Vu Thuy Dan Nguyen

TRANSMISSION STRATEGIES FOR FULL-DUPEX MULTIUSER MIMO COMMUNICATIONS SYSTEMS
VU THUY DAN NGUYEN

TRANSMISSION STRATEGIES FOR FULL-DUPELEX MULTIUSER MIMO COMMUNICATIONS SYSTEMS

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Technology and Natural Sciences of the University of Oulu for public defence in the OP auditorium (L10), Linnanmaa, on 1 April 2016, at 12 noon

UNIVERSITY OF OULU, OULU 2016
Nguyen, Vu Thuy Dan, Transmission strategies for full-duplex multiuser MIMO communications systems.
University of Oulu Graduate School; University of Oulu, Faculty of Information Technology and Electrical Engineering
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Abstract

This thesis considers data transmission in a full-duplex (FD) multiuser multiple-input multiple-output (MU-MIMO) system, where a FD capable base station (BS) bidirectionally communicates with multiple half-duplex (HD) users in downlink (DL) and uplink (UL) channels using the same radio resources. The main challenge in FD communications is how to deal with the self-interference (SI) between transmit and receive antennas at the BS. The work carried out in the thesis is motivated by recent advanced techniques in hardware design demonstrating that the SI can be suppressed to a degree that possibly allows for FD transmission in cellular networks. In particular, this thesis attempts to explore the potential gains in terms of the spectral efficiency (SE) and energy efficiency (EE) that can be brought by the FD MU-MIMO model. As the first of its kinds, the thesis aims to present a solid mathematical framework and report interesting results that foster research on wireless communications in general and FD communications in particular.

For the FD system of interest the major challenge is due to the SI and co-channel interference from users in the UL channel to the ones in the DL channel, resulting in the coupling between the two channels. As a result we are concerned with the problem of joint transmit processing design to maximize the SE and EE subject to certain power constraints. Since the design problems are naturally non-convex, it is difficult to find the globally optimal solutions or even when possible it is not practically appealing. Our contributions to solving these design problems are on the development of several iterative algorithms that can obtain locally optimal solutions. The proposed algorithms are built upon a framework of local optimization strategies such as the sequential parametric convex approximation and the Frank-Wolfe methods. In special cases closed-form designs are also presented.

The reported results show that when the SI is sufficiently suppressed the considered FD MU-MIMO system with the proposed SE designs achieves a significantly better SE but consumes more energy, compared to the HD counterpart. In terms of EE the proposed EE scheme is superior to the proposed SE design. Moreover, in the low transmit power region, the EE design achieves a worse EE than the HD system but a better one in the high transmit power regime when the SI power is low.

Keywords: D.C. program, energy efficiency, full-duplex, linear precoding, MIMO, multiuser transmission, self-interference, semidefinite programming, spectral efficiency, transmit beamforming
Nguyen, Vu Thuy Dan, Lähetyssstrategioita samanaikaisesti kaksisuuntaa suosintoja käyttäen moniantennin yhden
Oulun yliopiston tutkijakoulu; Oulun yliopisto, Tieto- ja sähkötekniikan tiedekunta
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Tiivistelmä
Tämä väitöskirja käsittelee datansiirtoa samanaikaisesti kaksisuuntaisessa (full-duplex, FD) use-
an käyttäjän moniantennijärjestelmässä (MU-MIMO), jossa FD-kykyinen tukiasema on yhtä
aikaa vuorosuuntaisten (half-duplex, HD) käyttäjien kanssa laskevalla (DL) ja nouse-
valalla (UL) siirtotietä käyttäen samoja radioressursseja. FD-kommunikation suurin haaste liittyy
lähetyss- ja vastaanottoantennien välisen omahäiriön (SI) hallintaan. Tässä työssä hyödynnetään
suurten tulosten selvityä tutkimustuloksia, joissa edistyneillä häiriönvaimennustekniikoilla on kyetty vaihto-
maalainen omahäiriö tasolle, jolla FD-lähetyys solukoverkoissa on toteutuskelpoista. Tässä työssä
suoritetaan erilaisia HD-MU-MIMO -järjestelmän tuomiaспектrockäytön tehokkuudesta (SE) ja energiate-
hokkuudesta (EE). Väitöskirjalla on uutuusar-
voa matemaattisessa suorituskykyyyrin ja työn
mielenkiinto iset tulokset edistävät jatko-
utkimustia aiheen ympärillä.

Tutkittavan FD-järjestelmän merkittävänä haasteena on omahäiriön ja muiden käyttäjien siir-
tosuuntien välisen samankanan häiriön yhteisvaikutus, jonka johdosta siirtosuunnat kykenevät
vaihtoutua toisiin. Tämä johtaa lähetysprosessoinnin yhteisoptimoimiseen, jossa spektri- ja energi-
hokkuus pyritään maksimoimaan määritellyillä tehoroajoituksilla. Nämä suunnitteluongelmia
vaihtavat usealta aloilta arvioituja, joilla saavutetaan paikallisesti optimoimiala, kuin
suurten tulostentoimintataparannuksien ja Frank-Wolfe –menetelmiä. Erityistapauks-
issa suljetun luokan ratkaisut on myös esitetty.

Raportoidut tulokset osoittavat, että omahäiriön ollessa riitävästi vaimennettu mallinnettu
järjestelmällä saavutetaan spektrinkäytön optimointimelessä luomattavaa etta HD-verrokin
lisääntyneen energian kulutukseen kustannuksella. Energiatehokkuuden optimointin ja
sytyttää strategiaan puolestaan päästään suurempia teknologiankulutuksiin. Pienillä lähetysteholilla
energiatehokkuus voi kuitenkin olla HD-järjestelmää alempi, mutta vastaavasti suurten lähetysteho-
jen alueella tilanne on päinvastainen kun omahäiriön teho on tarpeeksi alhainen.

Asiasanat: D.C. ohjelma, energiatehokkuus, lineaarinen esikoodaus, lähetyskänitysmuotoilu, MIMO, omahäiriö, semidefiniittii ohjelmointi, spektritehokkuus, täysdupleksi, usean käyttäjän lähetyss
To my family
Acknowledgements

The results of research in this thesis have been carried out at the Department of Communications Engineering, University of Oulu, Finland, during the years 2011-2016. I would like to give my many thanks to all the people in the Department who have inspired me a lot.

First and foremost I would like to express my deepest gratitude to my supervisor Prof. Matti Latva-aho. He provided me the opportunity to do my PhD degree at the RAT group and the freedom to develop my research interests as well as choose on my own research topic. Without his continuous support, guidance and encouragement, my PhD thesis could not have been completed.

My special thanks are also given to Dr. Le-Nam Tran for his patience, motivation and immense knowledge. He not only is my husband but has been my dedicated advisor on research. His guidance helped me in all the time of research and writing of this thesis. Additionally, my gratitude is extended to Dr. Pekka Pirinen, who has also been my great advisor, for inviting me to join the RAT group, giving me valuable comments and fully supporting me anytime I need.

I would also like to thank my follow-up group, Prof. Markku Juntti, Dr. Pradeep Kumar and Dr. Vinh Van Phan for their insightful advices and discussions during my doctoral program. I also extend my thanks to the pre-examiners, Prof. Robert Schober and Dr. Himal A. Suraweera. Their valuable comments and suggestions have greatly improved the quality of my thesis.

I am also very grateful to the whole administrative staff from the Department, specifically Elina Komminaho, Hanna Saarela, Jari Sillanpää, Kirsi Ojutkangas, Eija Pajunen and many other. Besides, my thanks go to Infotech Graduate School for financial support of this thesis.

My sincere thanks also goes to Kien Giang Nguyen and Hirley Alves who have helped me do practical procedures related to the thesis defense, especially during the time I was not physically present in the Department. During my doctoral program, I have been very lucky to make friend with friendly, cheerful as well as helpful colleagues at the Department, especially Trung Kien Vu, Kien Giang Nguyen, Ayotunde Laiyemo, Petri Luoto, Simon Scott, Sumudu Samarakoon, Inosha Sugathapala, Manosha Kapuruhamy, Tachporn Sanguanpuak, and many others. We had many joyful social gatherings as
well as shared many unforgettable memories during my long journey in finishing this
doctoral program. During my stay in Oulu, I got many helps and precious advices from
Mrs. Phuong and Mr. Bao. Thus, I would like to give special thanks to them.

Last but not the least, I would like to thank my family: my parents, my parents-in-law
and my sister for supporting me spiritually throughout my PhD program and my life in
general.

Vu Thuy Dan Nguyen
Oulu, April 1, 2016
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>3GPP</td>
<td>3rd Generation Partnership Project</td>
</tr>
<tr>
<td>ADC</td>
<td>Analog-to-digital converter</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive white Gaussian noise</td>
</tr>
<tr>
<td>BCA</td>
<td>Block coordinate ascent</td>
</tr>
<tr>
<td>BD</td>
<td>Block diagonalization</td>
</tr>
<tr>
<td>BS</td>
<td>Base station</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative distribution function</td>
</tr>
<tr>
<td>CCI</td>
<td>Co-channel interference</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel state information</td>
</tr>
<tr>
<td>CW</td>
<td>Continuous wave</td>
</tr>
<tr>
<td>D</td>
<td>Destination</td>
</tr>
<tr>
<td>D.C</td>
<td>Difference of convex</td>
</tr>
<tr>
<td>DAC</td>
<td>Digital-to-analog converter</td>
</tr>
<tr>
<td>DPC</td>
<td>Dirty paper coding</td>
</tr>
<tr>
<td>DL</td>
<td>Downlink</td>
</tr>
<tr>
<td>EE</td>
<td>Energy efficiency</td>
</tr>
<tr>
<td>FD</td>
<td>Full-duplex</td>
</tr>
<tr>
<td>FDD</td>
<td>Frequency division duplex</td>
</tr>
<tr>
<td>HD</td>
<td>Half-duplex</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush Kuhn Tucker</td>
</tr>
<tr>
<td>LB</td>
<td>Lower bound</td>
</tr>
<tr>
<td>LNA</td>
<td>Low-noise amplifier</td>
</tr>
<tr>
<td>LOS</td>
<td>Line-of-sight</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium access control</td>
</tr>
<tr>
<td>MAXDET</td>
<td>Determinant maximization</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum mean square error</td>
</tr>
<tr>
<td>MMSE-SIC</td>
<td>Minimum mean square error and successive interference cancellation</td>
</tr>
<tr>
<td>MU-MIMO</td>
<td>Multiuser multiple-input multiple-output</td>
</tr>
<tr>
<td>MUI</td>
<td>Multiuser interference</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-input multiple-output</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MISO</td>
<td>Multiple-input single-output</td>
</tr>
<tr>
<td>NLOS</td>
<td>Non line-of-sight</td>
</tr>
<tr>
<td>NP-hard</td>
<td>Non-deterministic polynomial-time hard</td>
</tr>
<tr>
<td>P</td>
<td>Problem</td>
</tr>
<tr>
<td>P2P</td>
<td>Point to point</td>
</tr>
<tr>
<td>PA</td>
<td>Power amplifier</td>
</tr>
<tr>
<td>PUPCs</td>
<td>Per user power constraints</td>
</tr>
<tr>
<td>RF</td>
<td>Radio frequency</td>
</tr>
<tr>
<td>S</td>
<td>Source</td>
</tr>
<tr>
<td>SCA</td>
<td>Successive convex approximation</td>
</tr>
<tr>
<td>SDP</td>
<td>Semidefinite program</td>
</tr>
<tr>
<td>SE</td>
<td>Spectral efficiency</td>
</tr>
<tr>
<td>SI</td>
<td>Self-interference</td>
</tr>
<tr>
<td>SPC</td>
<td>Sum power constraint</td>
</tr>
<tr>
<td>SPCA</td>
<td>Sequential parametric convex approximation</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal to interference plus noise ratio</td>
</tr>
<tr>
<td>SOC</td>
<td>Second order cone</td>
</tr>
<tr>
<td>SU-MIMO</td>
<td>Single user multiple-input multiple-output</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular value decomposition</td>
</tr>
<tr>
<td>TDD</td>
<td>Time division duplex</td>
</tr>
<tr>
<td>UL</td>
<td>Uplink</td>
</tr>
<tr>
<td>WiMAX</td>
<td>Worldwide Interoperability for Microwave Access</td>
</tr>
<tr>
<td>ZF</td>
<td>Zero forcing</td>
</tr>
</tbody>
</table>
List of symbols

\(d\) Distance (in kilometers) between the BS and a specific user
\(d_{\text{CCI}}\) Distance (in kilometers) from a user in the uplink transmission to another user in the downlink direction
\(\sigma_{\text{SI}}^2\) Self-interference strength
\(\varepsilon\) Power amplifier efficiency
\(\lambda\) Dual variable
\(\mathbf{I}_N\) Identity matrix with size \(N\)
\(P_{\text{BS}}\) Maximum transmit power of the base station
\(q_{U_j}\) Power allocation or the maximum transmit power of user \(U_j\)
\(D_i\) The \(i^{th}\) user in the downlink channel
\(U_i\) The \(i^{th}\) user in the uplink channel
\(N\) A number of antennas at the base station
\(N_T\) A number of transmit antennas at the full-duplex base station
\(N_R\) A number of receive antennas at the full-duplex base station
\(N_{D_i}\) A number of antennas at user \(D_i\)
\(N_{U_i}\) A number of antennas at user \(U_i\)
\(K_D\) A number of users in the downlink channel
\(K_U\) A number of users in the uplink channel
\(\bar{q}_{U_j}\) Power constraint at user \(U_j\)
\(\mathbf{Q}_{U_i}\) Input covariance matrix of the vector of data symbols of user \(U_i\)
\(\sigma_{n_{D_i}}^2\) Variance of \(n_{D_i}\) or each element of \(n_{D_i}/n_0\)
\(P_{\text{BS}}^{\text{Tx}}\) Transmit power determined by linear precoders
\(P_{\text{BS}}^{\text{cir}}\) Circuit power at the base station
\(P_{\text{BS}}^{\text{dyn}}\) Dynamic circuit power at the base station
\(P_{\text{BS}}^{\text{sta}}\) Static circuit power at the base station
\(q_{U_i}^{\text{Tx}}\) Transmit power at user \(U_i\)
\(q_{U_i}^{\text{cir}}\) Circuit power at user \(U_i\)
\(q_{U_i}^{\text{dyn}}\) Dynamic circuit power at user \(U_i\)
\(q_{U_i}^{\text{sta}}\) Static circuit power at user \(U_i\)
\(P_{\text{cir}}^{\text{sum}}\) Overall circuit power of a full-duplex system
\(N_0\) Single-sided noise spectral density
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{D_i}$</td>
<td>Transmitted data symbol for user $D_i$</td>
</tr>
<tr>
<td>$s_{U_i}$</td>
<td>Transmitted data symbol for user $U_i$</td>
</tr>
<tr>
<td>$s_{D_i}$</td>
<td>The vector of transmitted signal of user $D_i$</td>
</tr>
<tr>
<td>$s_{U_i}$</td>
<td>The vector of transmitted signal of user $U_i$</td>
</tr>
<tr>
<td>$w_{D_i}$</td>
<td>The linear beamforming vector of user $D_i$</td>
</tr>
<tr>
<td>$W_{D_i}$</td>
<td>The linear precoder of user $D_i$</td>
</tr>
<tr>
<td>$h_{D_i}$</td>
<td>The $N \times 1$ complex channel vector from the base station to user $D_i$</td>
</tr>
<tr>
<td>$h_{U_i}$</td>
<td>The $N \times 1$ complex channel vector from the base station to user $U_i$</td>
</tr>
<tr>
<td>$H_{D_i}$</td>
<td>The channel matrix from user $D_i$ to the base station</td>
</tr>
<tr>
<td>$H_{U_i}$</td>
<td>The channel matrix from the base station to user $D_i$</td>
</tr>
<tr>
<td>$G_{SI}$</td>
<td>The self-interference channel from the transmit antennas to the receive ones at the base station</td>
</tr>
<tr>
<td>$g_{ji}$</td>
<td>Complex channel coefficient from $U_j$ to $D_i$</td>
</tr>
<tr>
<td>$n_{D_i}$</td>
<td>Background noise at user $D_i$, which is assumed to be AWGN</td>
</tr>
<tr>
<td>$n_U$</td>
<td>Background noise at the base station, which is assumed to be AWGN</td>
</tr>
<tr>
<td>$\mathcal{CN}(x, y)$</td>
<td>A complex Gaussian random variable with mean $x$ and variance $y$</td>
</tr>
<tr>
<td>$E(\cdot)$</td>
<td>Expectation operator</td>
</tr>
<tr>
<td>$\log(\cdot)$</td>
<td>Natural logarithm</td>
</tr>
<tr>
<td>$\text{diag}(v)$</td>
<td>An operator that returns a square diagonal matrix with the elements of vector $v$ on the main diagonal and all entries outside the main diagonal equal to zero.</td>
</tr>
<tr>
<td>${\cdot}$</td>
<td>A set of variables</td>
</tr>
<tr>
<td>$[x]^+$</td>
<td>Positive part of scalar $x$, i.e., $\max(x, 0)$</td>
</tr>
<tr>
<td>$\sim$</td>
<td>Distributed according to</td>
</tr>
<tr>
<td>$</td>
<td>x</td>
</tr>
<tr>
<td>$X \otimes Y$</td>
<td>Kronecker product of two matrices $X$ and $Y$</td>
</tr>
<tr>
<td>$</td>
<td></td>
</tr>
<tr>
<td>$H^H$</td>
<td>Hermitian transpose of a matrix $H$</td>
</tr>
<tr>
<td>$H^T$</td>
<td>Transpose of a matrix $H$</td>
</tr>
<tr>
<td>$\text{Tr}(H)$</td>
<td>Trace of a matrix $H$</td>
</tr>
<tr>
<td>$</td>
<td>H</td>
</tr>
<tr>
<td>$H \succeq 0$</td>
<td>A matrix $H$ is a positive semidefinite matrix</td>
</tr>
<tr>
<td>$\text{rank}(H)$</td>
<td>Rank of a matrix $H$</td>
</tr>
<tr>
<td>$\nabla_X f(X)$</td>
<td>Gradient of $f(X)$</td>
</tr>
<tr>
<td>$\mathbb{C}$</td>
<td>Set of complex numbers</td>
</tr>
</tbody>
</table>
$\mathbb{C}^m$ Set of complex $m$-vectors

$\mathbb{C}^{m\times n}$ Set of complex $m \times n$ matrices
## Contents

Abstract

Tiivistelmä

Acknowledgements

List of abbreviations

List of symbols

Contents

1 Introduction

1.1 Motivation

1.2 Objectives, contributions, and outline of the thesis

1.3 The author’s contribution to the publications

2 Literature review and mathematical preliminaries

2.1 Full-duplex communications

2.2 Multiuser MIMO communications

2.2.1 Downlink transmission

2.2.2 Uplink transmission

2.3 Mathematical preliminaries

2.3.1 Sequential parametric convex approximation

2.3.2 Frank-Wolfe algorithm

2.3.3 Block coordinate ascent

2.3.4 Dual decomposition method

2.3.5 Concave-convex fractional program

3 Spectral and energy efficiency maximization approaches for full-duplex multiuser MIMO systems

3.1 System model and linear precoding techniques

3.2 Spectral efficiency maximization

3.2.1 Proposed joint design for SE maximization

3.2.2 Alternating optimization

3.3 Energy efficiency maximization

3.3.1 Power consumption model

3.3.2 Problem formulation

3.3.3 Proposed joint design for EE maximization
3.4 Implementation complexity .............................................. 58
3.5 Performance evaluation .................................................. 58
   3.5.1 System parameters ............................................... 58
   3.5.2 Convergence results ............................................. 59
   3.5.3 Spectral efficiency ................................................ 59
   3.5.4 Energy efficiency .................................................. 62
3.6 Summary and discussion ................................................ 65

4 Efficient beamforming solutions for full-duplex assisted small-cell
   networks ................................................................. 67
   4.1 System model and problem formulation .......................... 68
   4.2 Proposed beamformer designs .................................... 71
      4.2.1 Iterative MAXDET-based algorithm ......................... 71
      4.2.2 Iterative SDP-based algorithm .............................. 74
   4.3 Numerical results .................................................... 80
      4.3.1 Convergence and complexity comparison .................. 80
      4.3.2 Spectral efficiency performance ........................... 81
   4.4 Summary and discussion ............................................. 94

5 Conclusion and future work ............................................. 97
   5.1 Conclusion ............................................................ 97
   5.2 Future work .......................................................... 98

References ................................................................. 101

Appendices ................................................................. 109
1 Introduction

1.1 Motivation

In order to address the ever increasing demand of high data rates and the explosive growth of both mobile and fixed data traffic over finite radio resources, many modern communications technologies have been studied. Among those, the use of antenna arrays deployed at both ends of a communications link, referred to the multiple input multiple-output (MIMO) communications technique [1], is a well-known method to offer high spectral efficiency (SE). In fact, by adding the extra spatial dimensions to a wireless link, data rate can be greatly increased. This is done by transmitting simultaneously independent information streams on different antennas, called spatial multiplexing. Another benefit brought by MIMO is transmit/receive diversity gains which mitigate channel fading to significantly enhance link reliability. Therefore, MIMO has gradually become a core component to many wireless communications standards such as LTE [2], WLAN [3] and WiMAX [4]. In current and emerging cellular networks, MIMO is realized in the form of multiuser MIMO (MU-MIMO) in the downlink (DL) and uplink (UL) channels in which a set of users is scheduled to communicate with a base station (BS) at the same time. This is done by exploiting the spatial degrees of freedom which is defined as the number of independent information streams onto MIMO channels. MU-MIMO techniques can take advantages of both spatial multiplexing gain and multiuser diversity by choosing a group of users with good channel conditions for transmission [5].

