Department of Signal Processing and Acoustics

Multi-Antenna Techniques for Next Generation Cellular Communications

Umut Ugurlu

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 156/2016

Multi-Antenna Techniques for Next Generation Cellular Communications

Umut Ugurlu

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Electrical Engineering, at a public examination held at the lecture hall S2 of the school on September 2, 2016 at 12 pm.

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics

Supervising professor
Prof. Risto Wichman

Preliminary examiners
Prof. Tapani Ristaniemi, University of Jyväskylä, Finland
Dr. Ömer Bulakçı, Huawei European Research Center, Germany

Opponent
Prof. Nils Torbjörn Ekman, Norwegian University of Science and Technology, Norway

Aalto University publication series
DOCTORAL DISSERTATIONS 156/2016

ISBN 978-952-60-6953-1 (printed)
ISBN 978-952-60-6952-4 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-6952-4

Unigrafia Oy
Helsinki 2016

Finland

441 697
Printed matter

Aalto University

Abstract

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Author
Umut Ugurlu

Name of the doctoral dissertation
Multi-Antenna Techniques for Next Generation Cellular Communications

Publisher School of Electrical Engineering

Unit Department of Signal Processing and Acoustics

Series Aalto University publication series DOCTORAL DISSERTATIONS 156/2016

Field of research Signal Processing for Communications

Manuscript submitted 9 November 2015 **Date of the defence** 2 September 2016

Permission to publish granted (date) 18 February 2016 **Language** English

☐ **Monograph** ☒ **Article dissertation** ☐ **Essay dissertation**

Abstract

Future cellular communications are expected to offer substantial improvements for the pre-existing mobile services with higher data rates and lower latency as well as pioneer new types of applications that must comply with strict demands from a wider range of user types. All of these tasks require utmost efficiency in the use of spectral resources. Deploying multiple antennas introduces an additional signal dimension to wireless data transmissions, which provides a significant alternative solution against the plateauing capacity issue of the limited available spectrum. Multi-antenna techniques and the associated key enabling technologies possess unquestionable potential to play a key role in the evolution of next generation cellular systems.

Spectral efficiency can be improved on downlink by concurrently serving multiple users with high-rate data connections on shared resources. In this thesis optimized multi-user multi-input multi-output (MIMO) transmissions are investigated on downlink from both filter design and resource allocation/assignment points of view. Regarding filter design, a joint baseband processing method is proposed specifically for high signal-to-noise ratio (SNR) conditions, where the necessary signaling overhead can be compensated for. Regarding resource scheduling, greedy- and genetic-based algorithms are proposed that demand lower complexity with large number of resource blocks relative to prior implementations.

Channel estimation techniques are investigated for massive MIMO technology. In case of channel reciprocity, this thesis proposes an overhead reduction scheme for the signaling of user channel state information (CSI) feedback during a relative antenna calibration. In addition, a multi-cell coordination method is proposed for subspace-based blind estimators on uplink, which can be implicitly translated to downlink CSI in the presence of ideal reciprocity. Regarding non-reciprocal channels, a novel estimation technique is proposed based on reconstructing full downlink CSI from a select number of dominant propagation paths. The proposed method offers drastic compressions in user feedback reports and requires much simpler downlink training processes.

Full-duplex technology can provide up to twice the spectral efficiency of conventional resource divisions. This thesis considers a full-duplex two-hop link with a MIMO relay and investigates mitigation techniques against the inherent loop-interference. Spatial-domain suppression schemes are developed for the optimization of full-duplex MIMO relaying in a coverage extension scenario on downlink. The proposed methods are demonstrated to generate data rates that closely approximate their global bounds.

Keywords Baseband filter design, resource scheduling, cellular systems, massive MIMO, full-duplex, multi-antenna calibration, RF transceivers

ISBN (printed) 978-952-60-6953-1 **ISBN (pdf)** 978-952-60-6952-4

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2016

Pages 215 **urn** http://urn.fi/URN:ISBN:978-952-60-6952-4

Preface

The research, of which this doctoral thesis presents the findings, has been carried out at the Department of Signal Processing and Acoustics, Aalto University, Finland during 2013–2014 under the supervision of Prof. Risto Wichman. I would like to express my thanks for the support and creative freedom he has provided me with throughout the research work as well as during the writing of this thesis.

To potential readers I say the following. You are highly encouraged to consume a cup of coffee during your reading. This should be understood as the optimal practice for following the presented discussions within this thesis as the warmth and scent of a fresh coffee brew will ease your throat and leave behind a pleasurable feeling. In addition, it is also believed to be the best way of being in sync with the author's state of mind, as I have probably consumed more coffee during the writing of this thesis than ever in the rest of my life.

Espoo, July 22, 2015,

Umut Ugurlu

“目标很明显不能达成时，不调整目标，就要调整行动步骤。”

孔夫子

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I U. Ugurlu, R. Wichman, C. Ribeiro, C, Wijting. A Novel Joint Precoder and Receiver Design with Imperfect CSI for Multi-User MIMO Downlink. In *IEEE 80th Vehicular Technology Conference (VTC Fall)*, Vancouver, BC, pp. 1–5, September 2014.

II U. Ugurlu, R. Wichman, C. Ribeiro, C. Wijting. Iterative Genetic and Greedy-Based Algorithms for Multi-Carrier Multi-User Scheduling in MIMO Systems with Successive Zero Forcing. In *IEEE 80th Vehicular Technology Conference (VTC Fall)*, Vancouver, BC, pp. 1–5, September 2014.

III U. Ugurlu, R. Wichman, C. Ribeiro, C. Wijting. Adaptive Channel Quantization on TDD Reciprocity Performance of Relative Antenna Calibration. In *IEEE 80th Vehicular Technology Conference (VTC Fall)*, Vancouver, BC, pp. 1–5, September 2014.

IV U. Ugurlu, R. Wichman, C. Ribeiro, C. Wijting. Coordinated Optimization of EVD-Based Channel Estimators in Multi-Cell Massive MIMO Networks. In *IEEE International Communications Workshops (ICC) on 5G & Beyond - Enabling Technologies and Applications*, London, UK, pp. 1–5, June 2015.

V U. Ugurlu, R. Wichman, C. Ribeiro, C. Wijting. A Multipath Extraction

Based CSI Acquisition Method for FDD Cellular Networks with Massive Antenna Arrays. *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, 2940–2953, December 2015.

VI U. Ugurlu, R. Wichman, C. Ribeiro, C. Wijting. Multipath Phase Indication (MPI) Feedback for CSI Acquisition in FDD Massive MIMO. In *IEEE 25th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Hong Kong, pp. 1–5, August 2015.

VII U. Ugurlu, T. Riihonen, R. Wichman. Optimized In-Band Full-Duplex MIMO Relay under Single-Stream Transmission. *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, 155–168, January 2015.

VIII U. Ugurlu, R. Wichman, T. Riihonen, C. Ribeiro, C. Wijting. Power Control and Beamformer Design for the Optimization of Full-Duplex MIMO Relays in a Dual-Hop MISO Link. In *IEEE 9th Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Oulu, Finland, pp. 545–549, June 2014.

The author of this thesis has had the full responsibility for the development of all above publications (i.e., Publications [I–VIII]). He has written the first and every draft of them and he has developed all of their contributions, including all of the presented technical ideas and all of the evaluation results.

List of Abbreviations

3GPP	Third Generation Partnership Project
ADC	Analog-to-Digital Converter
AF	Amplify-and-Forward
ARQ	Automatic Repeat Request
BD	Block Diagonalization
BPSK	Binary Phase-Shift-Keying
C-RAN	Cloud Radio Access Network
CCM	Channel Covariance Matrix
CDD	Cyclic Delay Diversity
CEPT	The European Conf. of Postal and Telecomm. Admin.
CoMP	Coordinated Multi-Point
COST	European Cooperation in Science and Technology
CQI	Channel Quality Indicator
CSI	Channel State Information
D2D	Device-To-Device
DAC	Digital-to-Analog Converter
DC	Direct Current
DCI	Downlink Control Information
DCT	Discrete Cosine Transform
DF	Decode-and-Forward
DFT	Discrete Fourier Transform
DLL	Delay-Locked Loop
DM-RS	De-Modulation Reference Signal
DoA	Direction of Arrival
DoD	Direction of Departure
DPC	Dirty Paper Coding
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
eNB	Evolved Node B

EPC	Evolved Packet Core
EVD	Eigen Vector Decomposition
FCC	U.S. Federal Communications Commission
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FFR	Fractional Frequency Reuse
FRF	Frequency Reuse Factor
FTP	File Transfer Protocol
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HSS	Home Subscription Server
I/Q	In-phase/Quadrature
ICI	Inter-Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IoT	Internet of Things
IP	Internet Protocol
ISI	Inter-Symbol Interference
ITU	International Telecommunication Union
KLT	Karhunen-Loéve Transform
LNA	Low-Noise Amplifier
LoS	Line-of-Sight
LTE	Long Term Evolution
LTE-A	Long Term Evolution
MAC	Medium Access Control
METIS	Mobile communications Enablers for the Twenty-twenty Info. Society
MIMO	Multi-Input Multi-Output
MISO	Multi-Input Single-Output
MME	Mobile Management Entity
MMSE	Minimum Mean Square Error
MPI	Multipath Phase Indication
MRT	Maximum Ratio Transmit
MSE	Mean Square Error
MTC	Machine-Type Communications
MUI	Multi-User Interference
NAS	Non Access Stratum
NLoS	Non Line-of-Sight
NMT	Nordic Mobile Telephone
OFCOM	Office of Communications (UK)

OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OTA	Over-The-Air
P-GW	Packet Data Network Gateway
PCRF	Policy Control and Charging Rules Function
PDCP	Packet Data Convergence Protocol
PHY	Physical Layer
PLL	Phase-Locked Loop
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase-Shift-Keying
RAT	Radio Access Technology
RF	Radio Frequency
RLC	Radio Link Control
RoF	Radio-Over-Fiber
RRC	Radio Resource Control
RRM	Radio Resource Management
RVQ	Random Vector Quantization
S-GW	Serving Gateway
SC-FDMA	Single Carrier Frequency Division Multiple Access
SCM	Spatial Channel Model
SDMA	Spatial Division Multiple Access
SINR	Signal-to-Interference plus Noise Ratio
SIR	Signal-to-Interference Ratio
SLNR	Signal-to-Leakage plus Noise Ratio
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
TDD	Time Division Duplex
TLS	Total Least Squares
TM	Transmission Mode
TSP	Traveling Salesman Problem
TTI	Transmission Time Interval
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
VLC	Visible Light Communications
VoIP	Voice-over-IP
WINNER	Wireless World Initiative New Radio
WRC	World Radiocommunication Conference

ZF	Zero Forcing

1. Introduction

Modern wireless communications have come a long way since James Clark Maxwell predicted the existence of radio waves in 1864 and Heinrich Rudolph Hertz, for the first time, demonstrated them in 1885 [1]. Guglielmo Marconi kicked off wireless communications when he invented the first wireless telegraph in 1895. The ever-growing evolution of wireless communications is illustrated by a timeline in Fig. 1.1 [2, 3].

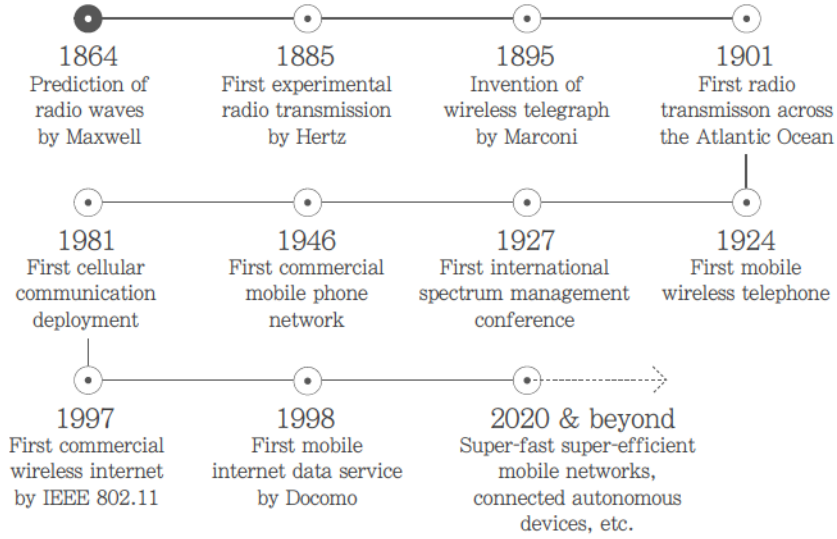

Figure 1.1. Evolution of wireless radio communications.

The first commercial mobile telephone system was developed in 1946 at the Bell labs with a high-power transmitter that could cover up to 50 miles to serve a very limited number of rich early adopters on low bandwidth. Due to the constraints on bandwidth, the Bell labs designed for the first time the concept of cellular technology and introduced it to the world. The idea behind the cellular concept was to divide a service zone to smaller unit areas and apply *frequency reuse*, which could theoretically

provide with more reliable service, lower transmission power, and better utilization of spectrum. The only major problem with the principle of cellular division was the necessity to hand over an established connection of a moving user over a cell border.

The cellular principle was first commercially deployed by Nordic Mobile Telephone (NMT) in 1981, which initiated a new era of mobile communications that we now call 1G cellular technology. Despite being a milestone, 1G systems had two fundamental problems. Analog communications proved to be susceptible to interference and the lack of standards led to incompatible technologies throughout the world. To address some of the earlier limitations, 2G cellular radio access (e.g., Global System for Mobile Communications (GSM)) was developed with improvements, such as overhead reduction, robust source and channel coding, a new text messaging service, better hand-off control, etc. Later, 2G systems were expanded to support data communications and internet service. Since then, further evolutions of cellular communications continued improving system performance and robustness with a principal focus on enabling faster mobile internet, which is still an ongoing endeavor to provide billions of users over the world with the reliability and comfort of technology in their palms that will one day melt away in our human lives and hopefully make the world marginally a better place.

1.1 Motivation

Although the initial attempts to develop a true mobile communication device for the mass market dates back to 1960s with pagers, cellular mobile phones have changed our lives tremendously, particularly since late 1990s. The speed of transformation for mobile phones has been unparalleled to any other product in technology. Particularly, the invent of mobile internet together with the absorption of other conveniences by cellular phones (i.e., a digital camera, fitness–health tracker, media player, navigator, game console, etc.) has made them one of the most influential disruptors of our information age. This unrivaled advancement of mobile devices gave birth to utmost demands for reliable, affordable, fast, and secure service delivery.

Cellular networks have been deployed almost everywhere with dense human population. Their benefits have been almost second nature during our daily lives, making it difficult for us to even be perceptive of their ex-

istence, of course only until we experience a service outage. It should be acknowledged that planning to deploy cellular coverage in all locations of the populated world is probably one of the most ambitious goals of mankind in any commercial industry. From the point of view of a network provider or operator, the goal is to improve customer satisfaction with minimum financial expenditure, therefore growing the net income. On the other hand, a user's desire is to receive a certain level of quality of service (QoS) determined by the individual's perception. Hence, network designers, in a nutshell, try to meet such expected QoS metrics with the highest possible efficiency.

By 2020 the commercial mobile data traffic per month is expected to increase by over 9-fold with 9.2 billion global mobile subscribers [4]. To gain a higher share of the market growth operators will try to attract new subscribers, decrease the churn rate, and lure their existing customers into more premium services. To achieve that, cellular network designers will aim to improve user QoS, as one of the most relevant indicators for these goals. Undeniably, QoS is a very broad term which is determined by various technical and economic aspects depending on individual use cases. For example, the expected quality of an offered service depends on its type and the associated priority metrics. A real-time based service, such as video conferencing or voice-over-IP (VoIP), should satisfy a minimum tolerable delay metric from the users' point of view. On the contrary, a non-real-time based service, such as file transfer protocol (FTP), should deliver high throughput for a guaranteed QoS. Offering a wide range of subscriptions with different price points allows operators to capture a larger market share from different types of users (e.g., human or machine) with different needs.

Motivated by the vitality of these benefits, this thesis revolves around the question of how equipping multiple antennas at a base station or at a mobile terminal device can help improve the user QoS in next generation cellular communications. Roughly speaking, the main QoS parameters in this thesis for exploring an answer to the above question are based on network-wide throughput or link-specific signal-to-interference plus noise ratio (SINR) metrics. In addition, system complexity is also considered an integral design aspect of this thesis from both hardware simplicity and computational load angles as a robust service provision with limited financial expenditure is a vital focus area in cellular network deployments as in any commercial venture. Potential improvements are studied, devel-

oped, and discussed in this thesis by means of introducing new methods, optimizing current state-of-the-art schemes, or integrating other technology concepts to cellular systems.

1.2 Scope, challenges, and beyond

First and foremost, next generation cellular networks are expected to introduce new types of services as well as offer improvements to the preexisting ones in terms of data rate, user capacity, scalability, cost efficiency, and flexibility. However, this thesis only focuses on the efficient utilization of multiple antenna deployments in cellular radio access with no particular attention to any service type. That being said, four major technology components are chosen as the field of discussion throughout this thesis for the investigation of various multi-antenna techniques. In addition, this thesis also touches on several other concepts in cellular communications and even discusses some of those concepts in detail while the attention is still on drawing the design framework for one of the said four main topics.

The first main topic of this thesis is downlink multi-user multi-input multi-output (MIMO) transmission, which includes transmit/receive baseband filter design as well as resource allocation/assignment problems. Although this is a well-studied topic in the last several years, ever-growing desire for higher spectral efficiency justifies the importance of optimizing baseband processing methods for multiple users who are sharing the same air interface resources. In particular, the goal of the thesis is to investigate the opportunities within the state-of-the-art linear/non-linear schemes for downlink precoding and multi-carrier multi-user scheduling. Consequently, low-complexity solutions can be developed with reduced transmit power and improved SINR. In addition, this thesis also focuses on the sensitivity of potential solutions to inevitable MIMO link deterrents, such as channel state information (CSI) errors and thermal noise effects.

The second topic is related to massive MIMO systems with reciprocal duplex channel response. More specifically, this thesis investigates popular estimation procedures for downlink CSI at base stations where the channel exhibits reciprocal characteristics. These methods mostly rely on first acquiring uplink channel responses and converting them to downlink based on the reciprocity property of the wireless propagation medium. However, such reciprocal propagation characteristics are often disturbed

in the overall channel response due to internal mismatches in the radio frequency (RF) transceivers. Therefore, the first part of the discussions on this topic is dedicated to dealing with antenna calibration methods that measure and recover from such non-reciprocal behaviors of an RF transceiver. The second half of the discussions includes analyzing and improving blind estimation methods for the simultaneous acquisition of uplink CSI from multiple users. Such blind estimators often rely on extracting the second-order channel properties from non-orthogonal data symbols as standard training based estimation is often susceptible to pilot contamination on uplink massive MIMO.

The third main topic is about non-reciprocity based CSI estimation on a massive MIMO link. Downlink channel acquisition is tedious in massive MIMO due to two major challenges. The first issue is the requirement for a very long time duration during downlink training whereas the second problem is related to the need for oversized signaling overhead during the feedback of user CSI reports. A very large antenna array increases the number of subchannels in the associated CSI vector. This requires long orthogonal pilot sequences to isolate different subchannels on downlink, which leads to long time periods of training. To restrict the duration, the antenna array size needs to be limited; however, such limitations can support only a small number of served users to preserve the asymptotical orthogonality among those users' CSI vectors. Moreover, the number of feedback bits needs to be increased proportionally to achieve a sufficient quantization granularity in a user's CSI report. Otherwise, the quantization resolution per channel coefficient decreases with more transmit antennas. This thesis investigates potential approaches that tackle the above challenges in three categories. The categories are established based on specific propagation characteristics which enable compression methods to reduce the size of CSI reports by removing some of the redundant information. These propagation characteristics are temporal channel correlations, spatial channel correlations, and channel sparsity. The goal of this thesis is to shed light on promising solutions in each category and justify the circumstances that create the favorable propagation conditions for CSI compression in a non-reciprocal massive MIMO propagation.

The fourth topic of this thesis is full-duplex in-band relaying. As the fundamental limitation of full-duplex relaying is caused by its inherent loop-interference, this topic is investigated by categorizing the types of mitigation techniques for the said interference signal. In particular, this

thesis focuses on spatial-domain suppression methods, which operate by means of multi-antenna baseband processing so as to nullify the bulk of the loop-interference in digital-domain. The goal of this thesis is to investigate such optimal suppression techniques and explore the feasibility of full-duplex deployments on a relay link. In addition, RF components of a full-duplex MIMO relay are also taken into consideration in the thesis as the accuracy of an interference mitigation scheme strongly depends on its sensitivity to imperfections in hardware.

The thesis investigates the above described topics in a cellular network architecture although the developed and discussed schemes can be applicable to other wireless radio technologies as well. All the discussions and emphases of the thesis are channeled toward next generation cellular communications. However, the specifications of the considered network models, when necessary, are often based on the state-of-the-art cellular communication standards, i.e., LTE-A. In addition, it should be also disclaimed that the communication models are occasionally simplified in this thesis to focus on the investigated multi-antenna schemes.

This thesis also takes an optimistic view toward some of the present obstacles that prevent a feasible real-world implementation of some of the state-of-the-art multi-antenna techniques. These challenges may include demanding computational requirements, high financial costs, insufficiency in backend connection speeds, reliance on short latency and long battery lives. Although current network, device, and equipment capabilities may be inadequate to fulfill most of these conditions, technological advancements in other fields are expected to overcome these shortcomings over time.

One of the upfront challenges toward next generation cellular systems is related to the physical scarcity of available RF spectra. As the allocated frequency bands have already been heavily used in the existing cellular systems, achieving better spectral efficiency is a key motivation for network designers developing next generation cellular systems. This thesis also puts this motivation to the spotlight in the context of multi-antenna techniques in spatial-domain as an additional signal dimension. In addition to higher efficiency, another popular means to overcome the said spectral scarcity is introducing higher frequency bands in the range of 60 GHz, which also serves as an enabler for other potential technology components of next generation systems [5, 6].

Another design challenge is related to reducing the environmental harms

caused by cellular deployments. Recent reports indicate that base stations are responsible for above 70% of the overall energy consumption of a cellular network operator [7]. In addition to financial savings, minimizing energy consumption also leads to greener wireless communications that can contribute to a significant reduction in CO_2 emission. For instance, deploying indoor communication technologies can be a significant derivative of higher energy efficiency. To be specific, liberating an outdoor macrocell evolved Node B (eNB) from the burden of serving far away users who are in non-favorable channel conditions helps lowering the transmit powers by means of avoiding unnecessary energy losses through wall penetrations. In addition, next generation systems also target harnessing higher frequency bands by millimeter-wave and visible light communication (VLC) technologies that will lead to better energy efficiency [8].

High user mobility notoriously leads to limited channel coherence time due to significant Doppler shifts, which shrink affordable eNB reaction time as well as cause difficulties in both acquiring up-to-date CSI and sustaining reliable connections. Although current mobile networks can support mobility up to 250 km/h, the availability of high-speed trains reaching approximately 500 km/h will demand a further support by the next generation cellular systems for higher user mobility scenarios [9].

Another dominant source of evolution for cellular systems is considered network densification [10]. In principle, denser deployments increase the spatial load factor per area and therefore improve the network capacity. Moreover, small cells lead to reduced pathloss, which boosts the strength of both desired and interference signals while the impact of thermal noise disappears. Therefore, advanced interference management concepts become integral to next generation system realizations. In addition, cell densification requires adaptive resource coordination, advanced interference cancelers at receivers, intelligent self-organizing network deployments, as well as commensurate densifications at the backhaul.

Last but not least, next generation cellular communications will rely on many applications that deserve recognition. This part is dedicated to mentioning the most popular of those applications as well as the associated enabling technologies, although they are beyond the purview of this thesis. Next generation cellular systems will support more types of connected devices with the introduction of *Internet of Things* (IoT) [11], connected car and connected home applications, *Tactile Internet* [12], etc. This variety will lead to fragmentations in user experience, demanding

more specific diversifications in user QoS requirements. For example, machine-type communications (MTC) will depend on the availability of an efficient low-bandwidth sporadic traffic medium in order to satisfy the expectations of long lifespan, low power, and low latency devices [13]. Wireless power transfer is a promising technology that can overcome the issue of limited battery life at user equipment (UE) in next generation cellular networks [14]. This technology may rely on delivering the energy transfer via microwave radiations emitted from the, so-called, *power beacons* that can be employed at eNB stations. Such wireless transfer of energy can potentially supply an infinite battery lifespan for always-on autonomous MTC users, e.g., portable measurement devices. Next generation cellular systems will continue evolving also at the network level by leveraging more virtualization and automation with cloud IT infrastructure and self-organization. Cloud IT refers to a software-defined network design that can offer virtualization and programmability for the network functions in order to reduce the operational costs and inflexibility [15]. Self-organizing networks refer to an automation technology that uses intelligence and autonomous adaptivity in order to introduce simplicity, speed, and flexibility for configuring, managing, and optimizing mobile network architectures [16]. Lastly, it is important to be aware that governments and regulatory bodies will continue playing even more prominent roles in the evolution of next generation cellular systems. The World Radiocommunication Conference (WRC) in 2019 will address the allocation of millimeter-wave bands (i.e., above 6 GHz) for 5G according to the current agenda of International Telecommunication Union (ITU). Governments will assign their regulatory bodies (such as FCC in the U.S., OFCOM in the UK, CEPT in Europe) in making the distribution of the established and newly introduced spectrum bands among national operators for various purposes, as an essential aspect of the next generation system will be to harmoniously combine different bands for different services within a unified global framework.

1.3 Contributions

This thesis conveys comprehensive discussions of multi-antenna processing techniques in different use cases for the next generation cellular communications. Several methods and new concepts are proposed under four main topics that are investigated throughout the thesis. The topics are

selected based on their suitability for providing a sufficient field of discussion on various multi-antenna methods as well as their potential to appear in future cellular systems. Therefore, this thesis offers an ample view of the state-of-the-art multi-antenna techniques and procedures presently in consideration for cellular communications, despite not necessarily being able to cover every known multi-antenna scheme.

The major contributions of the thesis are summarized as follows:

1. A joint transmit and receive filter design mechanism is developed for high signal-to-noise ratio (SNR) environments in the presence of CSI errors on the downlink of a multi-user MIMO communication.

2. Novel multi-user multi-carrier resource schedulers are designed based on the greedy and genetic based algorithms on downlink.

3. An adaptive quantization scheme is designed for reducing the total required signaling overhead during the operation of a relative antenna calibration procedure.

4. A multi-cellular coordination algorithm is proposed for the optimization of blind channel estimators on the uplink of a massive MIMO network.

5. A new downlink CSI acquisition mechanism is proposed for massive MIMO, which is based on estimating the characteristics of a select number of dominant paths and reconstructing the collective information.

6. A new type of user feedback report is proposed for realizing the above CSI reconstruction mechanism by means of signaling the multipath phase information attached to the targeted dominant paths on downlink.

7. An iterative optimization procedure is proposed for full-duplex dual-hop relaying setups, which adapts its solution based on the instantaneous channel gains.

8. For a relay-supported coverage extension scenario, the optimal pairs of power allocations are derived on full-duplex downlink with non-negligible direct link between the end connection points. In addition, an optimiza-

tion scheme is proposed and evaluated with multi-antenna processing.

1.3.1 Summary of the publications

This thesis is a compilation of the research findings and discussions from a collection of eight original publications. The author of this thesis, alone, has written the first and every draft of all the said publications and has developed all the technical contributions including all of the presented ideas, numerical evaluations and theoretical analyses. Each of these eight original publications are outlined below:

- Publication I develops a novel joint transmit and receive baseband filtering scheme under imperfect CSI knowledge for downlink multi-user MIMO transmissions. The contribution is distinguished from the prior art by its suitability for high SNR conditions as such a joint design typically requires sizable signaling between a base station and multiple user terminals, which can only be compensated for when the channel quality is sufficiently high. The simulation results verify the superior performance of the proposed design over other state-of-the-art schemes.

- Publication II designs multi-resource schedulers on downlink based on the well-known greedy and genetic based algorithms in a two-tier network. The schedulers deal with the problem of resource allocation, assignment, as well as multi-user ordering for shared resources based on maximizing the proportional fair and maximum throughput metrics during time- and frequency-domain packet scheduling, respectively. In contrast to the conventional solutions, the proposed schedulers adopt an iterative approach, where at each iteration one additional user is attempted to be co-scheduled next to the primarily scheduled users. Such approach leads to a new chromosome structure for the genetic-based algorithm. The numerical evaluations compare and contrast the proposed schedulers and the publication concludes that the genetic algorithm performs well in crowded networks whereas the greedy algorithm can be preferable with less co-scheduled users.

- Publication III proposes a novel adaptive quantization scheme for the necessary user reports of downlink CSI during a relative antenna calibration process. The adaptive scheme is based on adjusting the num-

ber of quantization levels on-the-go for every CSI measurement at the user terminal. The publication also proposes that the optimal number of feedback bits should be derived at the base station as it has access to measuring instantaneous uplink channel quality, and that the base station should subsequently request the CSI feedback report from the terminal according to the indicated number of quantization bits. The numerical results demonstrate that such an adaptive quantization approach can offer considerable reduction in signaling overhead.

