h}_i \mathbf{h}_i^H + \mathbf{I}_M \right)^{-1} \mathbf{h}_k. \quad (2.23)$$

2.4.2 Linear Precoders (in the Downlink)

In the downlink, with linear precoding techniques, the signal transmitted from M antennas, \mathbf{x} , is a linear combination of the symbols intended for the K users. Let q_k , $\mathbb{E}\{|q_k|^2\} = 1$, be the symbol intended for the k th user. Then, the linearly precoded signal vector \mathbf{x} is

$$\mathbf{x} = \sqrt{\alpha} \mathbf{W} \mathbf{q}, \quad (2.24)$$

where $\mathbf{q} \triangleq [q_1 \ q_2 \ \dots \ q_K]^T$, $\mathbf{W} \in \mathbb{C}^{M \times K}$ is the precoding matrix, and α is a normalization constant chosen to satisfy the power constraint $\mathbb{E}\{\|\mathbf{x}\|^2\} = 1$. Thus,

$$\alpha = \frac{1}{\mathbb{E}\{\text{tr}(\mathbf{W} \mathbf{W}^H)\}}. \quad (2.25)$$

A block diagram of the linear precoder at the BS is shown in Figure 2.3.

Plugging (2.24) into (2.4), we obtain

$$y_{\text{dl},k} = \sqrt{\alpha p_d} \mathbf{h}_k^T \mathbf{W} \mathbf{q} + z_k \quad (2.26)$$

$$= \sqrt{\alpha p_d} \mathbf{h}_k^T \mathbf{w}_k q_k + \sqrt{\alpha p_d} \sum_{k' \neq k}^K \mathbf{h}_k^T \mathbf{w}_{k'} q_{k'} + z_k. \quad (2.27)$$

Therefore, the SINR of the transmission from the BS to the k th user is

$$\text{SINR}_k = \frac{\alpha p_d \left| \mathbf{h}_k^T \mathbf{w}_k \right|^2}{\alpha p_d \sum_{k' \neq k}^K \left| \mathbf{h}_k^T \mathbf{w}_{k'} \right|^2 + 1}. \quad (2.28)$$

Three conventional linear precoders are maximum-ratio transmission (MRT) (also called conjugate beamforming), ZF, and MMSE precoders. These precoders have similar operational meanings and properties as MRC, ZF, MMSE receivers, respectively. Thus, here we just provide the final formulas for these precoders, i.e.,

$$\mathbf{W} = \begin{cases} \mathbf{H}^*, & \text{for MRT} \\ \mathbf{H}^* (\mathbf{H}^T \mathbf{H}^*)^{-1}, & \text{for ZF} \\ \mathbf{H}^* (\mathbf{H}^T \mathbf{H}^* + \frac{K}{p_d} \mathbf{I}_K)^{-1}, & \text{for MMSE} \end{cases} \quad (2.29)$$

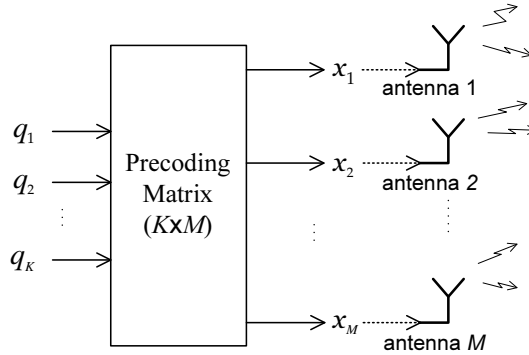


Figure 2.3: Block diagram of the linear precoders at the BS.

Figures 2.4 and 2.5 show the achievable sum rates for the uplink and the downlink transmission, respectively, with different linear processing schemes, versus $\text{SNR} \triangleq p_u$ for the uplink and $\text{SNR} \triangleq p_d$ for the downlink, with $M = 6$ and $K = 4$. The sum rate is defined as $\sum_{k=1}^K \mathbb{E} \{\log_2 (1 + \text{SINR}_k)\}$, where SINR_k is the SINR of the k th user which is given in the previous discussion. As expected, MMSE performs strictly better than ZF and MRC over the entire range of SNRs. In the low SNR regime, MRC is better than ZF, and vice versa in the high SNR regime.

2.5 Channel Estimation

We have assumed so far that the BS and the users have perfect CSI. However, in practice, this CSI has to be estimated. Depending on the system duplexing mode (TDD or FDD), the channel estimation schemes are very different.

2.5.1 Channel Estimation in TDD Systems

In a TDD system, the uplink and downlink transmissions use the same frequency spectrum, but different time slots. The uplink and downlink channels are reciprocal.¹ Thus, the CSI can be obtained by using following scheme (see Figure 2.6):

- For the uplink transmission: the BS needs CSI to detect the signals transmitted from the K users. This CSI is estimated at the BS. More precisely, the K users send K orthogonal pilot sequences to the BS on the uplink. Then the BS estimates the channels based on the received pilot signals. This process requires a minimum of K channel uses.

¹In practice, the uplink and downlink channels are not perfectly reciprocal due to mismatches of the hardware chains. This non-reciprocity can be removed by calibration [15, 29, 30]. In our work, we assume that the hardware chain calibration is perfect.

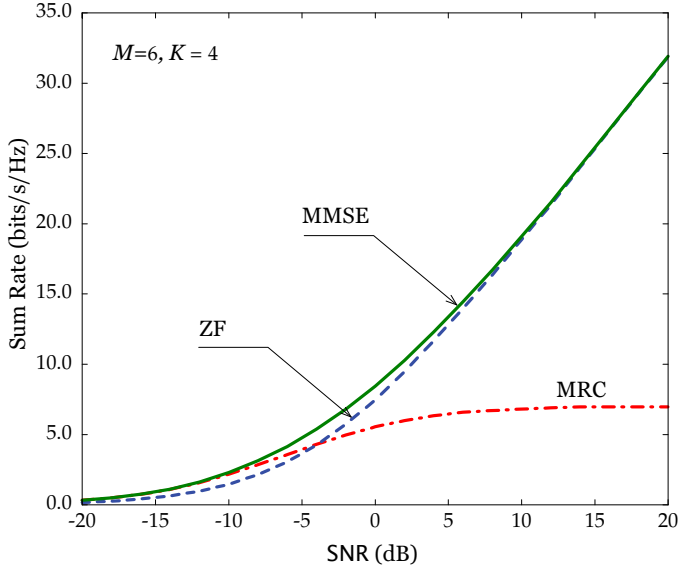


Figure 2.4: Performance of linear receivers in the uplink.