Although MU-MIMO has been shown to be a promising approach to increasing SE as well as the system throughput, we seem to reach the limit of capacity that MU-MIMO can provide in reality in not-so-far future. The fact is that we cannot integrate as many antennas as we want in both ends of a communications link due to several practical limitations. Also, many other issues will arise when expanding the operating bandwidth or transmit power. For example, increasing the bandwidth requires extra cost for operators in order to acquire the additional licensed spectrum. Recall that in current wireless communications systems the DL and UL channels are designed to run orthogonally in time domain (time division duplex, TDD) or in frequency domain (frequency division duplex, FDD), referred to as half-duplex (HD) mode. For example, cellular networks with TDD allocate the same frequency band, but different time slots,
to DL and UL channels. On the other hand, cellular networks with FDD allow DL and UL transmissions to take place in the same time slot, but over distinct frequencies. So, we are still able to enhance the system capacity of cellular networks by allowing the DL and UL channels to work simultaneously over the same radio resources, referred to as full-duplex (FD) transmission. Such a design, if possible, is capable of doubling the system capacity of the HD systems and also resolving many problems of existing wireless communications networks such as reducing hidden terminals, congestion due to medium access control scheduling, and large delays [6–9].

Though the gains of FD systems can be easily foreseen, practical implementations of such FD systems pose many challenges still needed to be solved before we can see the first trial commercial deployment on system level. The crucial barrier in implementing FD systems resides in the self-interference (SI), which is also referred to as loopback interference in the literature, from the transmit antennas to receive antennas at a FD wireless transceiver. More explicitly, the radiated power of the DL transmission interferes with its own desired signals at the receive antennas. Since the SI power is typically much stronger (>100 dB) than the desired signal in cellular networks, FD transmission was thought infeasible in the past. The reason is that the SI power, if not efficiently suppressed, dramatically raises the background noise at receive antennas which exceeds a limited dynamic range of the analog-to-digital converter (ADC) in the receiver circuit [10]. This results in the impossibility of decoding the received signal of interest even if the SI signal is perfectly known.

In recent years, many breakthroughs in hardware design for SI cancellation techniques have been reported, e.g., in [7, 8, 11–15]. Especially, these studies demonstrate the feasibility of FD transmission for short to medium range wireless communications by suppressing the SI to a degree that possibly allows for FD operation. Motivated by the fact that in such systems transmit powers and path losses are rather decreased compared to traditional wireless systems. As a result the power differential between SI from a FD transceiver’s own transmissions and the received desired signal sending from a distant node in the UL channel is rather small, making SI cancellation much more manageable. Since FD transmission exploits the radio resources in two dimensions, i.e., the same frequency and time, a new network protocol as well as radio resource management are required to attain the maximum system sum rate which can be offered by FD. Hence, several studies focusing on these research areas have been carried out in a variety of contexts such as point-to-point (P2P) MIMO [16–18], FD MIMO Interference channel [19], MIMO relay [20–22], and cognitive radio [23]. Prior to the research efforts made
in the thesis, no attempt has been made to investigate potential gains of a single cell FD MU-MIMO system, where a FD capable BS communicates with multiple HD users in both DL and UL directions at the same time slot over the same frequency band.

In the literature, many transmission schemes under various criteria were originally proposed for HD counterparts (see [24–28] and references therein). But a direct application of these schemes to FD MU-MIMO systems would result in a poor performance. An obvious reason is that besides multiuser interference (MUI) in the DL channel there still exists a small, but not negligible, amount of the SI even after an advanced SI cancellation technique is applied, which is known as residual SI. Moreover, the considered FD system also suffers from co-channel interference (CCI) caused by users in the UL channel to the ones in the DL channel. This is due to the fact that the transmitted signals of UL users to the FD BS interfere with the received signals of DL users from the same FD BS. Simply put, the above discussions mean a new class of transmission strategies for FD MU-MIMO networks needs to be developed.

1.2 Objectives, contributions, and outline of the thesis

Early experimental achievements on SI cancellation [7, 8, 12, 13] have made FD transmission become a reality. To accelerate the use of FD techniques in future wireless networks we need to answer the two basic questions (i) what are the potential gains that FD transmission can offer? (ii) how can we achieve them? Of course the answers to the two above questions depend on the system of interest. This can be done easily for simple scenarios such as P2P communications which was done by [18, 29–31], but the results may lead to a too optimistic conclusion. To have a more complete picture of FD technologies this thesis considers a more general scenario which is referred to as a single cell FD MU-MIMO system as shown in Figs. 3 and 10. Note that single cell systems are one of the fundamental communications systems that allow us to investigate the potential gains offered by FD techniques and can be also useful to find the upper bounds of multi cell scenarios.

In communications system, SE is traditionally used to evaluate the system performance. In fact SE defined as data rate per bandwidth unit has been a main focus for wireless communications design. Very recently energy efficiency (EE) defined as the number of bits transmitted per an energy unit has received significant attention due to growing interest in green communications, where energy consumption is the ultimate goal to be optimized [32–36]. Therefore, the purpose of this thesis is to study
performance of the FD system of interest in terms of SE and EE. To do this, we propose
transmission schemes to maximize the SE and EE under some power constraints for
different scenario settings of the considered single cell FD network. Since there still
exists residual SI after SI suppression appearing at the receiver circuit, the problem of
transmission design become more challenging. Moreover, the difficulty of the design
problem is increased further by the CCI from UL users to DL ones. By this very nature,
a joint design of the DL and UL transmissions would offer the best solution due to
the coupling of both channels. To solve the considered transmission design problems,
advanced techniques of algebra and mathematical optimization theory are utilized. The
following is the outline of the thesis and our contributions included in each chapter.

– **Chapter 2** We make a survey on the SI cancellation mechanisms as well as the
applications of the FD techniques in literature. In addition, we also review some
proposed transmission and detection schemes for HD DL and UL MU-MIMO systems.
Finally, some mathematical tools used in this thesis are briefly presented.

– **Chapter 3** We consider a FD MU-MIMO scenario with multiple antennas at both a
FD BS and HD users. Further to this, we assume that the users in the DL and UL
channels are geographically distant enough that CCI can be ignored, meaning that only
SI is considered in the design for simplicity. In such a case, our contributions are to
propose low-complexity precoder schemes for maximizing SE and EE and to evaluate
their performance numerically. Since the SE and EE maximization problems are
non-convex at hand, we rely on a rank relaxation technique to tackle the non-convexity
of the problem at the first step. However, the relaxed problem is still non-convex
and thus we then propose iterative algorithms based on the concept of a sequential
parametric convex approximation (SPCA). Specifically, for the SE maximization
problem, the non-convex relaxed problem is approximated by a convex program in
each iteration. To do this, a lower bound of the SE in each iteration, which turns out
to be convex, is derived. Then, linear precoders, which have closed forms, are found
to maximize the lower bound by solving the dual problem with the block coordinate
ascent (BCA) and dual decomposition approaches. This lower bound iteratively
increases until the proposed iterative algorithm is converged. For the EE maximization
problem, using the same method, we first transform the relaxed precoder design
problem into a concave-convex fractional program, which then can be reformulated
as a convex program using the parametric approach. The resulting problem then
can be solved similarly to the SE maximization problem. Our transmission designs
show that, compared to a HD system, the FD system of interest with the proposed SE scheme yields a better SE and a considerably smaller EE. The proposed EE algorithm outperforms the SE one with regard to EE. However, the EE design is inferior to the HD system in the low transmit power regime while superior to the HD one when the SI power is small in the high transmit power region.

– **Chapter 4** We are particularly interested in exploring the SE performance of FD systems in the context of small-cell networks which are identified as a potential use case that benefits FD transmission. To do so we consider a more complex system model in the sense that CCI is taken into consideration. Assuming a FD BS with multiple antennas and HD users with single antenna, we devise a joint beamformer design to maximize the SE. The design problem is first formulated as a rank-constrained optimization problem, and the rank relaxation method is then applied to arrive at a relaxed problem. Due to the non-convexity of the relaxed problem, optimal solutions are hard to find. Herein, we propose two provably convergent algorithms to obtain suboptimal solutions. Based on the concept of the Frank-Wolfe algorithm, we approximate the design problem by a determinant maximization program in each iteration in the first algorithm. The second method is built upon the SPCA method, which allows us to transform the relaxed problem into a semidefinite program (SDP) in each iteration. Extensive numerical experiments under small-cell setups illustrate that the FD system with the proposed algorithms can obtain a large gain over the HD system.

– **Chapter 5** The conclusion and discussion about open problems of future research directions are given.

### 1.3 The author’s contribution to the publications

The thesis is written as a monograph based on two journal papers [37] and [38] and one conference paper [39]. All of these contributions have already been published. In [37], the linear precoding schemes are developed and the SE and EE performances are evaluated. In [38], the linear beamformers are designed and extensive numerical experiments under small-cell setups are performed to evaluate the SE.

The author of this thesis had the main responsibility on developing the original ideas, deriving the equations as well as algorithms, simulating the algorithms using MATLAB, generating the numerical results, and writing the papers. The role of other authors was
to give guidance, support and comments on developing the ideas/algorithms as well as writing the papers.
2 Literature review and mathematical preliminaries

This chapter reviews the main components on which the thesis are based. Specifically, we first outline the historical development of FD communications and the state-of-the-art techniques on SI cancellation. Then, transmission and detection schemes for conventional MU-MIMO communications are described. Finally, we provide an introduction of optimization techniques which are the key to solving the design problems considered in Chapters 3 and 4.

2.1 Full-duplex communications

The idea of concurrently transmitting and receiving signals of a transceiver over the same frequency band has been noticed for years. For example, it has been first known in continuous wave radar systems at least since the 1940s [40]. Subsequently, FD techniques were realized in the context of a wireless relay system [22, 41–46] in which a wireless FD repeater, called a FD relay, receives a signal, processes it, and then retransmits it to a destination over the same time-frequency resources, as depicted in Fig. 1b. Such systems have the potential to extend cell coverage and enhance the cell-edge throughput. Note that in most of FD relay topologies only the relay is configured to operate in the FD mode while the remaining nodes such as source and destination do not need to operate in the FD one. Recently, a related but different application of FD relay systems has been gained significant attention in both academia and industry, called FD bi-directional or P2P wireless networks [7, 8, 10, 12, 13, 16–18, 30, 47] over short ranges as seen in Fig. 1a. In the Fig. 1a, two nodes with FD operation exchange messages at the same time over the same frequency. More recently, a more general model called FD MU-MIMO systems or FD BS topologies [37–39, 48–51] have been investigated, in which a FD BS is allowed to communicate with multiple users in the DL and UL channels on the same radio resources.

To take full advantage of FD communications, it is necessary to reduce the SI power to a level so that the residual SI does not overwhelm the information bearing signals transmitted in the UL channel. To achieve this, a numerous variety of SI suppression schemes on hardware designs have been proposed in literature [7, 8, 12, 13]. In general these SI cancellation schemes are based on shared-antenna and separate-antenna architectures and can be categorised into passive and active techniques. In this thesis we
Fig. 1. FD scenarios. In the figure, S and D stands for source and destination, respectively.
assume a separate-antenna platform due to its more maturity of the technique. To what
follows we provide a brief introduction of the two architectures and the SI cancellation
techniques.

- In shared-antenna architecture transmitter and receiver RF circuits shares the same
  antenna and a circulator is utilized to isolate the transmitted signal and the received
  one (see [52, 53] for more details). But in practice the circulator does not completely
  separate the outgoing signal from the incoming one and the experiment in [13] showed
  that only 15 dB of isolation is provided. Thus, the strong SI directly leaks into the
  receiver RF circuit through the circulator.
- In separate-antenna structure, transmitter/receiver processing chains interface with
  their own antennas to radiate/sense the outgoing/incoming signals, respectively
  [52, 53]. In this case the SI is the transmitted signal directly propagating from transmit
  antennas to receive ones of the FD transceiver.

For the both architecture mentioned above, efficient SI reduction mechanisms are
done combining active and passive techniques which are described below [52–54].

- For passive suppression the SI signal with a huge dynamic range is maximally
  attenuated by increasing path loss over the SI propagation channel before entering
  into receiver processing chain. The aim of the technique is to mitigate the processing
  burden on the analog-to-digital converter (ADC). In shared-antenna structure, this
  is accomplished by a circulator. In separate-antenna architecture, many passive
  methods were proposed including antenna separation, antenna directionality, and
  cross-polarization [52, 53].
- For active suppression, analog and/or digital cancellation mechanisms [52, 53] are
  applied to the received SI signal in receiver RF circuit. Mostly, analog cancellation is
  first implemented on the estimated SI signal in analog domain before the ADC since
  the dynamic range of ADC limits the amount of SI cancellation. Then, the resulting
  residual SI signal is further suppressed in digital domain after the ADC by digital
  cancellation.

It is also worth mentioning that there is also a more sophisticated type of SI
Cancellation techniques, called spatial domain suppression [37–39, 42, 55–57], which
have been gained much attention recently. The spatial domain cancellation method is
actually carried out by antenna selection, linear precoding, null-space projection, and
minimum mean square error (MMSE) filters. This is done by exploiting degrees of freedom provided by multiple antennas installed at the transceivers.

Although tremendous advancements have been made to SI cancellation, SI always exists but may be as the same level as background noise. For single shared-antenna scenarios the best known result was reported in [13] which demonstrated that SI can be suppressed by 110 dB to the receiver noise floor for standard WiFi signals. However, only 95 dB of the SI cancellation is at best provided for WiFi systems with two separate antennas [52].

Since FD exploits time and frequency simultaneously, conventional resource allocation algorithms need to be redesigned. In fact, a new class of resource allocation for FD communications have been proposed under different contexts. For example, a number of studies on evaluating achievable gains of FD P2P [18, 29–31] and FD relay [22, 58–60] systems have been carried out, assuming the perfect SI cancellation or residual SI. Besides, the lower and upper bounds of achievable sum rate under the effects of channel estimation error and limited dynamic range resulting on residual SI were also derived in [18] and [22] for FD P2P and FD relay, respectively. The authors in [61, 62] proposed the antenna selection algorithm to maximize the sum rate and minimize symbol-error-rate for FD P2P MIMO while the works in [63, 64] studied a joint antenna and relay selection scheme so that the signal-to-interference plus noise ratio (SINR) of FD relay systems is maximized. In [65], a subcarrier allocation algorithm for FD orthogonal frequency division multiple access (OFDMA) was reported considering both the residual SI and CCI.

2.2 Multiuser MIMO communications

Fig. 2. HD MU-MIMO system model.
It is known that the capacity of a MIMO system scales linearly with the minimum of the number of antennas at transmitter and receiver. In cellular networks, the number of antennas at BSs is larger than that at users. Thus, if we increase the number of antennas at BSs but keep the one at users unchanged, the capacity is not likely greatly improved. To fully leverage the potential of MIMO communications, the idea is to serve many users at the same time which results in MU-MIMO and this is the main motivation for the consideration of a FD MU-MIMO in this thesis.

Basically, the FD MU-MIMO model of interest is a HD MU-MIMO one except that the base station is configured to operate in the FD mode. Hence, this subsection briefly discusses some common transmission schemes for the DL and UL of a conventional MU-MIMO system which is referred to as a HD MU-MIMO system. Specifically, we consider a HD MU-MIMO system including one BS equipped with $N$ antennas, $K^D$ users in the DL channel and $K^U$ users in the UL channel, as shown in Fig. 2. The set of users in the DL is denoted by $\{D_1, D_2, \ldots, D_{K^D}\}$ and the set of users in the UL is $\{U_1, U_2, \ldots, U_{K^U}\}$. In addition, the notations $N^D_i$ and $N^U_i$ refer to the number of antennas at user $D_i$ and user $U_i$, respectively. Additionally, we assume that the channels are flat fading and channel state information (CSI) is perfectly known at both the BS and the users in both the DL and UL channels. Two different system settings are considered: (1) single-antenna users and (2) multi-antenna users. Each system model has their own of importance. The case of single-antenna users allows us to simplify the problem formulation and solutions involved. It can also be used to find a lower bound of the case of multi-antenna users by treating each antenna of a user as if it were a separate single-antenna user. Multi-antenna user scenarios are more general and they offer more deep understanding of the achievable performance. However, the resulting radio resource problems usually become more complex to solve and some proper simplification should be made.

### 2.2.1 Downlink transmission

It is now well known that the optimal transmit strategy for achieving capacity of the DL of MU-MIMO systems, also known as MIMO broadcast channel, is dirty paper coding (DPC) scheme [24]. In this scheme, MUI can be completely presubtracted at the BS, which is known when constructing the transmitted signal. By this way each user receives its desired signal without interference from signals of other users. In fact, the concept of DPC was first introduced by Costa [66]. Then it was considered by [67] for
the case of two transmit antennas and two single-antenna users and later extended for a general single cell MIMO case with multiple antennas at both transmitter and receivers [68, 69]. Although the authors in [24] proved that DPC is the optimal capacity achieving strategy across the entire rate region of the MIMO broadcast channel, it requires high computational complexity to implement in practice, especially when the number of user is large. Hence, we restrict ourselves to linear transmit processing techniques which have been widely used in the literature, e.g., in [25–28, 70], for simplicity. We note that the optimal linear processing design for the DL channel is challenging due to the presence of MUI, and thus many suboptimal schemes such as zero forcing (ZF) using pseudo inverse [71], ZF using block diagonalization (BD) [25], and MMSE [26] have been proposed. In the following subsections, we provide the overview of ZF methods, which are adopted in this thesis, due to its simplicity. The reason for this is that BD schemes can decompose a MU-MIMO system into a set of parallel interference-free single user MIMO channels which may allow for computationally efficient methods.

2.2.1.1 Single-antenna users:
Define $s_{Di}$ be the transmitted data symbol for user $Di$, which is normalized to $E(|s_{Di}|^2) = 1$. For linear beamforming, the data symbol $s_{Di}$ is multiplied by the beamforming vector $w_{Di} \in \mathbb{C}^{N \times 1}$ before transmission. Therefore, the transmitted signal from $N$ antennas at the BS is given by

$$x = \sum_{i=1}^{K_D} w_{Di}s_{Di}$$

(1)

and the received signal of user $Di$ is written as

$$y_{HDi} = h_{Di}^H w_{Di}s_{Di} + \sum_{k \neq i}^{K_D} h_{Di}^H w_{Di}s_{Di} + n_{Di}$$

(2)

where $h_{Di}$ is the $N \times 1$ complex channel vector from the BS to user $Di$, and $n_{Di} \sim \mathcal{C}\mathcal{N}(0,\sigma_n^2)$ is background noise assumed to be additive white Gaussian noise (AWGN). In (2), the first term is the signal of interest and the second one represents MUI in the DL channel. By treating MUI as the background noise, the SINR of user $Di$ can be written as

$$\gamma_{HDi} = \frac{|h_{Di}^H w_{Di}|^2}{\sigma_n^2 + \sum_{k \neq i}^{K_D} |h_{Di}^H w_{Di}|^2}$$

(3)

We note that the distribution of the MUI approaches that of a Gaussian noise if the number of interfering users is sufficiently large.