- Publication IV develops a multi-cellular optimization scheme for blind estimators on uplink massive MIMO based on iteratively and jointly optimizing CSI estimation accuracy. Such multi-cell optimization requires exchanging a subset of the available channel information at each base station toward a central backend in order to jointly extract the common information among the shared data. The optimization aims to minimize the estimation errors caused by the finite number of receive antennas and transmit symbols. The simulation results compare the performance of the proposed coordinated optimization scheme with conventional blind estimators and verifies that such multi-cell coordination helps reach lower symbol error rates.

- Publication V develops a new acquisition technique for downlink CSI based on multipath extraction in massive MIMO systems. The proposed method relies on extracting the reciprocal characteristics of the strongest dominant paths from uplink while the corresponding non-reciprocal path information is transferred from user terminals. Subsequently, final CSI is reconstructed at the base station by combining the available knowledge with the shared information. Such reconstruction offers several advantages over the conventional acquisition techniques, such as signaling overhead being immune to the size of the antenna array and removing the necessity of downlink training for full CSI estimation. The simulation results demonstrate that the proposed acquisition method is promising in massive MIMO communications thanks to its CSI accuracy and low signaling overhead.

- Publication VI is a continuation of Publication V. The main contribution is designing a new type of user feedback report for the realization of the said multipath acquisition method, where the reports are composed of

the phase information from the dominant paths that are targeted by the base station. In addition, Publication VI develops and discusses optimized selection and indication procedures for the extraction of the said dominant paths for CSI recovery. The publication concludes that the proposed technique demonstrates promising potential for future cellular communications with massive antenna arrays, particularly in large cells with highly elevated base station towers due to the dependency on receiving and transmitting planar wavefronts at the proximity of the massive antenna arrays.

- Publication VII designs a coherent scheme to approximate the globally optimal performance of a dual-hop link with a MIMO relay that operates in full-duplex and executes either one of the two most popular relaying protocols, i.e., amplify-and-forward (AF) and decode-and-forward (DF). Unlike most of the prior arts, the proposed scheme is based on optimizing the overall end-to-end link instead of isolating the interference mitigation at the relay. In addition, the proposed method is also based on a novel idea which adapts its optimal design strategy toward the baseband filters according to instantaneous channel gains. The numerical results verify that the proposed spatial suppression scheme very closely approximates its theoretical bound in terms of data rates.

- Publication VIII derives the optimal pair of transmit power allocations in a full-duplex downlink transmission with an interfering direct end-to-end connection as a coverage extension scenario. Moreover, an iterative algorithm is designed in baseband for the spatial suppressions of the loop-interference as well as the direct connection signal. Simulation results demonstrate that the proposed scheme achieves higher data rates when combined with the leakage-based transmit beamformers rather than those with conventional transmit filtering.

1.4 Structure of the thesis

The rest of this thesis is structured as follows. Chapter 2 provides an overview of a cellular network with emphasis on propagation channels, transceiver design components, wireless transmission modes, and performance impairments on multi-antenna links. Chapters 3, 4, 5, 6 deal with

each of the four main topics of the thesis, which are respectively: downlink multi-user transmission, reciprocity-based massive MIMO systems, non-reciprocal channel estimation in massive MIMO, and full-duplex multi-antenna relaying. Finally, Chapter 7 concludes this thesis by summarizing its contributions and hinting at potential future directions.

2. Overview of a Cellular Network

This chapter provides a brief overview of the state-of-the-art cellular network architecture as well as the current and future trends for its evolution. The focus is on giving a clear description of the enabling features and the limitations of a typical cellular communications system to establish an unbiased medium of assessment for the advanced multi-antenna signal processing techniques covered in the subsequent chapters of this thesis.

The most prominent example of the current cellular communication systems is commercially known as the 4G mobile networks, such as Long Term Evolution (LTE) and its evolved form LTE-Advanced (LTE-A). The next big advancement in radio access technology beyond 4G is directed toward the so-called fifth generation (5G) systems, which are expected to be commercially available by 2020 [17, 18]. In this chapter the description of the current cellular networks is mostly based on LTE and LTE-A as standardized by the Third Generation Partnership Project (3GPP). More detailed description of the LTE-A architecture can be found in [19, 20, 21, 22, 23, 24, 25]. In addition, major potential directions for the next wave of growth are also provided based on the concerted visions of the industry participants and ongoing alliance projects on 5G research. The initial timeline and technology clarifications of 5G can be found in [26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36].

In the following this chapter describes the multi-layered architecture of a cellular network, propagation characteristics of a typical multi-antenna wireless transmission, generic building blocks of a wireless RF transceiver, as well as the main sources of performance impairments in a standard multi-cellular multi-terminal mobile communication system.

2.1 Architecture

The network architecture of LTE is fundamentally composed of three major components: user equipment (UE), evolved UMTS Terrestrial Radio Access Network (E-UTRAN), and evolved packet core (EPC) [23, 37, 38]. The UE handles all the communication functions from and to a mobile user using the stored user-specific information. The E-UTRAN, also known as the *access network*, deals with the communications between UE and EPC, and consists of multiple eNB stations, each of which is connected to other eNBs by the X2 interface. The serving eNB handles the data transmissions to and from the connected UE, oversees the basic operations of UE by signaling control messages, and is also responsible for the radio resource management (RRM). The EPC, also known as the *core network*, is responsible for hosting operators' subscriber information, controls the functioning of high-level mobility management, communicates with packet data networks, and connects to the *services domain*. The EPC is composed of Mobile Management Entity (MME), Serving Gateway (S-GW), Packet Data Network Gateway (P-GW), Policy Control and Charging Rules Function (PCRF), and Home Subscription Server (HSS) units. The MME controls and processes the signaling between EPC and UE related

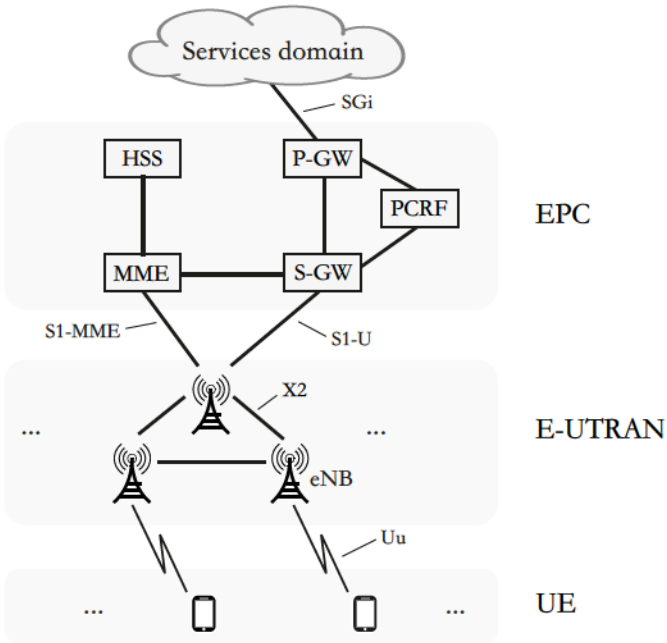

Figure 2.1. A typical cellular architecture (e.g., of LTE) with three major components.

to the non access stratum (NAS) protocols. The S-GW handles the transfer of user-plane packets and UE handovers between neighboring eNBs and different radio access technologies (RAT). The P-GW acts as a gateway and provides UE with connectivity to external packet data networks. The PCRF is responsible for policy control decision making and controlling the charging functionalities in the P-GW. The HSS holds information regarding users' subscription data and the packet data networks that are accessible by the users. The basic network architecture is illustrated in Fig. 2.1.

A complementary way of classifying the LTE architecture can be based on the control plane protocols, which comprise: physical layer (Layer-1), data link and access layer (Layer-2), control and application layer (Layer-3). The layer-1 handles the physical layer (PHY) operations, which include link adaptation (e.g., coding/decoding and modulation), power control, multi-antenna mapping, and initial cell search. During wireless signal transmissions the symbols are modulated by orthogonal frequency division multiple access (OFDMA) on downlink and single carrier frequency division multiple access (SC-FDMA) on uplink. The layer-2 consists of medium access control (MAC), radio link control (RLC), and packet data convergence protocol (PDCP). The MAC performs hybrid automatic repeat request (HARQ) retransmissions and scheduling in uplink/downlink. The RLC is responsible for the segmentation and concatenation of Internet Protocol (IP) packets, and also performs error correction through automatic repeat request (ARQ). The PDCP handles IP header compression/decompression, user/control plane data transfer, and integrity protection and encryption. The layer-3 contains radio resource control (RRC) and NAS. The NAS performs user authentication, location registration, bearer context activation/deactivation through messages directly between

Table 2.1. Control plane protocol stacks of UE in LTE.

	Protocol	**Control tasks**
Layer-3	NAS	User authentication, location registration
	RRC	Radio resource usage, paging, RRC connection
Layer-2	PDCP	IP header (de)compression, encryption
	RLC	Segmentation, concatenation, ARQ
	MAC	Scheduling, error correction (HARQ)
Layer-1	PHY	OFDMA and SC-FDMA related functions

UE and MME. The RRC controls the radio resource usage, and manages signaling, data connections. The RRC is also responsible for establishing, maintaining, and releasing of an RRC connection between UE and E-UTRAN. The protocol stacks and their main responsibilities are briefly listed in Table 2.1.

In the following, we mostly focus on the physical layer (Layer-1) aspects of cellular networks. First, the characteristics of a wireless propagation are described in detail. Then, we elaborate on RF transceivers and various factors for performance limitations on the physical layer.

2.2 Wireless propagation

The overall end-to-end channel response of a cellular communication link is composed of the effect of the electromagnetic channel response as well as the hardware responses of the transceivers.

The propagation of the electromagnetic waves is often characterized by multiple temporally dispersed paths through which the propagating waveforms travel. The paths that are impacted by scatterers, reflectors, or diffractors are called non line-of-sight (NLoS), whereas the paths that directly connect the end points of a wireless link are called line-of-sight (LoS). The LoS path of a wireless channel typically has the shortest delay and strongest propagation gain, although multipath channels may be composed of only NLoS paths if an obstruction blocks the line of sight between transmitter and receiver.

Wireless fading can be grouped into three types: pathloss, shadowing, and multipath fading. According to the pathloss model, the strength of a transmit signal from an omni-directional antenna attenuates inversely proportionally to the square of the distance traveled in free space due to the conservation of energy principle, as the surface area of a sphere is proportional to its square radius. In other propagation mediums, though, the RF power of a signal decays differently depending on the texture of the medium as well as the impact of the surrounding obstructions. Experimental studies have shown that the pathloss on our planet is usually proportional to $1/d^{\eta}$, where d is the travel distance and η is the attenuation coefficient that is often measured between 1.5[1] to 4. Shadowing

[1]At mobile communication frequencies, a LoS propagation inside a building may exhibit a lower pathloss exponent than the free space, depending on the geometry and material of the inner walls acting as a waveguide [39].

occurs due to the terrain characteristics, such as large obstacles like buildings or hills, and causes medium-scale fluctuations in the receive signal strength. Shadow fading generates non-uniformity in the coverage area, which creates difficulties for cost-efficient network planning. Multipath fading causes variations in the amplitude, phase, and direction-of-arrival (DoA) characteristics of a receive signal. The variation can be rapid and substantial due to the reflectors and diffractors creating many electromagnetic waves from each path, as they can be accumulated at a receiving terminal either constructively or destructively based on their altered DoA and delay characteristics [40].

2.3 RF transceivers

The response of a wireless channel is also impacted by the transmit and receive antenna radiation patterns, which are often distorted by the imperfections in RF chain circuitry. These RF impairments are mainly caused by the oscillator phase noise, direct current (DC) offset, in-phase/quadrature (I/Q) imbalance, and nonlinearities in power amplifiers.

The local oscillator of an RF chain is responsible for the phase noise. When an oscillator employs phase-locked loop (PLL), the resulting phase noise follows a random process with finite-power. On the other hand, without PLL the phase noise varies slowly; but it may accumulate over time [41, 42]. In case of time-division multiplexed switching with a single RF chain [43, 44], the impact of phase noise is particularly significant in multi-antenna communications due to distinct phase noise components for different channel coefficients. In such MIMO setups, the overall RF phase noise with zero crosstalk is often modeled as a diagonal matrix; whereas non-zero terms appear with non-negligible crosstalk at the non-diagonal places of the phase noise matrix [45]. The DC offset errors can be caused by the reflections of a local oscillator, which are mistakenly picked up and down-converted to DC; or by rapid changes in signal strength, which lead to short-period overloads [46]. In direct conversion architecture, the I/Q imbalance occurs due to the mismatches between the in-phase and quadrature channels of a transceiver when the relative phase difference within the local oscillator signal is not precisely $\pi/2$ radian [47, 48]. Nonlinearity in a power amplifier causes distortions in RF signals with high peak-to-mean power ratios, which may also lead to out-of-band emissions. The nonlinearity of a commercial power amplifier is

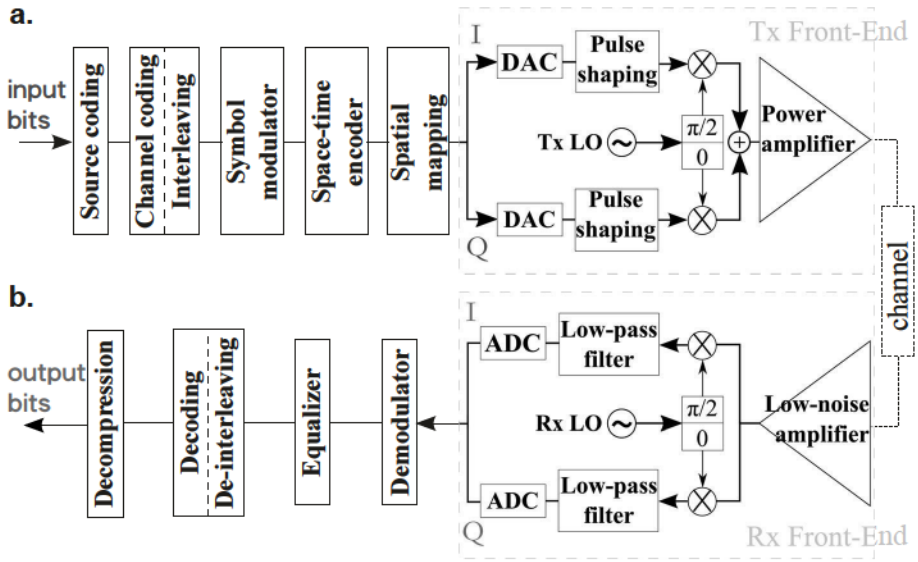

Figure 2.2. Direct conversion architecture of an RF circuitry for **a.** transmitters, **b.** receivers.

typically present due to the equipment manufacturers' desire for higher power efficiency, which requires extending the operation of an amplifier close to its saturation point within the available dynamic range. A common design practice is to strike a balance between higher power efficiency and better linearity.

Transmit signals of a typical transceiver are processed by source coding, channel coding/interleaving, modulation, space-time encoding, spatial mapping, digital-to-analog converter (DAC), pulse shaping, RF up-conversion, and power amplification; whereas receive signals go through low-noise amplifier (LNA), RF down-conversion, analog-to-digital converter (ADC), demodulation, equalization, decoding/de-interleaving, and decompression, as illustrated in Fig. 2.2.

From a deployment point of view one of the major challenges with very large antenna systems is related to their RF transceiver architectures. Massive MIMO systems, as a potential technology component in next generation cellular networks [31], are known to require extreme amounts of internal overhead due to data exchanges among very large number of antennas within the transceiver, which causes overwhelming hardware latency. In addition, transceivers should be composed of cheap RF components in order to reduce the expanding financial costs with more antenna elements. More detailed discussions on the RF hardware limitations of

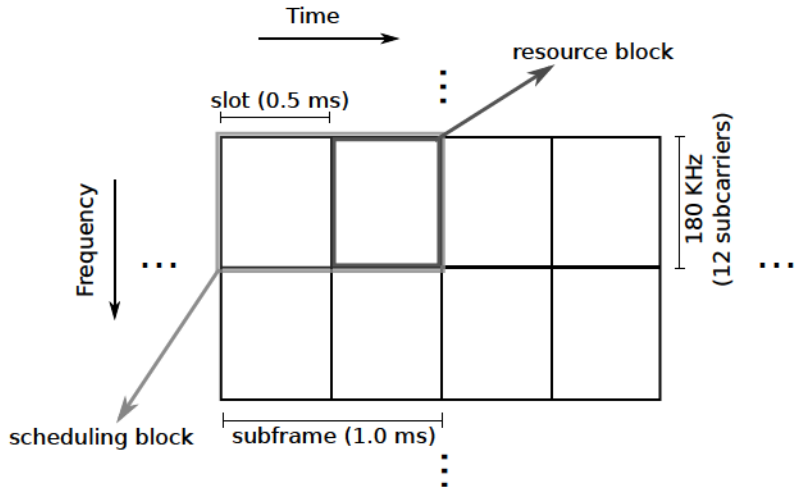

Figure 2.3. Illustration of an LTE scheduling block on downlink.

massive MIMO systems are available in Section 4.1.1.

Symbol transmissions on the downlink of LTE are based on OFDMA scheme. OFDMA is derived from OFDM modulation to enable simultaneous use of radio resources by multiple users, which helps achieve better multi-user diversity as well as flexibility in scheduling. Successive OFDM subcarriers are separated by 15 kHz and every group of 12 adjacent subcarriers forms a *resource block*. Hence, each resource block is 180 kHz and also has a time duration of 0.5 ms, which is called a *time slot*. A time slot is composed of either 6 or 7 OFDM symbols depending on the type of cyclic prefix. A pair of two successive resource blocks forms a *scheduling block*, which represents the minimum resource unit that can be assigned during resource allocation. An illustration of a scheduling block in LTE downlink is depicted in Fig. 2.3.

High data rate transmissions require the use of a wide bandwidth in order to tap into higher capacities whereas high user mobility leads to large Doppler spread. However, multipath channels with high-speed users often experience large temporal and spectral dispersions in wideband transmissions. Large delay spread causes inter-symbol interference (ISI) in time-domain and frequency-selectivity in frequency domain. Large frequency dispersion causes inter-carrier interference (ICI) in frequency domain and faster temporal variations of the propagation channel in time-domain. To provide support for high-rate and high-mobility scenarios in LTE, OFDM offers protection in wideband transmissions against these issues by means of applying cyclic prefix and multi-carrier modulation with

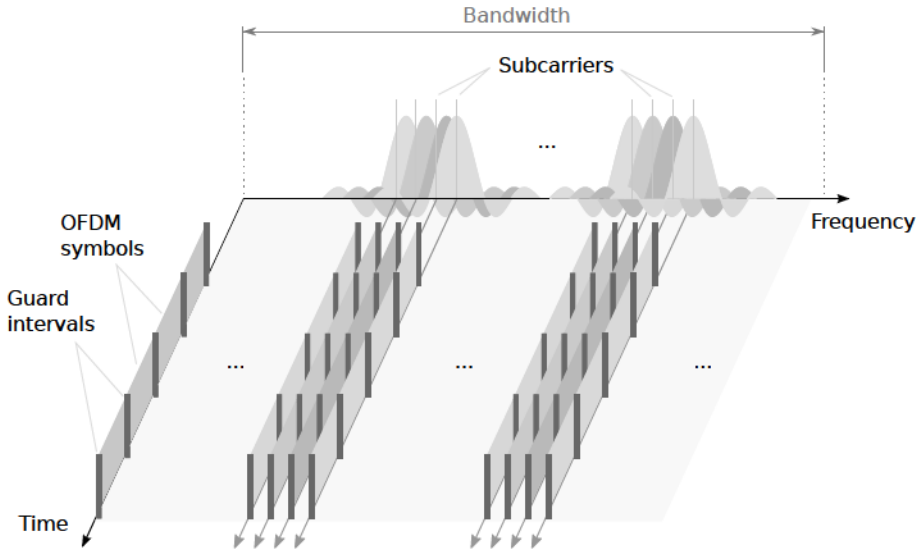

Figure 2.4. Time-frequency structure of an OFDMA signal.

pulse shaping, as illustrated in Fig. 2.4. Necessary channel information for a scheduling decision can be provided by separate CSI reports for different blocks due to resource division. As opposed to earlier cellular radio access technologies, LTE schedulers can assign both temporal and spectral resources to users, which unlocks additional flexibility that translates to better throughput and coverage performance.

Uplink of LTE embraces SC-FDMA scheme, which is mainly a precoded version of OFDMA. However, SC-FDMA facilitates lower power consumption and simpler amplifier design, both thanks to smaller peak-to-average power ratio, which makes SC-FDMA favorable for portable user terminals that require low power consumption and design cost. During signal modulation with SC-FDMA, transmit symbols are initially precoded by discrete Fourier transform (DFT) before being processed by a standard OFDM modulator, which involves baseband modulation, serial-to-parallel conversion, subcarrier mapping, inverse Fourier transform, and the addition of cyclic prefix.

Baseband modulation processes a sequence of binary bits into multilevel complex numbers based on the modulation format, such as binary phase-shift-keying (BPSK), quadrature phase-shift-keying (QPSK), or any 16/64/256 quadrature amplitude modulations (QAM). Modulated symbols are mapped to subcarriers and each block of these mapped data symbols are simultaneously transformed to time-domain samples by Inverse Dis-

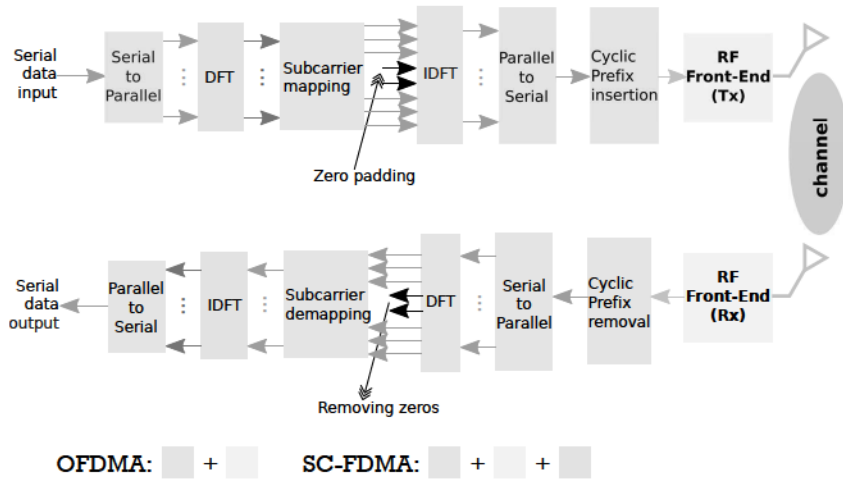

Figure 2.5. OFDMA and SC-FDMA communication architectures.

crete Fourier Transform (IDFT) after zero padding, which helps prevent potential aliasing that may be caused by the DAC. Next, a cyclic prefix block is copied from the end of these time-domain blocks of samples and added to the beginning. When transmit signals arrive at a receiver, the prefix part of the samples will have to be removed as it contains ISI from previous symbols. The additional DFT operation of SC-FDMA before subcarrier mapping allows each data symbol to be spread over multiple subcarriers, hence SC-FDMA offers higher frequency diversity gain compared to OFDMA due to such frequency spreading. The naming for SC-FDMA comes from the fact that data symbols are carried over a group of subcarriers which are transmitted simultaneously. In particular, the group of subcarriers carrying the same data symbol should be considered a single frequency band carrying the same information sequentially as a standard FDMA modulation. For signal reception, an OFDMA (or SC-FDMA) receiver processes demodulation by applying reversal operations (i.e., DFT, demapping, etc.) with respect to the steps of modulation, as illustrated in Fig. 2.5.

Deploying multiple antennas is being considered one of the key components for reaching the performance objectives in current and future cellular radio technologies by means of utilizing the following benefits. Multiple antennas provide receive and transmit diversities. The premise of receive diversity is related to that multiple antennas of a receiver acquire independent copies of the same transmit signal and the probability of all these copies simultaneously experiencing deep fading diminishes with

more antennas. In particular, receive diversity allows collecting more of the signal energy while avoiding channel fading. Transmit diversity also delivers resistance against multipath fading and decreases the variance of receiver-side SNR by space-time coding. For example, cyclic delay diversity (CDD) introduces virtual echos into the channel to artificially increase its frequency selectivity at the receiver [49].

In addition, multi-antenna communications can support beamforming, which performs directional transmission or reception of a signal. Beamforming is based on constructively combining multiple copies of the desired signal at a specific angle from a transmitter or receiver while creating interference nullifiers at other angles by means of introducing a phased array response to multiple antenna elements. The beamforming technique results in higher SINR and reduced power consumption via concentrated signal energy, particularly when carried out by a large antenna array.

The use of multiple antennas also allows spatial multiplexing on a MIMO link. When both ends of a link are equipped with multiple antennas, simultaneously transmitting multiple data streams usually leads to higher throughput, thanks to the improved dimensionality of the spatial domain. Multiple antennas are supported at both base stations and user terminals by most of the downlink transmission modes (TMs) in LTE-A [22]. Transmit diversity is used in TM2 and TM6. Spatial multiplexing is utilized in TM3, TM4, TM5, TM8, TM9, and TM10. Beamforming is applied in TM7 and TM8. Supported transmission modes and the required number of antennas by LTE-A are listed in Table 2.2.

2.4 Sources of impairments

This section is devoted to summarizing the major types of performance impairments in a typical cellular network apart from the hardware limitations of RF transceivers, which are already discussed in Section 2.3. The said performance impairments here indicate sources of interference signals, inadequate knowledge on the instantaneous channel status, connection issues during inter-cell hand-off, SNR loss by radio-over-fiber (RoF) dispersion, loss of LoS access, interference-unaware power control, limitations due to overhead or spectrum constraints, storage and battery limitations, necessary cost reduction for commercial feasibility, and so forth. In the following, we particularly focus on the sources of impairments that

Table 2.2. Transmission modes in downlink LTE-A [22].

TM	Task	Transmit antenna setup
TM1	Single transmit antenna	1 antenna (port 0)
TM2	Transmit diversity	2 or 4 antennas
TM3	Open-loop spatial multiplexing	2 or 4 antennas
TM4	Closed-loop spatial multiplexing	2 or 4 antennas
TM5	Multi-user MIMO	2 or 4 antennas
TM6	Closed-loop transmit diversity	2 or 4 antennas (1 layer)
TM7	Beamforming	multiple antennas (only port 5)
TM8	Dual-layer beamforming	multiple antennas (ports 7, 8)
TM9	Up to 8-layer spatial multiplexing	8 antennas (ports 7–14)
TM10	Up to 8-layer spatial multiplexing	8 antennas (ports 7–14)

are caused or can be recovered by multi-antenna deployments in a cellular network.

Inter-cell interference occurs due to the shared use of the same frequency band among multiple neighboring cells. Conventional prevention against inter-cell interference is based on dividing the whole spectrum into multiple orthogonal bands by *frequency reuse factor* (FRF). The separated bands are then assigned to cells in such a way that none of the adjacent cells use the same spectrum. However, due to the demand for higher spectral efficiency, more opportunistic adaptive schemes are considered in modern networks that can find a better balance between spectral reuse and efficiency, which also renders higher fairness versus throughput, or vice versa. These trade-offs are elaborated with more details in Chapter 3 regarding the multi-user multi-cell resource allocation problem on the downlink of a cellular network.

Multi-user interference (MUI) refers to the interference signal among a group of users sharing the same resource blocks in a downlink broadcast transmission. When multiple user terminals are co-scheduled on downlink, all data transmissions are unintentionally received at every user terminal, therefore causing severe reductions in user SINR metrics. Therefore, MUI mitigation is often desired without compromising spectral efficiency. To do so, a transmitting station must apply spatial suppression by means of pre-filtering the transmitted user data with respect to individual channel states. However, in case of inadequate knowledge of CSI as well as significant number of terminals receiving transmission on downlink,

an optimal suppression strategy requires detailed considerations. Chapter 3 addresses such transmission strategies in the presence of MUI on a multi-user MIMO downlink.

Self-interference is a general term that defines the unintentional leakage of a transmit signal toward its own transceiver. In this regard typical ISI and ICI during an OFDM transmission are considered self-interference due to the same signal stream interfering with itself [50]. Nevertheless, the self-interference term is more popularly attached to full-duplex communication systems, which transmit and receive on the same temporal and spectral resources through the same end-to-end connection. The self-interference signal on a full-duplex relay link is often called *loop-interference* due to the transmitted signal from one side of the relay curling toward its opposite end which houses the reception antenna equipment. Self-interference is also experienced in device-to-device (D2D) communications when operated in full-duplex mode. Chapter 6 further discusses self-interference signals, mitigation techniques, as well as associated imperfections on the topic of full-duplex relaying.

Multiple access interference is present in multicast transmissions during uplink when a plurality of users are simultaneously transmitting toward the same base station on the same spectral resources. Typically, multiple access interference can be avoided by inducing orthogonality in spatial domain among user links by advanced receiver techniques or by orthogonal code division. Conventionally, many multi-user detection schemes have been developed to combat multi access interference during uplink data transmissions [51]. More recently, a spatial division technique, called spatial division multiple access (SDMA), is studied to support multi-user access on uplink when a sufficient number of antennas is available at the base station. However, these techniques practice prevention schemes against multi access interference based on the available knowledge on the channel states. Multi access interference is also present during a blind channel estimation on uplink when a plurality of user terminals are simultaneously transmitting non-orthogonal signals to the same base station. Chapter 4 deals with such impairment caused by a multi access transmission on the uplink of a massive MIMO network during channel estimation.