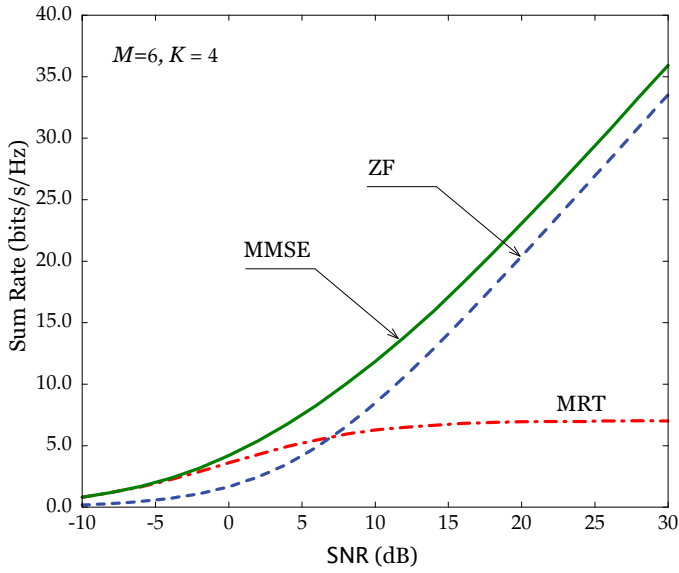


Figure 2.5: Performance of linear precoders in the downlink.

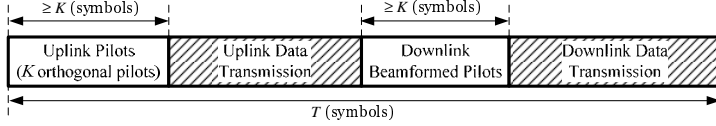


Figure 2.6: Slot structure and channel estimation in TDD systems.

- For the downlink: the BS needs CSI to precode the transmitted signals, while each user needs the effective channel gain to detect the desired signals. Due to the channel reciprocity, the channel estimated at the BS in the uplink can be used to precode the transmit symbols. To obtain knowledge of the effective channel gain, the BS can beamform pilots, and each user can estimate the effective channel gains based on the received pilot signals. This requires at least K channel uses.²

In total, the training process requires a minimum of $2K$ channel uses. We assume that the channel stays constant over T symbols. Thus, it is required that $2K < T$. An illustration of channel estimation in TDD systems is shown in Figure 2.6.

2.5.2 Channel Estimation in FDD Systems

In an FDD system, the uplink and downlink transmissions use different frequency spectrum, and hence, the uplink and downlink channels are not reciprocal. The channel knowledge at the BS and users can be obtained by using following training scheme:

- For the downlink transmission: the BS needs CSI to precode the symbols before transmitting to the K users. The M BS antennas transmit M orthogonal pilot sequences to K users. Each user will estimate the channel based on the received pilots. Then it feeds back its channel estimates (M channel estimates) to the BS through the uplink. This process requires at least M channel uses for the downlink and M channel uses for the uplink.
- For the uplink transmission: the BS needs CSI to decode the signals transmitted from the K users. One simple way is that the K users transmit K orthogonal pilot sequences to the BS. Then, the BS will estimate the channels based on the received pilot signals. This process requires at least K channel uses for the uplink.

²The effective channel gains at the users may be blindly estimated based on the received data, and hence, no pilots are required [31]. But, we do not discuss in detail about this possibility in this section.

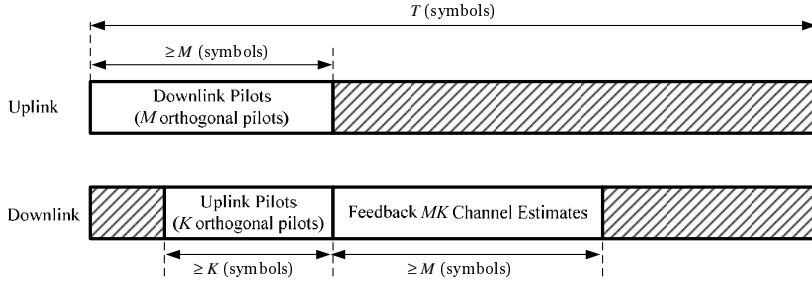


Figure 2.7: Slot structure and channel estimation in FDD systems.

Therefore, the entire channel estimation process requires at least $M + K$ channel uses in the uplink and M channel uses in the downlink. Assume that the lengths of the coherence intervals for the uplink and the downlink are the same and are equal to T . Then we have the constraints: $M < T$ and $M + K < T$. As a result $M + K < T$ is the constraint for FDD systems. An illustration of channel estimation in FDD systems is shown in Figure 2.7.

Chapter 3

Massive MIMO

3.1 What is Massive MIMO?

Massive MIMO is a form of MU-MIMO systems where the number of BS antennas and the numbers of users are large. In Massive MIMO, hundreds or thousands of BS antennas simultaneously serve tens or hundreds of users in the same frequency resource. Some main points of Massive MIMO are:

- TDD operation: as discussed in Section 2.5, with FDD, the channel estimation overhead depends on the number of BS antennas, M . By contrast, with TDD, the channel estimation overhead is independent of M . In Massive MIMO, M is large, and hence, TDD operation is preferable. For example, assume that the coherence interval is $T = 200$ symbols (corresponding to a coherence bandwidth of 200 kHz and a coherence time of 1 ms). Then, in FDD systems, the number of BS antennas and the number of users are constrained by $M + K < 200$, while in TDD systems, the constraint on M and K is $2K < 200$. Figure 3.1 shows the regions of feasible (M, K) in FDD and TDD systems. We can see that the FDD region is much smaller than the TDD region. With TDD, adding more antennas does not affect the resources needed for the channel estimation.
- Linear processing: since the number of BS antennas and the number of users are large, the signal processing at the terminal ends must deal with large dimensional matrices/vectors. Thus, simple signal processing is preferable. In Massive MIMO, linear processing (linear combining schemes in the uplink and linear precoding schemes in the downlink) is nearly optimal.

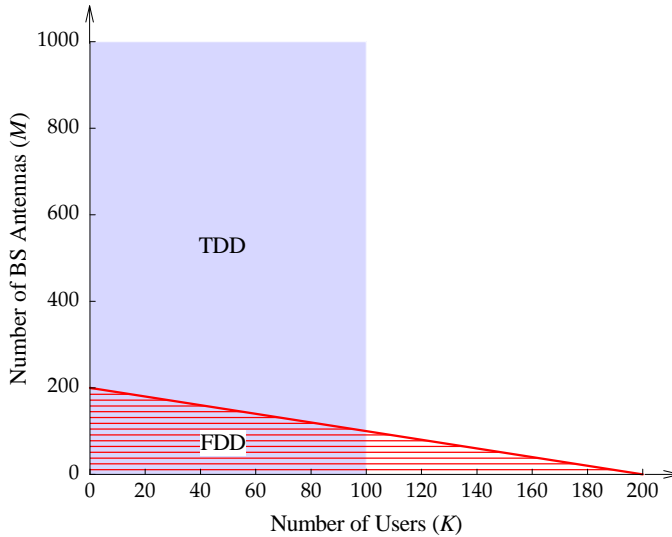


Figure 3.1: The regions of possible (M, K) in TDD and FDD systems, for a coherence interval of 200 symbols.

- Favorable propagation: favorable propagation means that the channel matrix between the BS antenna array and the users is well-conditioned. In Massive MIMO, under some conditions, the favorable propagation property holds due to the law of large numbers.
- A massive BS antenna array does not have to be physically large. For example consider a cylindrical array with 128 antennas, comprising four circles of 16 dual-polarized antenna elements. At 2.6 GHz, the distance between adjacent antennas is about 6 cm, which is half a wavelength, and hence, this array occupies only a physical size of $28\text{cm} \times 29\text{cm}$ [25].
- Massive MIMO is scalable: in Massive MIMO, the BS learns the channels via uplink training, under TDD operation. The time required for channel estimation is independent of the number of BS antennas. Therefore, the number of BS antennas can be made as large as desired with no increase in the channel estimation overhead. Furthermore, the signal processing at each user is very simple and does not depend on other users' existence, i.e., no multiplexing or de-multiplexing signal processing is performed at the users. Adding or dropping some users from service does not affect other users' activities.
- All the complexity is at the BS.