30
And the SE or sum rate in the DL direction is given by

\[ R_{D}^{HD} = \sum_{k=1}^{K_{D}} \log(1 + \gamma_{D_{k}}). \]  

(4)

Here, we solve the problem of finding \( \{w_{D_{i}}\}_{i=1}^{K_{D}} \) for maximizing the SE in (4) under the sum power constraint at the BS, which is formulated as

\[
\text{maximize} \quad \left\{ w_{D_{i}} : \sum_{i=1}^{K_{D}} ||w_{D_{i}}||_{2}^{2} \leq P_{BS} \right\} R_{D}^{HD}
\]

(5)

where \( P_{BS} \) is the maximum power at the BS. In the ZF scheme, beamformers are designed so that there is no MUI between users, i.e., \( h_{D_{i}}^{H}w_{D_{k}} = 0 \) for all \( i \neq k \). Let us define \( H_{D} \triangleq [h_{D_{1}}, \ldots, h_{D_{K_{D}}}]^{H} \) and \( W_{D} \triangleq [w_{D_{1}}, \ldots, w_{D_{K_{D}}}] \). The ZF beamformers are calculated by computing the pseudo inverse of \( H_{D} \) as

\[
W_{D} = H_{D}^{H}(H_{D}H_{D}^{H})^{-1}\text{diag}((\sqrt{p_{D_{1}}}), \ldots, (\sqrt{p_{D_{K_{D}}}})).
\]

(6)

Hence, the SE problem in (5) simplifies into

\[
\text{maximize} \quad \sum_{k=1}^{K_{D}} \log(1 + \frac{p_{D_{k}}}{\sigma_{n}^{2}}) \\
\text{subject to} \quad \sum_{k=1}^{K_{D}} p_{D_{k}} [(H_{D}H_{D}^{H})^{-1}]_{k,k} \leq P_{BS}
\]

(7)

which can be solved using water filling. In (7), \( p_{D_{k}} \) is the allocated power for user \( D_{k} \).

### 2.2.1.2 Multi-antenna users:

With linear precoding employed at the BS, the received signal by user \( D_{i} \) is written as

\[
y_{D_{i}}^{HD} = H_{D_{i}}W_{D_{i}}s_{D_{i}} + \sum_{k \neq i, k=1}^{K_{D}} H_{D_{i}}W_{D_{k}}s_{D_{k}} + n_{D_{i}}
\]

(8)

where \( s_{D_{i}} \in \mathbb{C}^{N_{D_{i}} \times 1} \), \( W_{D_{i}} \in \mathbb{C}^{N \times N_{D_{i}}} \), and \( H_{D_{i}} \in \mathbb{C}^{N_{D_{i}} \times N} \) are the vector of transmitted signal, the linear precoder and the channel matrix of user \( D_{i} \), respectively. Without loss of generality, we can further assume that \( E[s_{D_{i}}s_{D_{i}}^{H}] = \mathbf{I} \). The background noise

---

Footnote: Throughout the thesis, we use natural logarithm for the sake of mathematical convenience. However, the SE is calculated with logarithm to Base 2 in the numerical result sections of Chapter 3 and 4.
$n_{D_i} \in \mathbb{C}^{N_{D_i} \times 1}$ is assumed to be a zero-mean AWGN vector with $n_{D_i} \sim \mathcal{CN}(0,\sigma_n^2 I_{D_i})$. Since CSI is assumed to be perfectly known at the BS, the SE or sum rate in the DL channel is given by

$$R_{D_i}^{HD} = \sum_{i=1}^{K_{D}} \log \left[ \frac{\sigma_n^2 I + \sum_{j=1}^{K_{D}} H_{D_i} W_{D_j} W_{D_j}^H H_{D_i}^H}{\sigma_n^2 I + \sum_{j=1,j\neq i}^{K_{D}} H_{D_i} W_{D_j} W_{D_j}^H H_{D_i}^H} \right]$$  \hspace{1cm} (9)

In the similar manner of ZF beamforming in Subsection (2.2.1.1), the BD approach of [25] designs the linear precoders $\{W_{D_i}\}_{i=1}^{K_{D}}$ such that MUI is completely eliminated, i.e., $H_{D_i} W_{D_j} = 0$ for all $i \neq j$. In fact, the BD algorithm [25] is a generalization of the ZF precoder for receivers with multiple antennas. This decomposes a DL MU-MIMO system into a group of $K_D$ parallel SU-MIMO channels. To be specific, for user $D_i$, define $\tilde{H}_{D_i}$ as

$$\tilde{H}_{D_i} = [H_{D_1}^T \ldots H_{D_{i-1}}^T H_{D_{i+1}}^T \ldots H_{D_{K_D}}^T]^T$$  \hspace{1cm} (10)

and consider a singular value decomposition (SVD) of $\tilde{H}_{D_i}$ as

$$\tilde{H}_{D_i} = \tilde{U}_{D_i} \tilde{\Lambda}_{D_i} [\tilde{V}_{D_i}^{(1)} \tilde{V}_{D_i}^{(0)}]$$  \hspace{1cm} (11)

where $\tilde{V}_{D_i}^{(0)} \in \mathbb{C}^{N \times \tilde{N}_{D_i}}$, $\tilde{N}_{D_i} = N - \sum_{k \neq i} N_{D_k}$, an orthogonal basis of the null space of $\tilde{H}_{D_i}$. Since the ZF constraints in BD imply that $H_{D_i} W_{D_i} = 0$, for all $i$, we can write $W_{D_i} = \tilde{V}_{D_i}^{(0)} M_{D_i}$, where $M_{D_i} \in \mathbb{C}^{\tilde{N}_{D_i} \times N_{D_i}}$ and $R_{D_i}^{HD}$ reduces to

$$R_{D_i}^{HD} = \sum_{i=1}^{K_{D}} \log \left| I + \tilde{H}_{D_i} M_{D_i} M_{D_i}^H \tilde{H}_{D_i}^H \right|$$  \hspace{1cm} (12)

where $\tilde{H}_{D_i} = (\sigma_n^2 I)^{-1/2} H_{D_i} \tilde{V}_{D_i}^{(0)}$ is the effective channel matrix of user $D_i$. The problem of computing $M_{D_i}$ that maximizes $R_{D_i}^{HD}$ in (12) can be easily solved using the water-filling algorithm with the total power $P_{BS}$ to the non-zero eigenvalues of $\tilde{H}_{D_i} \tilde{H}_{D_i}^H$ [25].

### 2.2.2 Uplink transmission

In this subsection, we present MMSE and successive interference cancellation (MMSE-SIC) scheme which was proved to be optimal for multiuser UL transmission [72]. From a design perspective, the use of MMSE-SIC receiver in this thesis allows us to simplify the sum rate expression of the UL channel.
2.2.2.1 Single-antenna users:
The received signal vector at the BS can be expressed as

\[ y_{HD}^U = \sum_{j=1}^{K_U} h_{Uj} s_{Uj} + n_{Uj} \]  

(13)

where \( s_{Uj} \) is the data symbol transmitted by \( U_j \) in the UL direction, \( h_{Uj} \in \mathbb{C}^{N \times 1} \) is the complex channel vector from the BS to user \( U_j \) and \( n_{Uj} \sim \mathcal{CN}(0, \sigma_n^2 I_N) \). By applying MMSE-SIC decoder, the received SINR of user \( U_j \) can be written as 

\[ \gamma_{Uj} = q_{Uj} h_{Uj}^H \left( \sigma_n^2 I + \sum_{m>j}^{K_U} q_{Um} h_{Um} h_{Um}^H \right)^{-1} h_{Uj} \]  

(14)

where \( E(|s_{Uj}|^2) = q_{Uj}, j = 1, ..., K_U \), is power loading for user \( U_j \). In (14), we have assumed a decoding order from 1 to \( K_U \) and thus the achievable sum rate of the UL channel is given by

\[ R_{HD}^{UL} = \sum_{j=1}^{K_U} \log(1 + \gamma_{Uj}) \]  

(15a)

\[ = \sum_{j=1}^{K_U} \log \left( 1 + q_{Uj} h_{Uj}^H \left( \sigma_n^2 I + \sum_{m>j}^{K_U} q_{Um} h_{Um} h_{Um}^H \right)^{-1} h_{Uj} \right) \]  

(15b)

\[ = \log \left| 1 + \frac{1}{\sigma_n^2} \sum_{j=1}^{K_U} q_{Uj} h_{Uj} h_{Uj}^H \right|. \]  

(15c)

2.2.2.2 Multi-antenna users:
In this case, the received vector at the BS is given by

\[ y_{Uj} = \sum_{i=1}^{K_U} h_{Uij} s_{Ui} + n_{Uj} \]  

(16)

where \( H_{Uij} \in \mathbb{C}^{N \times N_{U_i}} \) is the channel matrix from the BS to user \( U_i \), \( s_{Ui} \in \mathbb{C}^{N_{U_i} \times 1} \) is the transmitted symbol vector of user \( U_i \), and \( n_{Uj} \) is an \( N \times 1 \) AWGN vector with \( n_{Uj} \sim \mathcal{CN}(0, \sigma_n^2 I_N) \). The SE in the UL channel achieved by a MMSE-SIC receiver is given by

\[ R_{HD}^{UL} = \log \left| 1 + \frac{1}{\sigma_n^2} \sum_{i=1}^{K_U} H_{Uij} Q_{Uij} h_{Uij}^H \right| \]  

(17)

where \( Q_{Uij} = E[s_{Ui} s_{Ui}^H] \) is the input covariance matrix of the vector of data symbols of user \( U_i \).
2.3 Mathematical preliminaries

The problems to be considered in the sequel of the thesis are nonconvex and thus are difficult to solve in general. The purpose of this subsection is to provide a brief introduction of some optimization techniques that have been proved powerful for solving non-convex optimization problems.

2.3.1 Sequential parametric convex approximation

Consider the non-convex inequality-constrained problem which is given by

\[(P): \begin{aligned}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad g_i(x) \leq 0, \quad \forall i = 1, \ldots, m, \\
& \quad x \in \mathbb{R}^n
\end{aligned} \tag{18}
\]

where \(f(x)\) is convex over \(\mathbb{R}^n\) and \(g_i(x) \leq 0, \forall i\) is a non-convex constraint which has the upper convex approximation function, i.e.,

\[g_i(x) \leq \tilde{g}_i(x, t) \quad \text{with} \quad x \in \mathbb{R}^n, \quad t \in \mathbb{R}^p. \tag{19}\]

In (19), \(t\) is called the parameter vector and \(\tilde{g}_i(x, t)\) is convex and differentiable for a fixed \(t\). Basically, the idea of SPCA is to iteratively replace the non-convex constraints by the convex upper estimates for a given \(t\) and solve the resulting problem until convergence. Formally, at iteration \((k+1)\) the problem \((P)\) is approximated as

\[(P_{k+1}): \begin{aligned}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad \tilde{g}_i(x, t^{(k)}) \leq 0, \quad \forall i = 1, \ldots, m, \\
& \quad x \in \mathbb{R}^n
\end{aligned} \tag{20}\]

In (20), \(t^{(k)} = h(x^{(k)})\) in which \(x^{(k)}\) is the optimal solution of \((P_k)\) and \(h(x^{(k)})\) is chosen such that the following properties for convergence of algorithm are satisfied

\[g_i(x) = \tilde{g}_i(x, t), \quad \nabla g_i(x) = \nabla \tilde{g}_i(x, t) \tag{21}\]

where \(\nabla g\) is the gradient of \(g\). Let \(x^*\) be the solution of \((P_{k+1})\). The variable \(x\) is updated by \(x^{(k+1)} := x^*\). In this way, \(x\) is updated in each iteration until convergence.
2.3.2 **Frank-Wolfe algorithm**

Let $S \subset \mathbb{R}^n$ be a compact convex set and $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function. The Frank-Wolfe algorithm in [73] solves the following minimization problem

\[
\begin{align*}
\min_{x} & \quad f(x) \quad (22) \\
\text{subject to} & \quad x \in S.
\end{align*}
\]

To do this, the Frank-Wolfe method linearly approximates the objective function in (22) in an iterative manner by the first-order Taylor expansion around $x_k$ as given by

\[
f(y) \geq f(x_k) + (y - x_k)^T \nabla f(x_k) \quad (23)
\]

and then searches for a direction that improves the objective in the same domain until convergence.

2.3.3 **Block coordinate ascent**

Consider the problem

\[
\begin{align*}
\min_{x} & \quad f(x) \quad (24) \\
\text{subject to} & \quad x_i \in X_i, \quad \forall i = 1, \ldots, n
\end{align*}
\]

where $x = [x_1^T, \ldots, x_n^T]^T$ and $X = X_1 \times \ldots \times X_n$ is the Cartesian product of closed convex sets. It can be seen that $x$ can be decomposed into $n$ block/variable, each block/variable $x_i$ is independently constrained to $x_i \in X_i$. From this property, the BCA algorithm [74, 75] is based on the idea of iteratively minimizing the objective of (24) with respect to one variable/block while the others are held fixed in a circular framework. In particular, let $k$ denote the iteration index. At the $(k+1)$th iteration, $x_i^{(k+1)}$ is found to be the solution of the partial minimization problem which is expressed by

\[
x_i^{(k+1)} = \arg\min_{x_i \in X_i} f(x_1^{(k+1)}, \ldots, x_i^{(k+1)}, x_{i+1}, x_{i+2}, \ldots, x_n^{(k)}).
\]

By this manner, the objective value is nonincreasing with the BCA update until the iterative procedure converges. Note that since the subproblem has lower dimensions than the original one, it can be solved more easily or even has a closed-form solution.
2.3.4 Dual decomposition method

Let us consider the primal problem

\[
\begin{align*}
\text{minimize} & \quad \sum_i f_i(x_i) \\
\text{subject to} & \quad \sum_i h_i(x_i) \leq c, \\
& \quad x_i \in X_i, \quad \forall i
\end{align*}
\]

(26a)

(26b)

where \(X_i\) is the feasible convex set, \(f_i(x_i)\) and \(h_i(x_i)\) are the convex functions. Note that \(\sum_i h_i(x_i) \leq c\) are coupling constraints associating all variables \(\{x_i\}\). Moreover, if the coupling constraints are removed, the optimization problem in (26) can be decomposed into independent subproblems, each of which can be solved separately in parallel or sequential fashion. To do this, by introducing dual variable \(\lambda \succeq 0\) only for the coupling constraints (26b), the partial Lagrangian of (26) is formed as follows

\[
L(\{x_i\}, \lambda) = \sum_i f_i(x_i) + \lambda^T (\sum_i h_i(x_i) - c)
\]

(27)

Given the fixed \(\lambda\), the Lagrange dual function of (26), denoted by \(g(\lambda)\), is the optimal value of the following problem

\[
g(\lambda) = \min_{\{x_i\}, x_i \in X_i} L(\{x_i\}, \lambda)
\]

(28)

which can be decoupled into the subproblems as

\[
g_i(\lambda) = \min_{x_i \in X_i} f_i(x_i) + \lambda^T h_i(x_i)
\]

(29)

And the dual problem of (26) is

\[
\lambda^* = \max_{\lambda \succeq 0} g(\lambda) = \max_{\lambda \succeq 0} \sum_i g_i(\lambda) - \lambda^T c.
\]

(30)

If \(g(\lambda)\) is differentiable, the dual problem can be solved using the gradient method. In case of the non-differential dual objective \(g(\lambda)\), the subgradient approach is applied to solve (30). In details, suppose that the subgradient of each \(-g_i(\lambda)\) is \(h_i(x_i^*(\lambda))\), where \(x_i^*(\lambda)\) is the optimal value of the Subproblem (29) for a given \(\lambda\). The subgradient of \(-g(\lambda)\) is then \(\sum_i h_i(x_i^*(\lambda)) - c\). The dual problem is solved based on this subgradient.
As stated in [74, 76] the dual decomposition method actually solves the dual problem (30) instead of the original problem (26). Therefore, to achieve an appropriate solution, strong duality of the original problem (26) must be held, i.e., duality gap is zero ($\lambda^* = p^*$ where $p^*$ is the optimal value of (26)).

2.3.5 Concave-convex fractional program

In the thesis, since EE is modeled as a ratio of SE to total power consumption, the design problem lends itself to a non-linear fractional program and then can be transformed to a concave-convex fractional one, which has the form

$$\max_{x \in S} \frac{f_1(x)}{f_2(x)} \quad (31)$$

where $S \subseteq \mathbb{R}^n$ is a convex set, $f_1(x)$ is concave, $f_2(x)$ is convex, and $f_1(x)$ and $f_2(x)$ are positive on $S$. In literature, there are several methods to solve the concave-convex fractional program in (31) such as parametric convex program, parameter-free convex program, etc. [77–79]. The parametric convex program is used in this thesis since it allows us to arrive at closed-form solutions in the considered problems in Chapter 3. Particularly, the epigraph form of (31) is given by

$$\max_{x \in S; t \in \mathbb{R}} t \quad (32a)$$
$$\text{subject to } \frac{f_1(x)}{f_2(x)} - t \geq 0 \quad (32b)$$

which is equivalent to

$$\max_{x \in S; t \in \mathbb{R}} t \quad (33a)$$
$$\text{subject to } f_1(x) - tf_2(x) \geq 0 \quad (33b)$$

Note that the constraint in (33b) is not concave in $x$ and $t$. However, for a fixed $t \geq 0$, we have the convex problem which is written as

$$F(t) = \max_{x \in S} f_1(x) - tf_2(x). \quad (34)$$

As stated in [78], the optimal solution $x^*$ of (34) is also the one of the problem in (31) if and only if $F(t) = 0$. Moreover, $F(t)$ is strictly decreasing in $t$. Thus, the optimal value of $t$ can be found by solving $F(t) = 0$ using the bisection method over $t$.

We have presented fundamentals of wireless communications and optimization techniques that will be used in the thesis. The next chapter considers the case of a
FD MU-MIMO system and applies these introduced components to investigate their performance gains.
In this chapter we focus on joint designs of linear precoders for the FD MU-MIMO as shown in Fig. 3, in which a FD BS can send data to its serving users in both DL and UL simultaneously. For this system model we aim to optimize SE and EE subject to a sum power constraint (SPC) in the DL channel and per user power constraints (PUPCs) in the UL channel. It is worth pointing out that in practice each antenna of a transceiver is subject to its own maximum power constraint due to the associated power amplifier, which leads to the so-called per-antenna power constraint (PAPC). In this thesis we assume a SPC which is more commonly considered in the literature. Moreover, if the total power budget is not so large, the SPC can satisfy the PAPCs automatically. To this end, the SE and EE maximization problems are first formulated as nonconvex problems. Then, we propose efficient algorithms based on the concept of SPCA [80]. Note that the design of linear processing for the FD MU-MIMO system of interest is challenging. The reason is that there always exists residual SI between transmit antennas and receive antennas at the FD BS after applying hardware cancellation techniques. A possibility is to design the DL and UL channels separately without accounting for the residual SI. However, such separate designs of the DL and UL channels become ineffective due to an amount of the induced residual SI [39]. In addition to the residual SI from the DL to the UL channel, the CCI caused by users in the UL channel to those in the DL channel should also be considered. For simplicity, we assume that the users in the DL and UL channels are geographically separated, meaning that CCI is ignored. A general design that takes both the residual SI and CCI into consideration is an interesting problem which is studied in the next chapter.

The contributions presented in this chapter include the following.

- For the problem of the SE maximization, a convex lower bound of the SE is derived by using the first order approximation in each iteration. Then, linear precoders, which have closed-form expressions, are found to maximize the lower bound by solving the dual problem with the BCA and dual decomposition methods [75]. The

---

This chapter has been modified from [37] ©2013 IEEE.
proposed iterative design, referred to as the SE-optimal design, iteratively increases this lower bound until convergence. Our numerical results show that the SE of the FD MU-MIMO system is superior to that of the HD counterpart when the power level of the residual SI is small enough.