Inaccurate channel estimation is another source of performance impairment. Significant CSI mismatches cause residual errors during interference suppression and result in inefficiency in resource scheduling. Chap-

ter 3 studies the impact of CSI mismatch on downlink multi-user MIMO performance. In addition, Chapters 4 and 5 deal with the problem of CSI estimation for massive MIMO systems operating in time division duplex (TDD) and frequency division duplex (FDD), respectively. Particularly, Chapter 4 is also concerned with RF impairments of the transceivers with massive antenna arrays and studies antenna calibration techniques for reciprocity recovery in TDD mode.

Downlink CSI mismatches at base stations also occur as a result of compressions during user feedback reporting. The impact of the compression rate is particularly significant when the reported CSI vectors are long due to a massive-sized antenna array setup. There have been several attempts for opportunistic schemes in literature based on efficiently transferring CSI by means of adaptive codebooks, iterative feedback reports, or exploiting pre-known channel characteristics. Chapter 5 elaborates on various CSI compression schemes and feedback reports in non-reciprocal FDD-mode massive MIMO communications.

3. Downlink Multi-User Transmission

MIMO technology has been originally adopted into cellular communications as a single-user point-to-point transmission technique. Later, it evolved into multi-user mode in the context of spatial division multiple access (SDMA) by means of exploiting the spatial signal domain to separate the users sharing the same spectrum. Such resource sharing among multiple users on the downlink of an OFDMA based transmission can offer higher achievable rates by optimized precoding and resource scheduling [21, 22].

Multi-user precoding techniques are commonly designed by improving the user signal quality while mitigating any undesired signal components on the link level [52, 53, 54]. Multi-carrier resource allocation techniques for multi-user scheduling aim to achieve higher spectral efficiency in a network while sustaining a desired level of fairness among users. Such resource schedulers are often designed based on either a heuristic approach [55, 56] or the maximization of a network-wide utility metric [57], such as proportional fair [58] or round-robin [59, 60]. Conceptually, an optimum joint design for multi-user resource allocation and precoding is considered too hard to attain in a typical cellular network as the search space is extremely large within the allotted time duration [55].

Many efficient algorithms have been studied in literature, which can intelligently generate high sum-throughput rates, network fairness, and flexibility for user priority. In this chapter multi-block resource allocation methods coupled with multi-antenna transceivers are discussed for forward link transmission. The contributions of this thesis are presented as well. Publication I presents a novel joint precoder and receiver method for a multi-user transmission scenario. The method is based on decoupling the filters with the help of introducing two target channel matrices which are designed for high SNR conditions. In addition, Publication II

explores multi-carrier multi-user schedulers with successive interference cancellation based on greedy and genetic based heuristic algorithms. The described schedulers optimize user selection, user ordering, as well as the assignment of scheduling blocks by heuristic iterations in order to maximize targeted performance metrics.

3.1 Precoding design

In 1983, Costa [61] demonstrated that the sum-capacity of a wireless transmission, even with additive interference, can be attained by dirty paper coding (DPC) as long as the knowledge of the interference signal is noncausally available at the transmitter. He showed that the optimal transmitter can adapt its signal to co-exist with the interference rather than trying to mitigate it. Specifically, if the interference signal at a receive terminal is known by its serving transmitter in advance, the achievable data rate can theoretically reach the system capacity of the interference-free transmission by means of adapting the transmit encoder according to the directional information of the future interference signal. This is illustrated by an analogy to writing some text on a dirty piece of paper; hence the naming of the optimal interference-aware encoder [61].

As such optimal encoding requires accurate knowledge at the transmitter about the receive-side interference, DPC is considered applicable to cellular communication systems mostly in the context of downlink multi-user transmissions as the interfering signal originates from the base station itself in the form of MUI[1]. In addition, as the served user terminals in a resource sharing based broadcast transmission are uninformed about the CSI of others, receiver-side MUI mitigation can only be performed sub-optimally. Although the proof of existence by Costa is elegant, the optimal encoding strategy for DPC is still unknown and the best attempts in literature are known to be computationally very demanding [62]. Moreover, DPC is sensitive to CSI errors, especially in multi-user downlink transmissions where accurate MUI prediction is hard.

Tomlinson-Harashima encoder is originally developed for single-stream transmissions to combat ISI by means of pre-subtracting the interfering components created by the previously transmitted symbols [63, 64]. However, this encoder is often considered for multi-user transmission sce-

[1]Or in the context of full-duplex in-band relaying as the *loop-interference* also originates from the same station point.

narios in the context of pre-subtracting the precoded symbols that are intended for other co-scheduled users. The performance of Tomlinson-Harashima precoder relies on very accurate CSI; any small mismatch can cause SNR saturation in response to increased transmit power [65].

The fundamental idea behind Tomlinson-Harashima precoding is illustrated in Fig. 3.1. The interference is pre-subtracted from the transmit symbols, after which a modulo-Δ operator is applied. Hence, the strengths of the transmit symbols are effectively constrained within $[-\Delta/2, \Delta/2)$. In case of complex-valued symbol modulations (such as, 16-QAM), separate modulo operations can be used for the real and complex parts of the symbols. The role of the modulo operator here is to keep the transmit power minimal. Finally, the receiver applies the same modulo operator to recover the intended symbols.

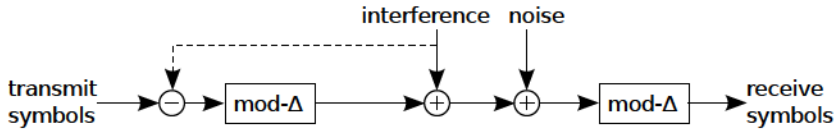

Figure 3.1. Tomlinson-Harashima encoding mechanism.

Trellis and convolutional coding methods are often applied together with the Tomlinson-Harashima encoders in order to overcome their inherent issues on shaping, modulo, and power limitations [66, 67]. Trellis precoding [68] combats the shaping loss at high SNR regime by taking into account both present and future interference sequences. Convolution decoding can recover the modulo and power losses present at low SNR regime when applied with partial interference presubtraction, which leaves out uniform-plus-Gaussian distributed residual noise [67]. Similarly, nested lattice structure [69, 70, 71] based methods can also be applied with the Tomlinson-Harashima encoders in order to combat the described inherent limitations.

However, these DPC based precoding techniques always require computationally heavy signal processing, which may not be feasible at the base station of a cellular network [72], e.g., due to limited time duration in the order of milliseconds. In addition, the necessary non-linear processing is often tailored for a specific symbol modulation scheme. On the other hand, linear filters are generically designed for any modulation method. In addition, they can still theoretically achieve the same multiplexing gain as DPC even with lower complexity [73, 74]. This can be realized by means

of spatial-domain MUI suppression, as elaborated in the following.

3.1.1 MUI suppression

A well-known spatial-domain suppression is zero-forcing; a linear precoder that is widely used and works particularly well under low CSI mismatch and high SNR conditions. The design is typically based on presuppressing the entire interference toward the unintended users by means of calculating the pseudo-inverse of the broadcast channel response at the base station. In case of multi-antenna terminals, block diagonalization (BD) is known to perform better while still removing out the MUI, which can be derived by either singular value decomposition (SVD) [52] or QR factorization [75]. The design process follows first eliminating the MUI for each user terminal by aligning its transmit filter to the nullspace of the right singular matrix of the aggregate leakage CSI, after which the decoupled effective channels of the MUI-free communication can be optimized by matched filtering. In case of massive transmit antenna arrays, MUI tends to disappear without any mitigation thanks to the asymptotic orthogonality, hence a matched filter can be directly applied at the transmitter[2]. More advanced linear precoders can be designed based on improving the SINR by means of minimum mean square error (MSE) [53] or maximum signal-to-leakage plus noise ratio (SLNR) criterion [76].

In case some or all user terminals employ multiple antennas, the optimal design of the filters becomes too hard to derive due to non-convexity. One approach forward is to decouple this joint design process of the transmit and receive filters, therefore establishing sub-optimal solutions. Specifically, MUI mitigation can be handled entirely at the transmit side while the receive filters are brought into the picture at the second stage of the optimization. For instance, transmit precoding can be considered a two-layer filtering operation, where the first filter is designed to mitigate the MUI alone while the second transmit filter as well as the receivers are derived to produce minimum MSE [75]. Such two-layer transmit filtering can strike a balance between MUI and noise mitigations. An alternative, yet similar, example can be an optimization process based on minimizing the interference-plus-noise signal as a whole by the regularized BD, after which the user data signals can be optimized disjointly [77]. Publication I demonstrates an optimized filter design by such interference-plus-noise

[2]More on this in Section 4.2.

minimization in the presence of CSI errors. In addition, MUI can also be mitigated in a successive manner for each user terminal, usually combined with a user selection mechanism in multi-block resource schedulers; as elaborated more in Section 3.2.

3.1.2 Impact of CSI mismatch and thermal noise

A significant requirement for high-performing multi-user transmissions is the availability of accurate CSI. In case of perfect channel knowledge, zero-forcing is well known to be an effective linear transmission scheme, specifically in high SNR conditions. When CSI mismatch is small, novel schemes can opportunistically allow negligible residual interference by means of imperfectly orthogonalizing user channels, therefore helping to achieve higher overall spectral efficiency via more efficient use of the spatial degrees of freedom [78]. In a realistic communication scenario, however, CSI mismatch usually occurs due to channel estimation and quantization errors, as well as temporal evolution between successive channel measurements due to mobility. Base stations, particularly in FDD mode, rely on uplink feedback from user terminals to acquire CSI. Due to the limited available control signaling, only a small number of bits can be signaled to convey the channel information during each feedback interval, hence generating quantized CSI reports. With the addition of channel estimation errors during downlink training, acquired CSI becomes less reliable at the base station. It is well known that CSI mismatches cause an *error floor* at high SNR. The general rule of thumb in literature for combatting such degradations of CSI mismatch is related to estimating its error variance, which can be accomplished via stochastic processes [79], quantization error estimation methods [80], or Shannon's rate-distortion theory [54]. The sensitivity of the transmit filter is then calibrated by a linear regularization approach based on the estimated variance of the CSI mismatch.

The response of a MIMO channel often contains spatial correlations as a result of the scattering environments and antenna coupling at either transmit or receive end of a link. Channel correlation is notorious for limiting the multiplexing gain in single-user MIMO transmissions, hence reducing the maximum spectral efficiency [81]. However, it has been also shown that correlated fading can be beneficial to multi-user MIMO communications as it can reduce the required amount of feedback overhead, particularly when multiple users experience the same transmit-side cor-

relators [82, 83]. Hence, channel correlations are occasionally considered an opportunity in certain use cases. With the recent popularity of massive MIMO system studies, various compression schemes have been investigated for limited CSI feedback overhead. Exploiting the spatial-domain correlations of a downlink channel is proven to be one of those promising compression techniques [84, 85], as discussed in Section 5.2.2. In essence, such two-sided exposition of spatially correlated channels indicates that the best way of approaching a challenging problem is sometimes to search for a solution from a different angle of view instead of giving up in the first place.

An optimal transmit filter in multi-user MIMO approaches zero-forcing (ZF) and maximum ratio transmit (MRT) schemes under low and high thermal noise conditions, respectively, whereas minimum MSE-based filter (also known as *transmit Wiener filter*) can strike an optimum balance between desired signal maximization and MUI suppression [86]. Although transmit-side operations have no direct impact on the user-side signal degradations that involve any thermal noise and inter-cell interference by external sources, the optimum balance of the transmit Wiener filter is known to be capable of influencing such degradations at user terminals by a *gain controller*. The design procedure is fundamentally based on harnessing the available transmit power at the base station to scale up the transmit signal amplitudes. In return, user terminals become exposed to additive noise and inter-cell interference with relatively smaller amplitudes than those of the incoming desired symbols. In essence, the core benefit of the gain controller is in relaxing the transmit-side filter design from any norm constraints, hence providing more flexibility to serve users under diverse SINR conditions. Capturing such an optimal balance between MUI and any other undesired signals by the described procedure is also known as *regularization* [87]. Of course, the thermal noise (or any external source of interference) at a user terminal cannot be observed by the base station. One approach for choosing the gain controller can be by means of deriving an estimate for the long-term average power of those noise-plus-interference signals based on the reported channel quality indicator (CQI) information. Another approach can be based on a local optimization at the base station, such as based on asymptotic SINR computations in case of large-scale MIMO [88].

The impacts of CSI mismatch and spatial correlations on multi-user downlink performance are studied by simulations in Fig. 3.2. The nu-

merical results are demonstrated by sum-rate plots and generated on an isolated cell with one 10-antenna base station and 5 user terminals each equipped with 2-antennas. The simulation setup is as follows. The propagation channels are composed of flat fading independent identically distributed elements and the thermal noise is modeled by zero-mean complex Gaussian random variables. The CSI mismatch is rendered via only channel estimation errors based on the Gauss-Markov uncertainty model with β characterizing the error coefficient as follows.

$$\mathbf{H}^{(\text{True})} = \sqrt{1-\beta^2}\mathbf{H}^{(\text{Est.})} + \beta\mathbf{E}, \qquad \text{where } 0 \leq \beta \leq 1, \tag{3.1}$$

for which $\mathbf{H}^{(\text{True})}$ and $\mathbf{H}^{(\text{Est.})}$ respectively denote the true and estimated CSI matrices; $\mathbf{E} \sim \mathbb{CN}(\mathbf{0}, \mathbf{I})$ denotes the random error matrix. Observe that $\beta = 0$ produces true CSI estimates and $\beta = 1$ indicates zero correlations between the estimated and true CSI matrices. Transmit-side channel correlations are generated by the Kronecker model with α representing the spatial correlation coefficient as follows.

$$\mathbf{H}^{(\text{True, Corr.})} = \mathbf{H}^{(\text{True})}\mathbf{R}^{1/2}, \quad \text{where } \mathbf{R}(i,j) = \begin{cases} \alpha^{|i-j|}, & \text{if } i \neq j, \\ 1, & \text{otherwise.} \end{cases} \tag{3.2}$$

for which $\mathbf{H}^{(\text{True, Corr.})}$ denotes the spatially correlated channel response, $\mathbf{R}$ denotes the transmit-side correlation matrix, and $0 \leq \alpha \leq 1$. Observe that $\alpha = 0$ renders spatially uncorrelated channels and $\alpha = 1$ causes full spatial correlations at the base station. In addition, channel SNR values are set to either 0 or 15 dB to assess the impact of the thermal noise on the performance. The numerical results compare the performance of standard linear precoding methods under channel correlations and CSI errors. As expected, MUI suppression proves to be less favorable under higher thermal noise, greater spatial correlation, and larger CSI mismatch conditions.

3.1.3 Joint transceiver design

When some or all of the terminals are equipped with multiple antennas, the optimum design for the transmit and receive filters becomes coupled with each other, hence creating a space to explore the optimal solution by a joint design approach [89, 90]. However, such joint design problem is known to be non-convex, therefore hard to solve efficiently. Low-complexity optimization algorithms have been investigated by many researchers over the years to attain close-to-optimal performance results.

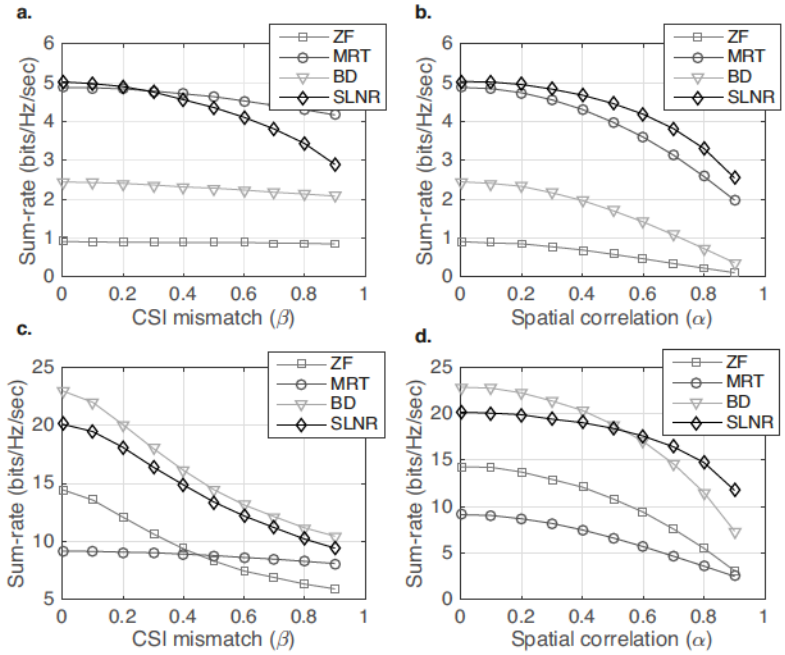

Figure 3.2. Impacts of CSI mismatch (*left-side*), spatial correlation (*right-side*), and thermal noise (*top vs. bottom*): **a.** SNR = 0 dB, $\alpha = 0$; **b.** SNR = 0 dB, $\beta = 0$; **c.** SNR = 15 dB, $\alpha = 0$; **d.** SNR = 15 dB, $\beta = 0$. Conclusions are: 1) Interference cancellation is more beneficial under low thermal noise; 2) Both CSI mismatch and spatial correlations devalue MUI suppression.

General strategies for these attempts can be classified into two categories. The first category is based on exploiting the available knowledge of the channel conditions. For example, when the amount of CSI mismatch or the thermal noise is estimated for a user terminal [79, 80], its linear filter can be approximated by a suitable unilateral closed-form scheme, hence simplifying the design process for other users. In other words, the goal revolves around reducing the extent of the joint solution set while ensuring minimum amount of diversion from the global optimum. The second strategy depends on decoupling the transmit and receive filters by fixing each of them, hence converting the joint optimization into two separate convex problems. Next, the jointly optimized filters can be approximated by an iterative algorithm, while its convergence may not be guaranteed.

Although such iterative algorithms are widely adopted in literature (i.e., often based on minimum MSE criterion), they usually share two major limitations. The first issue is that most of those algorithms focus on MUI cancellation while overlooking the impact of thermal noise on the perfor-

mance [91]. The second issue is related to the optimal power distribution among users. One simplified approach regarding the second issue has been applying per-user power constraints, such as unit power for each user [92], although optimal power allocations should be derived according to users' channel status. One attempt for such optimization has been designed based on an exhaustive search via the Lagrange method [93]. Despite its theoretical charm, such an exhaustive search is not feasible to operate in a short time window during an ongoing downlink transmission. On the other hand, a practical point of view toward the joint transceiver problem suggests that it is possible to compensate for the rate loss due to the additional control signaling, only when the channel conditions are favorable for high-rate transmissions. In the light of this opportunistic view, Publication I introduces two target channel matrices that are suitably designed for high SNR environments. A target channel matrix is here identified as an approximation of the desired effective channel response in a multi-antenna transmission.

The joint transceiver design of Publication I is based on integrating the optimal gain controller for power allocation into the minimum MSE expression and decoupling the transmit and receive filters with the help of the target channel matrices. Figure 3.3 compares the performance of this joint optimization with some of the most popular linear filtering methods from literature. One of them is known as the *generalized minimum mean square error (MMSE) channel inversion* [75] whereas the other is referred to as the *regularized block diagonalization* [77]. The described joint optimization outperforms the reference schemes under high SNR conditions. It also proves to be immune to large CSI mismatches as long as they are less significant than the degradations caused by the thermal noise. On the other hand, the proposed joint scheme exhibits slightly larger rate loss when exposed to extreme spatial correlations.

3.2 Multi-carrier resource allocation

As part of the RRM, eNB is responsible for allocating resource blocks, scheduling transmissions, and monitoring real-time traffic in order to meet the required QoS. The adoption of OFDMA on the downlink of cellular networks enables resource allocation techniques that are precise and multi-dimentional in the spatial, temporal, spectral, and even multi-user domains. With such design flexibility as well as the inevitable need for

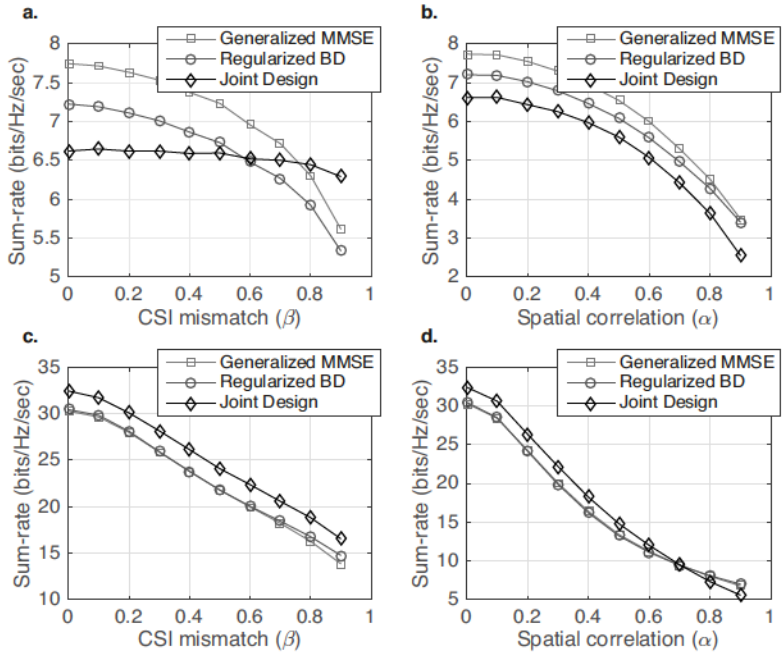

Figure 3.3. Impacts of CSI mismatch (*left-side*), spatial correlation (*right-side*), and thermal noise (*top vs. bottom*): **a.** SNR = 0 dB, $\alpha = 0$; **b.** SNR = 0 dB, $\beta = 0$; **c.** SNR = 15 dB, $\alpha = 0$; **d.** SNR = 15 dB, $\beta = 0$. Conclusions are: 1) Proposed joint design outperforms others under high SNR as intended; 2) Proposed joint design is also more immune to large CSI errors; 3) Extreme spatial correlations degrade the joint scheme comparatively more notably. Publication I, © 2014 IEEE.

spectrum efficiency, the problem of radio resource allocation has been receiving a lot of attention in the research community.

The main radio resources are frequency bands, time slots, and available transmit power. In case of multi-antenna transceivers at a base station, multiple users' connections can be simultaneously sustained using distinct spatial resources on the same spectral band and time slot. The distribution of all available resources among active connections can be optimized by either a utility based or a heuristic based approach. The utility based resource allocation provides robust and generalized solutions whereas heuristic methods offer faster and more efficient solutions in an opportunistic manner. In either case downlink resource allocation assigns the available scheduling blocks to a group of users in order to maximize a network-wide objective function.

Most OFDMA schedulers operate in two consecutive stages to reduce the complexity in time-domain and frequency-domain [94]. Time-domain

scheduling selects a list of candidate users to be scheduled whereas a frequency-domain scheduler subsequently determines the assignment of the available scheduling blocks to those users in the candidate list. Since a time-domain scheduler operates independently of the resource blocks, its decisions are determined based on wideband CQI reports, which can be considered binary-coded integers based on the observed SINR at user terminals [21]. On the contrary, both packet schedulers can target optimizing either the same or different utility metrics, such as system-wide throughput, fairness, minimum packet drop, etc., depending on the desired scheduling characteristics of the network. Apart from fulfilling the QoS requirements, the goal is to provide an optimal balance between spectral efficiency and fairness depending on the service type. For instance, fairness among users can be achieved by the time-domain scheduler while the frequency-domain scheduler focuses on maximizing the throughput metric [95, 96]. In real-time services, such as VoIP, packet schedulers often prioritize latency requirements [97]. Instead, both packet schedulers can employ *proportional fair* metric, which allocates resources to users with the highest instantaneous throughput relative to their average past throughput.

Frequency reuse among multiple cells is well known to be useful for improving the cell-edge user performance by avoiding severe co-channel interference, which is widely used in the conventional 2^{nd} generation GSM networks [98]. However, spectral efficiency can still be improved while preserving the cell-edge performance by fractional frequency reuse (FFR), which offers two different FRFs to every cell. The FFR can be implemented by assigning an FRF greater than unity to cell-edge regions and unit FRF to cell-center regions [99]. Major design problems on FFR is related to introducing an optimal division between the intra-cell regions with the corresponding optimal distribution of spectral bands, and also related to designing adaptive algorithms that can sense the instantaneous network conditions and proactively alter the FFR parameters. However, advancements in signal processing, more computational capabilities of user terminals, and especially the flexibility of fast frequency-domain schedulers allow advanced interference management techniques that can potentially render frequency reuse schemes obsolete. It is well known that non-unitary FRF in an OFDM system reduces the available number of resource blocks for scheduling, therefore causing reductions in achievable throughputs, particularly drastically for high-bandwidth services [100].

The main distinction between utility based and heuristic schedulers is as follows. Utility based methods evaluate the level of user satisfaction in case of a certain resource allocation by a quantifiable metric, such as throughput, fairness, or latency. The final scheduling decision is then determined by picking the set of allocation decisions that offer the maximum utility metric among all possible combinations. Utility based methods are reliable when the scheduling problem consists of a small search space. On the contrary, heuristic schedulers seek sub-optimal solutions through a *rational* search mechanism (instead of an exhaustive brute-force attempt) by means of following a systematic set of optimization steps, which exclude any unpromising combinations. Heuristic methods are usually beneficial when the optimal solution is practically unfeasible to attain as they perform iterative low-complexity computations. Utility based schedulers have been extensively studied for past cellular radio technologies as those schedulers are very suitable for less sophisticated network needs. On the other hand, growing demand for higher spectral efficiency increases the number of transmit antennas, targeted multi-user terminals, and available spectral resources, hence justifying the research interests in low-complexity heuristic schedulers more strongly than ever. In the following, the rest of this chapter focuses on the heuristic methods for downlink multi-user schedulers. In addition, the proposed genetic and greedy based heuristic schedulers by Publication II are also discussed and analyzed with numerical evaluations.

3.2.1 Heuristic-based scheduling

Allocation of available scheduling blocks among multiple user terminals is a combinatorial optimization problem with non-linear objective function and constraints. Hence, the optimum scheduling strategy is often attainable only via an exhaustive search, which is impractical in a large search space. On the other hand, heuristic algorithms are known to be capable of rendering acceptable, though sub-optimal, results with low computations. Therefore, network schedulers based on heuristic algorithms are considered favorable for practical deployments. Heuristic schedulers can operate in three distinct stages: The first step involves choosing the users that are to be served, which is conventionally addressed as *resource allocation* problem. The second step deals with the distribution of the available resources among the selected users, which is called *resource assignment*. The third step is related to ordering the allocated users on every schedul-

ing block with respect to either increasing or decreasing priorities in case the transmit station employs a successive type of MUI pre-cancellation.

Most popular heuristic methods for downlink schedulers include greedy algorithm [101], simulated annealing [102], tabu search [103], genetic evolutions [104], swarm optimization [105, 106], neural networks [107, 108], support vector machines [109], all of which are mainly developed in order to tackle *high time-complexity* or *premature local convergence* issues inherent to non-heuristic search algorithms. Arguably the most famous example in illustrating the gravity of the said issues is known as the *traveling salesman problem* (TSP). Imagine that a salesman has to visit n cities by starting from his home city and passing through each of the other cities only once until he returns his home again. Transportation is possible between any two cities and all transportation costs are known by the salesman prior to his journey. The question is in which order this salesman should visit the cities so that his total expenditure is minimized. Clearly, there are $(n-1)!$ possible travel plans. For $n = 15$ cities, an exhaustive search requires computations over more than 8×10^{10} combinations!

Greedy algorithms are based on picking a locally optimized choice at each stage of a large combinatorial optimization problem. The name of the algorithm comes from its selfish (or greedy) nature in always relying on the local optimality at any stage irrespective of other stages. Although a greedy algorithm always finds the best local optima, it does not necessarily render results close to the global optimum. In the context of the TSP a greedy algorithm may suggest the salesman every time to visit the city that will cost the least amount of expenditure for the short term. Since the rules disallow revisiting a city twice, the cost of visiting the subsequent city always tends to increase during the course of the salesman's travel. But, picking the cheapest visit at each step often does not lead to the most cost-efficient travel plan. Schedulers based on such greedy algorithms are known as low-complexity and high performance solutions for multi-user resource allocation problems. However, the required computations are usually more extensive due to the necessary metric calculations at every optimization stage, as compared with the computations of other heuristic methods, such as genetic algorithms. Greedy schedulers may introduce a simplified pre-selection mechanism or apply an intermediate user grouping to relax the complexity by means of limiting the local search space [110, 111].

The genetic algorithm is a heuristic search method based on evolutionary iterations of a random class of sample populations. The algorithm was initially developed by Holland [112] by mimicking the lifecycle of chromosomes, which involves reproduction, recombination, crossover, and mutation processes. At every generation a new set of individuals is created by reproduction followed by a natural selection that decides the survivors based on their *fitness* metrics. Similar to the natural evolution, this process leads to a gradual improvement of the population with fitter offspring that are more likely to survive than the members of the past generations. A genetic algorithm may resolve the TSP by randomly picking a sample number of travel plans among all possible combinations. Each complete path represents a unique chromosome designed as a sequence of binary digits and the corresponding travel costs are multiplicatively inverted to denote the associated fitness metrics. The chromosomes are processed by crossover and mutation, which respectively indicate randomly located exchanges and flips of some bits. At the end of continuous evolutions over sufficient number of generations, the path with the highest fitness metric is concluded as the best choice of travel plan for the salesman. A genetic based algorithm requires few computations besides negligible binary swap operations; and usually leads to satisfactory results, although sub-optimal, as long as the search space is large. Downlink schedulers based on the genetic algorithm are considered low-complexity methods that gradually approximate the optimum resource management. Such schedulers can be designed to simultaneously handle the allocation and assignment of resources, as well as successive user ordering within the evolution of each population.