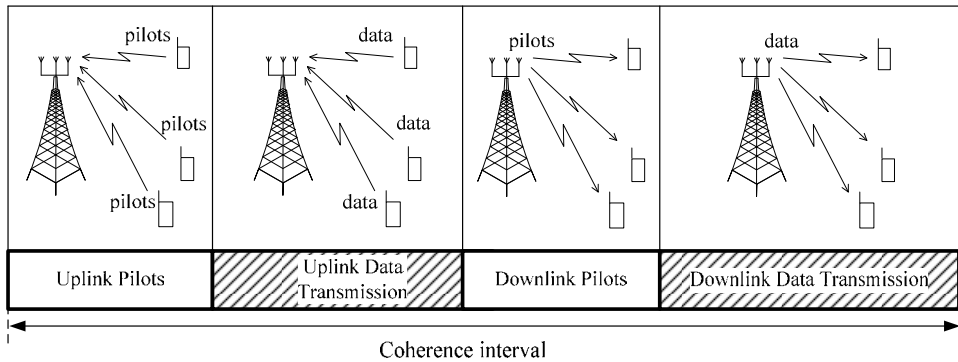


Figure 3.2: Transmission protocol of TDD Massive MIMO.

3.2 How Massive MIMO Works

In Massive MIMO, TDD operation is preferable. During a coherence interval, there are three operations: channel estimation (including the uplink training and the downlink training), uplink data transmission, and downlink data transmission. A TDD Massive MIMO protocol is shown in Figure 3.2.

3.2.1 Channel Estimation

The BS needs CSI to detect the signals transmitted from the users in the uplink, and to precode the signals in the downlink. This CSI is obtained through the uplink training. Each user is assigned an orthogonal pilot sequence, and sends this pilot sequence to the BS. The BS knows the pilots sequences transmitted from all users, and then estimates the channels based the received pilot signals. The estimation schemes were discussed in detail in Section 2.5.1.

Furthermore, each user may need partial knowledge of CSI to coherently detect the signals transmitted from the BS. This information can be acquired through downlink training or some blind channel estimation algorithm. Since the BS uses linear precoding techniques to beamform the signals to the users, the user needs only the effective channel gain (which is a scalar constant) to detect its desired signals. Therefore, the BS can spend a short time to beamform pilots in the downlink for CSI acquisition at the users.

3.2.2 Uplink Data Transmission

A part of the coherence interval is used for the uplink data transmission. In the uplink, all K users transmit their data to the BS in the same time-frequency resource. The BS then uses the channel estimates together with the linear combining

techniques to detect signals transmitted from all users. The detailed uplink data transmission was discussed in Section 2.2.

3.2.3 Downlink Data Transmission

In the downlink, the BS transmits signals to all K users in the same time-frequency resource. More specifically, the BS uses its channel estimates in combination with the symbols intended for the K users to create M precoded signals which are then fed to M antennas. The downlink data transmission was discussed in detail in Section 2.3.

3.3 Why Massive MIMO

The demand for wireless throughput and communication reliability as well as the user density will always increase. Future wireless communication requires new technologies in which many users can be simultaneously served with very high throughput. Massive MIMO can meet these demands. Consider the uplink transmission. (The same argument can be used for the downlink transmission.) From (2.3), under the conditions of favorable propagation (the channel vectors between the users and the BS are pairwise orthogonal), the sum-capacity of the uplink transmission is

$$C_{\text{sum}} = \log_2 \det(\mathbf{I}_K + p_u M \mathbf{I}_K) = K \log_2(1 + M p_u). \quad (3.1)$$

In (3.1), K is the multiplexing gain, and M represents the array gain. We can see that, we can obtain a huge spectral efficiency and energy efficiency when M and K are large. Without any increase in transmitted power per terminal, by increasing K and M , we can simultaneously serve more users in the same frequency band. At the same time the throughput per user also increases. Furthermore, by doubling the number of BS antennas, we can reduce the transmit power by 3 dB, while maintaining the original quality-of-service.

The above gains (multiplexing gain and array gain) are obtained under the conditions of favorable propagation and the use of optimal processing at the BS. One main question is: Will these gains still be obtained by using linear processing? Another question is: Why not use the conventional low dimensional point-to-point MIMO with complicated processing schemes instead of Massive MIMO with simple linear processing schemes? In Massive MIMO, when the number of BS antennas is large, due to the law of large numbers, the channels become favorable. As a result, linear processing is nearly optimal. The multiplexing gain and array gain can be obtained with simple linear processing. Also, by increasing the number of BS antennas and the number of users, we can always increase the throughput.

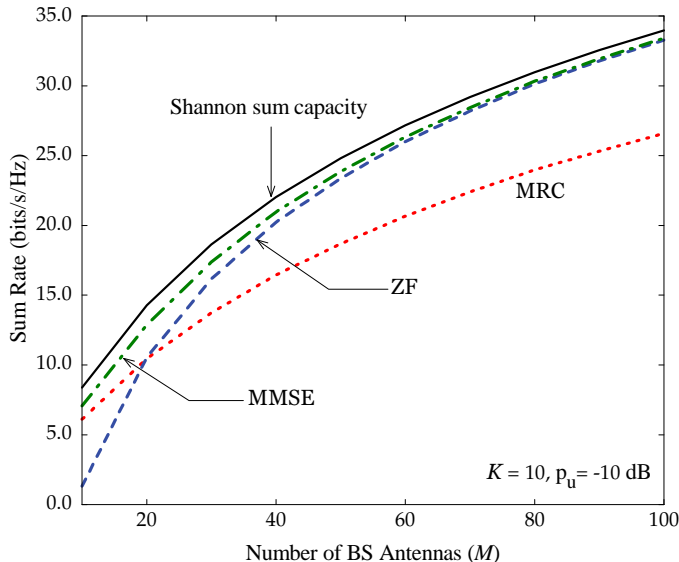


Figure 3.3: Uplink sum rate for different linear receivers and for the optimal receiver.

Figure 3.3 shows the sum rate versus the number of BS antennas with optimal receivers (the sum capacity is achieved) and linear receivers, at $K = 10$ and $p_u = -10$ dB. The sum capacity is computed from (2.3), while the sum rates for MRC, ZF, and MMSE are computed by using (2.12), (2.17), and (2.23), respectively. We can see that, when M is large, the sum rate with linear processing is very close to the sum capacity obtained by using optimal receivers. When $M = K = 10$, and with the optimal receiver, the maximum sum rate that we can obtain is 8.5 bits/s/Hz. By contrast, by using large M , say $M = 50$, with simple ZF receivers, we can obtain a sum rate of 24 bits/s/Hz.