- For the EE maximization problem, since the EE formulation belongs to the class of nonlinear fractional programs, the optimal solution is generally hard to find. Herein, we propose a suboptimal precoder design referred to as EE-optimal design using the framework of concave-convex fractional programming, presented in [77–79]. To this end, the EE maximization problem is approximated to be a concave-convex fractional program by using the convex relaxation method as employed in the SE-optimal design. Then the resulting problem is transformed into an equivalent convex program by applying some appropriate mathematical operations. The resulting convex programming for EE maximization has a similar structure to that for SE maximization, and thus the main steps in the SE-optimal design can be applied. As expected, the simulation results reveal that the EE-optimal design outperforms the SE-optimal design in terms of EE. Moreover, the FD system yields lower EE than the HD counterpart in the low transmit power regime but higher EE in the high transmit power region if the SI power is sufficiently cancelled.

3.1 System model and linear precoding techniques

Consider a single cell FD MU-MIMO system with a FD BS sending data to $K_D$ users in the DL channel, and receiving data from $K_U$ users in the UL channel at the same time on the same frequency, as shown in Fig. 3. The $i$th user in the DL channel is denoted by $D_i$, and the $j$th user in the UL one is $U_j$. We assume that the BS is equipped with $N$ antennas, of which $N_T$ antennas are used to transmit data in the DL channel and $N_R$ antennas are used to receive the data from the UL users, $N = N_T + N_R$. The number of antennas at user $D_i$, $i = 1, 2, \ldots, K_D$, in the DL channel is denoted by $N_{D_i}$, and that at user $U_i$, $i = 1, 2, \ldots, K_U$, in the UL channel is denoted by $N_{U_i}$.

Assuming linear precoding at the BS, the received signal at user $D_i$ is

$$y_{D_i} = \mathbf{H}_{D_i} \mathbf{W}_{D_i} \mathbf{s}_{D_i} + \sum_{k \neq i, k=1}^{K_D} \mathbf{H}_{D_i} \mathbf{W}_{D_k} \mathbf{s}_{D_k} + \mathbf{n}_{D_i}$$

(35)

where $\mathbf{s}_{D_i} \in \mathbb{C}^{N_{D_i} \times 1}$, $\mathbf{W}_{D_i} \in \mathbb{C}^{N_T \times N_{D_i}}$, and $\mathbf{H}_{D_i} \in \mathbb{C}^{N_{D_i} \times N_T}$ are the vector of transmitted signal, the linear precoder and the channel matrix of user $D_i$, respectively. Without
loss of generality, we can further assume that $E[s_D^H s_D] = I$. The background noise $n_D \in \mathbb{C}^{N_D \times 1}$ is assumed to be a zero-mean AWGN vector with the single-sided noise spectral density $N_0$. Assuming perfect CSI at both transmitter and receiver, the SE in the DL channel is given by

$$R_D = \sum_{i=1}^{K_D} \log \left( \frac{|N_0 I + \sum_{j=1,j\neq i}^{K_D} H_D^H W_D^H W_D H_D^H|}{|N_0 I + \sum_{j=1,j\neq i}^{K_D} H_D^H W_D^H W_D H_D^H|} \right).$$

(36)

Note that the optimal linear precoder design is hard to find even for the DL channel alone due to the MUI. In fact, the problem of precoder design for maximizing $R_D$ in (36) has been proved to be NP-hard in [81]. Hence, to simplify the design, we adopt the BD scheme described in the preceding chapter. Towards this end let us define $\hat{H}_D \triangleq [H_{D_1}^T \ldots H_{D_{i-1}}^T H_{D_{i+1}}^T \ldots H_{D_{K_D}}^T]^T$ for user $D_i$, and denote by $\tilde{V}_D \in \mathbb{C}^{N_T \times \tilde{N}_D}$, $\tilde{N}_D = N_T - \sum_{k \neq i} N_{D_k}$, an orthogonal basis of the null space of $\hat{H}_D$. Then we can write

Fig. 3. FD MU-MIMO system model. The number of transmit and receive antennas at the BS is $N_T$ and $N_R$, respectively. In the DL channel data vectors are weighted by linear precoders $W = [W_D^1, \ldots, W_D^{K_D}]$. The MMSE-SIC detector is adopted for the UL channel, [37] ©2013 IEEE.
\[ W_{D_i} = \tilde{V}_{D_i}M_{D_i}, \quad \text{where } M_{D_i} \in \mathbb{C}^{N_{D_i} \times N_{D_i}} \text{ and } R_0 \text{ reduces to } \]

\[
R_0 = \sum_{i=1}^{K_D} \log |I + \tilde{H}_{D_i}M_{D_i}M_{D_i}^H\tilde{H}_{D_i}^H| \\
= \sum_{i=1}^{K_D} \log |I + \tilde{H}_{D_i}Q_{D_i}\tilde{H}_{D_i}^H| 
\tag{37}
\]

where \( \tilde{H}_{D_i} = N^{-1/2}_{0}\tilde{H}_{D_i} \) is the effective channel matrix of user \( D_i \). In (37), we have denoted \( Q_{D_i} = M_{D_i}M_{D_i}^H \in \mathbb{C}^{N_{D_i} \times N_{D_i}} \). In order to recover \( M_{D_i} \) from \( Q_{D_i} \), we impose a rank constraint on \( Q_{D_i} \) that \( \text{rank}(Q_{D_i}) \leq N_{D_i} \). As we will show later, the adoption of BD in the DL channel allows us to arrive at a formulation that can be solved using closed-form expressions. This possibility eliminates the need of installing a generic optimization package.

Clearly, the BD scheme is feasible if \( N_T > \max_i \sum_{k=1, k \neq i}^{K_D} N_{D_k} \) due to the condition for the existence of the null space of \( \bar{H}_{D_i} \) for all \( i \). This dimensionality constraint means that the BS can simultaneously transmit data to a limited number of users. In case that \( K_D \) is larger than the supportable number of users, a group of users must be chosen by a user scheduling algorithm. We note that existing user scheduling algorithms for BD in the literature (see [82–84] and references therein) can be employed in combination with our proposed designs. However, these user scheduling algorithms, which were only devised for the DL channel, do not account for the residual SI. Obviously, a user selection algorithm that considers the residual SI may result in better performance and designing such an algorithm constitutes a rich area for future research.

For user \( U_i \) in the UL channel, let \( H_{U_i} \in \mathbb{C}^{N_{R} \times N_{U_i}} \) be the channel matrix and \( s_{U_i} \in \mathbb{C}^{N_{U_i} \times 1} \) be the transmitted symbols. Then, the received vector at the BS is given by

\[
y_U = \sum_{i=1}^{K_U} H_{U_i}s_{U_i} + \sum_{k=1}^{K_D} G_SW_{D_k}s_{D_k} + n_U \\
= \sum_{i=1}^{K_U} H_{U_i}s_{U_i} + \sum_{k=1}^{K_D} \tilde{G}_Ss_D + n_U 
\tag{38}
\]

where \( \tilde{G}_S = G_S\tilde{V}_{D_k} \) and \( n_U \) is an \( N_R \times 1 \) AWGN vector with \( n_U \sim \mathcal{CN}(0, N_0I) \). In (38), \( G_S \) specifies the residual SI channel from the transmit antennas to the receive ones at the BS, and its statistical properties depend on the effectiveness of the SI cancellation.
techniques. By treating the residual SI as the background noise and assuming perfect CSI at the BS and users, the SE of the UL users with a MMSE-SIC receiver is

\[ R_U = \log \left( \frac{N_0 I + \sum_{k=1}^{K_D} \tilde{G}_{S_k} Q_{D_k} \tilde{G}_{S_k}^H + \sum_{i=1}^{K_U} H_{U_i} Q_{U_i} H_{U_i}^H}{N_0 I + \sum_{k=1}^{K_D} \tilde{G}_{S_k} Q_{D_k} \tilde{G}_{S_k}^H} \right) \]  

(39)

where \( Q_{U_i} = E[s_{U_i} s_{U_i}^H] \) is the transmit covariance matrix of user \( U_i \).

Due to the term \( \sum_{k=1}^{K_D} \tilde{G}_{S_k} Q_{D_k} \tilde{G}_{S_k}^H \) in (39), the performance of the DL and UL channels is coupled. If one only concentrates on maximizing \( R_D \), this interference term can be so destructive that it can make the received signals of interest from UL users unreliable to detect. Thus, we consider a joint design of the DL and UL channels. Conventionally, we assume a SPC at the BS in the DL channel and PUPCs in the UL channel. Let \( f(\{Q_{D_i}\}, \{Q_{U_j}\}) \) be the performance measure of interest. Then, the problem of the joint design of \( \{Q_{D_i}\} \) and \( \{Q_{U_j}\} \) is generally expressed as

\[
\begin{align*}
\text{maximize} & \quad f(\{Q_{D_i}\}, \{Q_{U_j}\}) \\
\text{subject to} & \quad \sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) \leq P_{BS}, \\
& \quad \text{Tr}(Q_{U_j}) \leq q_{U_j}, \quad \forall j = 1, \ldots, K_U, \\
& \quad Q_{D_i} \succeq 0, \quad \forall i = 1, \ldots, K_D, \\
& \quad Q_{U_j} \succeq 0, \quad \forall j = 1, \ldots, K_U
\end{align*}
\]  

(40)

where \( P_{BS} \) is the maximum transmit power of the BS and \( q_{U_j} \) is the maximum transmit power of the user \( U_j \). In the following sections, we present joint designs for maximizing SE and EE.

### 3.2 Spectral efficiency maximization

In this section, we propose a low-complexity (i.e., closed-form) joint design for the SE maximization problem, referred to as the SE-optimal design. For this case,
\[ f((Q_{b_i}), (Q_{u_j})) = R_0 + R_U, \text{ and (40) becomes} \]

\[
\begin{align*}
\text{maximize} & \quad R_0 + R_U \\
\text{subject to} & \quad \sum_{i=1}^{K_D} \text{Tr}(Q_{b_i}) \leq P_{BS}, \\
& \quad \text{Tr}(Q_{u_j}) \leq q_{U_j}, \quad \forall j = 1, \ldots, K_U, \\
& \quad Q_{b_i} \succeq 0, \quad \forall i = 1, \ldots, K_D, \\
& \quad Q_{u_j} \succeq 0, \quad \forall j = 1, \ldots, K_U.
\end{align*}
\]

(41)

The optimization problem above is also known as the sum rate maximization problem which is one of the important ones in wireless communications design in general. In some cases where different preferences are imposed to DL or UL channel, we may associate \( R_0 \) and \( R_U \) with different weights, which results in the problem of weighted sum rate maximization of DL and UL. We remark that the proposed solutions presented in this thesis can be easily extended to deal with this general optimization problem.

It should be noted that all the constraints are convex with respect to \( \{Q_{b_i}\} \) and \( \{Q_{u_j}\} \). Moreover, \( R_0 \) and \( R_U \) are also concave with \( \{Q_{b_i}\} \) and \( \{Q_{u_j}\} \), respectively. However, due to the interference from the DL channel, \( R_U \) in (39) is neither convex nor concave with \( \{Q_{b_i}\} \). Consequently, problem (41) is a nonconvex program, which is difficult to solve in general.

Instead, we resort to a local optimization method, i.e., we propose a joint design that solves (41) locally. The proposed joint design is based on a SPCA of (41). More specifically, we employ the first order approximation of \( R_U \) to find a lower bound of the SE and update this lower bound iteratively until the algorithm converges. We note that the SPCA method has been shown to be efficient for the problem of beamformer design for multiple-input single-output (MISO) DL channels (see [70] and references therein).

3.2.1 Proposed joint design for SE maximization

Let \( Q_{b_i}^{(n)} \) denote the value of \( Q_{b_i} \) after \( n \) iterations. Then the SE after \( n \) iterations is bounded below as given by

\[
R_0 + R_U \geq \sum_{i=1}^{K_D} \log |I + \bar{H}_{b_i} Q_{b_i} \bar{H}_{b_i}^H| - \text{Tr}\left( (\Upsilon^{(n)})^{-1} \frac{1}{N_0} \sum_{i=1}^{K_D} \bar{G}_{S_i} Q_{b_i} \bar{G}_{S_i}^H \right) + \frac{1}{N_0} \sum_{i=1}^{K_D} \bar{G}_{S_i} Q_{b_i} \bar{G}_{S_i}^H + \frac{1}{N_0} \sum_{j=1}^{K_U} \bar{H}_{u_j} Q_{u_j} \bar{H}_{u_j}^H + \Psi^{(n)}
\]

(42)
where $\Upsilon^{(n)} = I + \frac{1}{N_0} \sum_{i=1}^{K_D} \bar{G}_n \mathbf{Q}_{b_i}^{(n)} \bar{G}_n^H$ and $\Psi^{(n)} = \text{Tr}( (\Upsilon^{(n)})^{-1} \frac{1}{N_0} \sum_{i=1}^{K_D} \bar{G}_n \mathbf{Q}_{b_i}^{(n)} \bar{G}_n^H ) - \log|\Upsilon^{(n)}|$. In (42), we have used the inequality $\log|I + X| \leq \log|I + X_0| + \text{Tr}(I + X_0)^{-1}(X - X_0))$ for $X \succeq 0$, where $X_0$ is an arbitrary operating point. This inequality is due to the concavity of logdet function [85]. In fact, the idea of iteratively maximizing lower bounds of a nonconvex function has been used in the literature for different contexts, e.g., in [86–88]. In the $(n+1)$th iteration of the proposed algorithm, $\{\mathbf{Q}_{b_i}^{(n+1)}\}$ are found to maximize the lower bound of the SE. Mathematically, $\{\mathbf{Q}_{b_i}^{(n+1)}\}$ are the solutions of the following problem

$$
\text{maximize} \quad R_{LB}^{(n)}(\{\mathbf{Q}_{b_i}\},\{\mathbf{Q}_{u_j}\})
$$

subject to

$$
\sum_{i=1}^{K_D} \text{Tr}(\mathbf{Q}_{b_i}) \leq P_{BS},
$$

$$
\text{Tr}(\mathbf{Q}_{u_j}) \leq q_{u_j}, \quad \forall j = 1, \ldots, K_U,
$$

$$
\mathbf{Q}_{b_i} \succeq 0, \quad \forall i = 1, \ldots, K_D,
$$

$$
\mathbf{Q}_{u_j} \succeq 0, \quad \forall j = 1, \ldots, K_U
$$

where $R_{LB}^{(n)}(\{\mathbf{Q}_{b_i}\},\{\mathbf{Q}_{u_j}\})$ is the right hand side of (42) with the constant $\Psi^{(n)}$ being ignored since it does not affect the optimization problem in (43). With the convex approximation of $R_0$, problem (43) now becomes a convex program in each iteration, which can be solved using general standard convex optimization packages, e.g., CVX [89] or YALMIP [90]. The values of $\{\mathbf{Q}_{b_i}^{(n+1)}\}$ are updated until the iterative procedure converges. Regarding the use of a generic method to solve (43), the rank constraints, $\text{rank}(\mathbf{Q}_{b_i}) \leq N_D$, for $i = 1, \ldots, K_D$, must be guaranteed so that we are able to extract $\mathbf{M}_{b_i}$ from $\mathbf{Q}_{b_i}$. The rank of the obtained solutions is further discussed in the Appendix 1.

Generic methods do not exploit the specific structure of the problem and thus require high computational complexity. Moreover, such methods do not offer useful insights into the optimal solutions. In this thesis, exploiting the specific expression of $R_{LB}(\{\mathbf{Q}_{b_i}\},\{\mathbf{Q}_{u_j}\})$, we develop an efficient iterative algorithm to solve (43), in which analytical expressions of $\mathbf{Q}_{b_i}$ and $\mathbf{Q}_{u_j}$ are found in each iteration. Moreover, the proposed precoder design guarantees that the rank constraints, $\text{rank}(\mathbf{Q}_{b_i}) \leq N_D$, for all
\( \mathbf{Q}_D \)'s, are satisfied automatically. To start with, we rewrite (43) as

\[
\begin{align*}
\text{maximize} & \quad \mathcal{R}^{(n)}_{LB}(\{\mathbf{Q}_D\}, \{\mathbf{Q}_U\}, \{\mathbf{p}_D\}) \\
\text{subject to} & \quad \text{Tr}(\mathbf{Q}_D) \leq \mathbf{p}_D, \quad \mathbf{Q}_D \succeq 0, \quad \forall i = 1, \ldots, K_D, \\
& \quad \text{Tr}(\mathbf{Q}_U) \leq q_U, \quad \mathbf{Q}_U \succeq 0, \quad \forall j = 1, \ldots, K_U, \\
& \quad \sum_{i=1}^{K_D} \mathbf{p}_D_i \leq P_{\text{BS}}, \\
& \quad \mathbf{p}_D_i \geq 0, \quad \forall i = 1, \ldots, K_D
\end{align*}
\]

(44)

where \( \{\mathbf{p}_D, \ldots, \mathbf{p}_{K_D}\} \) are newly introduced optimization variables. The partial Lagrangian of (44) is given by

\[
\mathcal{L}((\mathbf{Q}_D), \{\mathbf{Q}_U\}, \{\mathbf{p}_D\}, \lambda) = \mathcal{R}^{(n)}_{LB}(\{\mathbf{Q}_D\}, \{\mathbf{Q}_U\}, \{\mathbf{p}_D\}) - \lambda \left( \sum_{i=1}^{K_D} \mathbf{p}_D_i - P_{\text{BS}} \right)
\]

(45)

where \( \lambda \geq 0 \) is the dual variable associated with the constraint \( \sum_{i=1}^{K_D} \mathbf{p}_D_i \leq P_{\text{BS}} \). Consequently, the dual objective of (44), denoted by \( g(\lambda) \), is the optimal value of the following problem

\[
\begin{align*}
\text{maximize} & \quad \mathcal{L}((\mathbf{Q}_D), \{\mathbf{Q}_U\}, \{\mathbf{p}_D\}, \lambda) \\
\text{subject to} & \quad \text{Tr}(\mathbf{Q}_D) \leq \mathbf{p}_D, \quad \mathbf{Q}_D \succeq 0, \quad \forall i = 1, \ldots, K_D, \\
& \quad \text{Tr}(\mathbf{Q}_U) \leq q_U, \quad \mathbf{Q}_U \succeq 0, \quad \forall j = 1, \ldots, K_U, \\
& \quad \mathbf{p}_D_i \geq 0, \quad \forall i = 1, \ldots, K_D
\end{align*}
\]

(46)

and the dual problem of (44) is

\[
\begin{align*}
\text{minimize} & \quad g(\lambda) \\
\text{subject to} & \quad \lambda \geq 0.
\end{align*}
\]

(47)

It can be easily seen that strong duality holds for the problem (43), and thus the optimal solution of (43) can be found by solving its dual problem in (47). Due to the fact that a subgradient of \( g(\lambda) \) is given by

\[
h_{g(\lambda)} = P_{\text{BS}} - \sum_{i=1}^{K_D} \mathbf{p}_D_i
\]

(48)

and that \( \lambda \) is a scalar, the minimization of \( g(\lambda) \) over \( \lambda \geq 0 \) can be carried out efficiently using one dimension search method, e.g., bisection method [91]. Thus, solving (43) boils down to finding an efficient algorithm to evaluate \( g(\lambda) \), i.e., solving (46) efficiently.
Algorithm 1 The proposed SE-optimal design for the considered FD system. [37] ©2013 IEEE.