Publication II designs greedy and genetic based schedulers for a multi-carrier multi-user downlink transmission and analyzes their performance in a system-level simulation setup. Instead of running multiple unconnected algorithms for each scheduling block as commonly embraced in literature [104, 113], both of these schedulers here simultaneously operate on all available scheduling blocks and successively introduce a new user assignment at each iteration, as described in more details below.

The greedy-based scheduler of Publication II functions as follows. First, the time-domain scheduler selects a pre-defined number of users among all active users based on a wideband utility function. Second, the frequency-domain scheduler evaluates the selected users' utility metrics on every scheduling block and assigns the users with the highest metrics. As these

Table 3.1. Greedy-based Multi-Carrier and Multi-User Scheduler

Step#1:	*Time-domain scheduler* — Selection of candidate users:
	▶ A pre-defined number of users are selected among those requesting service based on maximum wideband utility
Step#2:	*Frequency-domain scheduler* — Primary user assignment:
	▶ The user with the highest metric is assigned on every scheduling block as its primary user
	▶ The primary users are associated to hermitian precoders for every resource block
Step#3:	*Frequency-domain scheduler* — Secondary user assignment:
	▶ On each scheduling block the benefit of assigning an extra user is evaluated based on the utility metrics
	▶ If a candidate user among the selected group can improve the collective utility of a scheduling block,
	• The best candidate is assigned as a secondary user on that resource block
	• The precoder for the new assignee is based on the zero-leakage rule toward the earlier assignees
	• *Step#3* is repeated on that scheduling block for the next secondary user
	▶ Otherwise, the scheduling decision is finalized on that resource block

assignees are the primary users of their scheduling blocks, the precoders should be matched filters. Next, the utility metrics of the initially selected users are re-evaluated on top of the last user assignment and the ones reaching the highest metrics are assigned as the secondary users for that scheduling block if the additional user assignment improves the overall utility metric on the associated scheduling block. The precoders for these secondary users are based on the zero-leakage criterion toward the previous assignees of that block so that the latest user assignment does not temper with the earlier scheduling decisions. Such successive multi-user assignment ends once the utility reduction for the previous assignees outweighs the benefit of assigning an additional user. The steps of the greedy-based scheduler are briefly listed in Table 3.1.

The proposed genetic-based scheduler of Publication II operates through the standard steps of a genetic algorithm, as explained below.

Initialization involves constructing a set of chromosomes as binary sequences based on a set of pre-defined rules. These rules describe the required chromosome structure and depend on the scheduler design. Publication II constructs two types of chromosomes for the primary and secondary user assignments. The primary chromosome structure is composed of *head* and *tail*. Each bit in the head represents one of the active users. A "binary–1" bit indicates *selection* for the corresponding user, whereas a "binary–0" indicates *no selection*. Hence, the head takes over the tasks of a time-domain scheduler. The tail allocates a group of binary digits for every scheduling block so that each group can assign a unique number for every selected user, which indicates the scheduling decision for primary user assignments. On the contrary, secondary chromosome structure is composed of only tail and preserves a group of binary digits for each scheduling block. Each of these binary groups can uniquely indicate an index of one candidate user for a secondary assignment on the corresponding resource block unless the binary group contains "all binary–0"s, which indicates *no more secondary assignments*. Hence, the tails mimic a frequency-domain scheduler by the first and second chromosome structures, respectively, for the primary and secondary user assignments. The proposed pair of chromosome structures is illustrated in Fig. 3.4a. As a comparison, some other multi-carrier schedulers in literature consider a single chromosome structure for multi-user MIMO resource allocation by simply extending the length of the chromosome sequences designed for one resource block [114] (See Fig. 3.4b for the illustration). However, such a chromosome structure renders longer evolutions over more generations until the population reaches a satisfactory fitness. Instead, assigning an additional user at every iteration shortens the chromosomes, therefore reducing the runtime of the algorithm by faster convergence. Alternatively, there are also so-called multi-carrier schedulers which run independent instances of a genetic algorithm designed for a single resource block [104] (See Fig. 3.4c for the illustration). However, such methods are inefficient with a large number of resource blocks and are also unable to fulfill user-based multi-carrier scheduling restrictions.

Selection process evaluates the fitness of every chromosome in a current generation. Two chromosomes are picked randomly based on their fitness metrics. The picked chromosomes are then paired for breeding.

Breeding operation produces child chromosomes from the selected parents. A random bit index is picked with uniform distribution and all the

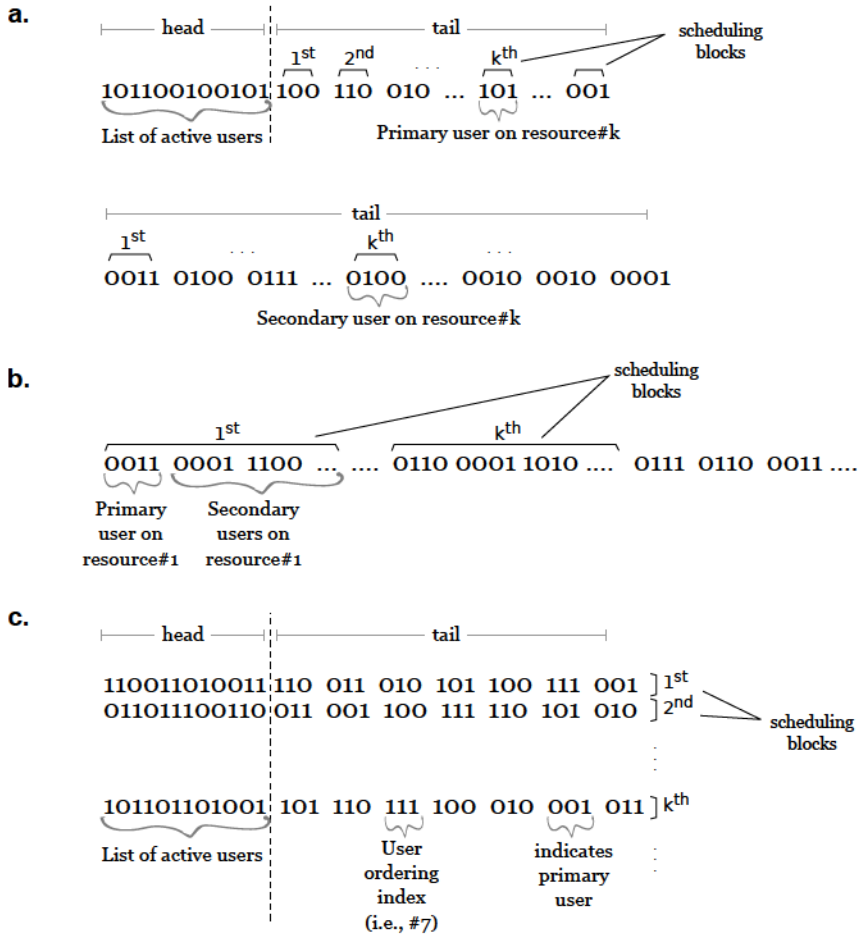

Figure 3.4. Alternative binary-coded multi-carrier chromosome structures for genetic-based schedulers: **a.** Former (*top*) and latter (*bottom*) structures as proposed in Publication II, **b.** Elongated form of a typical single-carrier based chromosome, **c.** Structure for multiple instances of a single-carrier based genetic scheduling. Publication II, © 2014 IEEE.

following bits till the end of the tails are swapped between the paired chromosomes. This is called *one-point crossover*. Next, every bit of the produced child chromosomes are toggled with a low probability which is inversely proportional to their fitness, hence providing robustness for the evolutionary adaptation. This operation is known as *mutation*. Also, the chromosome sequences are repaired in case of a violation in their structure. The selection and breeding processes are repeated multiple times until an equal number of child chromosomes are produced for the next generation, in addition to the fittest two chromosomes from the current generation that are directly transferred to the next generation.

Iteration suggests initiating the algorithm using the former chromo-

Table 3.2. Genetic-based Multi-Carrier and Multi-User Scheduler

Step#1:	*Time-domain scheduler* — Selection of candidate users:
	▶ The head sequences are constructed based on the former chromosome structure as a representation of user selection
Step#2:	*Frequency-domain scheduler* — Primary user assignment:
	▶ The tail sequences in the former chromosome structure represent the decisions on primary user assignments
	▶ The head and tail sequences of the chromosomes are optimized in combination over multiple generations
	▶ The primary users are associated with hermitian precoders on every resource block
Step#3:	*Frequency-domain scheduler* — Secondary user assignment:
	▶ Brand new chromosomes are initialized based on the latter chromosome structure
	▶ Chromosomes are optimized over multiple generations
	▶ If a new secondary user is assigned on a scheduling block,
	• The precoder for the new assignee is granted based on the zero-leakage principle toward the earlier assignees
	• *Step#3* is repeated on that scheduling block for the next secondary user
	▶ Otherwise, the scheduling decision is finalized on that resource block, and hence its corresponding binary sequence is fixed to "all binary–0"s for future iterations

some structure for the selection and assignment of primary users. After having produced a new population by multiple selection and breeding processes, the whole algorithm repeats itself in order to foster a certain number of generations. The fittest chromosome from the last generation fulfills the primary user assignments. Next, a new initialization process starts using the latter chromosome structure and the fitness metrics are determined based on the prior decisions on primary user assignments. The following selection, breeding, and iteration processes foster the targeted number of generations. The fittest chromosome from the latest generation indicates the scheduling decisions on the initial secondary user assignments. Then, another initialization starts using the latter chromosome structure to determine the subsequent secondary user assignments. More users are attempted to be assigned as long as the collective utility

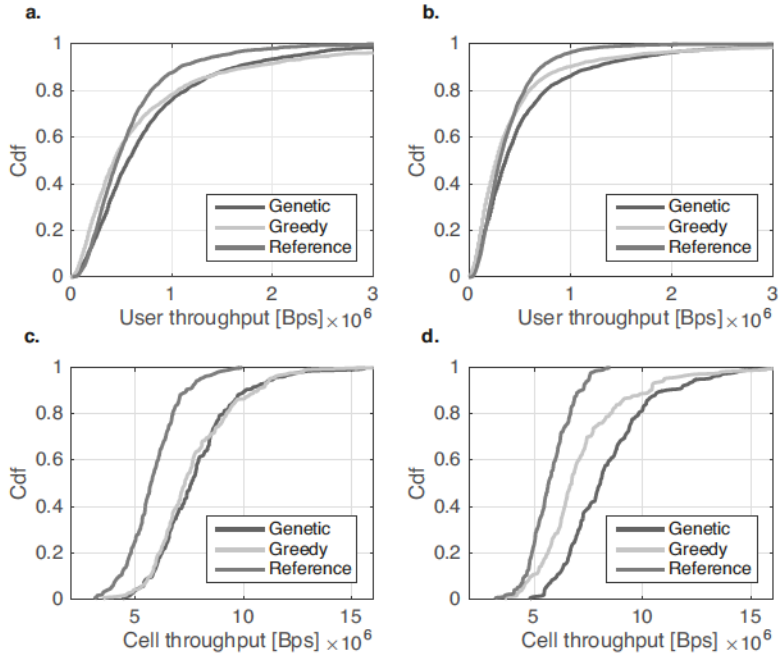

Figure 3.5. Comparison of throughput distributions by the multi-carrier schedulers: **a.** User throughputs with 10 UEs/cell, **b.** User throughputs with 15 UEs/cell, **c.** Cell throughputs with 10 UEs/cell, **d.** Cell throughputs with 15 UEs/cell. Conclusions are: 1) Genetic-based scheduler delivers higher throughput in more crowded networks; 2) Greedy-based scheduler more strongly favors the users at the cell-centers; 3) Genetic algorithm renders higher network fairness than the greedy algorithm. Publication II, © 2014 IEEE.

metrics can be improved on the corresponding resource blocks. The proposed genetic-based scheduler is summarized in Table 3.2.

Figure 3.5 demonstrates the numerical results of the described greedy and genetic based schedulers. The system-level simulation setup comprises a 2-tier wrap-around network with 3 cells in each site. Downlink channel model is based on WINNER-II [115]. There are 200 transmission time intervals (TTI) and 25 non-overlapping scheduling blocks each experiencing flat fading. Every base station employs 8 vertically polarized antennas whereas each user has 2 antennas. Downlink data transmission can be carried through either one or two signal streams toward any user terminal depending on its channel conditions, hence a base station can co-schedule up to 8 users on the same shared resources. The base stations are allowed to co-schedule any number of users receiving transmissions through any number of signal streams as long as the total number of transmitted streams is not more than the available number of antennas

at the base station. The greedy and genetic-based schedulers are evaluated in terms of user and cell throughput performance in a network setup with either 10 or 15 UEs per cell. In addition, evaluations also include a *reference* scheduler for comparison, which relies on the minimum average waiting time criterion between successive scheduling decisions for the selection of a candidate subset of users during the secondary user scheduling. All of these three schedulers employ the proportional fair and maximum throughput metrics, respectively, for the time- and frequency-domain scheduling operations. Based on the results in Fig. 3.5 one can draw the following inferences. The genetic-based scheduler provides better network fairness than the greedy-based scheduler as the genetic algorithm tends to assign more users on the shared resources while the greedy algorithm prioritizes users at the center of the cells. This also naturally leads to higher throughput performance by the genetic-based scheduler when the network is more crowded with 15 UEs/cell.

3.3 Discussion

This chapter discussed optimized resource sharing based multi-user transmission strategies on the downlink of a cellular radio system and presented the contributions of the thesis based on the original works of Publication I and Publication II. The first part of the chapter dealt with baseband signal processing methods whereas the second part investigated multi-block resource scheduling strategies.

A joint transceiver design can deliver higher performance in multi-user MIMO transmissions. However, an additional information exchange needs to be handled between a serving base station and user terminals. Such information retrieval should satisfy the required level of transparency for users and comply with the demodulation reference signal (DM-RS), control signaling, and CSI feedback configurations. Transparency clearly consumes less control signaling whereas non-transparent multi-user MIMO requires an increase in the downlink control information (DCI) size. Transparency here indicates that each user terminal is informed about only its own transmission rank and DM-RS ports while being unaware of sharing the same resources with other users. Despite such restrictions on the control information, transparent multi-user MIMO can still provide transmit-side support for receiver design at user terminals in the context of either joint methods or advanced non-linear processing, and can also

inform users about potential DM-RS collisions [116].

Spatial multiplexing is beneficial in cellular MIMO transmissions particularly at improving the peak spectral efficiency whereas user terminals in bad reception areas may experience less than satisfactory service due to intolerable co-channel interference levels. Hence, multi-cell cooperation techniques and joint power allocation schemes should also be integrated to multi-user MIMO transmission strategies to protect cell-edge users from unbearable interference. Such cooperative transmissions should rely on multi-eNB communications by either partial or full information sharing through the backhaul as well as opportunistic optimization schemes on downlink.

Although this chapter focused on utilizing multiple antenna deployments through the spatial dimension, downlink optimization techniques can be generalized based on exploiting multiple signal dimensions from the code, space, time, and frequency domains. Multi-antenna OFDM systems can perform coding to multiplexed symbols over the spatial, temporal, and spectral dimensions by means of *tensor decomposition* in order to achieve higher data rates or link reliability [117]. A tensor represents a multi-dimensional signal structure that can encapsulate the signal properties in any of the said dimensions [118]. In particular, optimization by tensor decomposition can prescribe a generalized procedure for concurrent multiuser separation, multi-symbol encoding/decoding, interference suppression, signal precoding/equalization, etc.

Schedulers are responsible for the allocation and assignment of the available resource blocks to achieve high performance and fairness in the network. In addition, such schedulers should also take into account the impact of adjacent channel interference on their performance. Heterogeneous network deployments or coverage area intersections among multiple operators can cause signal leakage between adjacent bands, leading to unrecoverable performance drops for some user terminals. Proactive awareness by resource schedulers can provide protection against such interference for those users as well as increased spectral availability and lower energy consumption in the network. Furthermore, various other radio technologies, such as D2D communications and cognitive radio, can potentially improve the spectral efficiency of schedulers by means of configuring allocation decisions on-the-go based on dynamical network conditions.

4. Reciprocity-Based CSI Acquisition in Massive MIMO

Base stations equipped with very large number of antennas can offer crucial benefits in cellular radio networks; for instance, adverse effects of uncorrelated noise and fast fading disappear, increased orthogonality among users' CSI allows simpler encoding/decoding, and the required transmitted energy per bit shrinks at user terminals [119]. However, deploying a very large antenna array at a base station increases the length of the associated CSI vector, and accordingly, necessitates larger codebook and more signaling bits in user feedback reports in order to preserve a sufficient quantization granularity.

This inherent acquisition problem of oversized CSI can be avoided in TDD systems by exploiting the duplex reciprocity of electromagnetic propagation channels in an *open-loop* approach as the channel coherence bandwidth is typically large. In particular, the channel response can be estimated at the base station directly on reverse link and this estimate can be interpreted as downlink CSI without any further dedicated aid from user terminals. However, such CSI acquisition is susceptible to mainly three types of predicaments: the transceiver impairments that are caused by the analog RF circuitry, the difficulty with uplink channel estimation due to notorious pilot contamination in a massive MIMO setup, and offset errors within the reciprocity of the wireless duplex channel as a result of its time evolution during successive measurements. Antenna calibration can resolve duplex channel mismatches due to hardware impairments and data-aided CSI estimators can help overcome pilot contamination, whereas the time-varying nature of the channel limits the available computational duration for the necessary antenna calibration and data-aided estimation to a tiny fraction of the channel coherence time.

This chapter explores such reciprocity-based CSI acquisition techniques in massive MIMO networks operating in TDD mode. The first part of the

chapter investigates the reciprocity characteristics of a channel, which can be impaired by the RF circuitry, and discusses the prominent recovery techniques by multi-antenna calibration. The second part is dedicated to data-aided blind channel estimators on reverse link. The contributions of this thesis are also highlighted in this chapter, which can be summarized as follows. Publication III presents an adaptive channel quantization algorithm that improves the efficiency of relative antenna calibration processes by means of reduced signaling overhead. Publication IV demonstrates a multi-cell coordination based optimization method for blind CSI estimators on uplink based on approximating the error sources that are inherent to standard eigenvector decomposition (EVD) related computations in massive MIMO setups with multiple transmit signal sources.

4.1 RF circuitry calibration

Wireless TDD systems use the same frequency bands on forward and reverse links. In addition, electromagnetic waves traveling at exactly opposite directions undergo the same physical obstructions. As a result, such electromagnetic propagation channels demonstrate identical frequency-domain responses in opposite directions as long as duplex transmissions occur within the same coherence time window [120]. However, transmit signal waveforms also travel through the transceiver circuitries at both ends of a wireless link, hence the duplex reciprocity of the overall wireless channel cannot be present in case of any dissimilarity in the frequency-domain responses of the transmit and receive components within the same transceiver at either side of the link.

Main causes for such dissimilarity can be explained by internal clocking structures (e.g., dividers, multipliers, and PLLs) or static effects depending on the hardware manufacturer. In addition, any mismatches of mutual coupling between the receive and transmit antennas further damage the reciprocity [121, 122]. Any asymmetries introduced by the RF hardware of a transceiver cannot be resolved during manufacturing or assembly stage by means of delicately eliminating all the differences between each of the transmit and receive circuitry components as the employed commercial electronics always contain unavoidable physical limitations, partly as a result of the efforts for keeping the total cost of massive MIMO deployments feasible [6]. Moreover, a terminal and its serving base station have different oscillators, which leads to relative channel drifting

over time. In that case, frequent recovery by antenna calibration is required to continually remove out any transceiver mismatches, and hence restore the channel reciprocity. For antenna calibration practices, hardware mismatches of a base station are considered the major issue whereas any terminal-side mismatches have minor impacts on the system performance [121]. In particular, when a base station employs linear precoding, the operating performance becomes insensitive to any transceiver imperfections at user terminals [123].

There are two types of RF chain calibration methods: self-calibration and over-the-air (OTA) calibration. Self-calibration methods require dedicated equipment prior to deployment or special RF circuitry for on-the-go local calibration. Self-calibration is insensitive to channel conditions and does not rely on signaling overhead on the active communication link. Nevertheless, self-calibration requires higher implementation costs and more sophisticated design [124]. On the other hand, OTA calibration is based on intermittent inter-node signaling during the operation time of a transceiver. As OTA calibration requires no special hardware, the implementation allows simpler and more cost-efficient designs. Yet, OTA calibration is based on relying on signaling resources on a communication link, therefore the calibration performance can be limited by the link conditions. Although large signaling overhead leads to reduced throughput performance, there is usually a point at which extensive signaling and feedback from a user terminal can improve the restored channel reciprocity while the reduction in system throughput is still smaller than the gain obtained through reciprocity.

Calibration by OTA method can be realized either absolutely or relatively [125]. Absolute OTA calibration compensates for the imperfections of each RF chain at a transceiver based on independent signaling from every active user terminal [123]. On the contrary, relative OTA calibration is based on measuring the non-symmetric characteristics of a transceiver relatively by means of signaling only from one *reference* user terminal [126, 127]. Although relative calibration causes complex-valued offsets in the CSI estimates for other terminals, their performance is unaffected as long as the base station employs linear processing for data transmissions.

4.1.1 Deployment challenges under massive arrays

From a theoretical point of view massive MIMO technology offers very large system capacities and resource efficient communications between a base station and multiple UEs. In addition, radiating power per antenna reduces drastically due to the distribution of transmit power among many antennas. As a result, a base station can equip cheaper power amplifiers and RF filters, and even skip the active cooling process [128]; which not only reduces power consumption but also simplifies design complexity. Similarly, user terminals can also enjoy lower transmit power consumption and therefore improved battery lives, which is particularly appreciated in portable handheld devices. However, a transceiver design with a very large number of antennas raises a few hardware challenges that need to be resolved in a real-world implementation of a base station. These problems are mainly related to physical size limitations, baseband processing, clock distribution, and synchronization.

The physical distance between two neighboring antenna elements should be at least half of the carrier wavelength to avoid strong correlations among the associated incoming/ongoing signal streams. As the linear arrangement of many antenna elements needs larger space of occupancy, at least two dimensional deployments have to be considered, which allows refined beamforming in the elevation plane as well [129, 130]. For an efficient design, beamforming weights should be applied at the radio front-ends where they can be supported by a common databus for downlink transmission. On the contrary, linear precoding during an uplink transmission is based on combining all the receive samples, therefore the databus can simply support a constant bandwidth. In addition, massive MIMO transceivers cannot be deployed with dedicated RF amplifiers for each antenna element in order to preserve economic viability. Hence, the baseband signal processing is limited. But, cheaper phase and amplitude adders can be attached to each element after carrier modulation [131].

4.1.2 Relative OTA calibration

In order to measure and compensate for the hardware mismatches at every radio front-end, a base station needs to acquire full downlink CSI. Random phase and amplitude responses of a transceiver can be approximated by the relative offset between uplink and downlink channel estimates, which is called *calibration coefficient*. After duplex channel mea-

surements, calibration coefficients can be applied to future uplink channel estimates to implicitly acquire the downlink CSI without any explicit user feedback reports in TDD. Notably, the same calibration coefficients can be applied to other user links in a relative manner, as a complex-valued random scaling offset does not degrade the link performance with a linear encoder. Also, properties of the transceiver parameters are known to be slowly time-varying relative to the electromagnetic channel response [132, 133, 134], therefore the amount of feedback overhead can be tolerable under very large antenna arrays due to the infrequent need for future recalibrations.

The main challenges for relative OTA calibration can be considered twofold under massive MIMO. First, duplex channel measurements must be taken during a very short time period, although consecutive pairs of measurements can be sporadic with longer delays. These time delays till the successive recalibration depends on the quality of the hardware components and manufacturing processes. Potentially, a transceiver may require recalibration every few minutes [132] or just once a day [128]. Second, CSI reports in UE feedback require large overhead due to massive number of antennas; however, the signaling overhead can be distributed among many UEs as long as every base station antenna is measured by the same user terminal. In case the signaling overhead is intolerable, an *internal calibration* procedure can be operated locally on every antenna element of a base station. Once the calibration coefficients are measured in a similar fashion, they can be applied to user links relatively. Such process requires even less frequent re-recalibration as the same clocking is most probably used internally at the base station [128]. However, as noted above, such internal calibration requires dedicated equipment at the base station. In addition, the performance is typically degraded by the limited dynamic range of the local power amplifier.

On the contrary, when relative OTA calibration is operated instead of such internal calibration, the size of the required signaling overhead can still be diminished as follows. Publication III proposes an adaptive quantization algorithm which alters the CSI granularity at a user terminal in response to the base station's request so that the overall feedback overhead is reduced by means of more efficient use of signaling. To put in other words, such adaptive quantization can deliver higher calibration performance using the same amount of signaling overhead. As the optimal adaptation of CSI quantization for a single measurement during a

calibration process depends on its earlier and later measurements, the decision criterion should be evaluated at the base station and shared with the UE by a feedforward request. However, any relative changes in the reliability of the downlink channel cannot be foreseen during an uplink measurement prior to user feedback, hence the global optimality is impractical to derive for such an adaptive CSI quantization scheme. The optimization can merely rely on a simplified approximation based on the current and previous uplink channel measurements. In particular, improved accuracy of downlink CSI may return small or large calibration gains depending on the corresponding quality of uplink channel. Accurate uplink estimation encourages for finer granular quantization of downlink CSI whereas inaccurate uplink measurements suggest reducing the signaling overhead. More specifically, the adaptive quantization algorithm of Publication III operates as follows.

Let us denote the measurement index number by k and the vector of optimized quantization bits by q. When $k = 1$, the first CSI measurement on downlink is quantized by an initial number of bits, denoted by q_0. After the initial duplex channel measurements, the base station evaluates the estimation accuracy for the next uplink channel measurement based on the channel gain and estimator sample variance. The accuracy is stored in a vector at the current measurement index (e.g., $k = 2$ for the second measurement), denoted by $\kappa(k)$, which can be calculated as follows.

$$\kappa(k) = \sum_{m=1}^{M} \frac{|\hat{\mathrm{h}}_{\mathrm{ul}}(m)|^2}{e_k(m)}, \tag{4.1}$$

where M denotes the number of channel coefficients, $\hat{\mathrm{h}}_{\mathrm{ul}}$ denotes the uplink channel estimate, and $e_k(m)$ denotes the sample variance of the least square estimator for the mth channel coefficient, which can be derived by the square-distance of the weighted least square method in pilot-aided estimators [135]. For each subsequent channel measurement, if the current uplink accuracy is better than the previous one, the number of quantization bits for the next downlink CSI feedback is increased by s^+; otherwise, it is reduced by s^-. More measurements are taken till the total allotted feedback bits (the number of which is denoted by B) are depleted. In addition, one can also define minimum and maximum number of allowed feedback bits for the quantization of any CSI report, denoted here by $q_{\min}$ and $q_{\max}$, respectively. The necessary steps of the proposed adaptive quantization algorithm are listed in Table 4.1.

The numerical performance results of the proposed adaptive quantiza-

Table 4.1. Adaptive Quantization of CSI Feedback for Relative Calibration

Step#1:	Initialize vectors κ, q of length $\lceil B/q_{\min} \rceil$, and set $k = 1$	
Step#2:	Calculate and store $\kappa(k)$	
Step#3:	**if** $k = 1$,	store $\mathrm{q}(1) = q_0$
	elseif $\kappa(k) < \kappa(k-1)$,	store $\mathrm{q}(k) = \max(\mathrm{q}(k-1) - s^-,\ q_{\min})$
	elseif $\kappa(k) > \kappa(k-1)$,	store $\mathrm{q}(k) = \min(\mathrm{q}(k-1) + s^+,\ q_{\max})$
	else,	store $\mathrm{q}(k) = \mathrm{q}(k-1)$
	end	
Step#4:	**if** $B - \sum_{j=1}^{k} \mathrm{q}(j) \leq 0$	
	if $\mathrm{q}(k) \leq q_0/2$,	store $\mathrm{q}(k) = 0$, and **exit**
	else,	store $\mathrm{q}(k) = q_0$, and **exit**
	end	
	else	
	set $k = k + 1$ and go to *Step#2*	
	end	

tion scheme are generated in a multi-cellular network simulation and demonstrated by Fig. 4.1. The simulation operates on a 7-site wrap-around network with each site being composed of three 120° cells. Each cell has one eNB and 10 UEs which are dropped uniformly. Each eNB employs 4 antennas whereas all UEs are equipped with single antennas. The antenna crosstalk is considered negligible at the eNBs. The channel is modeled based on WINNER-II [115] and the simulation involves 50 subcarriers and 100 TTIs. The user mobility is 3 km/h for all UEs, and hence, the channel model exhibits time evolutions at every TTI. Multiple duplex channel measurements are taken at every tenth TTI and optimized by minimizing the total least squares (TLS) errors. The adaptive quantization parameters from Table 4.1 are selected as follows: $q_0 = 16$, $q_{\min} = 8$, $q_{\max} = 24$, $s^+ = s^- = 4$, and $B = 64$. Figure 4.1 compares the number of quantization bits used by each CSI feedback report to the instantaneous uplink estimation error. Also, the number of duplex channel measurements at every tenth TTI is visible from the figure. The results demonstrate that the number of quantization bits consistently adapts to the changes in uplink MSE between the previous and current TTIs.