3.4 Challenges in Massive MIMO

Despite the huge advantages of Massive MIMO, many issues still need to be tackled. The main challenges of Massive MIMO are listed as follows:

3.4.1 Pilot Contamination

In previous sections, we considered single-cell setups. However, practical cellular networks consist of many cells. Owing to the limited availability of frequency spectrum, many cells have to share the same time-frequency resources. Thus, multicell

setups should be considered. In multicell systems, we cannot assign orthogonal pilot sequences for all users in all cells, due to the limitation of the channel coherence interval. Orthogonal pilot sequences have to be reused from cell to cell. Therefore, the channel estimate obtained in a given cell will be contaminated by pilots transmitted by users in other cells. This effect, called “pilot contamination”, reduces the system performance [32]. The effect of pilot contamination is major inherent limitation of Massive MIMO. It does not vanish even when the number of BS antennas grows without bound. Considerable efforts have been made to reduce this effect. The eigenvalue-decomposition-based channel estimation, pilot decontamination, as well as pilot contamination precoding schemes are proposed in [33–35]. In [36], the authors shown that, under certain conditions of the channel covariance, by using a covariance aware pilot assignment scheme among the cells, pilot contamination can be efficiently mitigated. There is much ongoing research on this topic.

3.4.2 Unfavorable Propagation

Massive MIMO works under favorable propagation environments. However, in practice, there may be propagation environments where the channels are not favorable. For example, in propagation environments where the numbers of the scatterers is small compared to the numbers of users, or the channels from different users to the BS share some common scatterers, the channel is not favorable [31]. One possibility to tackle this problem is to distribute the BS antennas over a large area.

3.4.3 New Standards and Designs are Required

It will be very efficient if Massive MIMO can be deployed in current systems such as LTE. However, the LTE standard only allows for up to 8 antenna ports at the BS [4]. Furthermore, LTE uses the channel information that is “assumed”. For example, one option of the downlink in LTE is that the BS transmits the reference signals through several fixed beams. Then the users report back to the BS the strongest beam. The BS will use this beam for the downlink transmission. By contrast, Massive MIMO uses the channel information that is estimated (measured). Therefore, to reduce Massive MIMO to practice, new standards are required. On a different note, with Massive MIMO, a costly 40 Watt transceiver should be replaced by a large number of low-power and inexpensive antennas. Related hardware designs should also be considered. This requires a huge effort from both academia and industry.

Chapter 4

Mathematical Preliminaries

4.1 Random Matrix Theory

We now review some useful limit results about very long random vectors [37] which will be used for the analysis in the rest of the thesis.

- Let $\mathbf{p} \triangleq [p_1 \dots p_n]^T$ and $\mathbf{q} \triangleq [q_1 \dots q_n]^T$ be $n \times 1$ vectors whose elements are independent identically distributed (i.i.d.) random variables (RVs) with $\mathbb{E}\{p_i\} = \mathbb{E}\{q_i\} = 0$, $\mathbb{E}\{|p_i|^2\} = \sigma_p^2$, and $\mathbb{E}\{|q_i|^2\} = \sigma_q^2$, $i = 1, 2, \dots, n$. Assume that \mathbf{p} and \mathbf{q} are independent.

Applying the law of large numbers, we obtain

$$\frac{1}{n} \mathbf{p}^H \mathbf{p} \xrightarrow{a.s.} \sigma_p^2, \text{ as } n \rightarrow \infty, \quad (4.1)$$

$$\frac{1}{n} \mathbf{p}^H \mathbf{q} \xrightarrow{a.s.} 0, \text{ as } n \rightarrow \infty, \quad (4.2)$$

where $\xrightarrow{a.s.}$ denotes almost sure convergence.

Applying the Lindeberg-Lévy central limit theorem, we obtain

$$\frac{1}{\sqrt{n}} \mathbf{p}^H \mathbf{q} \xrightarrow{d} \mathcal{CN}(0, \sigma_p^2 \sigma_q^2), \text{ as } n \rightarrow \infty, \quad (4.3)$$

where \xrightarrow{d} denotes convergence in distribution.

- Let X_1, X_2, \dots be a sequence of independent circularly symmetric complex RVs, such that X_i has zero mean and variance σ_i^2 . Further assume that the following conditions are satisfied: 1) $s_n^2 = \sum_{i=1}^n \sigma_i^2 \rightarrow \infty$, as $n \rightarrow \infty$; and 2) $\sigma_i/s_n \rightarrow 0$, as $n \rightarrow \infty$. Then by applying the Cramér's central limit theorem, we have

$$\frac{\sum_{i=1}^n X_i}{s_n} \xrightarrow{d} \mathcal{CN}(0, 1), \text{ as } n \rightarrow \infty. \quad (4.4)$$

- Let X_1, X_2, \dots, X_n be independent RVs such that $\mathbb{E}\{X_i\} = \mu_i$ and $\text{Var}(X_i) < c < \infty, \forall i = 1, \dots, n$. Then by applying the Tchebyshev's theorem, we have

$$\frac{1}{n} (X_1 + X_2 + \dots + X_n) - \frac{1}{n} (\mu_1 + \mu_2 + \dots + \mu_n) \xrightarrow{P} 0, \quad (4.5)$$

where \xrightarrow{P} denotes convergence in probability.

4.2 Capacity Lower Bounds

In this section, we derive a capacity lower bound for a SISO channel with interference and with partial/perfect CSI at the receivers. The channel model is:

$$y_k = \underbrace{a_k s_k}_{\text{desired signal}} + \underbrace{\sum_{k' \neq k}^K a_{k'} s_{k'}}_{\text{interference}} + \underbrace{n_k}_{\text{noise}}, \quad (4.6)$$

where $a_k, k = 1, \dots, K$, are the effective channel gains, s_k , is the transmitted signals from the k th sources, and n_k is the additive noise. Assume that $s_k, k = 1, \dots, K$, and n_k are independent RVs with zero-mean and unit variance.

Since in this thesis, we consider linear processing, the end-to-end channel can be considered as an interference SISO channel, and can be modeled as in (4.6). For example, consider the downlink transmission discussed in Section 2.4. The received signal at the k th user is given by (2.27) which precisely matches with the model (4.6), where $a_{k'}$ is $\sqrt{\alpha p_d} \mathbf{h}_k^T \mathbf{w}_{k'}, k' = 1, \dots, K$. The capacity lower bound derived in this section will be used throughout the thesis.

Let \mathcal{C} be the channel state information (CSI) available at the receiver. We assume that $s_k \sim \mathcal{CN}(0, 1)$. In general, Gaussian signaling is not optimal, and hence, this assumption yields a lower bound on the capacity:

$$C_k = I(s_k; y_k, \mathcal{C}) = h(s_k) - h(s_k | y_k, \mathcal{C}) \quad (4.7)$$

$$\stackrel{(a)}{=} \log_2(\pi e) - h(s_k - \alpha y_k | y_k, \mathcal{C}) \quad (4.8)$$

$$\stackrel{(b)}{\geq} \log_2(\pi e) - h(s_k - \alpha y_k | \mathcal{C}), \quad (4.9)$$

where (a) holds for any α , and (b) follows from the fact that conditioning reduces the entropy. Note that, in (4.7), $I(x; y)$ and $h(x)$ denote the mutual information between x and y , and the differential entropy of x , respectively.