1: Generate initial points for \( \{Q_{D_i}^{(0)}\}_{i=1}^{K_D} \); tolerance \( \epsilon > 0 \).
2: \( n := 0 \).
3: repeat
   4: Generate initial points for \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \).
   5: while \( (\lambda_{\text{max}} - \lambda_{\text{min}}) > \epsilon \) do
      6: \( \lambda = (\lambda_{\text{min}} + \lambda_{\text{max}})/2 \)
      7: Solve (46) to find optimal solutions \( \{Q_{D_i}^*, P_{D_i}^*\}_{i=1}^{K_D} \) and \( \{Q_{U_j}^*\}_{j=1}^{K_U} \) using the alternating optimization algorithm presented in Subsection 3.2.2.
      8: if \( \sum_{i=1}^{K_D} p_{D_i}^* < P_{\text{BS}} \) then
         9: \( \lambda_{\text{max}} = \lambda \)
      else
         10: \( \lambda_{\text{min}} = \lambda \)
      end if
   11: end while
   12: \( n := n + 1 \).
   13: Update covariance matrices: \( \{Q_{D_i}^{(n)}\}_{i=1}^{K_D} := \{Q_{D_i}^*\}_{i=1}^{K_D} \).
14: \( \text{until convergence.} \)
15: Apply the Cholesky decomposition to \( Q_{D_i}^{(n)} \) to find \( M_{D_i} \), and calculate the precoder \( W_{D_i} = \tilde{V}_{D_i} M_{D_i} \) for each user in the DL channel.

Before proceeding further, we outline the proposed joint design of the DL and UL channels for the SE maximization problem in Algorithm 1.

The proposed SE-optimal design is guaranteed to converge since the lower bound is increased after every iteration, and the total SE of system is bounded above due to the power constraints. Moreover, due to the fact that the first order approximation is employed, Algorithm 1 converges to a stationary point, i.e., a point that satisfies the Karush–Kuhn–Tucker (KKT) conditions of (41). The detailed proof is provided in [80].

### 3.2.2 Alternating optimization

As the core of Algorithm 1, we now present an efficient algorithm to solve (46). We observe that the problem formulation in (46) lends itself to the BCA method since the constraints for individual \( \{Q_{D_i}\} \) and \( \{Q_{U_j}\} \) are separable. Moreover, as we will show
shortly, the optimization of one variable, when others are fixed, can be expressed through analytical forms. Particularizing the BCA method to solve (46) leads to two different cases. In the first case, for each $k \in \{1, 2, \ldots, K_0\}$, we find $Q_{uk}$ that solves problem (46) by treating others $\{Q_{u1}, \ldots, Q_{uk-1}, Q_{uk+1}, \ldots, Q_{uK_0}\}$ and $\{(Q_{d1}, p_{d1}), \ldots, (Q_{dK_0}, p_{dK_0})\}$ as constants. Explicitly, we need to solve the following problem

$$(P1): \begin{cases} \text{maximize} & \log|I + \tilde{H}_{u_k} Q_{u_k} \tilde{H}_{u_k}^H| + \zeta_{(P1)} \\ \text{subject to} & \text{Tr}(Q_{u_k}) \leq q_{uk}, \ Q_{u_k} \succeq 0 \end{cases}$$

where

$$\tilde{H}_{u_k} = (N_0 Z_{u_k})^{-1/2} H_{u_k},$$

$$\zeta_{(P1)} = \sum_{j=1}^{K_0} \log|I + \tilde{H}_{d_j} Q_{d_j} \tilde{H}_{d_j}^H| - \frac{1}{N_0^2} \text{Tr}((\mathcal{T}(n))^{-1} \sum_{j=1}^{K_0} G_{d_j} Q_{d_j} G_{d_j}^H)
+ \log|Z_{u_k}| - \lambda(\sum_{j=1}^{K_0} p_{d_j} - P_{BS}).$$

$$Z_{u_k} = I + \frac{1}{N_0} \sum_{j=1}^{K_0} G_{d_j} Q_{d_j} G_{d_j}^H + \frac{1}{N_0} \sum_{i=1,i \neq k}^{K_0} H_{u_i} Q_{u_i} H_{u_i}^H.$$

It is easy to see that problem (P1) admits a solution based on the water-filling procedure. To be specific, let $r_{uk}$ and $\sigma_i^2, \forall i \in \{1, 2, \ldots, r_{uk}\}$ be the rank and singular values of $\tilde{H}_{u_k}$, respectively. Then the optimal solution of (P1) is given by

$$Q_{u_k}^* = \tilde{V}_{u_k} \text{diag}(q_1^*, \ldots, q_{r_{uk}}^*) \tilde{V}_{u_k}^H$$

where

$$q_i^* = \left[\mu - \frac{1}{\sigma_i^2} \right]^{+}, \ \forall i = 1, \ldots, r_{uk}$$

$\tilde{V}_{u_k}$ consists of right singular vectors of $\tilde{H}_{u_k}$, and the water level $\mu$ is chosen to meet the power constraint, i.e., $\sum_{i=1}^{r_{uk}} q_i^* = q_{uk}$.

In the remaining case, for each $k \in \{1, 2, \ldots, K_0\}$, the problem is to update $\{(Q_{d1}, p_{d1})\}$ while others $\{(Q_{d1}, p_{d1}), \ldots, (Q_{dK_0}, p_{dK_0})\}$ and $\{(Q_{d1}, p_{d1}), \ldots, (Q_{dK_0}, p_{dK_0})\}$ are held fixed. This amounts to solving the following problem

$$(P2): \begin{cases} \text{maximize} & \log|I + \tilde{H}_{d_k} Q_{d_k} \tilde{H}_{d_k}^H| + \log|I + \tilde{G}_{S_k} Q_{d_k} \tilde{G}_{S_k}^H| \\
- \text{Tr}(\tilde{G}_{S_k} Q_{d_k} \tilde{G}_{S_k}^H) - \lambda p_{d_k} + \zeta_{(P2)} \\
\text{subject to} & \text{Tr}(Q_{d_k}) \leq p_{d_k}, \ Q_{d_k} \succeq 0, \ p_{d_k} \geq 0 \end{cases}$$

$$48$$
where
\[
\hat{G}_{Sk} = (N_0 \Upsilon(n))^{-1/2} \tilde{G}_{Sk},
\]
\[
\tilde{G}_{Sk} = (N_0 Z_{Dk})^{-1/2} \bar{G}_{Sk},
\]
\[
\zeta(P2) = \sum_{j=1, j \neq k}^{K_S} \log |I + H_{Dj} Q_{Dj} H_{Dj}^H| - \frac{1}{N_0} \text{Tr}((\Upsilon(n)^{-1}) \sum_{j=1, j \neq k}^{K_S} G_{Sj} Q_{Dj} G_{Sj}^H)
\]
\[
+ \log |Z_{Dk}| - \lambda \sum_{i \neq k}^{K_U} P_{B_i} + \lambda P_{BS},
\]
\[
Z_{Dk} = 1 + \frac{1}{N_0} \sum_{i=1}^{K_U} H_{Dj} Q_{Dj} H_{Dj}^H + \frac{1}{N_0} \sum_{i=1, i \neq k}^{K_S} G_{Sj} Q_{Dj} G_{Sj}^H.
\]

The alternating procedure to solve (46) is outlined in Algorithm 2.

\begin{algorithm}
\caption{The proposed alternating optimization algorithm. [37] ©2013 IEEE.}
\begin{algorithmic}[1]
\State Initialize: $\{Q_{D_i}, p_{D_i}\}_{i=1}^{K_U} = 0$; $\{Q_{D_i}\}_{j=1}^{K_U} = 0$.
\Repeat
\For{$k = 1$ to $K_U$}
\State Solve (\textit{P1}) using the water-filling algorithm to find $Q_{Dk}$ while keeping all other variables fixed.
\EndFor
\For{$l = 1$ to $K_U$}
\State Solve (\textit{P2}) using Algorithm 3 to find $(Q_{D_l}, p_{D_l})$ while keeping all other variables fixed.
\EndFor
\Until desired accuracy is reached.
\end{algorithmic}
\end{algorithm}

Since each iteration of the alternating optimization algorithm increases the objective of (46), it converges to a locally optimal solution of problem (46), which is also the optimal solution of problem (43) due to its convexity.

As shown in Line 7 of Algorithm 2, we now present Algorithm 3 that can solve (\textit{P2}) efficiently, i.e., via analytical expressions. First, under the framework of dual
decomposition method, we rewrite (52) as

$$\begin{align*}
\text{maximize} & \quad f_1(Y_{D_k,1}, p_{D_k}) + f_2(Y_{D_k,2}) \\
\text{subject to} & \quad Y_{D_k,1} = Y_{D_k,2}, \\
& \quad \text{Tr}(Y_{D_k,1}) \leq p_{D_k}, \\
& \quad Y_{D_k,1} \succeq 0, Y_{D_k,2} \succeq 0, p_{D_k} \geq 0
\end{align*}$$

(53)

where

$$\begin{align*}
f_1(Y_{D_k,1}, p_{D_k}) & \triangleq \log |I + \hat{H}_{D_k} Y_{D_k,1} \hat{H}_{D_k}^H| - \lambda p_{D_k} \\
f_2(Y_{D_k,2}) & \triangleq \log |I + \tilde{G}_{S_k} Y_{D_k,2} \tilde{G}_{S_k}^H| - \text{Tr}(\tilde{G}_{S_k} Y_{D_k,2} \tilde{G}_{S_k}^H).
\end{align*}$$

In (53), we have introduced two variables $Y_{D_k,1}$ and $Y_{D_k,2}$ and imposed the equality constraint $Y_{D_k,1} = Y_{D_k,2}$ to make (53) equivalent to (52). In the context of the dual decomposition method, $Y_{D_k,1}$ and $Y_{D_k,2}$ are local versions of the complicating variable $Q_{D_k}$, along with a consistency constraint $Y_{D_k,1} = Y_{D_k,2}$ that requires the two local versions to be equal. We also omit the constant $\zeta_{(P_2)}$ in (52) since it does not affect the optimization of $(P_2)$. Next, let $\Phi$ be the Lagrange multiplier associated with the consistency constraint. Then, the partial Lagrangian function of (53) is written as

$$\begin{align*}
\mathcal{L}(Y_{D_k,1}, Y_{D_k,2}, p_{D_k}, \Phi) & = f_1(Y_{D_k,1}, p_{D_k}) + f_2(Y_{D_k,2}) + \text{Tr} \left[ (Y_{D_k,1} - Y_{D_k,2}) \Phi \right] \\
& = \log |I + \hat{H}_{D_k} Y_{D_k,1} \hat{H}_{D_k}^H| - \lambda p_{D_k} + \text{Tr}(\Phi Y_{D_k,1}) \\
& \quad + \log |I + \tilde{G}_{S_k} Y_{D_k,2} \tilde{G}_{S_k}^H| - \text{Tr} \left[ (\tilde{G}_{S_k}^H \tilde{G}_{S_k} + \Phi) Y_{D_k,2} \right]
\end{align*}$$

(54)

and the dual function $g(\Phi)$ is the optimal value of the following problem

$$\begin{align*}
\text{maximize} & \quad \mathcal{L}(Y_{D_k,1}, Y_{D_k,2}, p_{D_k}, \Phi) \\
\text{subject to} & \quad \text{Tr}(Y_{D_k,1}) \leq p_{D_k}, \\
& \quad Y_{D_k,1} \succeq 0, Y_{D_k,2} \succeq 0, p_{D_k} \geq 0.
\end{align*}$$

(55)

Since the objective function and all constraints in (55) are separable, it can be decomposed into dual subproblems 1 and 2 as follows

$$\begin{align*}
\text{Subproblem 1} & \triangleq \left\{ \begin{array}{l}
\text{maximize} \quad \log |I + \hat{H}_{D_k} Y_{D_k,1} \hat{H}_{D_k}^H| - \lambda p_{D_k} + \text{Tr}(\Phi Y_{D_k,1}) \\
\text{subject to} \quad \text{Tr}(Y_{D_k,1}) \leq p_{D_k}, \\
& \quad Y_{D_k,1} \succeq 0, p_{D_k} \geq 0
\end{array} \right. \\
& \text{subject to} \quad \text{Tr}(Y_{D_k,1}) \leq p_{D_k}, \\
& \quad Y_{D_k,1} \succeq 0, p_{D_k} \geq 0
\end{align*}$$

(56)
and

Subproblem 2 \(\doteq\) \[
\begin{align*}
\text{maximize} & \quad \log|I + \hat{G}_{S_k} Y_{D_k,2} \hat{G}_{S_k}^H| - \text{Tr}\left(\left(\hat{G}_{S_k}^H \hat{G}_{S_k} + \Phi\right) Y_{D_k,2}\right) \\
\text{subject to} & \quad Y_{D_k,2} \geq 0
\end{align*}
\]

with optimal values \(g_1(\Phi)\) and \(g_2(\Phi)\), respectively. We will show shortly that Subproblems 1 and 2 can be solved efficiently by analytical forms. Firstly, we consider the Lagrangian function of Subproblem 1 that is given by

\[
L(Y_{D_k,1}, p_{D_k}, \nu) = \log|I + \hat{H}_{D_k} Y_{D_k,1} \hat{H}_{D_k}^H| - \lambda p_{D_k} + \text{Tr}(\Phi Y_{D_k,1}) - \nu(\text{Tr}(Y_{D_k,1}) - p_{D_k})
\]

\[
= \log|I + \hat{H}_{D_k} Y_{D_k,1} \hat{H}_{D_k}^H| - \text{Tr}((\nu I - \Phi) Y_{D_k,1}) + (\nu - \lambda)p_{D_k}
\]

\[
= \log|I + \hat{H}_{D_k} Y_{D_k,1} \hat{H}_{D_k}^H| - \text{Tr}(Y_{D_k,1}') + (\nu - \lambda)p_{D_k}
\]

where \(\nu\) is the dual variable associated with the constraint, \(\Phi = \nu I - \Phi\), \(\hat{H}_{D_k} = \hat{H}_{D_k} \Phi^{-1/2}\), \(Y_{D_k,1}' = \Phi^{1/2} Y_{D_k,1} \Phi^{-1/2}\). Let \(\hat{H}_{D_k} = U_{D_k} \text{diag}(\gamma_{1,1}, \gamma_{1,2}, \ldots, \gamma_{r_{D_k},r_{D_k}}) \tilde{V}_{D_k}^H\) be an SVD of \(\hat{H}_{D_k}\), where \(r_{D_k} = \text{rank}(\hat{H}_{D_k} \hat{H}_{D_k}^H)\). Then, the optimal solution to Subproblem 1 can be found as

\[
\nu^* = \lambda,
\]

\[
p_{D_k}^* = \sum_{i=1}^{r_{D_k}} \left[1 - \frac{1}{\gamma_{i,i}^2}\right] m_i,
\]

\[
Y_{D_k,1}^* = \Phi^{1/2} U_{D_k} D_{D_k} \tilde{V}_{D_k}^H \Phi^{-1/2}
\]

(59)

where \(m_i = \left[\tilde{V}_{D_k}^H \Phi^{-1} \tilde{V}_{D_k}\right]_{i,i}\), and \(D_{D_k} = \text{diag}(\left[1 - 1/\gamma_{1,i}^2\right]^+, \ldots, \left[1 - 1/\gamma_{r_{D_k},i}^2\right]^+)\). We note that for Subproblem 1 to be solvable, \(\Phi\) must satisfy the constraint \(\Phi < \lambda I\).

Similarly, let \(\Phi_2 = \hat{G}_{S_k}^H \hat{G}_{S_k} + \Phi\), \(\hat{G}_{S_k} = \hat{G}_{S_k} \Phi_2^{-1/2}\) and \(Y_{D_k,2} = \Phi_2^{1/2} Y_{D_k,2} \Phi_2^{-1/2}\), the objective of (57) is rewritten as

\[
\log|I + \hat{G}_{S_k} Y_{D_k,2} \hat{G}_{S_k}^H| - \text{Tr}(Y_{D_k,2}')
\]

(60)

and thus the optimal \(Y_{D_k,2}^*\) of Subproblem 2 is simply given by

\[
Y_{D_k,2}^* = \Phi_2^{1/2} U_{S_k} D_{S_k} \tilde{V}_{S_k}^H \Phi_2^{-1/2}
\]

(61)

where \(\hat{G}_{S_k} = U_{S_k} \text{diag}(\gamma_{2,1}, \ldots, \gamma_{2,r_{S_k}}) \tilde{V}_{S_k}^H\) is the SVD of \(\hat{G}_{S_k}\), \(r_{S_k} = \text{rank}(\hat{G}_{S_k}^H \hat{G}_{S_k})\), and \(D_{S_k} = \text{diag}(\left[1 - 1/\gamma_{2,i}^2\right]^+, \ldots, \left[1 - 1/\gamma_{r_{S_k},i}^2\right]^+)\). We also note that Subproblem 2 is solvable if \(\Phi > -\hat{G}_{S_k}^H \hat{G}_{S_k}\).
We have shown that Subproblems 1 and 2 can be solved efficiently using the water-filling algorithm with the fixed water level. The problem now is to find the optimal $\Phi$ which minimizes the dual problem of (53), which is formulated as

$$\begin{align*}
\text{minimize} & \quad g(\Phi) = g_1(\Phi) + g_2(\Phi) \\
\text{subject to} & \quad F_1(\Phi) < 0, \\
& \quad F_2(\Phi) < 0
\end{align*}$$

where $F_1(\Phi) \doteq \Phi - \lambda I$ and $F_2(\Phi) \doteq -\Phi - \hat{G}_S^H \hat{G}_S$. The two constraints in (62) are actually the ones to make Subproblems 1 and 2 solvable as mentioned earlier. To solve the problem (62), we apply the subgradient method for constrained optimization presented in [92]. In this way, $\Phi$ is iteratively updated until convergence is reached. To be specific, at the $\kappa + 1$st iteration, $\Phi^{(\kappa+1)}$ is expressed as

$$\Phi^{(\kappa+1)} = \Phi^{(\kappa)} - \delta^{(\kappa)} \Delta \Phi$$

where $\delta^{(\kappa)}$ is a positive step size, and $\Delta \Phi$ is a proper subgradient. Explicitly, if $\Phi^{(\kappa)}$ is feasible, $\Delta \Phi$ is a subgradient of $g(\Phi)$, i.e., $\Delta \Phi = Y_{d_k,1} - Y_{d_k,2}$. Otherwise, $\Delta \Phi$ is set to be a subgradient of the violated constraint function at $\Phi^{(\kappa)}$ as follows [92].

$$\Delta \Phi = \begin{cases} 
    Y_{d_k,1} - Y_{d_k,2} & \text{if } -\hat{G}_S^H \hat{G}_S < \Phi^{(\kappa)} - \lambda I \\
    I & \text{if } \Phi^{(\kappa)} \geq \lambda I \\
    -I & \text{if } \Phi^{(\kappa)} \leq -\hat{G}_S^H \hat{G}_S.
\end{cases}$$

The step size $\delta^{(\kappa)}$ rule is chosen in many ways. Particularly, the value of our step size is chosen as [92]

$$\delta^{(\kappa)} = \begin{cases} 
    \omega & \text{if } -\hat{G}_S^H \hat{G}_S < \Phi^{(\kappa)} - \lambda I, \\
    \frac{\phi_{\max}(F_1(\Phi^{(\kappa)})) + \theta}{N_{d_k}} & \text{if } \Phi^{(\kappa)} \geq \lambda I, \\
    \frac{\phi_{\max}(F_2(\Phi^{(\kappa)})) + \theta}{N_{d_k}} & \text{if } \Phi^{(\kappa)} \leq -\hat{G}_S^H \hat{G}_S.
\end{cases}$$

where $\omega$ is the constant that can be chosen sufficiently small, $\phi_{\max}(F_i(\Phi^{(\kappa)}))$ is the largest eigenvalue of $F_i(\Phi^{(\kappa)})$ at $\Phi^{(\kappa)}$, and $\theta$ is a small positive margin. The proposed dual decomposition method to solve problem ($P_2$) is summarized in Algorithm 3.
Algorithm 3 A dual decomposition method for solving (P2). [37] ©2013 IEEE.

1: Initialize $\Phi^{(0)}$ such that $-\hat{G}_S^H \hat{G}_S < \Phi^{(0)} < \lambda I$.
2: $\kappa := 0$.
3: repeat
4: Compute $Y^{(\kappa)}_{d_k,1}$ and $Y^{(\kappa)}_{d_k,2}$ using (59) and (61).
5: Update $\Delta \Phi$ and $\delta^{(\kappa)}$ using (64) and (65).
6: $\kappa := \kappa + 1$.
7: until convergence
8: $Q^*_{d_k} = Y^*_{d_k,1} = Y^*_{d_k,2}$.