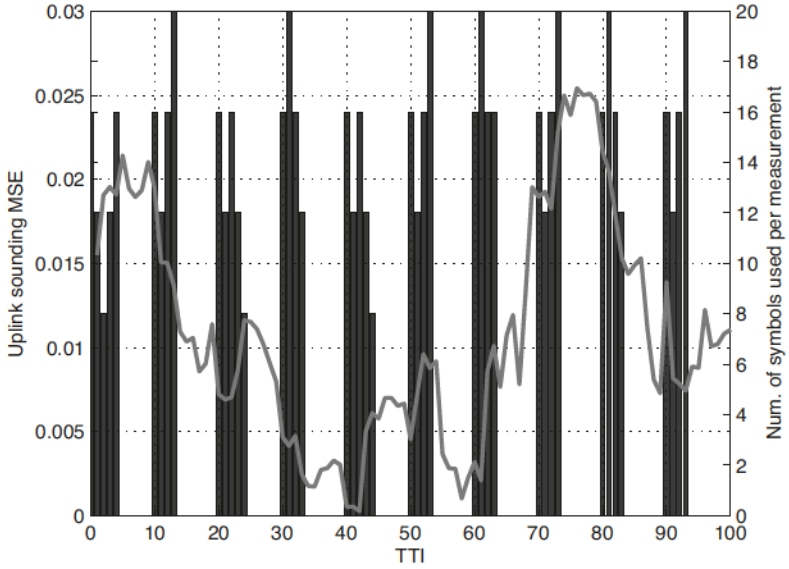

Figure 4.1. Comparison results between the number of adaptive quantization bits (bar plot) and instantaneous MSE uplink estimation (line plot). Notice that the number of duplex channel measurements is indicated by the number of bars at every tenth TTI. Publication III, © 2014 IEEE.

4.2 Blind CSI acquisition

The reverse link in a massive MIMO network can handle a large number of simultaneous user transmissions thanks to the available spatial degrees of freedom. Moreover, the length of a pilot sequence cannot exceed the channel coherence time and that the number of elements in an orthogonal binary sequence set is limited by its dimension. Hence, some of the training sequences, which are uniquely assigned to individual users, inevitably correlate with others, especially in a congested and perfectly synchronized multi-cellular network operating on the reverse link. That leads to impairments in uplink CSI measurements for some users who have been assigned those contaminated pilots [119, 136]. Various optimization techniques can utilize the limited training sequences efficiently to some extent by means of either exploiting instantaneous channel properties or adopting coordinated pilot assignment schemes among multiple eNBs [137, 138, 139, 140, 141]. However, the performance degradations by pilot contamination cannot be fully avoided, particularly when the number of active users is overwhelming.

Subspace-based CSI acquisition techniques can extract the necessary

channel information from receive sample symbols that are not necessarily composed of orthogonal sequences, therefore providing an alternative workaround against pilot contamination. Such acquisition techniques based on the subspace partitioning of receive samples have been extensively studied for the last few decades. The EVD based estimation is one of those methods which use the second-order statistics of the receive symbols at a base station to estimate the channel response vectors blindly [142, 143, 144, 145, 146, 147]. Such *blind* estimators rely on computing the eigenvectors of an average receive autocorrelation matrix and assigning one of them that corresponds to the largest eigenvalue as the CSI vector. These blind multi-antenna estimators can be optimized to achieve reduced time averaging window, eliminate potential phase ambiguities in the final CSI, or acquire more accurate channel information [148, 149, 150, 151, 152]. In particular, the accuracy of a blind CSI estimator can be improved by using both the receive samples and their complex conjugates, which is referred to as *widely linear subspace estimation* in literature [153, 154]. Such techniques help represent the channel vectors in real form, hence converting the phase ambiguity to a mere sign ambiguity [154, 155]. In addition, widely linear methods render longer vector representations of CSI, leading to reduced user channel distortions inside the cumulative signal originating from multiple sources on the uplink of a massive MIMO system [155]. On the other hand, a widely linear estimator can only function with improper receive symbols [154, 156], which limits its applicability to specific signal constellations. Improper signals are characterized by a non-zero pseudo-covariance, which, as an example, can be generated by real-valued modulation schemes, such as BPSK.

When blind estimators are applied in a massive MIMO setup, it has been shown that CSI from multiple signal sources (e.g., user terminals) can be extracted from the same receive autocorrelation matrix with small distortions even though the transmitted signals are composed of non-orthogonal data sequences [155, 157]. When a plurality of user terminals simultaneously transmit on shared resources during a multicast transmission on uplink, the user channel responses generate asymptotically orthogonal inner products due to the law of large numbers and Lindeberg-Lévy central limit theorem [158], as elaborated more in the following section.

4.2.1 Separation of user channel estimates

According to the random matrix theory, when the number of transmit antennas approaches infinity, propagation vectors of different user terminals exhibit asymptotic orthogonality with each other [119]. In particular, the mutual orthogonality between any two user channel responses increases with more antenna elements [158]. Hence the distinct eigenvectors of a blind estimator can be matched to different users with respect to their large-scale channel characteristics [157].

Upon receiving user data symbols from multiple terminals, a base station computes the receive autocorrelation matrix by temporal averaging. Assuming that the channel and additive noise elements are independent and identically distributed with zero mean, each eigenvector of the autocorrelation matrix corresponds to a user CSI. However, the acquired CSI contains residual errors in practice due to the finite size of the receive antenna array and insufficient receive symbols [155]. Furthermore, blind estimators are known to be impacted by a *multiplicative ambiguity* in their CSI estimates. Although this inherent ambiguity has no impact on the performance of a linear-type encoder, the CSI estimates still hold unknown amplitude and phase offsets relative to the true channel responses, which can potentially cause large impairments with non-linear encoding or by joint optimization algorithms coupled with symbol demodulations.

When the applied symbol modulation scheme is known, CSI estimates by the eigenvectors can be improved by joint iterations during demodulation. The initial demodulation of symbols can be derived by minimizing the least squares based on the constellation scheme and eigenvector estimates. The demodulated symbols are then used to correct the CSI estimates. Such multiple iterations are known to reduce the residual CSI errors [157]. In addition, multiplicative ambiguities can also be removed by transmitting a short orthogonal training sequence from each user [154]. However, the other two inherent error sources, which are based on the finite size of eNB antenna arrays and insufficient data symbols as explained above, cannot be suppressed either by the described iterative processes here or by any other known method from prior art. To address this deficiency in massive MIMO literature, Publication IV proposes a multi-cellular coordination scheme which mitigates the impacts of the said two error sources inherent to blind CSI estimators on the uplink of massive MIMO systems. The details of the proposed multi-cell coordination are

explained in the following part.

4.2.2 Multi-cellular coordination

The coordination algorithm in Publication IV is based on iteratively improving the uplink CSI accuracy at a group of neighboring eNBs. Each iteration jointly processes the available channel information from all eNBs, updates the approximated models for the inherent error sources due to the finite number of antennas and symbols, and subsequently applies the latest error models into separate CSI estimation mechanisms at every eNB. The major contributions of Publication IV involve proposing for the first time both such joint processing for a blind CSI estimator and such approximation for the said error sources.

The detailed description of the proposed multi-cell coordination scheme can be presented as follows. First, each eNB in the network estimates the CSI from all associated UEs by means of computing the eigenvectors of the mean receive autocorrelation. Next, multiplicative ambiguities are resolved by a short training sequence, as typical with standard semi-blind estimators. The ambiguity factors can be estimated at the lth base station as follows.

$$\hat{\mathbf{\Gamma}}_l = \arg\min_{\mathbf{\Gamma}_l} \left\| \mathbf{Y}_l - \sqrt{p_\mathrm{u}} \bar{\mathbf{H}}_l \mathbf{\Gamma}_l \mathbf{D}_l^{1/2} \mathbf{X} \right\|_\mathrm{F}^2 \tag{4.2}$$

where $\hat{\mathbf{\Gamma}}_l$ and $\mathbf{\Gamma}_l$ are the estimated and true diagonal matrices of multiplicative ambiguity factors, $\mathbf{Y}_l$ denotes the received training data, p_u denotes the uplink transmit power, and $\mathbf{X}$ denotes the aggregate matrix of transmitted training symbols. Also, $\bar{\mathbf{H}}_l \mathbf{D}_l^{1/2}$ represents the aggregate channel response on uplink, where $\bar{\mathbf{H}}_l$ and $\mathbf{D}_l^{1/2}$ respectively denote the fast fading and large-scale fading components. Second, all CSI estimates as well as a subset of the receive symbols from each eNB are collected through backend at a centralized processor, where the available information is jointly processed by exploiting the fact that the same transmitted symbols from all UEs are, although unintentionally, received at every eNB. Such joint optimization can be formulated based on minimum total least squares as follows.

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X} \in \mathcal{S}} \sum_{l=1}^{L} \left\| \hat{\mathbf{D}}_l^{-1/2} \hat{\mathbf{\Gamma}}_l^{-1} \hat{\mathbf{H}}_l^\dagger \mathbf{Y}_l - \sqrt{p_\mathrm{u}} \mathbf{X} \right\|_\mathrm{F}^2 \tag{4.3}$$

$$\hat{\mathbf{H}}_l = \frac{1}{\sqrt{p_\mathrm{u}}} \hat{\mathbf{D}}_l^{-1/2} \mathbf{Y}_l \hat{\mathbf{X}}^\dagger \tag{4.4}$$

where $\hat{\mathbf{X}}$ is the aggregate matrix of transmit symbol estimates, $\mathcal{S}$ denotes

Table 4.2. Joint Multi-Cell Coordination for Blind Estimators

	Initialize iteration index $i = 1$ & max. number of iterations $\mathcal{I}$
Step#1:	**for** each eNB
	Compute EVD-based CSI
	Match CSI to corresponding UEs
	Resolve ambiguities via short training
	end
Step#2:	CSI and receive samples are collected via backend
	Joint coordination derives new CSI
	CSI is signaled back to each corresponding eNB
Step#3:	**for** each eNB
	MSE-based error sources are approximated for every UE
	Recover local CSI from error sources
	end
	if $i < \mathcal{I}$, $i \rightarrow i+1$ & go to *Step#2*
	else, terminate
	end

the symbol constellation for $\mathbf{X}$, and $\hat{\mathbf{H}}_l \hat{\mathbf{D}}_l^{1/2}$ denotes the aggregate channel estimate. The iterations between (4.3) and (4.4) can be repeated multiple times for convergence. After joint optimization, the generated set of new CSI is signaled back to each corresponding eNB through the backhaul. Third, each eNB approximates the inherent errors by means of substituting the jointly optimized and transferred CSI data into asymptotical MSE expressions. Next, the accuracy of the initial CSI at each eNB is improved by recovering from the inherent error sources based on their approximated MSE models. The final CSI at every eNB is once again transferred back to the central processor via backend signaling, which kicks off the second iteration of the algorithm for the joint coordination. The algorithm can be terminated after a certain number of iterations is completed, or when the relative change of CSI between successive iterations drops below an acceptable margin. The described steps of the proposed joint coordination algorithm are summarized in Table 4.2.

It should be noticed that the jointly optimized CSI is only used for approximating the error sources while the final channel estimation is mainly based on the local estimates at each eNB. The logic behind this design choice can be explained as follows. Due to the limited amount of signaling

resources at the backend, only a subset of the receive samples are transferred which leads to reduced estimation accuracy relative to the local CSI. Moreover, backend signaling also introduces quantization errors to both initial CSI estimates and receive samples.

4.3 Discussion

This chapter investigated the issue of downlink CSI acquisition in massive MIMO systems with reciprocal duplex propagation channels and presented the contributions of the thesis based on the original works of Publication III and Publication IV. The first part of the chapter was concerned with multi-antenna calibration methods whereas the second part dealt with subspace-based blind CSI estimation methods.

Massive MIMO technology is an exciting, new enabler for current and next generation cellular radio access systems in achieving advancements in power efficiency and interference avoidance by sharp beams and directed energy transmissions. Nevertheless, the task of downlink CSI acquisition becomes non-tedious with massive antenna arrays due to the extended size of the conventional UE feedback reports. In case of channel reciprocity, a base station can estimate the downlink CSI from the reverse link although such reciprocity is often distorted by the imperfections in transceiver hardware. To overcome these CSI errors, antenna calibration techniques can measure and compensate for such distortions. Nevertheless, perfect restoration of channel reciprocity is usually improbable with affordable RF equipment. Moreover, estimation on uplink can be unreliable under massive antenna arrays due to pilot contamination. As a remedy for the latter, blind estimation techniques with non-orthogonal receive signal sequences can be employed for CSI acquisition from multiple user terminals. The residual non-orthogonality among (both intra- and inter-cell) user channels may drastically diminish CSI accuracy, which can be avoided to a certain extent by multi-cellular coordination.

However, such multi-cell coordination requires significant backend signaling, at least, initially whereas subsequent iterations may rely on much smaller backhaul overhead due to their gradual convergence. Moreover, the amount of signaling can be minimized by associating multiple groups of eNBs to separated clusters, as it is a common optimization practice in coordinated multipoint (CoMP) networks. Although the necessary amount of signaling can be handled by the standard X2 interface equipped with

fiber or microwave that can provide up to 1Gb/s, the compatibility to cloud-RAN (C-RAN) requires further investigation for the next generation systems. That being said, architectural features of C-RAN that are already in consideration seem to suggest the feasibility of such multi-cell coordination among blind CSI estimators. The main motivation for C-RAN is to provide smooth network access in densely populated urban areas by centrally controlling all available deployments of transceiver stations in order to improve the efficiency of interference management and reduce financial costs. Clearly, such centralized intelligence should simplify the operation of the proposed coordination process for CSI acquisition as well.

Due to the discussed residual errors in both antenna calibration and blind estimation practices, the investigation of downlink CSI acquisition via conventional UE feedback reports should still be considered a pivotal element for the adoption of the massive MIMO technology into cellular communications. That being the case, the focus of the thesis in the following chapter is switched to non-reciprocal propagation channels for CSI estimation at massive eNB antenna arrays.

5. Non-Reciprocal Channel Estimation in Massive MIMO

As the majority of commercial cellular networks operate in FDD, the feasibility of massive MIMO in FDD mode is considered of utmost importance for backward compatibility, cost efficiency, and interoperability [159]. However, FDD systems exhibit non-reciprocal channel characteristics due to separated carrier frequencies in uplink and downlink, which renders implementation issues, particularly regarding downlink CSI acquisition at base stations and heavy computational requirements at user terminals.

This chapter discusses these challenges and demonstrates the state-of-the-art methods and concepts for enabling massive MIMO in FDD mode mainly by reducing the training duration on downlink, compressing UE feedback overhead, and offloading the bulk of the necessary computational complexity from user terminals to the base station whenever possible. As most of the known solutions rely on exploiting certain characteristics of a propagation channel in cellular deployments, the investigated techniques in this chapter are grouped into three categories based on the following classes of channel properties: temporal channel correlations, spatial channel correlations, and multipath sparsity. In addition, the current state of progress on the feasible implementations of FDD massive MIMO is also discussed within the timeline of the current and next generation cellular systems.

Regarding the contributions of this thesis on FDD massive MIMO, this chapter presents the following. Publication V designs a new CSI acquisition method based on estimating the reciprocal characteristics of a propagation channel at a serving base station directly from the reverse link whereas the active user terminals are required to estimate only the non-reciprocal fast-fading properties of a select number of dominant paths. Publication VI proposes that the estimated multipath phase information

at user terminals is subsequently transferred to the serving base station as a new type of feedback report for full CSI acquisition. Numerical evaluation results and feasibility discussions are also presented from a realistic deployment point of view at the end of the chapter.

5.1 Challenges for FDD-mode massive MIMO

Besides the inherent hardware challenges for equipping a base station with very large antenna arrays, which are covered in Section 4.1.1, FDD operation mode particularly manifests certain issues that need to be resolved before realizing a feasible deployment framework. The fundamental problem is related to forward-link CSI acquisition at the base station. Apparently, uplink CSI is not convertible to its downlink counterpart as the coherence bandwidth is typically much shorter than the offset between duplex carrier frequencies. Therefore, CSI acquisition is typically realized within a closed-loop by quantized CSI feedback reports from user terminals in most state-of-the-art FDD radio access technologies, including cellular communications. However, the challenges for such feedback based acquisition stems from two major causes in massive MIMO systems, as elaborated below.

First and foremost, the necessary time duration during channel training on forward link is proportional to the number of transmit antennas. When orthogonal pilots are used among multiple base station antennas to isolate different subchannel coefficients, the number of active terminals simultaneously processing these training sequences needs to be limited to maintain the perks of massive MIMO deployment [119]. In case of multiple neighboring cells simultaneously performing channel training, this limitation on the number of served users becomes even more drastic. The second issue is related to the fact that very large antenna arrays generate long CSI vectors. Once the channel training is accomplished on downlink, user terminals are still required to transmit their CSI estimates on the reverse link. This feedback transmission has to be based on either an enormous codebook or a fine-granularity scheme. In either case, the transmission consumes large amounts of signaling overhead, which renders conventional CSI feedback techniques impractical. In addition, processing long CSI vectors increases the computational complexity, which is primarily critical for user terminals due to their low signal processing capabilities and limited energy consumptions. What is more, the required

storage size of an extended CSI codebook is expected to be overwhelming, which may challenge the local memory limitations at user terminals.

Because of all these difficulties, there is a widespread belief that future research efforts should be put into the TDD-mode massive MIMO [160, 161, 162, 163, 164]. On the other hand, there are those who think that enabling the feasibility of massive MIMO in FDD will be worth the time and efforts [165, 166, 167, 168, 169, 170]. In favor of the latter group's view, the rest of this chapter reviews the state-of-the-art techniques for overcoming the design issues on CSI acquisition and discusses potential future directions.

5.2 Channel characteristics as enablers of CSI compression

Realistic propagation channels most often exhibit characteristics that allow opportunities for developing intelligent CSI compression schemes. In the following, well-known techniques within the theoretical area of FDD massive MIMO research are grouped into three categories based on the types of exploited channel properties, which are namely temporal correlations, spatial correlations, and multipath sparsity.

5.2.1 Temporal correlations

Temporally correlated channels exist as a result of multipath propagation, Doppler spread, transmit pulse shaping and receive filtering [171]. Specifically, temporal correlation is the outcome of any amplitude correlations between different excess delays of a channel impulse response profile. Although temporal correlation is often present in all wireless communications, it is rarely exploited for improving CSI estimation in real-world commercial deployments due to the difficulties of conformance testing. In theoretical evaluation studies, simulation models for temporally correlated channels are usually designed based on either first-order Gauss-Markov process [172] or Jakes' model [173].

One common approach in literature for the utilization of temporal correlations is related to jointly estimating/tracking the fast-fading channel responses and time correlations at user terminals via Kalman filtering [174, 175, 176]. In a massive MIMO setup temporal correlations can be further exploited to improve the CSI quality at a base station. The issue of limited time period during downlink channel training can

be handled in a closed-loop by means of iteratively adapting the pilot sequences to current channel conditions based on the available knowledge on past CSI [177]. As the optimization by such a closed-loop mechanism informs the base station about the *best* training sequence, CSI can be implicitly extracted from this information at the base station without any dedicated quantized CSI feedback. Such implicit acquisition can be realized by designing the optimum beamformer that either minimizes MSE or maximizes SNR for the past training signals [178]. However, this type of pilot-assisted CSI acquisition approach can cause severe dimensionality loss due to large amounts of signaling overhead in the closed-loop, particularly if the level of temporal correlations is lower than an acceptable margin.

Another possible approach for improving downlink training is based on a sequentially optimized beam design by the Kalman filter and its prediction error covariance which exploits temporal channel correlations [179]. Due to such sequential optimization, user terminals are only required to perform low-complexity processing tasks. Historically, differential codebooks are also well known to utilize temporal correlations for successively improving the quantization accuracy of CSI vectors [180, 181]. Such compression performance within feedback reports can also be realized in a massive MIMO setup if the temporal channel correlations are sufficient [182].

5.2.2 Spatial correlations

The level of spatial correlations in a wireless channel depends on two factors: antenna geometry and scattering environment. The former factor causes spatial correlations often due to short antenna spacing, large mutual coupling among multiple antenna elements, or small angular spread. On the other hand, the latter causes spatial correlations when there is a lack of local scatterers at the proximity of a transceiver. In a typical cellular network, base stations are usually elevated and rather isolated from the surrounding scatterers whereas scattering in the vicinity of user terminals is a common phenomenon. When the correlation among antenna elements is strong, their transmit signals experience fading at the same time instances on downlink. In addition, strong correlations render multi-antenna processing methods ineffective in capturing the spatial structure of a channel. As a result, spatially correlated fading channels are well known to limit the diversity techniques at the receiver of a user terminal.

On the other hand, in a massive MIMO system spatial channel correlations are also considered beneficial for CSI compression. Regarding theoretical performance evaluations, the Kronecker correlation model [183] is often used for simulating spatially correlated channels in massive MIMO studies [184, 185, 186]. In case of negligible terminal-side correlations, the one-ring channel [187] is another popular model for more simplified analyses.

In massive MIMO systems, a two-layer transmit beamforming can effectively utilize spatial correlations. In such a design, the first beamformer can provide user grouping based on pairwise correlations while the second beamformer is applied to the effective channel response with reduced dimensionality [184]. This technique can offer significant CSI compressions without drastically compromising the performance. In addition, channel covariance matrices (CCM) are also widely utilized for various CSI compression techniques with multiple antennas as the channel covariance can be directly measured from the reverse link, even in FDD modulation mode with the help of a carrier frequency transformation [188]. With a very large antenna array setup spatial channel correlations render low-rank CCMs, therefore creating an optimization space for downlink training and user CSI feedback reports thanks to the reduced channel dimensionality [84].

Another means of reduction for the channel dimension can be based on grouping highly correlated antenna elements into a single representation with respect to pre-defined patterns at a base station [189]. In addition, spatial channel correlations can also be utilized for an optimized codebok design in massive MIMO. The port modulation technique can improve the downlink performance without instantaneous CSI by matching the *virtual precoder* with respect to the spatial correlations of the antenna array [185]. Moreover, codebooks for a spatially correlated channel can be designed more efficiently by taking into account also the temporal correlations [182, 186].

5.2.3 Sparsity by finite dominant paths

A very large number of transmit antennas produces a sparse multipath structure in wireless channels with limited number of scatterers due to the extended signal space dimension [190, 191]. Specifically, the majority of the channel eigenmode strengths drops below the noise level as a result of power spreading. Aside from very large antenna array setups,

other causes for such an extended dimension could be very large bandwidth or very long symbol durations [192]. For performance evaluations, sparse channel models are often designed based on simulating a few local scatterers or based on defining spatially correlated channels which render low-rank frequency responses as a requirement for channel sparsification. Realistic evaluation studies can also employ advanced geometry-based stochastic channel models with many more propagation paths, such as WINNER-II [115], COST 2100 [193], or 3GPP spatial channel model (SCM) [194].

The concept of compressive sensing is a well known signal processing technique that can provide accurate estimations of a wireless channel with small amounts of measurements when its propagation characteristics exhibit sparsity in the signal space [195, 196]. Compressive sensing methods do not require any prior knowledge on CSI as the encoding is based on random projections. The main design procedure among such compressive sensing based schemes is related to transforming channel responses by a *sparsifying basis*, which is most commonly derived by the DFT, Discrete Cosine Transform (DCT), or Karhunen-Loéve Transform (KLT).

In the most broad terms, the compressive sensing technique can enable high-rate data transmissions in massive MIMO communications with only small amounts of CSI measurements. User terminals transform the available CSI knowledge according to a pre-known transformation basis and then transfer the compressed output in feedback reports. The receiving base station decodes the transformed information to full CSI with negligible loss thanks to the sparse characteristics of the channel. The transformation basis can also be selected adaptively based on instantaneous CSI conditions that allow dynamic configurations on user feedback. Such adaptive optimization can be designed by adjusting the compression ratio according to the sensitivity of downlink performance to the recovered CSI accuracy at the base station [85]. Furthermore, CSI compression can be designed to additionally exploit the joint sparsity structure among multiple user channel responses due to common terminal-side local scatterers. Such joint optimization can be realized by a distributed compression scheme that collects the encoded copies of the local CSI measurements from multiple users and performs joint recovery at the base station [197]. Moreover, channel sparsity also allows reductions in signaling overhead while tracking continuous-time channel parameters, such as multipath

direction-of-departure (DoD) angles and their evolutions due to terminal mobility [131].

Alternatively to these compressive sensing schemes, channel sparsity can also be exploited to extract only the set of dominant path characteristics in a multipath propagation channel via parameter estimation, as proposed in Publication V. When a base station receives planar wavefronts on uplink, multipath properties are similarly represented by the elements of a CSI vector, therefore facilitating the estimation of some multipath characteristics from uplink, such as DoA angles and path amplitudes. The planarity of receive wavefronts depends on the physical distance between the neighboring elements of an antenna array relative to the end-to-end link distance. Specifically, when the distance between the base station and a user terminal is much longer than the physical size of the antenna arrays, the properties of the multipath components are almost identically observed at different antenna elements. This manifestation is known as the *far-field approximation principle*. Therefore, the CSI acquisition method of Publication V is mainly targeted for macro cell networks with highly elevated base station platforms. On the other hand, it should also work well in small cells operating at higher spectrum bands as they allow more densely-packed antenna arrays.

Publication V designs such multipath extraction based CSI acquisition method as follows. A base station acquires the time-delays of the multipath propagation response from a plurality of user terminals on uplink. The estimation can be realized by minimum least-squares via standard channel training or by a sliding cross-correlator mechanism in case of a multi-selective rake receiver setup [198]. It should be noted that such dedicated time-delay estimations are necessary only when the successive delays of a receive signal are separated by shorter time offsets than the width of the receive autocorrelation; otherwise, multipath delays can be detected simply by the locations of the autocorrelation peaks [200]. After multipath time-delay and amplitude estimations, base stations can acquire the incoming DoA angles from uplink and transform them to downlink assuming that the duplex time elapse between the time frames is insignificant [199]. For instance, the DoA angles can be estimated by a subspace-based estimator based on the following receive autocorrelation.

$$\dot{\mathbf{R}}_{t,kl} = \frac{1}{N} \sum_{n=1}^{N} \hat{\mathbf{y}}_{kl}\left(t + (n-1)\mathcal{T}_{\mathrm{s}}\right) \hat{\mathbf{y}}_{kl}\left(t + (n-1)\mathcal{T}_{\mathrm{s}}\right)^{H}, \tag{5.1}$$

where N denotes the number of transmit symbols, $\hat{\mathbf{y}}_{kl}$ denotes the receive

Table 5.1. Channel parameter estimation at a base station

Step#1:	**Multipath time-delay estimation**
	∘ If the time offsets between successive paths are longer than the receive signal autocorrelation:
	☐ The delays are detected at the autocorrelation peaks
	∘ Otherwise:
	☐ If the base station is equipped with a multi-selective rake receiver:
	→ A sliding-cross correlator mechanism derives the path delays [198]
	☐ Otherwise:
	→ The delays are estimated by minimum least-squares by channel training
Step#2:	**The estimation of DoA angles**
	∘ If the number of resolvable paths is overwhelming:
	→ The estimated time-delays and fading parameters are processed for each dominant path individually
	∘ Otherwise:
	→ A subspace-based DoA estimator processes the aggregate receive autocorrelation [199]
Step#3:	**The design of a single-beam downlink training**
	∘ The optimum direction can be determined by:
	→ Orthogonal pilots that are precoded based on the second-order channel correlation statistics
	→ Iterations in a closed-loop by user feedback [177]

signal vector from the kth user's lth path on uplink, and $\mathcal{T}_s$ denotes the transmission time of a symbol. Notice that the receive signal components $\hat{\mathbf{y}}_{kl}$ can be formulated for each individual path thanks to the estimated multipath delay and fading terms. In addition, It should be also noted that the eigenvectors of $\dot{\mathbf{R}}_{t,kl}$ correspond to distinct multipath components due to the law of large numbers and Lindeberg-Lévy central limit theorem [155, 158]. Upon DoA estimation, downlink training sequences can be transmitted on a single beam from the base station, which overcomes the issue of long training duration inherent to massive MIMO. The optimum direction of the training beam can be derived by iterations in a closed-loop mechanism [177]. Table 5.1 summarizes the necessary operations of

Table 5.2. Channel parameter estimation at a user terminal

Step#1:	**Multipath time-delay estimation & tracking**
	∘ The initial estimation:
	→ The sample channel correlation matrix is structured on the pilot subcarriers
	→ The multipath time-delays are acquired by a subspace-based estimator
	∘ The subsequent tracking:
	→ Each path is modeled based on its last delay and amplitude estimates and the model is processed by a delay-locked loop (DLL) structure
	→ The output is cross-correlated with the early and late versions of the reference pilot symbols
	→ The difference between the squared outputs of the cross-correlations produces a tracking error, which indicates the new delay estimate
Step#2:	**The design of the MPI feedback**
	→ Downlink path coefficients are approximated by a minimum least-squares error detector
	→ MPI vector is formed based on the phase information of the indicated dominant path coefficients

a base station for the estimation of uplink channel parameters.

After channel training on downlink, each user terminal can estimate and track its corresponding multipath delays via the second-order channel statistics derived from the OFDM symbols of the received training sequence [201]. In addition, user terminals can approximate their downlink path coefficients and design *multipath phase indication* (MPI) vectors formed by the complex-valued fading coefficients of the selected dominant paths, as proposed for the first time in Publication VI. Table 5.2 lists the necessary operations at a user terminal for channel parameter estimation on downlink and the subsequent feedback.