Since the differential entropy of a RV with fixed variance is maximized when the RV is Gaussian, we obtain

$$C_k \geq \log_2(\pi e) - \mathbb{E} \left\{ \log_2 \left(\pi e \mathbb{E} \left\{ |s_k - \alpha y_k|^2 \middle| \mathcal{C} \right\} \right) \right\}, \quad (4.10)$$

which leads to

$$C_k \geq \mathbb{E} \left\{ \log_2 \left(\frac{1}{\mathbb{E} \left\{ |s_k - \alpha y_k|^2 \middle| \mathcal{C} \right\}} \right) \right\}. \quad (4.11)$$

To obtain the tightest bound, we choose $\alpha = \alpha_0$ so that $\mathbb{E} \left\{ |x_k - \alpha_0 y_k|^2 \middle| \mathcal{C} \right\}$ is minimized:

$$\alpha_0 = \arg \min_{\alpha} \mathbb{E} \left\{ |s_k - \alpha y_k|^2 \middle| \mathcal{C} \right\}. \quad (4.12)$$

We can see that $\alpha_0 y_k$ is the LMMSE estimate of s_k . Therefore,

$$\alpha_0 = \frac{\mathbb{E} \{ y_k^* s_k | \mathcal{C} \}}{\mathbb{E} \{ |y_k|^2 | \mathcal{C} \}} = \frac{\mathbb{E} \{ a_k^* | \mathcal{C} \}}{\mathbb{E} \{ |y_k|^2 | \mathcal{C} \}}. \quad (4.13)$$

We have

$$\begin{aligned} \mathbb{E} \left\{ |s_k - \alpha y_k|^2 \middle| \mathcal{C} \right\} &= \mathbb{E} \left\{ |s_k|^2 \right\} - \alpha^* \mathbb{E} \{ s_k y_k^* | \mathcal{C} \} - \alpha \mathbb{E} \{ s_k^* y_k | \mathcal{C} \} + |\alpha|^2 \mathbb{E} \left\{ |y_k|^2 \middle| \mathcal{C} \right\} \\ &= 1 - \alpha^* \mathbb{E} \{ a_k^* | \mathcal{C} \} - \alpha \mathbb{E} \{ a_k | \mathcal{C} \} + |\alpha|^2 \mathbb{E} \left\{ |y_k|^2 \middle| \mathcal{C} \right\}. \end{aligned} \quad (4.14)$$

When $\alpha = \alpha_0$, substituting (4.13) into (4.14), we get

$$\mathbb{E} \left\{ |x_k - \alpha y_k|^2 \middle| \mathcal{C} \right\} = 1 - \frac{|\mathbb{E} \{ a_k^* | \mathcal{C} \}|^2}{\sum_{k=1}^K \mathbb{E} \left\{ |a_k|^2 \middle| \mathcal{C} \right\} + 1}. \quad (4.15)$$

Plugging (4.15) into (11), we obtain the following lower bound on the capacity of (4.6):

$$C_k \geq \mathbb{E} \left\{ \log_2 \left(1 + \frac{|\mathbb{E} \{ a_k | \mathcal{C} \}|^2}{\sum_{k'=1}^K \mathbb{E} \left\{ |a_{k'}|^2 \middle| \mathcal{C} \right\} - |\mathbb{E} \{ a_k | \mathcal{C} \}|^2 + 1} \right) \right\}. \quad (4.16)$$

We next consider two special cases:

1. No instantaneous CSI: for this case, $\mathcal{C} = \emptyset$. Therefore, from (2), we obtain

$$C_k \geq \log_2 \left(1 + \frac{|\mathbb{E}\{a_k\}|^2}{\mathbb{E}\{|a_k - \mathbb{E}\{a_k\}|^2\} + \sum_{k' \neq k}^K \mathbb{E}\{|a_{k'}|^2\} + 1} \right). \quad (4.17)$$

This bound is often used in Massive MIMO research since it has a simple closed-form solution. Furthermore, in most propagation environments, when the number of BS antennas is large, the channel hardens (the effective channel gains become deterministic), and hence, this bound is very tight.

2. Full CSI: in this case, $\mathcal{C} = \{a_1, \dots, a_K\}$. Substituting $\mathcal{C} = \{a_1, \dots, a_K\}$ into (2), we obtain

$$C_k \geq \mathbb{E} \left\{ \log_2 \left(1 + \frac{|a_k|^2}{\sum_{k' \neq k}^K |a_{k'}|^2 + 1} \right) \right\}. \quad (4.18)$$

Chapter 5

Summary of Specific Contributions of the Dissertation

This dissertation consists of two parts. Firstly, we study fundamentals of Massive MIMO. The performance of Massive MIMO systems is analysed in terms of spectral efficiency and energy efficiency. Effects of pilot contamination and finite-dimensional channel models are also analysed. In addition, some aspects of favorable propagation in Massive MIMO are investigated. Secondly, we propose some system designs for Massive MIMO. Specifically, the optimal power as well as training duration allocation is studied. Applications of Massive MIMO in relay channels are also considered.

5.1 Included Papers

Brief summaries of the papers included in this dissertation are as follows:

Paper A: Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems

Authored by Hien Quoc Ngo, Erik G. Larsson, and Thomas L. Marzetta.

Published in the IEEE Transactions on Communications, 2013. This work is an extension of the conference paper [38].

A multiplicity of autonomous terminals simultaneously transmits data streams to a compact array of antennas. The array uses imperfect channel-state information derived from transmitted pilots to extract the individual data streams. The power radiated by the terminals can be made inversely proportional to the square-root of the number of base station antennas with no reduction in performance. In contrast if perfect channel-state information were available the power could be made inversely proportional to the number of antennas. Lower capacity bounds for maximum-ratio combining (MRC), zero-forcing (ZF) and minimum mean-square error (MMSE) detection are derived. A MRC receiver normally performs worse than ZF and MMSE. However as power levels are reduced, the cross-talk introduced by the inferior maximum-ratio receiver eventually falls below the noise level and this simple receiver becomes a viable option. The tradeoff between the energy efficiency (as measured in bits/J) and spectral efficiency (as measured in bits/channel use/terminal) is quantified for a channel model that includes small-scale fading but not large-scale fading. It is shown that the use of moderately large antenna arrays can improve the spectral and energy efficiency with orders of magnitude compared to a single-antenna system.

Paper B: The Multicell Multiuser MIMO Uplink with Very Large Antenna Arrays and a Finite-Dimensional Channel

Authored by Hien Quoc Ngo, Erik G. Larsson, and Thomas L. Marzetta.

Published in the IEEE Transactions on Communications, 2013. This work is an extension of the conference paper [39].

We consider multicell multiuser MIMO systems with a very large number of antennas at the base station (BS). We assume that the channel is estimated by using uplink training. We further consider a physical channel model where the angular domain is separated into a finite number of distinct directions. We analyze the so-called pilot contamination effect discovered in previous work, and show that this effect persists under the finite-dimensional channel model that we consider. In particular, we consider a uniform array at the BS. For this scenario, we show that when the number of BS antennas goes to infinity, the system performance under a finite-dimensional channel model with P angular bins is the same as the performance under an uncorrelated channel model with P antennas. We further derive a lower bound on the achievable rate of uplink data transmission with a linear detector at the BS. We then specialize this lower bound to the cases of maximum-ratio combining (MRC) and zero-forcing (ZF) receivers, for a finite and an infinite number of BS antennas. Numerical results corroborate our analysis and show a comparison between the performances of MRC and ZF in terms of sum-rate.