3.3 Energy efficiency maximization

Clearly, the main use of the FD transmission in wireless communications systems is to improve their SE performance. While the SE has been recognized from the early development of wireless communications as one of the most important design criteria, the EE, defined in Section 1.2, is another metric that has drawn much attention recently due to increasing interest in green wireless networks. Thus, it is interesting to investigate the EE of the FD system considered in this chapter. We note that since the DL and UL channels in the FD system operate simultaneously in an active mode, the FD system consumes more energy than a traditional HD one. Thus, the EE of the FD system can be much smaller than that of the HD system. For this purpose, we continue with the problem of precoder design for maximizing EE of the FD system with the BD scheme being used in the DL channel and evaluate its performance numerically in the next subsection.

3.3.1 Power consumption model

The energy spent to send data comes from many hardware components involved in the data transmission [93–95]. According to [94] and [96], apart from data-dependent transmit power, the circuit power consumption, dissipated in all other electronics devices for signal processing (such as mixer, filter, analog-to-digital converter (ADC), digital-to-analog converter (DAC), low-noise amplifier (LNA), etc.) also plays an important role in EE performance. Noticeably, for short range communications such as micro or femto cells, the circuit power consumption can be comparable to or even dominate the actual transmit power for the data transmission. In addition, the EE can be
expressed as [78]

\[ \eta = \frac{SE \text{ or } \text{Sum rate}}{\text{Total power consumption in transmit duration}}. \] (66)

Therefore, a power consumption model of wireless transmission is very important and has a primary effect on EE evaluation. However, there is no study on a power consumption model for FD transceivers. From a modeling perspective, it is reasonable to assume the power consumption model of FD transceivers is quite similar to that of traditional ones, except that the value of each component of the model such as the circuit power consumption is different because of the SI cancellation mechanism. Herein, we adopt linear power model as presented in [94] and [96] for simplicity. In particular, the total power consumption at the BS in the DL transmission is modeled as

\[ P_{\text{Total}}^{\text{BS}} = \frac{1}{\varepsilon} P_{\text{TX}}^{\text{BS}} + P_{\text{CIR}}^{\text{BS}} \] (67)

where \( \varepsilon \in [0, 1] \) is the power amplifier efficiency which depends on the design and implementation of the power amplifier, \( P_{\text{TX}}^{\text{BS}} \) is the transmit power determined by linear precoders, i.e., \( P_{\text{TX}}^{\text{BS}} = \sum_{i=1}^{K} \text{Tr}(Q_{D_i}) \). In (67), \( P_{\text{CIR}}^{\text{BS}} = N_l P_{\text{DYN}}^{\text{BS}} + P_{\text{STA}}^{\text{BS}} \) is called the circuit power, where \( P_{\text{DYN}}^{\text{BS}} \) is the dynamic circuit power consumption, corresponding to the power dissipation of all circuit blocks and proportional to the number of the transmit antennas, and \( P_{\text{STA}}^{\text{BS}} \) is the static circuit power spent by cooling system, power supply, etc. Similarly, the total power consumed at the transmitter of the \( i \)th user in the UL channel is denoted as

\[ q_{\text{Total}}^{U_i} = \frac{1}{\varepsilon} q_{\text{TX}}^{U_i} + q_{\text{CIR}}^{U_i} \] (68)

where \( q_{\text{TX}}^{U_i} = \text{Tr}(Q_{U_i}) \) is the transmit power allocated for data transmission, and \( q_{\text{CIR}}^{U_i} = N_{U_i} q_{\text{DYN}}^{U_i} + q_{\text{STA}}^{U_i} \) is the circuit power.

### 3.3.2 Problem formulation

In this subsection, our key objective is to optimize the EE by jointly designing linear precoders under SPC in the DL channel and PUPCs in the UL channel for the considered FD MU-MIMO system, referred to as the EE-optimal design. For simplicity, we only consider EE at the transmitter side. Since the DL and UL channels operate simultaneously, we propose a definition of the overall EE of the FD MU-MIMO system.
as

\[ \eta_{\text{sum}} = \frac{R_0 + R_d}{\frac{1}{e} \sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) + \frac{1}{e} \sum_{j=1}^{K_U} \text{Tr}(Q_{U_j}) + P_{\text{cir}}^c} \]  

(69)

where \( P_{\text{cir}}^c = P_{\text{BS}} + \sum_{j=1}^{K_U} q_{U_j}^c \) is the overall circuit power of the system, assumed to be constant for simplicity. The overall EE given in (69) is the actual EE that the FD can achieve and does not take into account the fairness of EE among users and the BS. For this problem we may consider the problem of maximizing the weighted sum of individual EE of all the agents in the system. However, this general case is not considered in the thesis and left as future work.

The performance measure now is \( f((Q_{D_i}), (Q_{U_j})) = \eta_{\text{sum}} \) and the problem of interest (40) can be reformulated as

\[
\begin{aligned}
\text{maximize} & \quad \eta_{\text{sum}} \\
\text{subject to} & \quad \sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) \leq P_{\text{BS}}, \\
& \quad \text{Tr}(Q_{U_j}) \leq q_{U_j}, \quad \forall j = 1, \ldots, K_U, \\
& \quad Q_{D_i} \succeq 0, \quad \forall i = 1, \ldots, K_D, \\
& \quad Q_{U_j} \succeq 0, \quad \forall j = 1, \ldots, K_U.
\end{aligned}
\]  

(70)

We note that, like (41), problem in (70) is also nonconvex. Fortunately, we will show shortly that the convex approximation method presented in Subsection 3.2.1 for maximizing SE is useful to obtain a locally optimal solution to (70).

3.3.3 Proposed joint design for EE maximization

We first observe that the denominator in (69) is a linear function of \( (Q_{D_i}) \) and \( (Q_{U_j}) \). From results reported in [77–79], problem (70) becomes tractable if the denominator, i.e., \( R_0 + R_d \) in (69) is a concave function of \( (Q_{D_i}) \) and \( (Q_{U_j}) \). Motivated by this observation, we propose to iteratively replace \( R_0 + R_d \) by the lower bound given in (42). In this way, after \( n \) iterations, the EE of the considered FD MU-MIMO system is lower bounded by

\[ \eta_{\text{sum}} \geq \frac{R_{\text{LB}}^n}{\frac{1}{e} \sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) + \frac{1}{e} \sum_{j=1}^{K_U} \text{Tr}(Q_{U_j}) + P_{\text{cir}}^c} \]  

(71)

\(^4\)In the thesis, we use the unit of EE as nats/J for the sake of mathematical convenience.
where $R_{lb}^{(n)}$ is the right hand side of (42). Consequently, the values of $\{Q_{bi}\}$ and $\{Q_{uj}\}$ in the $(n+1)$th iteration are the solution to the following problem

\[
\begin{align*}
\text{maximize} & \quad R_{lb}^{(n)} \\
\text{subject to} & \quad \frac{1}{\epsilon} \sum_{i=1}^{K_b} \text{Tr}(Q_{bi}) + \frac{1}{\epsilon} \sum_{j=1}^{K_u} \text{Tr}(Q_{uj}) + \frac{P_{\text{sum}}}{\epsilon} \\
& \quad \sum_{i=1}^{K_b} \text{Tr}(Q_{bi}) \leq P_{\text{BS}}, \\
& \quad \text{Tr}(Q_{uj}) \leq q_{uj}, \quad \forall j = 1, \ldots, K_u, \\
& \quad Q_{bi} \geq 0, \quad \forall i = 1, \ldots, K_b, \\
& \quad Q_{uj} \geq 0, \quad \forall j = 1, \ldots, K_u.
\end{align*}
\] (72)

For mathematical convenience, let us define $f_1(\{Q_{bi}\}, \{Q_{uj}\}) = R_{lb}^{(n)}$, $f_2(\{Q_{bi}\}, \{Q_{uj}\}) = \frac{1}{\epsilon} \sum_{i=1}^{K_b} \text{Tr}(Q_{bi}) + \frac{1}{\epsilon} \sum_{j=1}^{K_u} \text{Tr}(Q_{uj}) + \frac{P_{\text{sum}}}{\epsilon}$, and rewrite (72) as

\[
\begin{align*}
\text{maximize} & \quad f_1(Q_{bi}, Q_{uj}) \\
\text{subject to} & \quad \frac{1}{\epsilon} \sum_{i=1}^{K_b} \text{Tr}(Q_{bi}) \leq P_{\text{BS}}, \\
& \quad \text{Tr}(Q_{uj}) \leq q_{uj}, \quad \forall j = 1, \ldots, K_u, \\
& \quad Q_{bi} \geq 0, \quad \forall i = 1, \ldots, K_b, \\
& \quad Q_{uj} \geq 0, \quad \forall j = 1, \ldots, K_u.
\end{align*}
\] (73)

We recall that $f_1(\{Q_{bi}\}, \{Q_{uj}\})$ is a concave function of $\{Q_{bi}\}$ and $\{Q_{uj}\}$, and $f_2(\{Q_{bi}\}, \{Q_{uj}\})$ is an affine function of $\{Q_{bi}\}$ and $\{Q_{uj}\}$. Accordingly, problem (73) belongs to the class of concave-convex fractional programs. There are several algorithms to globally solve concave-convex fractional programs such as the parametric convex program, parameter-free convex program or duality program, etc. [77–79]. For simplicity, we adopt the parametric convex approach to solve problem (73). In this way, for a fixed $\alpha > 0$, we define the following concave function $\phi_\alpha(\{Q_{bi}\}, \{Q_{uj}\})$ as

\[
\phi_\alpha(\{Q_{bi}\}, \{Q_{uj}\}) = f_1(\{Q_{bi}\}, \{Q_{uj}\}) - \alpha f_2(\{Q_{bi}\}, \{Q_{uj}\})
\] (74)

and consider the following problem

\[
\begin{align*}
\text{maximize} & \quad \phi_\alpha(\{Q_{bi}\}, \{Q_{uj}\}) \\
\text{subject to} & \quad \frac{1}{\epsilon} \sum_{i=1}^{K_b} \text{Tr}(Q_{bi}) \leq P_{\text{BS}}, \\
& \quad \text{Tr}(Q_{uj}) \leq q_{uj}, \quad \forall j = 1, \ldots, K_u, \\
& \quad Q_{bi} \geq 0, \quad \forall i = 1, \ldots, K_b, \\
& \quad Q_{uj} \geq 0, \quad \forall j = 1, \ldots, K_u.
\end{align*}
\] (75)
Algorithm 4 The proposed EE-optimal design. \cite{37} ©2013 IEEE.

1: Randomly initialize \(\{Q_{D_i}^{(0)}\}_{i=1}^{K_D}\); tolerance \(\epsilon > 0\).
2: \(n := 0\).
3: repeat
4: Randomly initialize \(\alpha_{\text{min}}\) and \(\alpha_{\text{max}}\) such that \(J(\alpha_{\text{min}}) < J(\alpha_{\text{max}}) < 0\).
5: while \((\alpha_{\text{max}} - \alpha_{\text{min}}) > \epsilon\) do
6: \(\alpha := \frac{1}{2}(\alpha_{\text{min}} + \alpha_{\text{max}})\)
7: Solve (75) using the proposed algorithm presented in Section 3.2 to find optimal solutions \(\{Q_{D_i}^*\}_{i=1}^{K_D}\), \(\{Q_{U_j}^*\}_{j=1}^{K_U}\) and \(J(\alpha)\).
8: if \(J(\alpha)J(\alpha_{\text{min}}) < 0\) then
9: \(\alpha_{\text{max}} = \alpha\)
10: else
11: \(\alpha_{\text{min}} = \alpha\)
12: end if
13: end while
14: \(n := n + 1\).
15: Update covariance matrices: \(\{Q_{D_i}^{(n)}\}_{i=1}^{K_D} = \{Q_{D_i}^*\}_{i=1}^{K_D}\).
16: until convergence.
17: Calculate the linear precoder for the user \(i\) in the DL channel: \(W_{D_i} = \bar{V}_{D_i} M_{D_i}\), where \(M_{D_i}\) is obtained from Cholesky decomposition to \(Q_{D_i}^{(n)}\).

The optimal value of (75) for a given \(\alpha\) is denoted by \(J(\alpha)\). A result of \cite{79} states that \(\{Q_{D_i}\}\) and \(\{Q_{U_j}\}\) are optimal solutions to (73) if and only if \(J(\alpha) = 0\). Moreover, it can be shown that \(J(\alpha)\) is strictly decreasing in \(\alpha\), meaning that the equation \(J(\alpha) = 0\) can be solved efficiently using bisection method over \(\alpha\).

From the expression of \(\phi_\alpha(\{Q_{D_i}\}, \{Q_{U_j}\})\) in (74), it is clear that the convex problem in (75) is similar to the formulation of the SE maximization in (43). Hence, our proposed algorithm for SE maximization presented in Section 3.2 can be slightly modified to solve (75) for a fixed \(\alpha\). The proposed algorithm for EE maximization is outlined in Algorithm 4.
3.4 Implementation complexity

We now provide some remarks concerning the implementation issues of the proposed solutions in this chapter. For the considered system model the BS can attain the CSI of UL users directly. The CSI of DL users can be made available to the BS by two ways. In the first way, DL users can estimate the CSI and report the obtained value to the BS by a feedback control channel, which certainly incurs feedback overhead and the quality of the CSI depends on the capacity of the feedback channel. In the second one, given that the channel coherence time is larger than the transmission time, the BS can estimate CSI itself by exploiting the channel reciprocity. Thus there is no signaling overhead to acquire CSI at the BS in this case.

From a signal processing viewpoint the computational complexity of proposed solutions is mainly determined by solving the considered optimization problems. We note that the proposed solutions are iterative methods and the complexity in each iteration is dominated by the SVD which takes $O(N_R^2 N_U^2 + N_U^3)$ floating-point operations (flops), where $N_U = \max_{1 \leq i \leq K_U} N_U$. We remark that their total complexity depends on the number of iterations that is required to converge. However, it is difficult, if not impossible, to analytically find the precise (or even an upper bound) number of iterations for the proposed solutions to satisfy a desired accuracy level. In the following section we will numerically evaluate the convergence rate of the proposed algorithms.

3.5 Performance evaluation

3.5.1 System parameters

In this subsection, we carry out numerical simulations to investigate the potential gains of the FD MU-MIMO system with the proposed designs over the conventional HD MU-MIMO scheme in terms of SE and EE. The channels are assumed to be standard Rayleigh fading of unit variance, i.e., the channel coefficients of $H_D$ and $H_U$ are modeled as i.i.d. complex Gaussian random variables with zero mean and unit variance. The entries of $G_S$ are simply generated as $CN(0, \sigma_{SI}^2)$, where $\sigma_{SI}^2$ represents the capability of the SI cancellation technique. Unless otherwise mentioned, the number of users in the DL and UL channels is set to $K_D = K_U = 2$ with two antennas for each user in the DL and UL channels, and the BS is equipped with $N = 8$ antennas, four of which are used in the DL and UL channels, respectively, i.e., $N_T = 4$ and $N_R = 4$. The
initial values for \( \{Q_k^{(0)}\} \) in the proposed iterative method are generated randomly for each channel realization. For the sake of simplicity, we assume \( q_k = q_U, q_k^{\text{dyn}} = q_U^{\text{dyn}}, \) and \( q_k^{\text{sta}} = q_U^{\text{sta}} \) for all \( k \) and \( q = \beta P_{BS}, \) and \( q^{\text{dyn}} = \beta P_{BS}^{\text{dyn}}, \) where \( \beta \in [0, 1]. \) In addition, the noise power of each DL user and BS is taken as unity, i.e., \( N_0 = 1. \) The parameters in the linear power model are set to \( \varepsilon = 0.2, \beta = 0.7, q^{\text{sta}} = 20 \text{ dBm} \) and \( P^{\text{sta}} = 27 \text{ dBm}. \) For a fair comparison, the BS in the HD counterpart is allowed to use all the same antennas to transmit or receive data one at a time, i.e., \( N = N_T + N_R. \) While this assumption can enable the HD counterpart to exploit all the available degrees of freedom provided by the system model, we remark that it comes at a cost that the number of RF chains in the HD counterpart is doubled. We also note that the performance of the HD system is obtained based on maximizing its SE.

### 3.5.2 Convergence results

Fig. 4 shows the convergence rate of our proposed SE- and EE-optimal algorithms for a random channel realization. At the BS, the maximum transmit power is set to \( P_{BS} = 33 \text{ dBm}, \) dynamic circuit power to \( P^{\text{dyn}}_{BS} = 38 \text{ dBm}. \) The \( \sigma_{SI}^2 \) is fixed at 0 dB. Each point in Figs. 4a and 4b is computed using Algorithms 2 and 3, in which the error tolerance for convergence is set to \( 10^{-6}. \) For Algorithm 3, we use a constant step size \( \omega = 0.00125. \) It is observed that our proposed algorithms for maximizing SE (Algorithm 1) and EE (Algorithm 4) exhibit monotonic convergence to a locally optimal point within a few iterations. This convergence rate is the same for other channel realizations.

### 3.5.3 Spectral efficiency

In Fig. 5, we plot the overall SE of the FD and conventional HD systems with the SE-optimal designs versus SI strength \( \sigma_{SI}^2 \) for a fixed maximum transmit power at the BS \( P_{BS} = 42 \text{ dBm} \) for two scenarios.\(^5\) In the first setup, the number of antennas at the BS is set to \( N_T = N_R = 4, \) and the DL and UL channels serve two users each. In the second one, the number of the BS is reduced to \( N_T = N_R = 2, \) and the DL and UL channels serve one user each. The purpose of considering the two cases is to study the spatial multiplexing gain provided by the FD MU-MIMO, i.e., how the SE varies with \( N_T \) and \( N_R. \)

\(^5\)For the SE-optimal design of the HD system, the conventional water-filling procedure is employed to maximize the SE of the BD scheme in the DL channel [25], and the iterative water filling algorithm is applied to obtain the SE of the UL channel [97].
Fig. 4. Convergence rate of Algorithm 1 (for SE maximization) and Algorithm 4 (for EE maximization).

[37] © 2013 IEEE.
As can be seen in Fig. 5, the SE of the FD system with Algorithm 1 is greatly improved, about 25% when $\sigma_{SI}^2$ is relatively small. It is worth mentioning that even when the residual SI is negligible, the FD system cannot attain a double gain of SE as compared to the HD system. This is due to the fact that the BS of the HD system is allowed to use all antennas to send data in the DL channel (while that of the FD system only uses half of them). Moreover, the SE of the FD system is degraded and becomes lower than that of HD system as the residual SI increases. The main reason is that if the SI suppression is not efficient enough, the high power from the transmitted signals in the DL channel overwhelms the received signals in the UL channel, which makes it erroneous to recover the signals from users in the UL channel. This results in a loss of the SE of the FD system. We also observe that the SE of the FD system for the case $N_T = N_R = 4$ nearly doubles that for the case $N_T = N_R = 2$, implying that the proposed precoder design for the FD system can exploit the spatial multiplexing gain.

Fig. 6 compares the SE of the FD and conventional HD systems for a fixed $\sigma_{SI}^2 = -20$ dB, but in this case we vary the maximum transmit power at the BS, $P_{BS}$ (the maximum
transmit power at each user in the UL channel, i.e., $q_0$, is changed accordingly due to their relationship mentioned earlier). We can see that increasing $P_{BS}$ also leads to a significant gain of the SE of the FD over that of the HD system.

### 3.5.4 Energy efficiency

In Fig. 7, we show the comparison between the EE of the considered FD system with both the EE-optimal and SE-optimal schemes and that of the HD one as a function of $\sigma_{SI}^2$. At the BS, we set the maximum transmit power to be $P_{BS} = 42$ dBm and the dynamic circuit power to $P_{BS}^{dyn} = 38$ dBm. An interesting observation is that although the FD system with the SE-optimal design increases the SE but it consumes more energy compared to the HD model as shown in Figs. 5 and 7. This can be explained by the fact that the DL and UL channels of the FD system are active continuously while those of the HD one operate in an on-off fashion. Hence, the EE-optimal scheme is proposed to enhance EE of the FD network. In Fig. 7, we have numerically found that the EE gain...
of the EE-optimal design is significantly improved compared to the SE-optimal one and even better than the HD system as the residual SI power is small.