The set of the selected dominant paths are proposed to be chosen and indicated by the base station for every user. Choosing only a subset of the propagation paths implies that the targeted CSI vector by the base station involves only those paths' responses instead of the whole propagation channel response, hence the naming of the multipath extraction

based CSI acquisition method. The optimal selection procedure for the dominant paths should depend on a number of parameters, such as path strengths, latency requirements, user priorities, pairwise path correlations, and prior estimation errors for the multipath parameters. Upon the selection of dominant paths, their indication can be realized by means of signaling only the number of paths as a single integer and relying on pre-defined path priorities based on either amplitude or delay information. An explicit indication of the selected paths can also be performed by a modified version of a time-reversal type precoding. With the received MPI feedback the base station acquires the necessary phase information on the targeted propagation paths. Typically, the number of dominant paths is much smaller than the number of antenna elements in a massive MIMO setup, therefore the signaling overhead by user feedback reports gets sharply reduced without sacrificing CSI accuracy.

The proposed CSI acquisition and MPI feedback schemes in Publication V and Publication VI offer the following benefits. First of all, only non-reciprocal multipath characteristics are required at user terminals while the rest of the channel parameters are estimated by the base station. Therefore, the issue of limited training duration can be avoided by a single beam transmission on downlink as full CSI vectors are no longer required at user terminals. What is more, the bulk of the necessary signal processing is handled at the base station, hence relaxing the computational complexities at user terminals. Both of these benefits are particularly welcome when compared with the compressive sensing methods from massive MIMO literature, which require not only full CSI knowledge but also higher computational loads at user terminals. Secondly, the signaling overhead by user feedback reports is also significantly reduced as the number of dominant paths is often measured to be between two to six in a typical mobile macro cell with elevated base station platforms [201, 202]. Furthermore, the proposed acquisition technique also allows flexible implementation with easily controllable feedback size and CSI accuracy.

Figure 5.1 presents the numerical evaluation results for the proposed multipath extraction based CSI acquisition scheme by means of average symbol error rate and sum-rate performance. The simulation incorporates a single cell with one eNB and 8 randomly dropped UEs. The eNB is equipped with 100 vertically polarized antennas in a 10×10 array structure while each UE has a single antenna. All transmit symbols are modu-

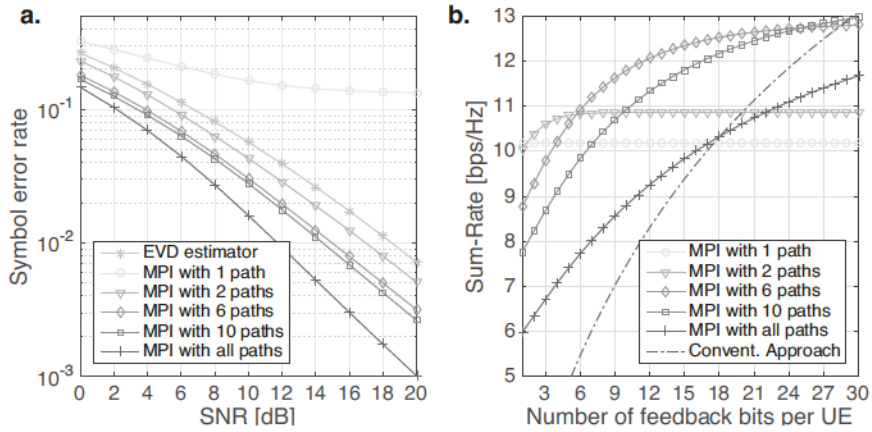

Figure 5.1. Performance comparison results of the MPI feedback based on 1, 2, 6, 10 of the most dominant propagation paths as the targets by **a.** average symbol error probabilities and **b.** average sum-rates. Publication VI, © 2015 IEEE.

lated by QPSK and the propagation channels are generated by WINNER+ model with the urban macro settings [203]. Downlink and uplink carrier frequencies are respectively defined as 2 and 2.5 GHz. The simulation lasts for 100 TTIs and operates on 50 subcarrier blocks. In addition, all user codebooks are simulated based on the random vector quantization (RVQ) based codebook scheme, as described by [204].

The left-hand side of Fig. 5.1 demonstrates the symbol error rates with the MPI feedback based on the strongest 1, 2, 6, 10 dominant paths as the targets. In addition, the performance of the conventional EVD estimator with a carrier frequency transformation is also included as a reference scheme. For the MPI feedback reports, two bits are allocated for each vector element. One can observe that the performance of the CSI reconstruction by only one path is limited by the ceiling effect at high SNR levels. Moreover, incorporating more propagation paths into the MPI feedback reports gradually reduces the symbol error rates while the signaling overhead is also increased.

The right-hand side of Fig. 5.1 shows the average sum-rate results versus the total number of feedback bits per each user report. The MPI feedback is simulated again with the same numbers of dominant paths as before, in addition to another reference curve representing the theoretical bound for the conventional type of feedback reports based on true narrowband CSI quantizations, although the indicated bits are counted only once for all subcarriers. As equal numbers of bits are allocated for the MPI vectors of different lengths, extracting less paths provides better performance

as a result of higher granularity with the vector quantizations. However, as more and more bits are allocated along the horizontal axis, increasing the number of incorporated paths generates higher sum-rates for the MPI feedback scheme.

5.3 Discussion

This chapter investigated promising CSI acquisition techniques by user feedback reports for massive MIMO and explained the contributions of the thesis based on the original works of Publication V and Publication VI. Most popular acquisition methods have been described and analyzed in three explicit categories based on the exploited characteristics of propagation channels.

Enabling the massive MIMO technology in cellular communications for the most widely deployed duplexing mode (i.e., FDD) is highly desirable. However, the estimation of downlink CSI is problematic in massive MIMO, due to the expanded size of CSI feedback reports as well as more demanding signal processing at UEs. Nevertheless, these limitations can be overcome by exploiting certain propagation characteristics that are seemingly considered unfavorable in conventional wireless communication systems. For example, spatial channel correlations are often undesirable as they limit the achievable spatial diversity. However, spatially correlated channels can facilitate significant CSI compressions, which is desirable for compact user feedback reports in massive MIMO.

The proposed CSI acquisition technique in this chapter takes advantage of the channel sparsity in multipath propagations. Compared with compressive-sensing based methods, this acquisition technique does not require full CSI estimation at user terminals. Therefore, downlink training can be based on a single-beam pilot transmission. In addition, computational load is sharply reduced at user terminals as the necessary signal processing does not increase with respect to the number of eNB antenna elements. The applicability in a real-world deployment, however, relies on planar incoming/ongoing wavefronts to fulfill the far-field approximation principle. This requirement should be suitable for large coverage areas, such as stadiums, concerts, open air festivals, etc., which are determined as potential test cases for 5G systems by the METIS project [205]. Large numbers of users in such service environments are expected to consume and exchange high-quality multimedia contents which will lead to spikes

in data traffic during a short period of time. These scenarios can be addressed by massive MIMO deployments in a large coverage area with elevated base station towers that can serve many users with very high spectral efficiency. In addition, future advancements in manufacturing and the introduction of higher spectrum bands (e.g., millimeter-wave frequencies) to cellular communications will help employ more densely packed antenna arrays with much shorter physical distance between neighboring antenna elements. As a result, the far-field approximation can be valid in smaller cells as well, which should facilitate the feasibility of the proposed method in more coverage locations. In any case, technological progress and better cost efficiency in computational power, battery life, and manufacturing will remove some of the present limitations concerning massive MIMO deployments. When network designers also figure out how to solve the problem of downlink CSI acquisition in an efficient manner, this technology will likely be indispensable in future cellular communication systems.

6. Full-Duplex In-Band Multi-Antenna Relaying

Wireless relaying is a significant enhancement in cellular networks for coverage extension, cost-efficient backend support, capacity improvement, and low power consumption. Relays may have various applications in a cellular communication by means of establishing reliable point-to-point connections, taking part in a cooperative cognitive radio system [206, 207], providing affordable wireless backhaul, or assisting in a multiuser transmission [208, 209, 210]. Concerning all of these use cases of wireless relays, the full-duplex technology can offer higher QoS by more efficiently harnessing the precious air interface resources compared with resource division methods, therefore potentially delivering up to twice spectral efficiency with higher data rates.

However, multi-hop full-duplex communication links are well known to be susceptible to loop-interference between the transmit and receive sides of the intermediate wireless relays. The adverse impacts of this unfavorable byproduct can be tolerated by applying physical isolation, internal cancelation in RF and digital baseband, or spatial-domain suppression in case of multiple antenna deployments. In the following, this chapter presents the benefits of full-duplex transmissions and the state-of-the-art mitigation techniques for the loop-interference inherent to wireless full-duplex relay links. In addition, hardware limitations and practical deployment challenges are also touched upon, as well as the current and future directions of the wireless full-duplex technology in cellular communications. Regarding the contributions of this thesis on multi-antenna techniques, Publication VII presents an optimization procedure for spatial-domain interference suppression on a full-duplex point-to-point MIMO relay, which closely approximates its global optimality in a single-stream transmission. In addition, Publication VIII derives the optimal set of transmit powers and develops an iterative optimization algorithm by SLNR-

based spatial-domain filters in a two-hop multi-input single-output (MISO) link setup.

6.1 Merits of full-duplex over resource division

The conventional relaying operation requires a resource division in either time or frequency domain. The former is based on assigning dedicated time slots for the transmit and receive tasks of the relay on a shared frequency band whereas the latter splits the available spectrum between the transmit and receive operations that are occurring simultaneously. In either case, a conventional relay transceiver can only tap into half of the available resources during any given time window, hence the performance of the end-to-end link stays inherently limited. Such resource-division based relaying is known as *half duplexing*. On the other hand, full duplex relaying utilizes the entire allotted radio resources, therefore it can score up to twice as much throughput despite the inadvertent loop-interference signal arriving at the receive end of the relay, as illustrated in Fig. 6.1.

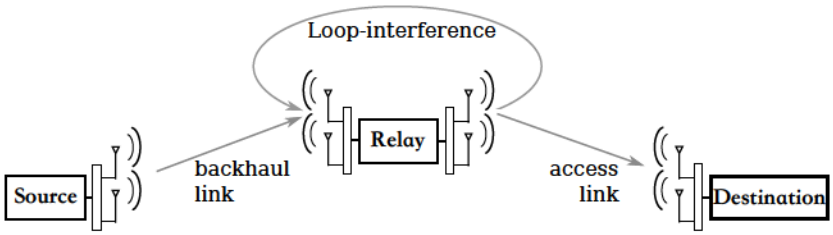

Figure 6.1. The transmit signal of a full-duplex relay interferes at its receive end.

Thanks to the full-duplex alleviating the inherent *rate loss*, wireless relays can be deployed for a wide range of use cases. First and foremost, relays can extend the service coverage by means of both overcoming shadow fading effects and reaching out farther locations with reduced transmit power [211]. Relays can operate either as a dumb transceiver which is controlled by its connected base station or as an intelligent entity with its dedicated scheduling and resource allocation decisions [212]. Relays can support multicast and broadcast transmission scenarios serving multiple terminals at a time. Massive antenna arrays can be deployed on a relay to act as a connected hub among a plurality of terminals [213, 214]. Relays can be beneficial in cognitive radio networks, where unlicensed users are allowed to use the bandwidth of primary users [215]. In addition, relays can also be utilized to serve as wireless backhaul for reduced finan-

cial costs during deployment and maintenance, or for overcoming terrain limitations that may prevent any wired backhaul installations [216].

6.2 Hardware limitations

In a full-duplex relay, the loop-interference emitted from the transmit side couples with the receive front-end, which may cause saturations due to both the transceiver's limited dynamic range and the inadequate knowledge of the incoming loop-interference signal. The issue of limited dynamic range is a typical result of the non-idealities in ADC, DAC, and power amplifiers whereas the inaccurate estimation of the loop-interference is mostly a result of CSI errors and RF imperfections [217]. Excessive CSI mismatches can also cause an instability by signal oscillations, particularly at non-regenerative relays. Full-duplex relays are also susceptible to a bias problem, which renders large correlations between the loop-interference and the desired signals at the receive-side in case of insufficient time delays until the repetitive arrival of the incoming signal stream through the transmit-side. In addition, other generic limitations, such as the thermal noise, quantization errors, phase noise, IQ imbalance, and keyhole effects, also degrade the performance of a full-duplex relay.

The characteristics of these hardware-related limitations usually depend on the type of relaying, which can be of either regenerative or non-regenerative nature. The most popular regenerative relaying protocol is DF [216]. Relays with DF decode and re-encode the source signal prior to its retransmission. Hence, they require complex design and heavy computations. To overcome these high demands, DF relays can employ only symbol-by-symbol decoding at the expense of performance, instead of fully decoding the entire source codeword [218]. Note that in a two-hop relaying setup only full error-free decoding with DF can theoretically reach the mutual information of the weaker hop. The most popular non-regenerative relaying is AF [219, 220]. The AF protocol operates by straightforwardly amplifying the incoming signal subject to a power constraint. On the other hand, network planning is more difficult with AF, compared to DF. In addition, due to lack of decoding, AF relays may amplify also the loop-interference, which may further lead to internal oscillations. Although AF performs sub-optimally, it is still considered a popular type of relaying due to its design flexibility and low complexity [221]. Although ADC and DAC are not necessary for the operation of AF, they are indispensable for

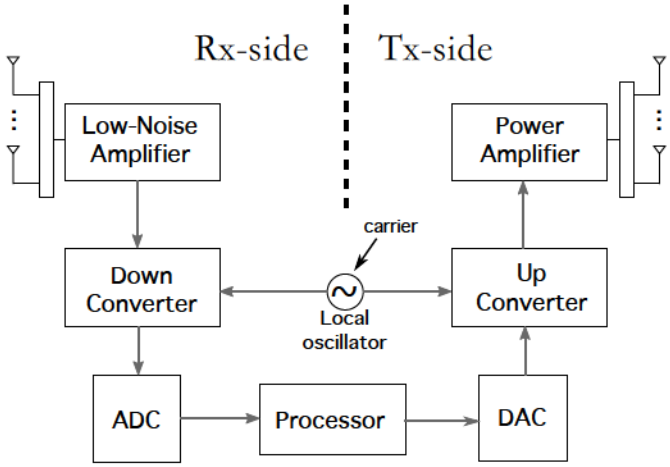

Figure 6.2. Block diagram of a MIMO full-duplex relay transceiver.

loop-interference mitigation in a full-duplex relay. The operation diagram of a MIMO full-duplex relay is depicted in Fig. 6.2.

In a two-hop link scenario, the destination node may receive the same signal stream twice, depending on the level of direct signal attenuation from the source node. If the gap between the arrival times of these OFDM-modulated signals emitted from the source and the relay toward the destination is shorter than the cyclic prefix, the signals will be combined constructively[1]. Otherwise, the weak signal directly originating from the source should be interpreted as interference at the destination.

In the following, the rest of this chapter fundamentally focuses on the mitigation of loop-interference as it is unavoidable under the full-duplex mode. In addition, other limitations (e.g., interference through the direct link, etc.) are also elaborated more in the context of end-to-end link optimization.

6.3 Loop-interference mitigation

In the part, the mitigation techniques for the loop-interference signal on a full-duplex relay link are grouped into two main categories: prevention/cancellation methods and multi-antenna spatial-domain suppression schemes. The former group comprises antenna/relay selection, physical isolation, analog RF cancellation, and digital cancellation by time-domain

[1]This is, however, unlikely in practice as full-duplex relays usually opt to add artificial, long delays to avoid any bias problem.

subtraction whereas the latter involves either full nullification or partial diminution of the loop-interference by multi-antenna processing in the spatial domain. Both of these mitigation categories are discussed in more details below.

6.3.1 Prevention/cancellation methods

Transmit and receive antenna selection is an effective, low-complexity way of preventing large loop-interference signals while requiring small amounts of CSI feedback overhead [222]. In case of multiple relay nodes in a network, the relay with the lowest loop-interference can be activated for the required end-to-end communication, hence proactively alleviating the issue of *self-interference* [223, 224]. In addition, a joint scheme can combine such antenna and relay selection strategies [225].

Physical isolation, also known as *passive suppression*, provides power attenuation for the loop-interference signal by means of blocking its propagation medium via physical separation between the transmit and receive sides of a full-duplex relay. Physical isolation techniques also often steer transmit and receive antenna directivities in such a way that the intersection between the main transmit and receive radiation lobes is minimized. Another means of isolation can be based on increasing the physical distance between the transmit and receive antenna arrays of a relay. In addition, physical isolation may also exploit the surrounding environment through obstructions, employ shielding plates between the antenna arrays, or polarize the transmit and receive waves orthogonally.

The cancellation of an RF interference signal can be realized by a standard noise canceler chip [226]. In a full-duplex relay such an RF canceler applies the necessary amplitude and phase corrections to the reference loop-interference, which is extracted from the previous transmission slots, so that it matches to the actual incoming loop-interference at the current time instant. After a successful matching, the RF canceler subtracts this reference signal in analog domain. Such RF signal cancellation is verified to be able to provide up to 30 dB reductions in loop-interference signals [227].

Loop-interference signals can also be canceled digitally in baseband by means of subtracting an estimate of the interference based on the previously transmitted baseband signal from the relay. Apparently, accurate modeling for the loop-interference also requires proper knowledge on the loop-interference channel. However, the feasibility of the necessary chan-

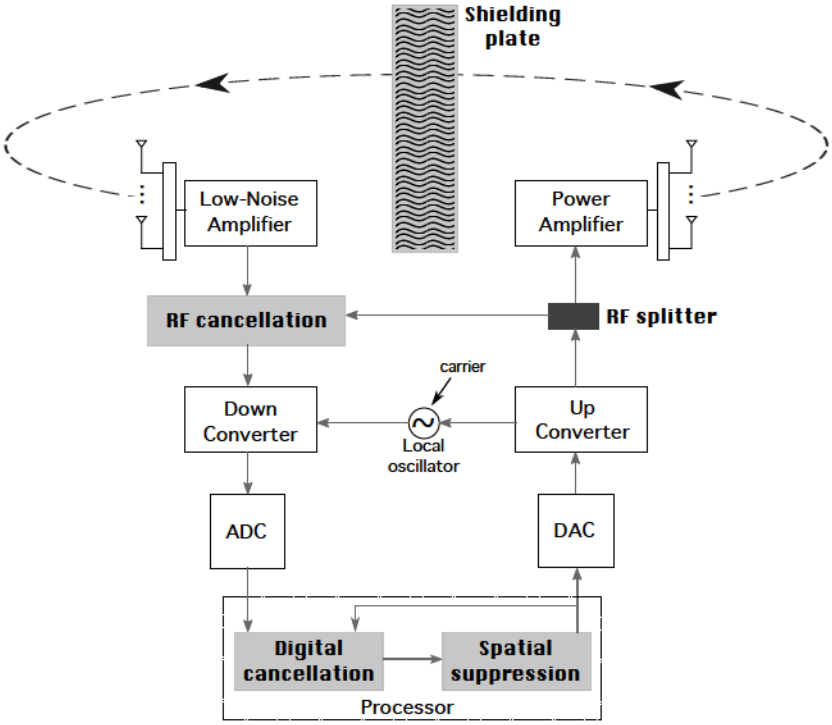

Figure 6.3. Illustration of the loop-interference cancellation techniques operated within a MIMO full-duplex relay, where each grey-colored block represents a different cancellation type.

nel estimation can be dubious due to phase noise errors [228] and limited dynamic range [229]. In addition, delay and phase alignments are also required on the interference model prior to the final subtraction, which can be computed by a standard detector that measures the peak of the signal correlations. The operational diagrams of these cancellation techniques are illustrated on a full-duplex MIMO relay model by Fig. 6.3.

6.3.2 Spatial suppression

Due to the discussed hardware limitations in Section 6.2, the prevention/cancellation methods for the loop-interference often exhibit limited performance, which leaves room for improvement by spatial-domain suppression. Specifically, analog RF cancellation leaves behind residual noise as a result of the imperfections in RF splitters and CSI estimates. Moreover, such analog cancellation requires a dedicated hardware chip, which may be missing in a transceiver. Digital cancellation is performed only after the receive signal goes through ADC, hence the issue of front-end saturation cannot be prevented. In addition, physical isolation and digital

cancellation together are experimentally shown to be insufficient for fully removing out the loop-interference when equipped with omnidirectional antennas in an indoor prototype cellular setup [230]. On the contrary, spatial-domain multi-antenna techniques can create spatial nulls at both transmit and receive antenna arrays, or particularly alleviate the receive front-end saturation via pre-suppressing the bulk of the loop-interference by transmit-side filtering.

An intermediate relay can establish a two-hop connection between a base station and a user terminal for coverage extension on downlink. In such scenarios, the direct link between the end points of a transmission is often considered weak. Therefore, the direct link can be interpreted either as negligible or as a cause of interference at the user terminal. During such downlink transmissions by a two-hop relaying, the serving base station acts as the *transmission source* whereas the user terminal is considered the *signal destination*, as depicted in Fig. 6.4. In a MIMO two-hop relaying, spatial suppression techniques can improve the performance of the end-to-end link by means of either full nullification of the loop-interference or by its partial mitigation based on global SINR maximization.

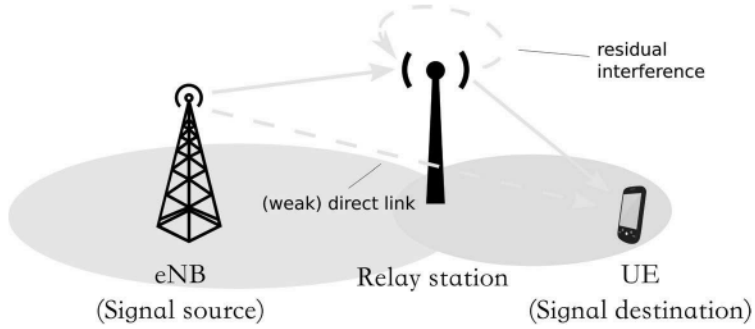

Figure 6.4. Coverage extension in a downlink transmission by full-duplex relaying.

Forming spatial nulls at the transmit and receive antenna arrays of a full-duplex relay is an effective measure that performs well in interference limited conditions. Such spatial nullifiers are usually computed according to the right and left singular matrices of the loop-interference channel, after which the remaining degrees of freedom are taken into account to improve the quality of the desired signal. However, the optimal solution for such procedure is hard to derive even if the ideal CSI is available. One promising approach for a sub-optimal derivation can be based on an iterative computation of the spatial filters, where each of

them is optimized separately in succession based on the orthogonal projection of the other filter from the last iteration [231]. Similarly, such optimization can also be realized by the gradient projection algorithm [232]. Another potential sub-optimal solution for full interference suppression can be based on individually nullifying the disjoint eigenbeam subsets of the loop-interference channel by means of computing its SVD. Although the feasible solution set is smaller than the former category of iterative methods, final spatial filters can be derived in closed forms by the latter approach due to its disjoint nullifications. Publication VII combines the orthogonal projection method with such a disjoint nullification approach in order to derive the spatial filters for the most generalized case, which incorporates any rank-deficient channels in addition to the full-rank condition presented in literature [233, 234]. Such rank-deficiency is expected to occur more frequently in a full-duplex operation when the same antenna array is not shared between the receive and transmit ends, which is usually the case with full-duplex relaying as opposed to half-duplex. In addition, note that the said rank-deficiency here also encapsulates any full-column or full-row rank conditions, which should frequently occur under a pair of receive and transmit arrays that are equipped with unequal numbers of antennas.

Regarding the global maximization of the end-to-end SINR metric, power control is an integral part to the feasibility of a full-duplex relay link. As its optimality requires striking an ideal balance between improving the desired signal and mitigating the loop-interference, the feasible solution set often entails only partially suppressed residual interference. However, there can be cases with the DF protocol where a full nullification indicates the globally optimal solution. Apparently, altering the transmit power of a full-duplex relay has a direct impact on the loop-interference. Higher transmit powers boost the interference signal, therefore causing higher rate losses at the first hop while improving the signal quality at the second hop. On the contrary, low transmit powers at the relay diminish the loop-interference at the receive front-end while also reducing the strength of the desired signal at the destination terminal. Hence, generalized spatial suppression techniques with residual loop-interference cannot be developed independently of the transmit power allocations in a full-duplex relay. Due to the sophistication and non-convexity of this global optimization problem, there is no known solution providing a generalized optimum spatial suppression strategy for full-duplex MIMO relays, although

a handful of attempts has been made in literature in order to approach the theoretical performance bound on a two-hop link. One such study is based on an iterative algorithm that derives suboptimal transmit spatial filters at the source and relay that is equipped with a digital interference canceller [235]. Another study is based on deriving the receive and transmit filters of a MIMO relay in succession based on maximum signal-to-interference ratio (SIR) [236] while the other filter is being disregarded. Instead of designing sub-optimal solutions, some studies have conducted theoretical analyses and derived lower and upper performance bounds for the end-to-end achievable rates on a full-duplex MIMO relay link [237].

In terms of the contributions of this thesis, Publication VII investigates the global optimization of a two-hop single-stream point-to-point link with a full-duplex MIMO relay by means of end-to-end SINR maximization. Since the end-to-end link exhibits dissimilar behavior based on the relationship between the first and second hop channel SNRs, the optimization problem is divided into three explicit subproblems. When the link quality between the source and relay stations is much better than that of the relay–destination link, the global optimality renders almost full loop-interference nullification. On the contrary, when the second hop is much stronger than the first, the optimal spatial filters converge to channel matched filters. Under the DF protocol both of these convergences reach to the global optimality as long as the strength of the stronger hop surpasses a certain threshold, which is derived in Publication VII based on the instantaneous channel SNR metrics. In case of AF, however, this convergence does not conclude during the search in the feasible design set. Such distinction is present between these protocols due to their different natures on operational linearity. Assuming full decoding/re-encoding without errors, the performance of a DF relay link is only tightly bounded by its weaker hop, independently of the other. On the contrary, the AF protocol exhibits a continuous rate of change in the overall performance, which indicates that the relative strengths of the hops with respect to each other do not lead to a performance saturation on the end-to-end link. Concerning the third subproblem where neither hop is much stronger than the other, the optimization relies on an iterative approximation between the transmit- and receive-side filters based on separately deriving one of them after having fixed the other. For a given transmit-side filter, the optimal receiver can be computed by minimum MSE criterion. For a given receive-side filter, the optimal transmitter can be obtained by a con-

vex combination between the full nullification and hermitian transpose schemes in a single-stream transmission.

Figure 6.5 presents the simulation results of the described optimization scheme with comparison to the maximum theoretical rates achievable by spatial-domain suppression. The simulation setup is a two-hop one-way relay link with negligible direct link between the end points. The propagation channels are modeled by flat-fading uncorrelated random variables and the thermal noise at the receivers is generated as zero-mean unit-variance independent Gaussian elements. The average gain of the channels between different communication nodes is defined as 15 dB whereas the average gain of the loop-interference channel is varied from -30 to 30 dB. The data transmission is carried over a single stream and the relay station is equipped with either 2 or 5 antennas at both receive and transmit sides. The source station employs maximum transmit power as it is the ideal strategy with negligible direct link between the source and destination. In addition, the simulation is modeled to be free from any hardware imperfections or CSI estimation errors.

The vertical axes in Fig. 6.5 represent the end-to-end transmission rates whereas the horizontal axes indicate the average strength of the loop-interference channel. The following conclusions can be drawn based on the transmission rate results of Fig. 6.5. A common observation among all antenna configurations is that the theoretical bound is distinctly more sensitive to the loop-interference when its strength is about -30 dB with the DF protocol. This phenomenon can be explained as follows. When the level of loop-interference is high, the optimal spatial suppression almost nullifies the interference, hence rendering the sensitivity of the transmission rates low along the horizontal axes. On the contrary, low loop-interference necessitates a more delicate balance between improving the desired signal and minimizing the interference, therefore the optimum spatial suppression leaves behind larger residual interference. This is in fact a common behavior in wireless communications, where the optimal multi-antenna filtering approaches the zero-forcing solution under high interference whereas the channel matched filter performs well when the interference is low. Another remarkable indication of this phenomenon is that it occurs as a side-effect due to insufficient spatial degrees-of-freedom. More specifically, the maximum achievable rate plots in Fig. 6.5 should theoretically produce the same transmission rates along the horizontal axes if the relay is considered to be equipped with unlimited num-

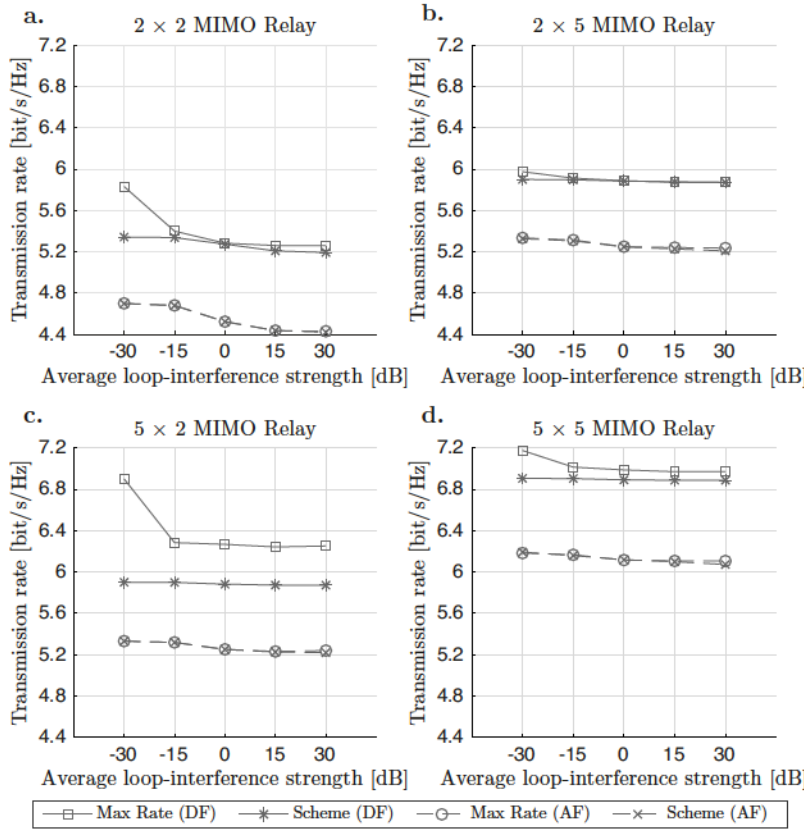

Figure 6.5. Achievable data rates by the proposed spatial suppression scheme (red lines) in comparison to their theoretical limits (blue lines) in a two-hop downlink scenario with a full-duplex **a.** 2×2, **b.** 2×5, **c.** 5×2, **d.** 5×5 MIMO relay. Publication VII, © 2014 IEEE.

ber of antennas.