Paper C: Aspects of Favorable Propagation in Massive MIMO

Authored by Hien Quoc Ngo, Erik G. Larsson, and Thomas L. Marzetta.

Published in the proceedings of the European Signal Processing Conference (EU-SIPCO), 2014 [40].

Favorable propagation, defined as mutual orthogonality among the vector-valued channels to the terminals, is one of the key properties of the radio channel that is exploited in Massive MIMO. However, there has been little work that studies this topic in detail. In this paper, we first show that favorable propagation offers the most desirable scenario in terms of maximizing the sum-capacity. One useful proxy for whether propagation is favorable or not is the channel condition number. However, this proxy is not good for the case where the norms of the channel vectors are not equal. For this case, to evaluate how favorable the propagation offered by the channel is, we propose a “distance from favorable propagation” measure, which is the gap between the sum-capacity and the maximum capacity obtained under favorable propagation. Secondly, we examine how favorable the channels can be for two extreme scenarios: i.i.d. Rayleigh fading and uniform random line-of-sight (UR-LoS). Both environments offer (nearly) favorable propagation. Furthermore, to analyze the UR-LoS model, we propose an urns-and-balls model. This model is simple and explains the singular value spread characteristic of the UR-LoS model well.

Paper D: EVD-Based Channel Estimations for Multicell Multiuser MIMO with Very Large Antenna Arrays

Authored by Hien Quoc Ngo and Erik G. Larsson.

Published in the proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012 [33].

This paper consider a multicell multiuser MIMO with very large antenna arrays at the base station. For this system, with channel state information estimated from pilots, the system performance is limited by pilot contamination and noise limitation as well as the spectral inefficiency discovered in previous work. To reduce these effects, we propose the eigenvalue-decomposition-based approach to estimate the channel directly from the received data. This approach is based on the orthogonality of the channel vectors between the users and the base station when the number of base station antennas grows large. We show that the channel can be estimated from the eigenvalue of the received covariance matrix excepting the multiplicative factor ambiguity. A short training sequence is required to solved this ambiguity. Furthermore, to improve the performance of our approach, we investigate the joint eigenvalue-decomposition-based approach and the Iterative Least-Square with Projection algorithm. The numerical results verify the effectiveness of our channel estimate approach.

Paper E: Massive MU-MIMO Downlink TDD Systems with Linear Precoding and Downlink Pilots

Authored by Hien Quoc Ngo, Erik G. Larsson, and Thomas L. Marzetta.

Published in the proceedings of the Allerton Conference on Communication, Control, and Computing, 2013 [41].

We consider a massive MU-MIMO downlink time-division duplex system where a base station (BS) equipped with many antennas serves several single-antenna users in the same time-frequency resource. We assume that the BS uses linear precoding for the transmission. To reliably decode the signals transmitted from the BS, each user should have an estimate of its channel. In this work, we consider an efficient channel estimation scheme to acquire CSI at each user, called beamforming training scheme. With the beamforming training scheme, the BS precodes the pilot sequences and forwards to all users. Then, based on the received pilots, each user uses minimum mean-square error channel estimation to estimate the effective channel gains. The channel estimation overhead of this scheme does not depend on the number of BS antennas, and is only proportional to the number of users. We then derive a lower bound on the capacity for maximum-ratio transmission and zero-forcing precoding techniques which enables us to evaluate the spectral efficiency taking into account the spectral efficiency loss associated with the transmission of the downlink pilots. Comparing with previous work where each user uses only the statistical channel properties to decode the transmitted signals, we see that the proposed beamforming training scheme is preferable for moderate and low-mobility environments.

Paper F: Blind Estimation of Effective Downlink Channel Gains in Massive MIMO

Authored by Hien Quoc Ngo and Erik G. Larsson.

Submitted to the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015 [31].

We consider the massive MIMO downlink with time-division duplex (TDD) operation and conjugate beamforming transmission. To reliably decode the desired signals, the users need to know the effective channel gain. In this paper, we propose a blind channel estimation method which can be applied at the users and which does not require any downlink pilots. We show that our proposed scheme can substantially outperform the case where each user has only statistical channel knowledge, and that the difference in performance is particularly large in certain types of channel, most notably keyhole channels. Compared to schemes that rely on downlink pilots (e.g., [41]), our proposed scheme yields more accurate channel estimates for a wide range of signal-to-noise ratios and avoid spending time-frequency resources on pilots.

Paper G: Massive MIMO with Optimal Power and Training Duration Allocation

Authored by Hien Quoc Ngo, Michail Matthaiou, and Erik G. Larsson.

Published in the IEEE Wireless Communications Letters, 2014 [42].

We consider the uplink of massive multicell multiple-input multiple-output systems, where the base stations (BSs), equipped with massive arrays, serve simultaneously several terminals in the same frequency band. We assume that the BS estimates the channel from uplink training, and then uses the maximum ratio combining technique to detect the signals transmitted from all terminals in its own cell. We propose an optimal resource allocation scheme which jointly selects the training duration, training signal power, and data signal power in order to maximize the sum spectral efficiency, for a given total energy budget spent in a coherence interval. Numerical results verify the benefits of the optimal resource allocation scheme. Furthermore, we show that more training signal power should be used at low signal-to-noise ratio (SNRs), and vice versa at high SNRs. Interestingly, for the entire SNR regime, the optimal training duration is equal to the number of terminals.

Paper H: Large-Scale Multipair Two-Way Relay Networks with Distributed AF Beamforming

Authored by Hien Quoc Ngo and Erik G. Larsson.

Published in the IEEE Communications Letters, 2013 [43].

We consider a multipair two-way relay network where multiple communication pairs simultaneously exchange information with the help of multiple relay nodes. All nodes are equipped with a single antenna and channel state information is available at the relay nodes. Each relay uses very simple signal processing in a distributed manner, called distributed amplify-and-forward (AF) relaying. A closed-form expression for the achievable rate is derived. We show that the distributed AF scheme outperforms conventional orthogonal relaying. When the number of relays is large, the distributed AF relaying scheme can achieve the capacity scaling given by the cut-set upper bound. Furthermore, when the number of relays grows large, the transmit powers of each terminal and of the relay can be made inversely proportional to the number of relays while maintaining a given quality-of-service. If the transmit power of each terminal is kept fixed, the transmit power of each relay can be scaled down inversely proportional to the square of the number of relays.

Paper I: Spectral Efficiency of the Multi-pair Two-Way Relay Channel with Massive Arrays

Authored by Hien Quoc Ngo and Erik G. Larsson.

Published in the proceedings of the Asilomar Conference on Signals, Systems, and Computer, 2013 [44].