In Fig. 8, we investigate the EE of the FD system with the SE- and EE-optimal designs as a function of the dynamic circuit power $P_{\text{Dyn}}^{\text{BS}}$ at the BS. The maximum transmit power of the BS is set to $P_{\text{BS}} = 42$ dBm and $\sigma_{\text{SI}}^2$ is fixed at $-20$ dB. As mentioned above, the circuit power model of the considered FD system is defined as $P_{\text{cir}} = P_{\text{cir}}^{\text{BS}} + \sum_{j=1}^{K_0} q_{j}^{\text{cir}}$ where $P_{\text{cir}}^{\text{BS}} = N_T P_{\text{Dyn}}^{\text{BS}} + P_{\text{sta}}^{\text{BS}}$, $q_{0}^{\text{cir}} = N_0 q_{0}^{\text{dyn}} + q_{0}^{\text{sta}}$, and $q_{0}^{\text{dyn}} = \beta P_{\text{Dyn}}^{\text{BS}}$. Clearly, the SE-optimal precoder design is inferior to the EE-optimal one. This is because the SE-optimal precoder design does not take into account the effect of other sources of power consumption. We observe that the EE of the two designs is degraded as $P_{\text{Dyn}}^{\text{BS}}$ increases. When $P_{\text{Dyn}}^{\text{BS}}$ is small, the EE-optimal precoder design yields the significant gain over the SE-optimal design. When $P_{\text{Dyn}}^{\text{BS}}$ is sufficiently large, the SE- and EE-optimal designs obtain nearly the same EE performance. We note that if $P_{\text{Dyn}}^{\text{BS}}$ is large, the totally consumed power is mostly constituted by the circuit power. Thus, from (69), we can see that maximizing the EE is equivalent to maximizing the SE.
To obtain further insights into the performance of the transmission designs, we evaluate the EE of three approaches with the maximum transmit power at the BS $P_{BS}$ as shown in Fig. 9. The dynamic circuit power of the BS $P_{BS}^{\text{dyn}}$ is fixed at 38 dBm and the residual SI channel gain is chosen to be $\sigma_{SI}^2 = -20$ dB. A general observation is that in terms of EE the HD paradigm outperforms the FD SE-optimal scheme. Nonetheless, compared to the HD design, the FD EE-optimal design yields a better EE gain in the high transmit power regime but a worse one in the low transmit power region. In addition, we see that in the low transmit power regime the EE of the two proposed designs and HD model increases as $P_{BS}$ increases. This is due to the fact that the power consumption is largely determined by the circuit power consumption in this region. An increase in transmit power leads to an improvement on the SE, and thus also on the EE. Particularly, the EE of the FD SE-optimal design and HD model achieves maximum EE for a certain transmit power, and decreases after that. This is because the SE-optimal/HD designs always transmit with full power to maximize the SE. However, the SE increases only logarithmically while the total power consumption grows up linearly with the transmit
power (in high transmit power regime). As a result, the EE of the SE-optimal/HD designs are reduced. On the other hand, the EE of the FD EE-optimal one remains constant after reaching its peak value. By the optimization mechanism of the EE-optimal design, the EE-optimal design can find an optimal transmit power \( P_{\text{BS}}^* \). If \( P_{\text{BS}}^* < P_{\text{BS}} \), the EE-optimal design will not transmit at full power, and thus its EE remains unchanged.

### 3.6 Summary and discussion

In this chapter, we have investigated potential gains of the FD MU-MIMO system in which the DL and UL channels are designed to operate in the FD mode and both the BS and users are equipped with multiple antennas. Taking the natural coupling of the DL and UL channels into consideration by the SI and assuming the CCI negligible, we have proposed joint designs of linear precoders to optimize SE and EE, referred to as the SE-optimal and EE-optimal designs, respectively. These linear precoding schemes were designed under a SPC in the DL channel and PUPCs in the UL channel. Particularly, for the SE-optimal design, since the problem is formulated as a nonconvex program, we have proposed an iterative algorithm to find the suboptimal solutions based on a convex
relaxation method. Particularly, in each iteration, the relaxed convex program is solved using the alternating and dual decomposition method, by which we obtain the analytical solutions for precoders with given dual variables. For the EE-optimal design, we first approximate the nonlinear fractional program of EE as a concave-convex fractional program, which is then transformed into a parametric convex program by applying the parametric approach. We have shown that the resulting optimization problem for EE maximization can be efficiently solved by applying the same techniques as in the SE-optimal design.

Numerical results indicate that the proposed SE-optimal design for the considered FD MU-MIMO system achieves better SE, but is not double as theoretically expected, when the power level of the residual SI is small compared to the HD one. The reason is that the BS of the FD system can only use half of the total number of antennas to transmit and receive data simultaneously, while the BS of the HD counterpart system can use all of them to send or receive data one at a time. Additionally, although the SE-optimal design for the FD network improves the SE of the HD one, it consumes much more energy. Hence, the EE-optimal design was proposed. The simulation results demonstrate that the EE-optimal design is superior to the SE-optimal scheme in terms of EE, and yields lower EE than the HD system in the low transmit power region. In contrast, the EE of the FD system with the EE-optimal design is better than the one of the HD design in the high transmit power regime as the residual SI power is low. In particular the FD system with the EE-optimal design can improve the EE of the HD system by 9% and that of the SE-optimal design by 45%.
4 Efficient beamforming solutions for full-duplex assisted small-cell networks

In the previous chapter linear precoder designs were proposed in the absence of CCI. Moreover, the performance evaluation was carried out under the assumption of small scale fading channels, i.e., the pathloss was not taken into account. Towards a more realistic study we consider in this chapter the FD radios under the specific context of small-cell networks which was identified as one of several potential deployment scenarios that may benefit from FD communications [98]. In fact, small-cell systems are considered to be especially suitable for deployment of FD technology due to low transmit powers, short transmission distances and low mobility.

To explore the achievable gains of the FD small-cell networks, this chapter is concerned with the problem of linear beamforming design taking into consideration both the residual SI and CCI. To this end, the total SE maximization problem is first formulated as a rank-constrained optimization problem for which it is difficult to find globally optimal solutions in general. The standard method of rank relaxation is then applied to arrive at a relaxed problem, which is still nonconvex. After solving the relaxed problem, the randomization technique presented in [99] is employed to find the beamformers for the original design problem.

To tackle the nonconvexity of the relaxed problem, we propose two iterative local optimization algorithms. The first proposed algorithm is a direct result of exploiting the ‘difference of convex’ (D.C) structure of the relaxed problem. To be specific, based on the idea of the Frank-Wolfe algorithm [73], we arrive at a determinant maximization (MAXDET) program at each iteration. The second design approach involves some transformations before invoking the framework of SPCA method [80], which has proven to be an effective tool for numerical solutions of nonconvex optimization problems [70, 80, 87]. In particular, we are able to approximate the relaxed problem as a SDP at each iteration in the second iterative algorithm. While the first design algorithm sticks to MAXDET problem solvers, the second one offers more flexibility in choosing optimization software and can take advantage of many state-of-the-art SDP solvers.

FD transmission, if successfully implemented, is clearly expected to improve the SE of small-cell communications systems. However, a quantitative answer on the potential

---

*This chapter has been modified from [38] ©2014 IEEE.*
Fig. 10. A small-cell FD wireless communications system. The number of transmit and receive antennas at the BS is $N_T$ and $N_R$, respectively. Linear beamforming is adopted for the DL channel, while MMSE-SIC for the UL channel. [38] ©2014 IEEE.

gains for some particular scenarios is still missing. For this purpose, the proposed algorithms are used to evaluate the performance of the FD system of consideration under the 3GPP LTE specifications for a small-cell system. The numerical experiments demonstrate that small-cell FD transmissions are superior to the conventional HD counterpart as long as the residual SI power is sufficiently reduced.

4.1 System model and problem formulation

We consider a small-cell FD wireless communications system in which a FD capable BS is designed to communicate with $K_D$ single-antenna users in the DL channel and $K_U$ single-antenna users in the UL channel at the same time over the same frequency band, as depicted in Fig. 10. As stated in Subsection 2.2, the notations $D_i$ and $U_j$ refer to the $i$th and $j$th user in the DL and UL channels, respectively. The total number of antennas at the FD BS is $N = N_T + N_R$, of which $N_T$ transmit antennas are used for data transmissions in the DL channel and $N_R$ receive antennas are dedicated to receiving data in the UL channel. Additionally, we assume that the DL and UL channels are flat fading and CSI is perfectly known at both the BS and users.
First, in the DL channel, let $s_D$ be the transmitted data symbol for $D_i$, which is normalized to $E(|s_D|^2) = 1$. For linear beamforming, the data symbol $s_D$ is multiplied by the beamforming vector $w_D \in \mathbb{C}^{N_T \times 1}$ before transmission, and the received signal of user $D_i$ is given by

$$y_{D_i} = h_{D_i}^H w_D s_D + \sum_{k \neq i}^K h_{D_i}^H w_D s_k + \sum_{j=1}^{K_U} g_{ji} s_{U_j} + n_{D_i}$$  \hspace{1cm} (76)$$

where $h_D$ is the $N_T \times 1$ complex channel vector from the BS to user $D_i$, $g_{ji}$ is the complex channel coefficient from $U_j$ to $D_i$, $s_{U_j}$ is the data symbol transmitted by $U_j$ in the UL direction, and $n_{D_i} \sim \mathcal{CN}(0, \sigma_n^2)$ is background noise assumed to be AWGN. In (76), the first and second summations represent MUI in the DL channel and CCI from the UL to the DL channels, respectively. The received SINR of user $D_i$ can be written as

$$\gamma_{D_i} = \frac{|h_{D_i}^H w_D|^2}{\sigma_n^2 + \sum_{k \neq i}^K |h_{D_i}^H w_D|^2 + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2} \hspace{1cm} (77)$$

where $E(|s_{U_j}|^2) = q_{U_j}$, $j = 1, \ldots, K_U$, is power loading for user $U_j$ in the UL direction; $Q_D = w_D w_D^H$, and rank$(Q_D) = 1$. Then, SE in the DL channel is given by

$$R_0 = \sum_{i=1}^{K_D} \log(1 + \gamma_{D_i}) \hspace{1cm} (78a)$$

and

$$= \sum_{i=1}^{K_D} \log \left( \frac{\sigma_n^2 + \sum_{k \neq i}^K |h_{D_i}^H Q_D h_D|^2 + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2}{\sigma_n^2 + \sum_{k \neq i}^K |h_{D_i}^H Q_D h_D|^2 + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2} \right) \hspace{1cm} (78b)$$

Next, for the UL transmission, we can express the received signal vector at the FD BS as

$$y_U = \sum_{j=1}^{K_U} h_{U_j} s_{U_j} + \sum_{i=1}^{K_D} H_{SI} w_D s_D + n_U$$ \hspace{1cm} (79)$$

where $h_{U_j} \in \mathbb{C}^{N_R \times 1}$ is the complex channel vector from the BS to $U_j$ and $n_U \sim \mathcal{CN}(0, \sigma_n^2 I_{N_R})$. The matrix $H_{SI}$ is called the residual SI channel at the FD BS. In this case, by treating
the residual SI as background noise and applying the MMSE-SIC decoder, we can write the received SINR of $U_j$ as [72]

$$\gamma_{U_j} = q_{U_j} h_{U_j}^H \left( \sigma_n^2 I + \sum_{m > j} q_{U_m} h_{U_m} h_{U_m}^H + \sum_{i=1}^{K_D} H_{SI} Q_{D_i} H_{SI}^H \right)^{-1} h_{U_j}$$

(80)

where we have assumed a decoding order from 1 to $K_U$. Consequently, the achievable SE in the UL channel is given by [72]

$$R_U = \sum_{j=1}^{K_U} \log \left( 1 + \gamma_{U_j} \right)$$

(81a)

$$= \sum_{j=1}^{K_U} \log \left( 1 + q_{U_j} h_{U_j}^H \left( \sigma_n^2 I + \sum_{m > j} q_{U_m} h_{U_m} h_{U_m}^H + \sum_{i=1}^{K_S} H_{SI} Q_{D_i} H_{SI}^H \right)^{-1} h_{U_j} \right)$$

(81b)

$$\geq \log \frac{\left| \sigma_n^2 I + \sum_{i=1}^{K_S} H_{SI} Q_{D_i} H_{SI}^H \right|}{\left| \sigma_n^2 I + \sum_{j=1}^{K_S} H_{SI} Q_{D_j} H_{SI}^H \right|}.$$  

(81c)

From (76) and (79), we observe that the DL and UL transmissions are coupled by the CCI and residual SI. This problem greatly impacts the performance of the system of interest. Herein, our main purpose is to jointly design beamformers so that the total system SE is maximized under the sum transmit power constraint in the DL channel and per-user power constraints in the UL one. Specifically, the total SE maximization problem is formulated as a rank-constrained optimization one as

maximize $R_0 + R_U$ 

subject to

- $0 \leq q_{U_j} \leq \bar{q}_{U_j}$, $\forall j = 1, \ldots, K_U$,  
- $\sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) \leq P_{BS}$,  
- $Q_{D_i} \succeq 0$, $\forall i = 1, \ldots, K_D$,  
- $\text{rank}(Q_{D_i}) = 1$, $\forall i = 1, \ldots, K_D$.

(82a-c)

where $P_{BS}$ is the maximum power at the BS and $\bar{q}_{U_j}$ is the power constraint at each user in the UL channel. We remark that due to the presence of CCI the expression of the SE in the DL channel will not be convex with the covariance matrices, even if BD is used as in Chapter 3. In other words, the use of BD does not reduce the nonconvexity of the

70
problem, and thus we consider general beamforming method in this chapter. Clearly, problem (82) is a nonconvex program, which is difficult to solve optimally in general. We also note that a simplified problem of (82), in which \( q_0 \) and \( R_0 \) are omitted (i.e., the SE maximization problem for the DL channel itself), was proved to be NP-hard [81]. Thus, we conjecture that the NP-hardness is carried over into our problem. Towards a tractable solution, we first apply the relaxation method to obtain a relaxed problem of (82) by dropping the rank-1 constraints (82e). Then, two efficient iterative algorithms proposed to solve the resulting problem are presented in the next section.

4.2 Proposed beamformer designs

Note that the relaxed problem of (82) after dropping the rank constraints is still nonconvex. Thus, computing its globally optimal solution is difficult and very computationally expensive in general. To the best of our knowledge, finding an optimal solution to the nonconvex problems similar to (82) is still an open problem. In this section, we present two reformulations of the relaxed problem, based on which two iterative algorithms of different level of complexity are developed.

4.2.1 Iterative MAXDET-based algorithm

The first beamforming algorithm is built upon an observation that the SE of the system at hand is a difference of two concave functions. Indeed, from (78b) and (81c), we can write

\[
R_D + R_U = h(Q, q) - g(Q, q)
\]

where

\[
h(Q, q) = \log\left(\sigma_n^2 I + \sum_{i=1}^{K_D} H_{SI} Q_D i H_{SI}^H + \sum_{j=1}^{K_U} q_U j h_{U j} h_{U j}^H\right)
\]

and

\[
g(Q, q) = \sum_{i=1}^{K_D} \log\left(\sigma_n^2 + \sum_{k \neq i} h_{D i}^H Q_D k h_{D i}\right) + \sum_{j=1}^{K_U} q_U j \left|g_{ji}\right|^2 + \log\left(\sigma_n^2 I + \sum_{i=1}^{K_D} H_{SI} Q_D i H_{SI}^H\right)
\]

and \( Q \) and \( q \) are the symbolic notations that denote the set of design variables \( \{Q_D\} \) and \( \{q_U\} \), respectively. It should be noted that the functions \( h(Q, q) \) and \( g(Q, q) \) are jointly concave with respect to \( Q \) and \( q \) [85]. Borrowing the concept of the Frank-Wolfe method...
which considers a linear approximation of the objective function and searches for a direction that improves the objective, we now present the first joint design algorithm to find \( Q \) and \( q \). First, the relaxed problem is reformulated as

\[
\begin{align*}
\text{maximize} & \quad h(Q,q) - g(Q,q) \\
\text{subject to} & \quad (82b), (82c), (82d).
\end{align*}
\]

Since the constraints (82b)-(82d) are convex, the difficulty in solving (85) lies in the component \(-g(Q,q)\). Suppose the value of \((Q,q)\) at iteration \( n \) is denoted by \((Q(n),q(n))\). To increase the objective in the next iteration we replace \( g(Q,q) \) by its affine majorization at a neighborhood of \((Q(n),q(n))\). Since \( g(Q,q) \) is concave and differentiable on the considered domain \([Q_0, q_0 : Q_0 \succeq 0, q_0 \succeq 0]\), one can easily find an affine majorization as a first order approximation as [85]

\[
g^{(n)}(Q,q) = g(Q(n),q(n)) + \sum_{i=1}^{K_h} \sum_{k=1,k\neq i}^{K_h} \left[(\theta_D^{(n)})^{-1} h_D^{(n)}(Q_{Dk} - Q_{Dk}^{(n)}) h_D \right]
\]

\[
+ \sum_{i=1}^{K_h} \sum_{j=1}^{K_h} (\theta_D^{(n)})^{-1} |g_{ij}| l_i^2 (q_{ij} - q_{ij}^{(n)}) + \sum_{i=1}^{K_h} \text{Tr} \left[ H_{SI} (\Theta^{(n)})^{-1} H_{SI} (Q_{Dk} - Q_{Dk}^{(n)}) \right]
\]

where \( \theta_D^{(n)} \) and \( \Theta^{(n)} \) are defined as

\[
\theta_D^{(n)} = \sigma_n^2 + \sum_{m \neq i}^{K_h} h_{Dm}^{(n)} Q_{Dm} h_D + \sum_{l=1}^{K_h} |g_{li}|^2,
\]

\[
\Theta^{(n)} = \sigma_n^2 I + \sum_{j=1}^{K_h} H_{SI} Q_{Dj}^{(n)} H_{SI}^H.
\]

To derive (86), we have used the fact \( \nabla_X \log |I + AXA^H| = A^H (I + AXA^H)^{-1} A \). \( \nabla_x \log(1 + ax) = a(1 + ax)^{-1} \), the inner product of two semidefinite matrices \( X \succeq 0 \) and \( Y \succeq 0 \) is \( \text{Tr}(XY) \), and the inner product of two vector is \( x^H y \) [100]. Now, we approximate problem (85) at iteration \( n + 1 \) by a convex program given by

\[
\begin{align*}
\text{maximize} & \quad h(Q,q) - g^{(n)}(Q,q) \\
\text{subject to} & \quad (82b), (82c), (82d).
\end{align*}
\]

The objective in (89) is in fact a lower bound of the SE of the FD system. We note that problem (89) is a MAXDET program, and hence the name of the first algorithm. Let \((Q^*, q^*)\) be the optimal value of \((Q,q)\) in (89). Then we update
Algorithm 5 Iterative MAXDET-based algorithm.[38] ©2014 IEEE.

Initialization:
1: Generate initial values for $Q^{0}_0$ for $i = 1, 2, \ldots, K_D$ and $q^{0}_j$ for $i = 1, 2, \ldots, K_U$.
2: Set $n := 0$.

Iterative procedure:
3: repeat
4: Solve (89) and denote the optimal solutions as $(Q^*, q^*)$.
5: Update: $Q^{n+1}_0 := Q^*; \text{and } q^{n+1}_j := q^*_j$.
6: Set $n := n + 1$.
7: until Convergence.

Finalization:
8: Perform randomization to extract a rank-1 solution if required.

$(Q^{(n+1)}, q^{(n+1)}) := (Q^*, q^*)$. In this way, the design variables are iteratively updated and the lower bound of the SE increases after every iteration. Since the SE is bounded above due to the power constraints (82b) and (82c), the iterative procedure is guaranteed to converge. The iterative MAXDET-based algorithm is outlined in Algorithm 5. The convergence properties of Algorithm 5 are stated in Theorem 1 (see Subsection 4.2.2, Page 79).