Moreover, the DF protocol, in contrast to AF, renders unequal rates with the antenna configurations $2{\times}5$ and $5{\times}2$ in Fig. 6.5. Such dissimilar behavior with the performance of DF can be explained as follows. At the relay, both the transmit power and the transmit-side filter can balance the optimization in a similar manner by improving the quality of the transmit signal as opposed to reducing loop-interference, or vice versa. On the other hand, the balancing between improved desired signal and reduced loop-interference at the relay is only impacted by the receive-side filtering among all design parameters. As a result, having a larger antenna array at the receive-side, rather than at the transmit-side, allows more refined optimization with the DF protocol, as its performance is tightly bounded by the weaker hop. Thus, employing more antennas at the receiver helps

to relax the limited freedom on the source–relay hop, particularly when the level of loop-interference is low. On the contrary, the same observation is not present with the AF protocol as it is more stable against unequal hop qualities. To be specific, the performance gap between DF and AF is maximized when the source–relay and relay–destination links deliver equal rates. On the contrary, when either of the hops theoretically delivers infinite data rate, DF and AF reach asymptotically identical throughputs.

Publication VIII presents an iterative optimization method for a full-duplex MIMO relay on MISO end-to-end link, where a base station is equipped with multiple antennas during a downlink transmission. The optimization in spatial domain involves designing baseband multi-antenna filters at the receive and transmit sides of the relay station as well as the transmit filter at the base station. The design is based on the assumptions that the relay contributes non-negligible processing delays and that the direct link between the end points is weak, hence interpreted as interference at the destination terminal. The optimization problem should also address developing optimum power allocations at both of the transmit stations, as constrained by their upper bounds. Publication VIII firstly derives for both DF and AF protocols the optimal pairs of transmit powers as well as the optimal receivers when the rest of the design parameters are known. Next, these findings are combined to a single closed-form expression as an iterative search between these separate expressions is not guaranteed to converge to the global optimum due to the non-convexity of the transmit power expressions. Second, both transmit filters are optimized by maximum SLNR condition for a fixed receive filter and transmit powers. Finally, an iterative algorithm is designed to jointly optimize all system parameters in baseband, as elucidated in the following.

Let us first denote the receive and transmit side baseband filters of the MIMO relay respectively by $\mathbf{g}_r$ and $\mathbf{g}_t$, and the transmit filter at the source station by $\mathbf{g}_s$. In addition, let us also denote the power allocations at the source and relay stations by p_{S} and p_{R}, respectively. The optimization algorithm is initiated by setting $\mathbf{g}_s$ and $\mathbf{g}_t$ to channel matched filters. At each iteration the optimal p_{S}, p_{R}, and $\mathbf{g}_r$ are derived as described above, and subsequently both $\mathbf{g}_s$ and $\mathbf{g}_t$ are re-computed based on the maximum SLNR. These computations are repeated at every iteration till the first drop in the end-to-end SINR, denoted by γ_{e2e}. Additionally, the algorithm can also be terminated after a pre-determined number of iterations, e.g., when $\mathcal{I} = 100$ iterations are completed. The required steps of this joint

Table 6.1. Joint Iterative Algorithm on a Full-Duplex MISO Dual-Hop Link

Step#1:	Initialize iteration index $i = 1$, set $\gamma_{\text{e2e}} = 0$, and
	define $\mathbf{g}_s^{(0)}$, $\mathbf{g}_t^{(0)}$ as channel hermitian transpose
Step#2:	Derive $p_{\text{S}}^{(i)}$, $p_{\text{R}}^{(i)}$, and $\mathbf{g}_r^{(i)}$ based on given $\mathbf{g}_s^{(i-1)}$ and $\mathbf{g}_t^{(i-1)}$
Step#3:	Compute $\mathbf{g}_s^{(i)}$ and $\mathbf{g}_t^{(i)}$ by max. SLNR
Step#4:	**if** $\gamma_{\text{e2e}}^{(i)} < \gamma_{\text{e2e}}^{(i-1)}$
	Set $p_{\text{S}} = p_{\text{S}}^{(i-1)}$, $p_{\text{R}} = p_{\text{R}}^{(i-1)}$, $\mathbf{g}_r = \mathbf{g}_r^{(i-1)}$, $\mathbf{g}_s = \mathbf{g}_s^{(i-1)}$, $\mathbf{g}_t = \mathbf{g}_t^{(i-1)}$,
	and **exit**
	elseif $i > \mathcal{I}$
	Set $p_{\text{S}} = p_{\text{S}}^{(i)}$, $p_{\text{R}} = p_{\text{R}}^{(i)}$, $\mathbf{g}_r = \mathbf{g}_r^{(i)}$, $\mathbf{g}_s = \mathbf{g}_s^{(i)}$, $\mathbf{g}_t = \mathbf{g}_t^{(i)}$,
	and **exit**
	else
	Set $i = i + 1$ and go to *Step#2*
	end

iterative algorithm are listed in Table 6.1.

Figure 6.6 demonstrates the numerical evaluation results of the described algorithm in Table 6.1. The system setup is a dual-hop MISO relay link with an additional direct end-to-end connection that is processed as interference at the destination terminal. Due to the full-duplex operation the system is also degraded by loop-interference at the relay receiver. The wireless channels are modeled with flat-fading independent and identically distributed elements. The system also contains thermal additive noise at both receivers, which is modeled by zero-mean independent Gaussian random variables. The mean strengths of the source–relay and relay–destination channels are 10 dB whereas the direct link and loop-interference channels have varying strengths from 0 to 10 dB while always being kept equal to each other. The source station is equipped with 4 antennas whereas the relay station employs 2 transmit and 2 receive antennas. The destination terminal has a single antenna, hence the end-to-end link is MISO. Also, the maximum transmit powers are constrained to one and the system is considered to experience neither hardware nor CSI imperfections. In addition, note that all the curves in Fig. 6.6 employ the optimal power control, although the spatial filters are designed differently. Last but not least, the number of maximum iterations (i.e., $\mathcal{I}$) is set to 100 to keep the simulation run-time short.

The vertical axes in Fig. 6.6 indicate the achievable data rates whereas

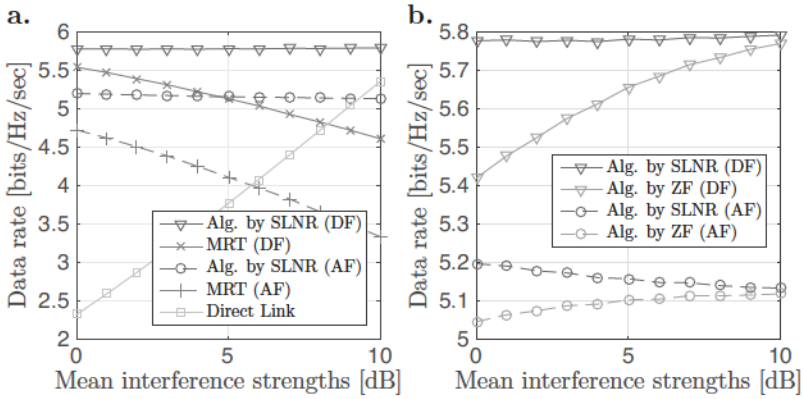

Figure 6.6. Achievable data rates by the joint iterative algorithm, as compared with: **a.** linear MRT filter and direct link transmission, and **b.** the same joint iterative algorithm, but with ZF transmitters instead of SLNR. Publication VII, © 2014 IEEE.

the horizontal axes represent the mean strengths of both interference channels, which are namely the loop-interference and direct link channels. One can draw the following conclusions based on the numerical results. On the left-hand side, Fig. 6.6a compares the performance of the iterative algorithm with that of MRT filtering and direct link connection under both AF and DF modes. The former scheme is realized by defining the transmit filters as channel matched filters while the optimal relay receiver is derived by the minimum MSE. The latter scheme is based on shutting off the relay station and relying on the direct link for data transmissions between the source station and destination terminal. Clearly, increasing the strengths of the loop-interference and direct link gradually degrades the MRT performance while enhancing the direct transmission link. However, the proposed iterative algorithm with SLNR outperforms others at any point along the horizontal axis. On the right-hand side, Fig. 6.6b investigates the impact of the SLNR transmitters on the proposed iterative algorithm by re-defining the transmitters as ZF filters. This is realized by means of first pre-nullifying the interference signals transmitted from the source and relay stations prior to employing the remaining degrees-of-freedom for enhancing the desired signal quality. The SLNR based transmitters outperform ZF at every point within the simulated range of interference strengths. In addition, one can notice the following counterintuitive phenomenon in Fig. 6.6b. The performance of the ZF-based schemes unconventionally improves with higher interference levels along the horizontal axis, which can be explained as

follows. The balancing problem between optimizing the desired signal and weakening the interference signals becomes a more delicate procedure at lower interference levels. Therefore the iterations in Table 6.1 start bouncing with small offsets relatively faster when the interference signals are weak. This leads to an early termination of the algorithm by satisfying the exit condition at *Step#4* of Table 6.1 fairly quickly.

6.4 Discussion

This chapter investigated the optimization of wireless full-duplex relay links and presented the contributions of the thesis based on the original works of Publication VII and Publication VIII. Specifically, the benefits of the full-duplex technology over half-duplex are highlighted; hardware imperfections and the concept of loop-interference signal are explained; and popular mitigation methods for residual loop-interference are discussed with particular emphasis on the multi-antenna spatial suppression techniques.

An important design target in cellular communications is providing reliable, high-capacity connections toward cell-edge users, which can be achieved with wireless relay deployments. AF relaying can be integrated to cellular systems without any standardization efforts as it is simply operated as a Layer-1 repeater. As a matter of fact, it has already been employed in the second and third generation cellular communication networks [238]. However, this type of relaying causes unintentional amplifications in the noise and inter-cell interference, therefore high-capacity transmission goals cannot be thoroughly fulfilled. On the contrary, the use of DF relaying requires standardization and, in fact, has been supported by LTE-A for fixed deployments since Release-10 [239]. This type of standardized DF relays, which operate on Layer-3, can achieve higher throughputs by combatting inter-cell interference and noise signals thanks to demodulation/decoding and modulation/re-encoding processes. In addition, the deployment of mobile relay nodes (e.g., on top of a high-speed bullet train) has also been in discussion by 3GPP for a while as a study item [240].

Although wireless relays are currently deployed in cellular networks, the configured operational modes are based on either out-band or in-band half-duplexing. Despite the available support for in-band full-duplexing by the cellular standards [241], the only successful commercial implemen-

tation of the wireless full-duplex technology has been in the field of digital terrestrial broadcasting so far, where the deployed full-duplex relay stations are called *on-channel repeaters* [242]. However, the full-duplex technology is expected to appear in future cellular networks as it offers not only higher spectral efficiency and better link reliability but also opens up new, exciting opportunities for many technology enablers toward next generation cellular services. As more diversified types of UEs and IoT terminals require incessant cellular connections, the increasing demand for improved access links can be provided by full-duplex relaying. With a higher user density, terminal devices can act as wireless relays to forward the cellular traffic as intermediate ad-hoc elements, particularly for indoor coverage. Vehicles can operate as moving relays for filling in coverage holes while driving around urban environments. In addition, relays can support green communications and help MTC applications in conserving terminal battery lives. It should be noted that future relaying applications are expected to be compatible to multitude of devices and services, therefore specified relay functionalities also need to be adjustable for different usage scenarios.

7. Conclusions

To address the ever-growing user demand for mobile communication technologies, service providers and designers of cellular systems want to attract more subscribers and seduce the existing ones toward more premium tiers by means of introducing new types of applications and establishing more aggressive improvements to the existing offerings at a more competitive pricing. These advancements push the technical boundaries of cellular communications further at each iteration. Future evolutions of cellular systems are expected to offer even higher data rates, lower latency, reduced power emissions in denser coverage areas, as well as initiate new application fields for more diverse usage scenarios. As a result of these objectives, many technology components are being developed and analyzed for feasibility and suitability studies within the timeline of each generation over the incessant evolution of cellular communication systems.

Multi-antenna deployments keep attracting significant interests from both academic and industrial research communities by creating vast opportunities for reaching the said objectives. Signal processing techniques for multiple antennas allow establishing a third dimension by the spatial domain next to the spectral and temporal domains. This third dimension helps researchers bend the rules of the game by enabling concurrent multi-layer communications, low-power concentrated transmissions, spatial orthogonalization, etc. without consuming auxiliary spectro-temporal resources. This thesis investigated state-of-the-art multi-antenna processing techniques and developed novel contributions for enabling some of the target objectives for next-generation cellular communication systems.

First and foremost, serving multiple users on shared resources is of utmost importance for the efficient sustainability of high-rate transmissions

on downlink. Deploying multiple antennas facilitates both multi-layer data delivery and orthogonalization among users. This thesis investigated linear and non-linear baseband processing strategies and explored their benefits and drawbacks in terms of performance, complexity, and signaling requirements in the presence of channel state information (CSI) errors and thermal noise. In particular, joint transmitter and receiver design methods are explored on a multi-user multi-input multi-output (MIMO) link as they can potentially offer higher performance results despite the associated optimization being a non-convex problem. In terms of the contributions, a novel joint design technique is proposed for high signal-to-noise ratio (SNR) channel conditions, where the captured performance gain can compensate for the signaling loss. Simulation results demonstrate that the proposed method outperforms state-of-the-art filtering schemes under low thermal noise levels, and also under high thermal noise as long as the CSI mismatch stays the primary performance limitation in the system.

As resource allocation is equally, if not more, important in the efficient utilization of the available spectrum, this thesis also dealt with the optimization of downlink schedulers. Allocation, assignment, and user ordering strategies are investigated on a multi-user MIMO downlink with the assumption of perfect orthogonalization among the resources. In particular, this thesis proposed two heuristic scheduling methods based on the greedy and genetic algorithms. The proposed schedulers apply an iterative method, where only one additional user is attempted to be co-scheduled at each iteration next to the set of primarily scheduled users on every resource block. Such approach also motivated for a new type of chromosome structure for the genetic algorithm. System-level simulation results indicate that the proposed genetic scheduler performs well with low complexity in crowded networks whereas the greedy based scheduler is preferable when less users are co-scheduled.

The future work on the proposed joint baseband processing can be related to its extension to a non-linear scheme. In addition, the necessary feedforward signaling for the distribution of the receive filter weights and the associated impact on the performance in terms of resource loss can be analyzed theoretically or evaluated in a simulation setup. Regarding the future work on multi-resource schedulers, the proposed algorithms can be evaluated in more sophisticated network setups, such as under a massive antenna array that allows tens of users on the same shared resources or

by taking into account the residual inter-carrier interference (ICI) among successive resource blocks. Also, the associated scheduler complexities can be derived by a theoretical analysis in a future work.

Second of all, the massive MIMO technology is well known to offer improved CSI orthogonality among multiple users and resistance against multipath fading, as well as help achieve reduced power loss at user terminals. However, the acquisition of a channel response is a concerning issue due to the requirements for extended user feedback in addition to the limitations in storage and computational capabilities. This thesis investigated the state-of-the-art CSI estimation strategies that exploit the reciprocity property of propagation channels on massive MIMO links operating in time division duplex (TDD) mode. To start with, the thesis discussed multi-antenna calibration techniques that can recover the reciprocal characteristics of the channel once the hardware imperfections are resolved at radio frequency (RF) transceivers. Regarding the contributions, this thesis proposed an adaptive quantization scheme for user CSI feedback reports during a relative antenna calibration procedure. The adaptation is based on measuring the instantaneous channel conditions of uplink and deriving the optimum amount of granularity for the subsequent CSI feedback report. Numerical results demonstrate considerable reductions in signaling overhead when the CSI feedback is quantized adaptively.

Assuming that the channel reciprocity is restored, this thesis also investigated blind channel estimators on the uplink of a multi-user massive MIMO connection. Although massive MIMO allows multi-source signal separation during a blind CSI estimation process, the residual multi-user distortions deteriorate the estimation accuracy as a result of the long channel vectors being only asymptotically orthogonal with each other. In addition, the limited sample size of the transmit symbols also reduces the CSI accuracy. This thesis proposed a multi-cellular coordination scheme for such multi-source blind estimators on the uplink of a massive MIMO connection based on jointly processing the available receive data from multiple base stations and iteratively approximating the sources of degradations due to limited receive symbols and finite antennas. Simulation results verify the improvements in terms of higher CSI accuracy and demonstrate the analyses on the speed of convergence for the proposed coordination scheme.

The future research work could be related to extending the proposed coordination algorithm for the pilot-based uplink training. Due to the poten-

tial risk for pilot contamination, only a small number of pilot sequences can be allowed in a dense network. The optimization could address firstly separating the users into a few groups. Then, all the users in each group can be simultaneously detected by a single blind estimator while different groups are separated from each other by orthogonal signal sequences.

Thirdly, as channel reciprocity is unavailable in frequency division duplex (FDD) systems due to much shorter coherence bandwidth relative to the duplex frequency gap, enabling massive MIMO in FDD operation requires brand new acquisition methods for downlink CSI. This thesis investigated the most popular techniques addressing the notorious challenges associated to downlink training and user feedback compression. The state-of-the-art schemes are grouped into three categories based on the exploited channel characteristics, which are temporal correlation, spatial correlation, and multipath sparsity. In terms of the contributions, this thesis proposed a novel acquisition technique based on reconstructing full CSI by a selected number of the strongest propagation paths. As some channel parameters are identical on forward and reverse links, a base station is proposed to estimate those reciprocal characteristics of the channel directly from uplink whereas the served user terminals are only required to measure and transfer the phase information associated to the targeted dominant paths for final CSI reconstruction.

The benefits of the proposed CSI acquisition technique include reduced feedback overhead by uncoupling its size from the number of base station antennas as well as removing the requirement for full CSI vector estimation at user terminals. Numerical results verify the potential of the proposed technique for FDD-mode massive MIMO connections, particularly in macro cells with elevated base station towers. Future research directions should include more detailed investigations on practical realizations as well as the applicability to smaller cells at millimeter-wave frequencies, which can enable more densely-packed antenna array structures.

Last but not least, the full-duplex technology has been receiving significant interests in wireless communications due to its ability for an economical use of resources. Instead of conventional duplex division schemes, a transceiver operating in full-duplex simultaneously receives and transmits on the same spectral band while the separation between incoming and outgoing streams are safeguarded by other means. This thesis investigated the full-duplex transmission technology on a relay link, which is a frequent application area for the full-duplex operation, and focused on

the state-of-the-art mitigation schemes for the loop-interference occurring at the receiver side of a full-duplex relay. The presented discussions also address the sensitivity of the mitigation techniques to hardware imperfections whereas the focus is still mostly on spatial-domain suppression methods, which can be performed thanks to multi-antenna deployments at such relays.

Regarding the contributions on full-duplex, this thesis proposed a coherent algorithm that approximates the global optimality on a dual-hop link with a MIMO full-duplex relay. The algorithm is based on designing baseband transmit and receive filters via piecewise analyses on the global optimization problem. In addition, optimal power allocations are derived on a multi-input single-output (MISO) link and the necessary baseband filters are optimized by means of swapping the associated leakage signals between the source and relay transmitters. Future work on full-duplex should focus on improving the resistance of the proposed schemes against a large CSI mismatch or a potential receiver-end saturation, which may occur as a result of extreme signal overloads with non-ideal hardware equipment. A mechanism can be developed that performs on-line analyses on the instantaneous stability conditions of the system and feeds the outcome into the proposed schemes based on an approximated metric on the CSI inaccuracy.

References

[1] V. J. Phillips, *Early Radio Wave Detectors*. London, UK: Peter Peregrinus Ltd., 1980.

[2] T. K. Sarkar, R. Mailloux, A. A. Oliner, M. Salazar-Palma, and D. L. Sengupta, *History of Wireless*. Hoboken, New Jersey: Wiley, 2006.

[3] G. Bussey, *Wireless, the Crucial Decade: History of the British Wireless Industry, 1924-34*. London, UK: Peter Peregrinus Ltd., 1990.

[4] Ericsson, "Ericsson mobility report," Ericsson, Tech. Rep., June 2015.

[5] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[6] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, February 2014.

[7] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo, "Green radio: radio techniques to enable energy-efficient wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, June 2011.

[8] S. Wu, H. Wang, and C.-H. Youn, "Visible light communications for 5G wireless networking systems: from fixed to mobile communications," *IEEE Network*, vol. 28, no. 6, pp. 41–45, November 2014.

[9] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, February 2014.

[10] N. Bhushan, J. Li, D. Malladi, R. Gilmore *et al.*, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, February 2014.

[11] Y. Ding, Y. Jin, L. Ren, and K. Hao, "An intelligent self-organization scheme for the Internet of things," *IEEE Computational Intelligence Magazine*, vol. 8, no. 3, pp. 41–53, August 2013.

[12] G. P. Fettweis, "The tactile Internet: Applications and challenges," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 64–70, March 2014.

[13] N. Michailow, M. Matthe, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, "Generalized frequency division multiplexing for 5th generation cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3045–3061, September 2014.

[14] K. Huang and V. K. N. Lau, "Enabling wireless power transfer in cellular networks architecture, modeling and deployment," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 902–912, January 2014.

[15] V. Yazici, U. C. Kozat, and M. O. Sunay, "A new control plane for 5G network architecture with a case study on unified handoff, mobility, and routing management," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 76–85, November 2014.

[16] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 336–361, February 2013.

[17] R. E. Hattachi and J. Erfanian, "5G white paper," NGMN Alliance, Final Deliverable (Approved), February 2015. [Online]. Available: http://www.3gpp.org/

[18] M. Gupta, "5G technologies and services," *Eurescom message*, Spring 2014. [Online]. Available: https://www.celticplus.eu/

[19] 3GPP TS 23.401 v13.2.0, "GPRS enhancements for E-UTRAN access (Release 13)," 3GPP, Tech. Rep., March 2015.

[20] 3GPP TS 23.402 v13.1.0, "Architecture enhancements for non-3GPP accesses (Release 13)," 3GPP, Tech. Rep., March 2015.

[21] 3GPP TS 36.213 v12.5.0, "Physical layer procedures (Release 12)," 3GPP, Tech. Rep., March 2015.

[22] 3GPP TS 36.211 v12.5.0, "Physical channels and modulation (Release 12)," 3GPP, Tech. Rep., March 2015.

[23] 3GPP TS 23.002 v13.1.0, "Network architecture (Release 13)," 3GPP, Tech. Rep., December 2014.

[24] 3GPP TR 36.814 v9.0.0, "Further advancements for EUTRA (physical layer aspects)," 3GPP, Tech. Rep., March 2010.

[25] 3GPP TS 36.300 v12.5.0, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); overall description; stage 2 (Release 12)," 3GPP, Tech. Rep., March 2015.

[26] B. Bangerter, S. Talwar, R. Arefi, and K. Stewart, "Networks and devices for the 5G era," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 90–96, February 2014.

[27] G. Wunder, P. Jung, M. Kasparick, T. Wild *et al.*, "5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, February 2014.

[28] S. Hong, J. Brand, J. Choi, M. Jain *et al.*, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114–121, February 2014.

[29] J. Mitola, J. Guerci, J. Reed, and Y.-D. Yao, "Accelerating 5G QoE via public-private spectrum sharing," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 77–85, May 2014.

[30] J. Qiao, X. Shen, J. Mark, and Q. Shen, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 209–215, January 2015.

[31] R. L. G. Cavalcante, S. Stanczak, M. Schubert, A. Eisenblaetter, and U. Tuerke, "Toward energy-efficient 5G wireless communications technologies: Tools for decoupling the scaling of networks from the growth of operating power," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 24–34, November 2014.

[32] D. Wubben, P. Rost, J. S. Bartelt, M. Lalam *et al.*, "Benefits and impact of cloud computing on 5G signal processing: Flexible centralization through cloud-RAN," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 35–44, November 2014.

[33] R. Q. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 94–101, May 2014.

[34] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, February 2014.

[35] A. Al-Dulaimi, S. Al-Rubaye, Q. Ni, and E. Sousa, "5G communications race: Pursuit of more capacity triggers LTE in unlicensed band," *IEEE Vehicular Technology Magazine*, vol. 10, no. 1, pp. 43–51, March 2015.

[36] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini *et al.*, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80–93, November 2014.

[37] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution: HSPA and LTE for mobile broadband*, 2nd ed. Burlington: Elsevier, 2009.

[38] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, 2nd ed. West Sussex: Wiley, 2011.

[39] D. Didascalou, J. Maurer, and W. Wiesbeck, "Subway tunnel guided electromagnetic wave propagation at mobile communications frequencies," *IEEE Transactions on Antennas and Propagation*, vol. 49, no. 11, pp. 1590–1596, November 2001.

[40] B. Sklar, "Rayleigh fading channels in mobile digital communication systems .I. Characterization," *IEEE Communications Magazine*, vol. 35, no. 7, pp. 90–100, July 1997.

[41] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL characterization and compensation," *IEEE Transactions on Communications*, vol. 55, no. 8, pp. 1607–1616, August 2007.

[42] P. Almers, S. Wyne, F. Tufvesson, and A. F. Molisch, "Effect of random walk phase noise on MIMO measurements," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, vol. 1, Stockholm, May 2005, pp. 141–145.

[43] D. S. Baum and H. Bolcskei, "Information-theoretic analysis of MIMO channel sounding," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7555–7577, November 2011.

[44] A. Taparugssanagorn and J. Ylitalo, "Characteristics of short-term phase noise of MIMO channel sounding and its effect on capacity estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 1, pp. 196–201, December 2008.

[45] D. S. Baum and H. Bolcskei, "Impact of phase noise on MIMO channel measurement accuracy," in *Proc. IEEE Vehicular Technology Conference (VTC Fall)*, vol. 3, Los Angeles, CA, September 2004, pp. 1614–1618.

[46] C.-H. Yih, "Analysis and compensation of DC offset in OFDM systems over frequency-selective rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3436–3446, February 2009.

[47] S. A. Bassam, S. Boumaiza, and F. M. Ghannouchi, "Block-wise estimation of and compensation for I/Q imbalance in direct-conversion transmitters," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4970–4973, June 2009.

[48] M. Inamori, A. M. Bostamam, Y. Sanada, and H. Minami, "IQ imbalance compensation scheme in the presence of frequency offset and dynamic DC offset for a direct conversion receiver," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2214–2220, May 2009.

[49] G. Bauch and J. S. Malik, "Cyclic delay diversity with bit-interleaved coded modulation in orthogonal frequency division multiple access," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2092–2100, September 2006.

[50] L. Zhang, J. Han, J. Huang, and M. Brandt-Pearce, "OFDM transmission over time-varying channel with self interference cancellation," in *Proc. IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, August 2014, pp. 743–746.

[51] J. Andrews, "Interference cancellation for cellular systems: A contemporary overview," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 19–29, April 2005.

[52] Q. H. Spencer, L. A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, February 2004.

[53] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, January 2004.

[54] A. D. Dabbagh and D. J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Transactions on Communications*, vol. 56, no. 11, pp. 1859–1868, November 2008.

[55] V. K. N. Lau, "Optimal downlink space-time scheduling design with convex utility functions–Multiple-antenna systems with orthogonal spatial multiplexing," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1322–1333, July 2005.

[56] M. Kobayashi and G. Caire, "Joint beamforming and scheduling for a multi-antenna downlink with imperfect transmitter channel knowledge," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1468–1477, September 2007.

[57] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150–154, February 2001.

[58] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. IEEE 51st Vehicular Technology Conference*, vol. 3, Tokyo, May 2000, pp. 1854–1858.

[59] O.-S. Shin and K. Bok, "Antenna-assisted round robin scheduling for MIMO cellular systems," *IEEE Communications Letters*, vol. 7, no. 3, pp. 109–111, March 2003.

[60] Z. Zhang and Y. H. . E. Chong, "Opportunistic downlink scheduling for multiuser OFDM systems," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 2, March 2005, pp. 1206–1212.

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

[62] J. Lee and N. Jindal, "Dirty paper coding vs. linear precoding for MIMO broadcast channels," in *Proc. 40th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2006, pp. 779–783.

[63] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, no. 5, pp. 138–139, March 1971.

[64] M. Miyakawa and H. Harashima, "A method of code conversion for a digital communication channel with intersymbol interference," *Trans. Inst. Electron. Commun. Eng. Japan*, vol. 52-A, pp. 272–273, June 1969.