We consider a multipair two-way relay channel where multiple communication pairs share the same time-frequency resource and a common relay node. We assume that all users have a single antenna, while the relay node is equipped with a very large antenna array. We consider two transmission schemes: (I) separate-training zero-forcing (ZF) and (II) a new proposed coupled-training ZF. For both schemes, the channels are estimated at the relay by using training sequences, assuming time-division duplex operation. The relay processes the received signals using ZF. With the separate-training ZF, the channels from all users are estimated separately. By contrast, with the coupled-training ZF, the relay estimates the sum of the channels from two users of a given communication pair. This reduces the amount of resources spent in the training phase. Self-interference reduction is also proposed for these schemes. When the number of relay antennas grows large, the effects of interpair interference and self-interference can be neglected. The transmit power of each user and of the relay can be made inversely proportional to the square root of the number of relay antennas while maintaining a given quality-of-service. We derive a lower bound on the capacity which enables us to evaluate the spectral efficiency. The coupled-training ZF scheme is preferable for the high-mobility environment, while the separate-training ZF scheme is preferable for the low-mobility environment

Paper J: Multipair Full-Duplex Relaying with Massive Arrays and Linear Processing

Authored by Hien Quoc Ngo, Himal A. Suraweera, Michail Matthaiou, and Erik G. Larsson.

Published in the IEEE Journal on Selected Areas in Communications, 2014 [45].

We consider a multipair decode-and-forward relay channel, where multiple sources transmit simultaneously their signals to multiple destinations with the help of a full-duplex relay station. We assume that the relay station is equipped with massive arrays, while all sources and destinations have a single antenna. The relay station uses channel estimates obtained from received pilots and zero-forcing (ZF) or maximum-ratio combining/maximum-ratio transmission (MRC/MRT) to process the signals. To significantly reduce the loop interference effect, we propose two techniques: i) using a massive receive antenna array; or ii) using a massive transmit antenna array together with very low transmit power at the relay station. We derive an exact achievable rate expression in closed-form for MRC/MRT processing and an analytical approximation of the achievable rate for ZF processing. This approximation is very tight, particularly for a large number of relay station antennas. These closed-form expressions enable us to determine the regions where the full-duplex mode outperforms the half-duplex mode, as well as to design an optimal power allocation scheme. This optimal power allocation scheme aims to maximize the energy efficiency for a given sum spectral efficiency and under peak power constraints at the relay station and sources. Numerical results verify the effectiveness of the optimal power allocation scheme. Furthermore, we show that, by doubling the number of transmit/receive antennas at the relay station, the transmit

power of each source and of the relay station can be reduced by 1.5 dB if the pilot power is equal to the signal power, and by 3 dB if the pilot power is kept fixed, while maintaining a given quality of service.

5.2 Not Included Papers

The following publications by the author are not included in the dissertation either because they do not fit within the main scope of the dissertation, or they were earlier versions of the journal publications included in the dissertation.

- H. Q. Ngo and E. G. Larsson, “Linear multihop amplify-and-forward relay channels: Error exponent and optimal number of hops,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3834–3842, Nov. 2011.
- M. Matthaiou, G. C. Alexandropoulos, H. Q. Ngo, and E. G. Larsson, “Analytic framework for the effective rate of MISO fading channels,” *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1741–1751, June 2012.
- T. Q. Duong, H. Q. Ngo, H.-J. Zepernick, and A. Nallanathan, “Distributed space-time coding in two-way fixed gain relay networks over Nakagami-m fading,” *IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June 2012.
- H. Q. Ngo, M. Matthaiou, and E. G. Larsson, “Performance analysis of large scale MU-MIMO with optimal linear receivers,” *Proceedings of the IEEE Swedish Communication Technologies Workshop (Swe-CTW)*, 2012.
- H. Q. Ngo, M. Matthaiou, T. Q. Duong, and E. G. Larsson, “Uplink performance analysis of multicell MU-SIMO systems with ZF receivers,” *IEEE Trans. Vehicular Techno.*, vol. 62, no. 9, pp. 4471–4483, Nov. 2013.
- H. A. Suraweera, H. Q. Ngo, T. Q. Duong, C. Yuen, and E. G. Larsson, “Multi-pair amplify-and-forward relaying with very large antenna arrays,” in *Proc. IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 2013.
- A. K. Papazafeiropoulos, H. Q. Ngo, M. Matthaiou, and T. Ratnarajah, “Uplink performance of conventional and massive MIMO cellular systems with delayed CSIT,” in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Washington, D.C., Sept. 2014.
- H. Phan, T. M. C. Chu, H.-J. Zepernick, H. Q. Ngo, “Performance of cognitive radio networks with finite buffer using multiple vacations and exhaustive service,” in *Proc. International Conference on Signal Processing and Communication (ICSPCS)*, Gold Coast, Australia, Dec. 2014.

Chapter 6

Future Research Directions

As discussed in Section 3.4, a number of issues and challenges in Massive MIMO remain to be investigated. There are many open research directions. Here is the list of possible research directions in Massive MIMO:

- Pilot contamination: pilot contamination is one of the inherent limitations of Massive MIMO which degrades the Massive MIMO performance significantly. This effect persists even when the number of BS antennas goes to infinity. Dealing with the pilot contamination effect is an important research direction. Pilot contamination arises due to interference from other cells during the training phase. Therefore, one way to reduce the pilot contamination effect is to use large frequency-reuse factors during the training phase. However, this will reduce the pre-log factor, and hence, will reduce the spectral efficiency. Another way is increasing the cell-size. With a large cell-size, due to the path loss, the power of the desired signal in a given cell is much stronger than the interference power from other cells. However, owing to the large cell-size and the effect of path loss, the users that are located around the cell edges could not receive a good quality-of-service. A suitable design of the cell-size, frequency-reuse factor during the training, and power control to reduce the pilot contamination effect should be investigated.
- Channel state information acquisition: the acquisition of CSI is very important in Massive MIMO. Channel estimation algorithms are attracting much attention. There is much ongoing research in this direction. Many questions are not still appropriately answered:
 - Can the channel be blindly estimated? Can payload data help improve channel estimation accuracy? How much can we gain from such schemes?

- Is the use of orthogonal pilot sequences among users optimal, especially in multicell systems where the pilot contamination effect occurs? Which pilot sequences should be used? How to optimally assign pilot sequences to the users, especially for new users which enter the system? We believe that the design and assignment of the pilot sequences is an important research direction.
- Should each user estimate the effective channel in the downlink? And how much gain we can obtain?
- System architecture: it would be good if Massive MIMO can combine with practical systems such as LTE. Furthermore, Massive MIMO, small cell, and millimeter wave technologies are promising candidates for 5G wireless systems. Designing new efficient systems with the combination of these technologies is a good research direction.
- In our work, Paper F, we considered the case where the transmit powers during the training and payload data transmissions are not equal, and are optimally chosen. The performance gain obtained by this optimal power allocation was studied. However, the cost of performing this optimal power allocation may be an increase in the peak-to-average ratio of the emitted waveform. This should be investigated in future work.
- In Papers G and H, amplify-and-forward relaying networks are studied. We assume that the destinations have no CSI, and an interpair interference reduction scheme is proposed by using the knowledge of its deterministic equivalent. This will work well if the interference hardens quickly in large systems. However, with amplifying-and-forward relaying, the interference contains the products of two channels, and hence, the interference hardens very slowly. The system performance may noticeably improve if the estimate of interference is considered.
- In current works, we assume that each user has a single antenna. It would be interesting to consider the case where each user is employed with several antennas. Note that in the current wireless systems (e.g. LTE), each users can have two antennas [48]. The transceiver designs (transmission schemes at the BS and detection schemes at the users) and performance analysis (achievable rate, outage probability, etc) of Massive MIMO systems with multiple-antenna users should be studied.
- Distributed Massive MIMO: in our work, we considered massive MIMO with colocated antenna arrays at the BS. Alternatively, massive BS antenna arrays can be distributed in a large area. Design and analysis of distributed Massive MIMO systems are of interest. Some related works are considered in the literature. For example, in [49], the authors proposed a distributed massive MIMO structure by clustering the cooperating BS and partitioning the users into groups. It was shown that the proposed scheme can achieve a spectral efficiency comparable with that of colocated Massive MIMO in [13] with