[65] W. Shi and R. Wesel, "The effect of mismatch on decision-feedback equalization and Tomlinson-Harashima precoding," in *Proc. 32nd Asilomar Conference on Signals, Systems and Computers*, vol. 2, Pacific Grove, CA, November 1998, pp. 1743–1747.

[66] S. Shamai and R. Laroia, "The intersymbol interference channel: lower bounds on capacity and channel precoding loss," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1388–1404, September 1996.

[67] W. Yu, D. P. Varodayan, and J. M. Cioffi, "Trellis and convolutional precoding for transmitter-based interference presubtraction," *IEEE Transactions on Communications*, vol. 53, no. 7, pp. 1220–1230, July 2005.

[68] J. G. D. Forney, "Trellis shaping," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 281–300, March 1992.

[69] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1250–1276, August 2002.

[70] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 3820–3833, November 2005.

[71] R. Fischer and C. A. Windpassinger, "Improved MIMO precoding for decentralized receivers resembling concepts from lattice reduction," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 4, December 2003, pp. 1852–1856.

[72] R. Ghaffar and R. Knopp, "Linear precoders for multiuser MIMO for finite constellations and a simplified receiver structure under controlled interference," in *Proc. 43rd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2009, pp. 1431–1435.

[73] F. Boccardi, F. Tosato, and G. Caire, "Precoding schemes for the MIMO-GBC," in *Proc. Int. Zurich Seminar on Communications*, Zurich, 2006, pp. 10–13.

[74] H. Sheng and Y. Li, "Iterative nonlinear precoding and link adaptation for MU-MIMO with CSI impairments," in *Proc. IEEE International Conference on Computing, Networking and Communications (ICNC)*, Maui, HI, February 2012, pp. 1092–1097.

[75] H. Sung, S. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 57, no. 11, pp. 3489–3499, November 2009.

[76] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1711–1721, May 2007.

[77] V. Stankovic and M. Haardt, "Generalized design of multi-user MIMO precoding matrices," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 953–961, March 2008.

[78] X. Hou, Z. Zhang, and H. Kayama, "QRD-based MU-MIMO transmission scheme towards evolved LTE TDD system," in *Proc. IEEE GLOBECOM Workshops*, Miami, FL, December 2010, pp. 1–6.

[79] M. B. Shenouda and T. N. Davidson, "Tomlinson-Harashima precoding for broadcast channels with uncertainty," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1380–1389, September 2007.

[80] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, November 2006.

[81] E. Lahetkangas, K. Pajukoski, E. Tiirola, J. Hamalainen, and Z. Zheng, "On the performance of LTE-Advanced MIMO: How to set and reach beyond 4G targets," in *Proc. 18th European Wireless Conference (EW 2012)*, Poznan, Poland, April 2012, pp. 1–6.

[82] B. Clerckx, G. Kim, and S. Kim, "MU-MIMO with channel statistics-based codebooks in spatially correlated channels," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, New Orleans, LO, December 2008, pp. 1–5.

[83] T. Y. Al-Naffouri, M. Sharif, and B. Hassibi, "How much does transmit correlation affect the sum-rate scaling of MIMO gaussian broadcast channels?" *IEEE Transactions on Communications*, vol. 57, no. 2, pp. 562–572, February 2009.

[84] Z. Jiang, A. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Transactions on Wireless Communications*, vol. PP, January 2015.

[85] P.-H. Kuo, H. T. Kung, and P.-A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, April 2012, pp. 492–497.

[86] M. Joham, K. Kusume, W. Utschick, and J. A. Nossek, "Transmit matched filter and transmit wiener filter for the downlink of FDD DS-CDMA systems," in *Proc. IEEE 13th International Symposium on Personal Indoor and Mobile Radio Communications*, vol. 5, Lisbon, September 2002, pp. 2312–2316.

[87] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, January 2005.

[88] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, New Orleans, LO, December 2008, pp. 1–4.

[89] J. W. Smith, "The joint optimization of transmitted signal and receiving filter for data transmission systems," *The Bell System Technical Journal*, vol. 44, no. 10, pp. 2363–2392, December 1965.

[90] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Transactions on Communications*, vol. 42, no. 12, pp. 3221–3231, December 1994.

[91] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint tx-rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, September 2003.

[92] A. J. Tenenbaum and R. S. Adve, "Joint multiuser transmit-receive optimization using linear processing," in *Proc. IEEE International Conference on Communications (ICC)*, Paris, France, June 2004, pp. 588–592.

[93] J. Zhang, Y. Wu, S. Zhou, and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Communications Letters*, vol. 9, no. 11, pp. 991–993, November 2005.

[94] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, May 2013.

[95] G. Monghal, K. I. Pedersen, I. Z. Kovacs, and P. E. Mogensen, "QoS oriented time and frequency domain packet schedulers for the UTRAN Long Term Evolution," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Singapore, May 2008, pp. 2532–2536.

[96] A. Pokhariyal, K. I. Pedersen, G. Monghal, I. Z. Kovacs, C. Rosa, T. E. Kolding, and P. E. Mogensen, "HARQ aware frequency domain packet scheduler with different degrees of fairness for the UTRAN Long Term Evolution," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Dublin, April 2007, pp. 2761–2765.

[97] A. R. Braga, E. B. Rodrigues, and F. R. P. Cavalcanti, "Packet scheduling for VOIP over HSDPA in mixed traffic scenarios," in *Proc. IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, Helsinki, September 2006, pp. 1–5.

[98] S. Engstrom, T. Johansson, F. Kronestedt, M. Larsson, S. Lidbrink, and H. Olofsson, "Multiple reuse patterns for frequency planning in GSM networks," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Ottawa, Ont., May 1998, pp. 2004–2008.

[99] T. Novlan, J. G. Andrews, I. Sohn, R. K. Ganti, and A. Ghosh, "Comparison of fractional frequency reuse approaches in the OFDMA cellular downlink," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, Miami, FL, December 2010, pp. 1–5.

[100] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4G networks," *IEEE Communications Magazine*, vol. 47, no. 4, pp. 74–81, May 2009.

[101] J. Jiang, R. M. Buehrer, and W. H. Tranter, "Greedy scheduling performance for a zero-forcing dirty-paper coded system," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 789–793, May 2006.

[102] M. E. Aydin, R. Kwan, J. Wu, and J. Zhang, "Multiuser scheduling on the LTE downlink with simulated annealing," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Budapest, May 2011, pp. 1–5.

[103] S. Jin, M. Li, Y. Huang, Y. Du, and X. Gao, "Pilot scheduling schemes for multi-cell massive multiple-input multiple-output transmission," *IET Communications*, pp. 689–700, April 2015.

[104] R. C. Elliott and W. A. Krzymien, "Downlink scheduling via genetic algorithms for multiuser single-carrier and multicarrier MIMO systems with dirty paper coding," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3247–3262, September 2009.

[105] Y. Zhao, X. Li, Y. Li, and H. Ji, "Resource allocation for high-speed railway downlink MIMO-OFDM system using quantum-behaved particle swarm optimization," in *Proc. IEEE International Conference on Communications (ICC)*, Budapest, June 2013, pp. 2343–2347.

[106] H. Ghazzai, E. Yaacoub, and M.-S. Alouini, "Optimized LTE cell planning for multiple user density subareas using meta-heuristic algorithms," in *Proc. IEEE Vehicular Technology Conference (VTC Fall)*, Vancouver, BC, September 2014, pp. 1–6.

[107] H. B. Zhang and X. X. Wang, "Resource allocation for downlink OFDM system using noisy chaotic neural network," *IET Electronics Letters*, pp. 1201–1202, October 2011.

[108] K. Das and S. D. Morgera, "Adaptive interference cancellation for DS-CDMA systems using neural network techniques," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 9, pp. 1774–1784, August 2002.

[109] M.-G. Garcia, J. L. Rojo-Alvarez, F. A. Atienza, and M. Martinez-Ramon, "Support vector machines for robust channel estimation in OFDM," *IEEE Signal Processing Letters*, vol. 13, no. 7, pp. 397–400, July 2006.

[110] S. Sigdel and W. A. Krzymien, "Simplified fair scheduling and antenna selection algorithms for multiuser MIMO orthogonal space-division multiplexing downlink," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1329–1344, March 2009.

[111] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3658–3663, September 2006.

[112] J. H. Holland, *Adaptation in Natural and Artificial Systems*, 1st ed. Ann Arbor: MI: Univ. of Michigan Press, 1975.

[113] E. Conte, S. Tomasin, and N. Benvenuto, "A simplified greedy algorithm for joint scheduling and beamforming in multiuser MIMO OFDM," *IEEE Communications Letters*, vol. 14, no. 5, pp. 381–383, May 2010.

[114] F. Sun, M. You, J. Liu, Z. Shi, P. Wen, and J. Liu, "Genetic algorithm based multiuser scheduling for single- and multi-cell systems with successive interference cancellation," in *Proc. IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications*, Istanbul, September 2010, pp. 1230–1235.

[115] P. Kyösti, J. Meinilä, L. Hentilä, X. Zhao, T. Jämsä, C. Schneider, M. Narandzić, M. Milojević, A. Hong, J. Ylitalo, V. M. Holappa, M. Alatossava, R. Bultitude, Y. de Jong, and T. Rautiainen. (2007, September) WINNER II channel models, ver. 1.1. [Online]. Available: https://www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf

[116] X. She, H. Taoka, J. Zhu, and L. Chen, "Investigation of control signaling and reference signal design for downlink MU-MIMO in LTE-Advanced," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Budapest, May 2011, pp. 1–6.

[117] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, "A constrained factor decomposition with application to MIMO antenna systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2429–2442, June 2008.

[118] L. Li and D. Boulware, "High-order tensor decomposition for large-scale data analysis," in *Proc. IEEE International Congress on Big Data (BigData Congress)*, New York, NY, June 2015, pp. 665–668.

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

[120] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, November 1997.

[121] Y. Zou, O. Raeesi, R. Wichman, A. Tolli, and M. Valkama, "Analysis of channel non-reciprocity due to transceiver and antenna coupling mismatches in TDD precoded multi-user MIMO-OFDM downlink," in *Proc. IEEE Vehicular Technology Conference (VTC Fall)*, Vancouver, BC, September 2014, pp. 1–7.

[122] O. Raeesi, Y. Zou, A. Tolli, and M. Valkama, "Closed-form analysis of channel non-reciprocity due to transceiver and antenna coupling mismatches in multi-user massive MIMO network," in *Proc. IEEE GLOBECOM Workshops*, Austin, TX, December 2014, pp. 333–339.

[123] A. Bourdoux, B. Come, and N. Khaled, "Non-reciprocal transceivers in OFDM/SDMA systems: Impact and mitigation," in *Proc. IEEE Radio and Wireless Conference (RAWCON)*, August 2003, pp. 183–186.

[124] H. Xiaolin, A. Harada, and H. Suda, "Experimental study of advanced MU-MIMO scheme with antenna calibration for the evolving LTE TDD system," in *Proc. IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications*, Sydney, September 2012, pp. 2443–2448.

[125] M. Guillaud, D. T. M. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration," in *Proc. Eighth International Symposium on Signal Processing and Its Applications*, vol. 1, August 2005, pp. 403–406.

[126] M. Petermann, M. Stefer, F. Ludwig, D. Wubben, M. Schneider, S. Paul, and K.-D. Kammeyer, "Multi-user pre-processing in multi-antenna OFDM TDD systems with non-reciprocal transceivers," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3781–3793, August 2013.

[127] F. Kaltenberger and M. Guillaud, "Exploitation of reciprocity in measured MIMO channels," in *Proc. COST 2100, 9th Management Committee Meeting*, Vienna, Austria, September 2009.

[128] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proc.18th Annual International Conference on Mobile Computing and Networking (Mobicom)*, Istanbul, August 2012, pp. 53–64.

[129] Y. Song, X. Yun, S. Nagata, and L. Chen, "Investigation on elevation beamforming for future LTE-Advanced," in *Proc. IEEE International Conference on Communications Workshops (ICC)*, Budapest, June 2013, pp. 106–110.

[130] Y.-H. Nam, Y. Li, and J. C. Zhang, "3D channel models for elevation beamforming and FD-MIMO in LTE-A and 5G," in *Proc. 48th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2014, pp. 805–809.

[131] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive tracking with 1000-element arrays: A framework for multi-Gbps mm wave cellular downlinks," in *Proc. 50th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, October 2012, pp. 690–697.

[132] M. Petermann, F. Ludwig, D. Wübben, S. Paul, and K.-D. Kammeyer, "Effects of downlink channel quantization on the performance of relative calibration in OFDM systems," in *Proc. 16th International OFDM-Workshop (InOWo 11)*, Hamburg, Germany, August 2011.

[133] M. Stefer, M. Petermann, M. Schneider, D. Wübben, and K.-D. Kammeyer, "Influence of non-reciprocal transceivers at 2.4 GHz in adaptive MIMO-OFDM systems," in *Proc. 14th International OFDM-Workshop 2009 (InOWo 09)*, Hamburg, Germany, September 2009.

[134] M. Guillaud and F. Kaltenberger, "Towards practical channel reciprocity exploitation: Relative calibration in the presence of frequency offset," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, April 2013, pp. 2525–2530.

[135] Z. Xiaofei, X. Dazhuan, and Y. Bei, "A novel adaptive channel estimation method in WCDMA system based on weighed least squares," in *Proc. 7th International Conference on Signal Processing (ICSP)*, vol. 2, September 2004, pp. 1699–1702.

[136] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, August 2011.

[137] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, February 2013.

[138] R. R. Muller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 773–786, March 2014.

[139] R. R. Muller, M. Vehkaperä, and L. Cottatellucci, "Analysis of blind pilot decontamination," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2013, pp. 1016–1020.

[140] V.-D. Nguyen, S. Dinh-Van, Y. Shin, W. C. Lee, and O.-S. Shin, "The problem of pilot contamination reduction in massive MIMO for physical channel models," in *Proc. IEEE 6th International Conference on Ubiquitous and Future Networks (ICUFN)*, Shanghai, July 2014, pp. 378–382.

[141] F. Xu, Y. Xiao, and D. Wang, "Adaptive semi-blind channel estimation for massive MIMO systems," in *Proc. 12th International Conference on Signal Processing (ICSP)*, Hangzhou, October 2014, pp. 1698–1702.

[142] J. P. Delmas and Y. Meurisse, "On the second-order statistics of the EVD of sample covariance matrices—Application to the detection of noncircular or/and non-Gaussian components," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 4017–4023, August 2011.

[143] A.-J. van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 173–190, January 1997.

[144] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1699–1712, July 2002.

[145] A. J. van der Veen, S. Talwar, and A. Paulraj, "Blind estimation of multiple digital signals transmitted over FIR channels," *IEEE Signal Processing Letters*, vol. 16, no. 5, pp. 99–102, May 1995.

[146] Z. Xu, "Asymptotic performance of subspace methods for synchronous multirate CDMA systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 2015–2026, August 2002.

[147] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous cochannel digital signals using an antenna array–Part I: Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1184–1197, May 1996.

[148] C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1539–1544, March 2010.

[149] S. Visuri and V. Koivunen, "Resolving ambiguities in subspace-based blind receiver for MIMO channels," in *Proc. 36th Asilomar Conference on Signals, Systems and Computers*, vol. 1, Pacific Grove, CA, November 2002, pp. 589–593.

[150] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5113–5129, November 2008.

[151] W. A. Gardner, "A new method of channel identification," *IEEE Transactions on Communications*, vol. 39, no. 6, pp. 813–817, June 1991.

[152] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 340–349, March 1994.

[153] D. Darsena, G. Gelli, L. Paura, and F. Verde, "Widely linear equalization and blind channel identification for interference-contaminated multicarrier systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 1163–1177, March 2005.

[154] S. Abdallah and I. N. Psaromiligkos, "Widely linear versus conventional subspace-based estimation of SIMO flat-fading channels: Mean squared error analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1307–1318, March 2012.

[155] K. Guo, Y. Guo, and G. Ascheid, "On the performance of EVD-based channel estimations in MU-Massive-MIMO systems," in *Proc. IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications*, London, September 2013, pp. 1376–1380.

[156] J. P. Delmas, Y. Meurisse, and P. Comon, "Performance limits of alphabet diversities for FIR SISO channel identification," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 73–82, January 2009.

[157] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, March 2012, pp. 3249–3252.

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

[159] 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 Processing Magazine*, vol. 30, no. 1, pp. 40–60, January 2013.

[160] H. Wang, Z. Pan, J. Ni, S. Wang, and C.-L. I, "A temporal domain based method against pilot contamination for multi-cell massive MIMO systems," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Seoul, May 2014, pp. 1–5.

[161] H. C. Papadopoulos, G. Caire, and S. A. Ramprashad, "Achieving large spectral efficiencies from MU-MIMO with tens of antennas: Location-adaptive TDD MU-MIMO design and user scheduling," in *Proc. 44th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2010, pp. 636–643.

[162] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving "Massive MIMO" Spectral Efficiency with a Not-so-Large Number of Antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3226–3239, September 2012.

[163] J. Hoydis, S. ten Brink, and M. Debbah, "Comparison of linear precoding schemes for downlink massive MIMO," in *Proc. IEEE International Conference on Communications (ICC)*, Ottawa, ON, June 2012, pp. 2135–2139.

[164] J. Zhang, X. Yuan, and L. Ping, "Hermitian Precoding for Distributed MIMO Systems with Individual Channel State Information," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 241–250, January 2013.

[165] X. Rao, V. K. N. Lau, and X. Kong, "CSIT estimation and feedback for FDD multi-user massive MIMO systems," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, May 2014, pp. 3157–3161.

[166] H. Yin, L. Cottatellucci, and D. Gesbert, "Enabling massive MIMO systems in the FDD mode thanks to D2D communications," in *Proc. 48th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 2014.

[167] A. Adhikary, E. A. Safadi, M. K. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, "Joint spatial division and multiplexing for mm-wave channels," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1239–1255, May 2014.

[168] S. H. Won, S. C. Chae, S. Y. Cho, I. Kim, and S. C. Bang, "Massive MIMO test-bed design for next-generation long term evolution (LTE) mobile systems in the frequency division duplex (FDD) mode," in *Proc. IEEE International Conference on Information and Communication Technology Convergence (ICTC)*, Busan, October 2014, pp. 841–844.

[169] S. Bazzi, G. Dietl, and W. Utschick, "Subspace precoding with limited feedback for the massive MIMO interference channel," in *Proc. IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, A Coruna, June 2014, pp. 277–280.

[170] J. Chen and V. K. N. Lau, "Two-tier precoding for FDD multi-cell massive MIMO time-varying interference networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1230–1238, June 2014.

[171] T. S. Rappaport, S. Y. Seidel, and K. Takamizawa, "Statistical channel impulse response models for factory and open plan building radio communicate system design," *IEEE Transactions on Communications*, vol. 39, no. 5, pp. 794–807, May 1991.

[172] R. H. Etkin and D. N. C. Tse, "Degrees of freedom in some underspread MIMO fading channels," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1576–1608, April 2006.

[173] W. C. Jakes and D. C. Cox, Eds., *Microwave mobile communication*, 2nd ed. Wiley-IEEE Press, 1994.

[174] X. Dai, W. Zhang, J. Xu, J. E. Mitchell, and Y. Yang, "Kalman interpolation filter for channel estimation of LTE downlink in high-mobility environments," *EURASIP Journal on Wireless Communications and Networking*, no. 1, pp. 1–14, 2012.

[175] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[176] K. Huber and S. Haykin, "Improved bayesian MIMO channel tracking for wireless communications: Incorporating a dynamical model," *IEEE Transactions on Wireless Communications*, vol. 5, no. 9, pp. 2458–2466, September 2006.

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

[178] D. J. Love, J. Choi, and P. Bidigare, "A closed-loop training approach for massive MIMO beamforming systems," in *Proc. 47th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, March 2013, pp. 1–5.

[179] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 787–801, May 2014.

[180] K. Huang, R. W. Heath, and J. G. Andrews, "Limited feedback beamforming over temporally-correlated channels," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1959–1975, February 2009.

[181] T. Kim, D. J. Love, and B. Clerckx, "MIMO systems with limited rate differential feedback in slowly varying channels," *IEEE Transactions on Communications*, vol. 59, no. 4, pp. 1175–1189, March 2011.

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

[183] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 637–650, August 2002.

[184] J. Nam, J.-Y. Ahn, A. Adhikary, and G. Caire, "Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information," in *Proc. IEEE 46th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, March 2012, pp. 1–6.

[185] H. Noh, Y. Kim, J. Lee, and C. Lee, "Codebook design of generalized space shift keying for FDD massive MIMO systems in spatially correlated channels," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 2, pp. 513–523, February 2015.

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

[187] D. shan Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, March 2000.

[188] H. Wu, C. Qiao, S. De, and O. Tonguz, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 2, pp. 222–232, February 2001.

[189] B. Lee, J. Choi, J.-Y. Seol, D. J. Love, and B. Shim, "Antenna grouping based feedback reduction for FDD-based massive MIMO systems," in *Proc. IEEE International Conference on Communications (ICC)*, Sydney, NSW, June 2014, pp. 4477–4482.

[190] N. Czink, X. Yin, H. Ozcelik, M. Herdin, E. Bonek, and B. H. Fleury, "Cluster characteristics in a MIMO indoor propagation environment," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1465–1475, April 2007.

[191] L. Vuokko, V.-M. Kolmonen, J. Salo, and P. Vainikainen, "Measurement of large-scale cluster power characteristics for geometric channel models," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 11, pp. 3361–3365, November 2007.

[192] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[193] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. D. Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Communications*, vol. 19, no. 6, pp. 92–99, December 2012.

[194] 3GPP TR 25.996 v12.0.0, "Spatial channel model for multiple input multiple output (MIMO) simulations," 3GPP, Tech. Rep., September 2014.

[195] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, July 2007.

[196] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

[197] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, June 2014.

[198] G. Bottomley, T. Ottosson, and Y.-P. Wang, "A generalized RAKE receiver for interference suppression," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 8, pp. 1536–1545, August 2000.

[199] G. Xu and H. Liu, "An effective transmission beamforming scheme for frequency-division-duplex digital wireless communication systems," in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, Detroit, MI, USA, May 1995, pp. 1729–1732.

[200] T. G. Manickam, R. J. Vaccaro, and D. W. Tufts, "A least-squares algorithm for multipath time-delay estimation," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3229–3233, November 1994.

[201] J.-T. Chen, A. Paulraj, and U. Reddy, "Multichannel maximum-likelihood sequence estimation (MLSE) equalizer for GSM using a parametric channel model," *IEEE Transactions on Communications*, vol. 47, no. 1, pp. 53–63, January 1999.

[202] M. Z. Win and R. A. Scholtz, "Energy capture vs. correlator resources in ultra-wide bandwidth indoor wireless communications channels," in *Proc. MILCOM 97*, vol. 3, Monterey, CA, November 1997, pp. 1277–1281.

[203] P. Kyösti and *et al.* (2010, June) WINNER+ final channel models, ver. 1.0. [Online]. Available: http://projects.celtic-initiative.org/winner+/WINNER+%20Deliverables/D5.3_v1.0.pdf

[204] C. K. Au-Yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 458–462, February 2007.

[205] M. Fallgren and B. Timus. (2013, April) Scenarios, requirements and KPIs for 5G mobile and wireless system. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D1.1_v1.pdf

[206] Z. Ding, I. Krikidis, B. Rong, J. S. Thompson, C. Wang, and S. Yang, "On combating the half-duplex constraint in modern cooperative networks: protocols and techniques," *IEEE Wireless Communications*, vol. 19, no. 6, pp. 20–27, December 2012.

[207] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2498–2511, May 2013.

[208] R. Zhang, C. C. Chai, and Y. C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 726–737, February 2009.

[209] C.-B. Chae, T. Tang, R. W. Heath, and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 727–738, February 2008.

[210] P. Rost, G. Fettweis, and J. N. Laneman, "Energy- and cost-efficient mobile communication using multi-cell MIMO and relaying," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3377–3387, July 2012.

[211] K. Doppler, C. Wijting, and K. Valkealahti, "On the benefits of relays in a metropolitan area network," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Singapore, May 2008, pp. 2301–2305.

[212] M. Shaqfeh and H. Alnuweiri, "Joint power and resource allocation forblock-fading relay-assisted broadcast channels," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1904–1913, April 2011.

[213] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multi-pair full-duplex relaying with massive arrays and linear processing," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1721–1737, June 2014.

[214] H. Cui, L. Song, and B. Jiao, "Multi-pair two-way amplify-and-forward relaying with very large number of relay antennas," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2636–2645, April 2014.

[215] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Stable throughput of cognitive radios with and without relaying capability," *IEEE Transactions on Communications*, vol. 55, no. 12, pp. 2351–2360, December 2007.

[216] C. Hoymann, W. Chen, J. Montojo, A. Golitschek, C. Koutsimanis, and X. Shen, "Relaying operation in 3GPP LTE: challenges and solutions," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 156–162, February 2012.

[217] A. Mohammadi and F. M. Ghannouchi, *RF Transceiver Design for MIMO Wireless Communications*, 1st ed. New York: Springer, 2012.

[218] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.

[219] R. Pabst, B. H. Walke, D. C. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Communications Magazine*, vol. 42, no. 9, pp. 80–89, September 2004.

[220] D. Soldani and S. Dixit, "Wireless relays for broadband access," *IEEE Communications Magazine*, vol. 46, no. 3, pp. 58–66, March 2008.

[221] S. Berger, M. Kuhn, A. Wittneben, T. Unger, and A. Klein, "Recent advances in amplify-and-forward two-hop relaying," *IEEE Communications Magazine*, vol. 47, no. 7, pp. 50–56, July 2009.

[222] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 913–927, January 2014.

[223] W. Yang, L. Li, G. Wu, H. Wang, and Y. Wang, "Joint uplink and downlink relay selection in cooperative cellular networks," in *Proc. IEEE Vehicular Technology Conference (VTC Fall)*, Ottawa, ON, September 2010, pp. 1–5.

[224] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4381–4393, December 2012.

[225] K. Yang, H. Cui, L. Song, and Y. Li, "Joint relay and antenna selection for full-duplex AF relay networks," in *Proc. IEEE International Conference on Communications (ICC)*, Sydney, NSW, June 2014, pp. 4454–4459.

[226] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, H. V. Balan, and G. Dejean, "Rethinking indoor wireless: Low power, low frequency, full-duplex," Microsoft Research, Tech. Rep., July 2009. [Online]. Available: http://research.microsoft.com/apps/pubs/default.aspx?id=104950

[227] A. Raghavan, E. Gebara, E. M. Tentzeris, and J. Laskar, "Analysis and design of an interference canceller for collocated radios," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 11, pp. 3498–3508, November 2005.

[228] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4494–4510, November 2013.

[229] D. W. Bliss, T. M. Hancock, and P. Schniter, "Hardware phenomenological effects on cochannel full-duplex MIMO relay performance," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2012, pp. 34–39.

[230] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *Proc. 45th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2011, pp. 2002–2006.

[231] D. Choi and D. Park, "Effective self interference cancellation in full duplex relay systems," *Electronics Letters*, vol. 48, no. 2, pp. 129–130, January 2012.

[232] B. Chun and H. Park, "A spatial-domain joint-nulling method of self-interference in full-duplex relays," *IEEE Communications Letters*, vol. 16, no. 4, pp. 436–438, April 2012.

[233] B. Chun, E. R. Jeong, J. Joung, Y. Oh, and Y. Lee, "Pre-nulling for self-interference suppression in full-duplex relays," in *Proc. APSIPA 2009 Annual Summit and Conference*, Sapporo, Japan, October 2009, pp. 91–97.

[234] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 3933–3938, August 2009.

[235] J. Sangiamwong, T. Asai, J. Hagiwara, Y. Okumura, and T. Ohya, "Joint multi-filter design for full-duplex MU-MIMO relaying," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Barcelona, April 2009, pp. 1–5.

[236] P. Lioliou, M. Viberg, M. Coldrey, and F. Athley, "Self-interference suppression in full-duplex MIMO relays," in *Proc. 44th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2010, pp. 658–662.

[237] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1541–1553, September 2012.

[238] S. Nagata, Y. Yan, X. Gao, A. Li, H. Kayama, T. Abe, and T. Nakamura, "Investigation on system performance of L1/L3 relays in LTE-Advanced downlink," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Yokohama, May 2011, pp. 1–5.

[239] 3GPP TS 36.216 v12.0.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer for relaying operation (Release 12)," 3GPP, Tech. Rep., September 2014.

[240] 3GPP TR 36.836 v12.0.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Study on mobile relay (Release 12)," 3GPP, Tech. Rep., June 2014.

[241] 3GPP TS 36.201 v12.2.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE physical layer; General description (Release 12)," 3GPP, Tech. Rep., March 2015.

[242] K. Shibuya, "Broadcast-wave relay technology for digital terrestrial television broadcasting," *Proceedings of the IEEE*, vol. 94, no. 1, pp. 269–273, January 2006.

Aalto-DD 156/2016

ISBN 978-952-60-6953-1 (printed)
ISBN 978-952-60-6952-4 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics
www.aalto.fi

BUSINESS + ECONOMY

ART + DESIGN + ARCHITECTURE

SCIENCE + TECHNOLOGY

CROSSOVER

DOCTORAL DISSERTATIONS