a much smaller number of active antennas. However, in [13], the authors considered conjugate beamforming. In future work, the comparison between [49] and [13] with ZF should be considered. It is also interesting to compare the energy efficiency between the system in [49] and the system in [13].

References

- [1] Qualcomm Inc., *The 100× Data Challenge*, Oct. 2013. [Online]. Available: <http://www.qualcomm.com/solutions/wireless-networks/technologies/1000x-data>.
- [2] Ericson, *5G Radio Access—Research and Vision*, June 2013. [Online]. Available: <http://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf>.
- [3] Cisco, “*Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018*,” Feb. 2014. [Online]. Available: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-520862.html>.
- [4] D. Gesbert, M. Kountouris, R. W. Heath Jr., C.-B. Chae, and T. Sälzer, “Shifting the MIMO paradigm,” *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.
- [5] M. Kobayashi, N. Jindal, and G. Caire, “Training and feedback optimization for multiuser MIMO downlink,” *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.
- [6] V. Stankovic and M. Haardt, “Generalized design of multiuser MIMO precoding matrices,” *IEEE Trans. Wireless Commun.*, vol. 7, pp. 953–961, Mar. 2008.
- [7] G. Caire and S. Shamai, “On the achievable throughput of a multi-antenna Gaussian broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
- [8] P. Viswanath and D. N. C. Tse, “Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality,” *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
- [9] T. Yoo and A. Goldsmith, “On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

- [10] S. Verdú, *Multiuser Detection*. Cambridge, UK: Cambridge University Press, 1998.
- [11] N. Jindal and A. Goldsmith, “Dirty-paper coding vs. TDMA for MIMO broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1783–1794, May 2005.
- [12] C. Suh, M. Ho, and D. N. C. Tse, “Downlink interference alignment,” *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2616–2626, Sep. 2011.
- [13] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [14] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, “Scaling up MIMO: Opportunities and challenges with very large arrays,” *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–46, Jan. 2013.
- [15] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [16] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [17] S. K. Mohammed and E. G. Larsson, “Per-antenna constant envelope precoding for large multi-user MIMO systems,” *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.
- [18] C. Shepard, H. Yu, N. Anand, L. E. Li, T. L. Marzetta, R. Yang, and L. Zhong, “Argos: Practical many-antenna base stations,” in *Proc. ACM Int. Conf. Mobile Computing and Networking (MobiCom)*, Aug. 2012.
- [19] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, “On the optimality of single-carrier transmission in large-scale antenna systems,” *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 276–279, Aug. 2012.
- [20] —, “Effect of oscillator phase noise on uplink performance of large MU-MIMO systems,” in *Proc. of the 50-th Annual Allerton Conference on Communication, Control, and Computing*, 2012.
- [21] W. Yang, G. Durisi, and E. Riegler, “On the capacity of large-MIMO block-fading channel,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013.
- [22] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A coordinated approach to channel estimation in large-scale multiple-antenna systems,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

- [23] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 303–313, Feb. 2013.
- [24] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Interference reduction on cellular networks with large antenna arrays," in *Proc. IEEE International Conference on Communications (ICC)*, Ottawa, Canada, Jun. 2012.
- [25] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Linear pre-coding performance in measured very-large MIMO channels," in *Proc. IEEE Vehicular Technology Conf. (VTC)*, San Francisco, CA, US, Sept. 2011.
- [26] S. Payami and F. Tufvesson, "Channel measurements and analysis for very large array systems at 2.6 GHz," in *Proc. 6th European Conference on Antennas and Propagation (EuCAP)*, Prague, Czech Republic, Mar. 2012.
- [27] A. Goldsmith and S. A. Jafar and N. Jindal and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.
- [28] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, UK: Cambridge University Press, 2005.
- [29] K. Nishimori, K. Cho, Y. Takatori, and T. Hori, "Automatic calibration method using transmitting signals of an adaptive array for TDD systems," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1636–1640, 2001.
- [30] R. Rogalin, O. Y. Bursalioglu, H. C. Papadopoulos, G. Caire, and A. F. Molisch, "Hardware-impairment compensation for enabling distributed large-scale MIMO," in *Proc. ITA Workshop*, San Diego, CA, USA, 2013.
- [31] H. Q. Ngo and E. G. Larsson, "Blind estimation of effective downlink channel gains in Massive MIMO," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, Australia, April 2015, submitted.
- [32] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [33] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimations for multicell multiuser MIMO with very large antenna arrays," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, 2012.
- [34] R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Sig. Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

- [35] A. Ashikhmin and T. L. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Cambridge, MA, Jul. 2012.
- [36] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- [37] H. Cramér, *Random Variables and Probability Distributions*. Cambridge, UK: Cambridge University Press, 1970.
- [38] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Uplink power efficiency of multiuser MIMO with very large antenna arrays," in *Proc. 49th Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, Illinois, US, Sep. 2011.
- [39] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'11)*, Prague, Czech Republic, May 2011, pp. 3464–3467.
- [40] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson, "Aspects of favorable propagation in Massive MIMO," in *Proc. European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, Sep. 2014.
- [41] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, Illinois, US, Oct. 2013.
- [42] H. Q. Ngo, M. Matthaiou, and E. G. Larsson, "Massive MIMO with optimal power and training duration allocation," *IEEE Wireless Commun. Lett.*, vol. 3, Dec. 2014.
- [43] H. Q. Ngo and E. Larsson, "Large-Scale Multipair Two-Way Relay Networks with Distributed AF Beamforming," *IEEE Commun. Lett.*, vol. 17, no. 12, pp. 2288–2291, Dec. 2013.
- [44] H. Q. Ngo and E. G. Larsson, "Spectral efficiency of the multi-pair two-way relay channel with massive arrays," in *Proc. Asilomar Conference on Signals, Systems, and Computer*, Pacific Grove, CA, Nov. 2013.
- [45] Hien Quoc Ngo, Hima A. Suraweera, Michail Matthaiou, and Erik G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.
- [46] H. Q. Ngo, E. Larsson, and T. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

- [47] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.
- [48] 3GPP TR 36.211 V 11.1.0 (Release 11), "Evolved universal terrestrial radio access (E-UTRA)," Dec. 2012.
- [49] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprasad, "Achieving "Massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.

Part II

Fundamentals of Massive MIMO

Papers

The articles 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-112780>