An important point to note here is that the iterative procedure in Algorithm 5 possibly returns a locally optimal solution to a relaxed problem of (82) at convergence. Obviously, if rank($Q^*_D$) = 1, then $Q^*_D$ is also feasible to (82) and the beamformer for $D_i$ can be immediately recovered from the eigenvalue decomposition of $Q^*_D$ [85]. However, this may not be the case since the rank-1 constraints are dropped. Thus, a method to extract the beamformer is required if a high-rank solution is obtained. For this purpose, we adopt the randomization technique presented in [99] which is mentioned in Line 8 of Algorithm 5 and briefly described as follows. We first generate a random (column) vector $v_D$ whose elements are independently and uniformly distributed on the unit circle in the complex plane, and then calculate the eigen-decomposition of $Q^*_D$ as $Q^*_D = U_D \Sigma_D U^H_D$. Next a beamformer is taken as $\tilde{w}_D = U_D \Sigma_D^{1/2} v_D$, which is feasible to the original design problem since $\|\tilde{w}_D\|^2 = \text{Tr}(U_D \Sigma_D^{1/2} v_D v^H_D \Sigma_D^{1/2} U^H_D) = \text{Tr}(Q^*_D)$ [99]. The obtained beamformer $\tilde{w}_D$ is then used to compute the resulting sum rate. We repeat this process for a number of randomization samples and pick up the one that offers the best sum rate. Our numerical results have shown that the high-rank solutions of $\{Q_D\}$ only occur when $\sigma^2_{SI}$ is sufficiently large, which is not of practical importance since this
is not the interesting case for the FD systems. When \( \text{rank}(\mathbf{Q}_b) > 1 \), we also observe that the largest eigenvalue significantly dominates the remaining ones. More explicitly, the largest eigenvalue is always 10 times larger than the second largest one, meaning that \( \mathbf{Q}_b \) is not far from a rank-1 matrix. This explains the fact that the beamforming vectors returned by the randomization method offer a performance very close to that of the relaxed problem. Explicitly, the extracted solutions achieve a SE performance always higher than 95% of the upper bound given by the relaxed problem.

Although the objective in (89) is not a linear function with respect to the design parameters as in the original work of [73], (89) can be equivalently transformed into the problem of maximizing an affine function over a convex set as \( \max \{ \omega - \| \mathbf{Q} \|_F \} \). Thus, Algorithm 5 can be considered as a variant of the Frank-Wolfe method. It has been reported in many studies that the type of Frank-Wolfe methods can efficiently exploit the hidden convexity of the problem [86, 87, 101]. Thus, the same results as the Frank-Wolfe type method can also be expected in the first proposed design algorithm. However, solvers for MAXDET programs are quite limited, compared to their counterparts for SDPs. Because none of the general convex program solvers are perfect for all problems, a more flexible choice of a problem solver is of practical importance.

\subsection{4.2.2 Iterative SDP-based algorithm}

Motivated by the discussion above, we propose in this subsection an iterative SDP-based algorithm to solve the relaxed problem of (82). Specifically, based on the general framework of the SPCA method and proper transformations, we can iteratively approximate the relaxed problem of (82) by an SDP in each iteration. The second approach allows us to take advantage of a wide class of SDP solvers which are more and more efficient due to continuing progress in semidefinite programming. To begin with, due to the monotonicity of the log function, we first reformulate the relaxed problem of (82) as

\[
\begin{align*}
\maximize_{\mathbf{Q}_b, h, \mathbf{q}_D} & \quad \prod_{i=1}^{K_b} (1 + \gamma_D) \
\text{subject to} & \quad (82b), (82c), (82d)
\end{align*}
\]

The dedicated solver for the MAXDET problem in (89) is SDPT3 [102]. In fact, CVX solves this type of problems using a succesive convex approximate method, allowing us to choose other SDP solvers, e.g., [103]. However, this method can be slow and is still in an experimental stage.

\[\text{(90)}\]
which then can be rewritten as

\[
\begin{aligned}
\text{maximize} & \quad \prod_{i=1}^{K_D} t_{D_i} \prod_{j=1}^{K_U} t_{U_j} \\
\text{subject to} & \quad 1 + \gamma_{D_i} \geq t_{D_i}, \quad i = 1, \ldots, K_D, \\
& \quad 1 + \gamma_{U_j} \geq t_{U_j}, \quad j = 1, \ldots, K_U, \\
& \quad t_{D_i} \geq 1, \quad \forall i; \\
& \quad t_{U_j} \geq 1, \quad \forall j, \\
& \quad (82b), (82c), (82d)
\end{aligned}
\]

(91a)

(91b)

(91c)

(91d)

(91e)

by using the epigraph form of (90) [85]. Note that maximizing a product of variables admits a SOC representation [70, 104]. Thus, we only need to deal with the nonconvex constraints in (91b) and (91c). Let us treat the constraint (91b) first. It is without loss of optimality to replace (91b) by following two constraints

\[
\begin{aligned}
\sigma_n^2 + \sum_{k=1}^{K_D} h_{D_k}^H Q_{D_k} h_{D_k} + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2 & \geq t_{D_i} \beta_{D_i}, \\
\sigma_n^2 + \sum_{k=1}^{K_D} h_{D_k}^H Q_{D_k} h_{D_k} + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2 & \leq \beta_{D_i}
\end{aligned}
\]

(92a)

(92b)

where \( \beta_{D_i} \) is newly introduced variable and can be considered as the soft interference threshold of \( D_i \). The equivalence between (91b) and the two inequalities in (92a) and (92b) follows the same arguments as in [70] which can be justified as follows. At optimum, suppose the constraint in (92b) holds with inequality, i.e., \( \sigma_n^2 + \sum_{k=1}^{K_D} h_{D_k}^H Q_{D_k} h_{D_k} + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2 < \beta_{D_i} \). Then, we form a new pair \( (\tilde{\beta}_{D_i}, \tilde{t}_{D_i}) \) as \( \tilde{\beta}_{D_i} = \beta_{D_i} / c \) and \( \tilde{t}_{D_i} = c t_{D_i} \) where \( c \) is a positive constant. Obviously, there exists a given \( c > 1 \) such that (92b) is still met when \( \beta_{D_i} \) is replaced by \( \tilde{\beta}_{D_i} \). Since \( \beta_{D_i} \tilde{t}_{D_i} = \beta_{D_i} t_{D_i} \), i.e., the right side of (92a) remains the same, the constraint in (92a) is still satisfied. However, since \( \tilde{t}_{D_i} > t_{D_i} \) with \( c > 1 \), a strictly higher objective of the design problem is obtained. This contradicts with the assumption that we already obtained the optimal objective. Likewise, we can decompose (91c) into

\[
\begin{aligned}
x_{U_j}^2 h_{U_j}^H X_{U_j}^{-1} h_{U_j} & \geq t_{U_j} - 1, \\
q_{U_j} & \geq x_{U_j}^2
\end{aligned}
\]

(93a)

(93b)

where \( X_{U_j} = \sigma_n^2 I + \sum_{m > j} q_{U_m} h_{U_m}^H h_{U_m} + \sum_{i=1}^{K_D} H_{SI} Q_{D_i} H_{SI}^H \) and \( x_{U_j} \) is an auxiliary variable. The purpose of introducing the slack variable \( x_{U_j} \) will be clear shortly when we show
that it is necessary to arrive at a SDP formulation. Now, we can equivalently transform (91) into a more tractable form as

\[
\begin{align*}
\text{maximize} & \quad \prod_{i=1}^{K_0} t_{D_i} \prod_{j=1}^{K_0} t_{Q_j} \\
\text{subject to} & \quad \sigma^2_n + \sum_{k=1}^{K_0} h^H_{D_i} Q_{D_i} h_{D_i} + \sum_{j=1}^{K_0} q_{Q_j} |g_{ji}|^2 \geq f(t_{D_i}, \beta_{D_i}), \forall i = 1, \ldots, K_0, \quad (94b) \\
& \quad \sigma^2_n + \sum_{k=1}^{K_0} h^H_{D_i} Q_{D_i} h_{D_i} + \sum_{j=1}^{K_0} q_{Q_j} |g_{ji}|^2 \leq \beta_{D_i}, \forall i = 1, \ldots, K_0, \quad (94c) \\
& \quad g(x_{D_i}^2, Q, q) \geq t_{Q_j} - 1, \forall j = 1, \ldots, K_0, \quad (94d) \\
& \quad q_{Q_j} \geq x_{D_i}^2, \forall j = 1, \ldots, K_0, \quad (94e) \\
& \quad (82b), (82c), (82d), (91d) \quad (94f)
\end{align*}
\]

where \( f(t_{D_i}, \beta_{D_i}) = t_{D_i} \beta_{D_i}, g(x_{D_i}^2, Q, q) = x_{D_i}^2 h^H_{D_i} X_{D_i}^{-1} h_{D_i} \), and \( Q, q, t_{D_i}, \beta_{D_i}, x_0 \) are the symbolic notations that denote the sets of optimization variables \( \{Q_{D_i}\}, \{q_{Q_j}\}, \{t_{D_i}\}, \{t_{Q_j}\}, \{\beta_{D_i}\}, \{x_{D_i}\} \), respectively.

We note that the constraints in (94c) and (94e) are linear and SOC ones, respectively. Consequently, the barrier to solving (94) is due to the nonconvexity in (94b) and (94d). In what follows, we will present a low-complexity approach that locally solves (94). Toward this end we resort to an iterative algorithm based on SPCA. To show this, let us tackle the nonconvex constraint (94b) first. Note that \( f(t_{D_i}, \beta_{D_i}) \) is neither a convex nor concave function of \( t_{D_i} \) and \( \beta_{D_i} \). Fortunately, in the spirit of [70, 80], we recall the following inequality

\[
\begin{align*}
f(t_{D_i}, \beta_{D_i}) \leq F(t_{D_i}, \beta_{D_i}, \psi_{D_i}^{(n)}) = \frac{1}{2\psi_{D_i}^{(n)} t_{D_i}^2} + \frac{\psi_{D_i}^{(n)}}{2} \beta_{D_i}^2
\end{align*}
\]

which holds for every \( \psi_{D_i}^{(n)} > 0 \). The right side of (95) is called a convex upper estimate of \( f(t_{D_i}, \beta_{D_i}) \). The approximation shown in (95) deserves some comments. First, it is straightforward to note that \( f(t_{D_i}, \beta_{D_i}) = F(t_{D_i}, \beta_{D_i}, \psi_{D_i}^{(n)}) \) when \( \psi_{D_i}^{(n)} = t_{D_i} / \beta_{D_i} \). Moreover, with this selection of \( \psi_{D_i}^{(n)} \), one can easily check that the first derivative of \( F(t_{D_i}, \beta_{D_i}, \psi_{D_i}^{(n)}) \) with respect to \( t_{D_i} \) or \( \beta_{D_i} \) is equal to that of \( f(t_{D_i}, \beta_{D_i}) \), i.e., \( \nabla F(t_{D_i}, \beta_{D_i}, \psi_{D_i}^{(n)}) = \nabla f(t_{D_i}, \beta_{D_i}) \). These two properties are important to establish the local convergence of the second iterative algorithm which is deferred to the Appendix 2.

\[\text{Since } t_{D_i} \geq 1 \text{ and } \beta_{D_i} \geq \sigma^2_n > 0 \text{ (from (92b)) and both of them are bounded above (i.e., } t_{D_i} < +\infty \text{ and } \beta_{D_i} < +\infty \text{ due to the transmit power constraint at the BS, the value of } \psi_{D_i}^{(n)} \text{ is well defined.} \]
Now we turn our attention to (94d), which is equivalent to \( t_{\bar{U}_j} - 1 - g(x^2_{\bar{U}_j}, Q, q) \leq 0 \). First, we note that \( g(x^2_{\bar{U}_j}, Q, q) \) is jointly convex in the involved variables. As a proof, consider the epigraph of \( g(x^2_{\bar{U}_j}, Q, q) \) which is given by [85]

\[
\{(\alpha, x^2_{\bar{U}_j}, Q, q) | \alpha \geq x^2_{\bar{U}_j} h^H_{\bar{U}_j} X^{-1}_{\bar{U}_j} h_{\bar{U}_j}\}. \tag{96}
\]

By Schur complement [100], (96) is equivalent to

\[
\begin{bmatrix}
\alpha \\
x_{\bar{U}_j} h_{\bar{U}_j} \\
X_{\bar{U}_j}
\end{bmatrix}
= \begin{bmatrix}
\alpha \\
x_{\bar{U}_j} h_{\bar{U}_j} \\
X_{\bar{U}_j}
\end{bmatrix}
\begin{bmatrix}
\sigma_n^2 I + \sum_{m>j} q_{\bar{U}m} h_{\bar{U}m} h^H_{\bar{U}m} + \sum_{i=1}^{K_\delta} Q_{\bar{b}} H^H_{\bar{S}_1} \\
x_{\bar{U}_j} h_{\bar{U}_j} \\
X_{\bar{U}_j}
\end{bmatrix} \succeq 0. \tag{97}
\]

Since the epigraph of \( g(x^2_{\bar{U}_j}, Q, q) \) is representable by linear matrix inequality which is a convex set, so is \( g(x^2_{\bar{U}_j}, Q, q) \) [85]. Now a convex upper bound of the term \(-g(x^2_{\bar{U}_j}, Q, q)\) in (94d) can be found as its first order approximation at a neighborhood of \((x_{\bar{U}_j}^{(n)}, Q^{(n)}, q^{(n)})\), i.e.,

\[
-g(x^2_{\bar{U}_j}, Q, q) \leq G(x_{\bar{U}_j}, Q, q, x_{\bar{U}_j}^{(n)}, Q^{(n)}, q^{(n)})
= -\left\{ g(x_{\bar{U}_j}^{(n)}, Q^{(n)}, q^{(n)}) + 2 x_{\bar{U}_j} h^H_{\bar{U}_j} (X_{\bar{U}_j}^{(n)})^{-1} h_{\bar{U}_j} (x_{\bar{U}_j} - x_{\bar{U}_j}^{(n)})
- \text{Tr}\left[\left((x_{\bar{U}_j}^{(n)})^2 (X_{\bar{U}_j})^{-1} h_{\bar{U}_j} h^H_{\bar{U}_j} (X_{\bar{U}_j}^{(n)})^{-1}\right) (X_{\bar{U}_j} - X_{\bar{U}_j}^{(n)})\right]\right\}. \tag{98}
\]

where \( X_{\bar{U}_j} \) is replaced by the affine function of \( Q \) and \( q \) defined below (93) and we have used the fact that \( \nabla A a^H A^{-1} b = -A^{-1} a b^H A^{-1} \) for \( A \succeq 0 \) [100].
Algorithm 6 Iterative SDP-based algorithm. [38] ©2014 IEEE.

Initialization:
1: Generate initial points for $\psi_{D_i}^{(0)}$ and $Q_{D_i}^{(0)}$ for $i = 1, \ldots, K_D$; and $q_{U_j}^{(0)}$ and $x_{U_j}^{(0)}$ for $j = 1, \ldots, K_U$.
2: Set $n := 0$.

Iterative procedure:
3: repeat
4: Solve (99) to find optimal solutions $Q_{D_i}^{*}$, $t_{D_i}^{*}$, and $\beta_{D_i}^{*}$ for $i = 1, \ldots, K_D$, and $q_{U_j}^{*}$, and $x_{U_j}^{*}$ for $j = 1, \ldots, K_U$.
5: Set $n := n + 1$.
6: Update : $\psi_{D_i}^{(n)} := t_{D_i}^{*} / \beta_{D_i}^{*}$; $x_{U_j}^{(n)} := x_{U_j}^{*}$; $Q_{D_i}^{(n)} := Q_{D_i}^{*}$; $q_{U_j}^{(n)} := q_{U_j}^{*}$.
7: until Convergence.

Finalization:
8: Perform randomization to extract a rank-1 solution as in Algorithm 5.

The mathematical discussions above imply that the convex approximate problem at iteration $(n+1)$ of the second iterative design approach is the following:

\begin{align}
\text{maximize} & \quad \prod_{i=1}^{K_D} t_{D_i} \prod_{j=1}^{K_U} t_{U_j} \\
\text{subject to} & \quad F(t_{D_i}, \beta_{D_i}, \psi_{D_i}^{(n)}) \leq \alpha_n^2 + \sum_{k=1}^{K_D} h_{D_i}^H Q_{D_i} h_{D_i} + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2, \forall i = 1, \ldots, K_D, \tag{99b}\ \\
& \quad G(x_{U_j}, Q, q, x_{U_j}^{(n)}, Q^{(n)}, q^{(n)}) \leq 1 - t_{U_j}, \forall j = 1, \ldots, K_U, \tag{99c}\ \\
& \quad \sigma_n^2 + \sum_{k=1}^{K_D} h_{D_i}^H Q_{D_i} h_{D_i} + \sum_{j=1}^{K_U} q_{U_j} |g_{ji}|^2 \leq \beta_{D_i}, \forall i = 1, \ldots, K_D, \tag{99d}\ \\
& \quad q_{U_j} \geq x_{U_j}^{2}, \forall j = 1, \ldots, K_U, \tag{99e}\ \\
& \quad 0 \leq q_{U_j} \leq \overline{q}_{U_j}, \forall j = 1, \ldots, K_U, \tag{99f}\ \\
& \quad \sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) \leq P_{BS}, \tag{99g}\ \\
& \quad Q_{D_i} \succeq 0, \forall i = 1, \ldots, K_D, \tag{99h}\ \\
& \quad t_{D_i} \geq 1, \forall i = 1, \ldots, K_D; t_{U_j} \geq 1, \forall j = 1, \ldots, K_U. \tag{99i}
\end{align}
After the iterative procedure terminates, the randomization trick may be applied to extract a rank-1 solution as in Algorithm 5. The proposed iterative SDP-based algorithm is summarized in Algorithm 6.

The convergence results of Algorithms 5 and 6 are stated in the following theorem whose proof is given in the Appendix 2.

**Theorem 1.** Algorithms 5 and 6 produce a sequence of solutions converging to a KKT point of (85) and (91), respectively.

As mentioned in [80], the SPCA method can start with an infeasible initial point. However, it is desired to generate initial values for \( Q^{(0)}_{D_i}, q^{(0)}_{U_j}, \psi^{(0)}_{D_i} \) and \( x^{(0)}_{U_j} \) such that Algorithm 6 is guaranteed to be solvable in the first iteration. For this purpose, we first randomly generate \( Q^{(0)}_{D_i} \geq 0 \) for \( i = 1, \ldots, K_D \) and \( q^{(0)}_{U_j} \) in the range from 0 to \( \tilde{q}_{U_j} \) for \( j = 1, \ldots, K_U \). If necessary, \( Q^{(0)}_{D_i} \) is scaled so that the constraint (99g) is satisfied. Then, \( x^{(0)}_{U_j} \) is calculated as \( x^{(0)}_{U_j} = \sqrt{q^{(0)}_{U_j}} \) and \( \psi^{(0)}_{D_i} \) is set to \( t^{(0)}_{D_i} / \beta^{(0)}_{D_i} \) where \( t^{(0)}_{D_i} \) and \( \beta^{(0)}_{D_i} \) are computed from (91b) and (99d) by setting the inequalities to equalities, respectively.

At the first look, the SDP solved at each iteration in Algorithm 6 has more optimization variables due to some slack variables introduced. Thus, the theoretical (worst case) computational complexity of Algorithm 6 could possibly be higher than that of Algorithm 5. We note that the complexity of the two proposed methods mainly depends on the semidefinite constraints \( Q_{D_i} \geq 0, \forall i = 1, \ldots, K_D \). That is to say, the per iteration complexity formulation used in Algorithm 6 just slightly requires higher complexity than Algorithm 5. Note again that the per iteration computational complexity of each step in Algorithm 6 is approximated as \( O(K_D^3 + N_T^2K_D^2 + K_DN_T^3) \), assuming \( K_D \approx K_U \) [105, ch. 4]. As aforementioned, the advantage of the second proposed algorithm is that it allows us to make use of efficient SDP solvers such as SEDUMI [103] and MOSEK [106]. Alternatively, we can use both proposed two approaches in parallel for solving the original problem. The solving process can be terminated if one of the algorithms has converged. It is also possible to solve the problem until both methods converge and choose the better solution. More insights on the computational complexity of the iterative MAXDET- and SDP-based algorithms are given in Section 4.3.
4.3 Numerical results

4.3.1 Convergence and complexity comparison

In the first experiment, we compare the complexity and the convergence rate of Algorithms 5 and 6 proposed in Section 4.2 for two cases, the first case for independent and identically distributed (i.i.d) channel model and the second case for realistic channel model generated in Subsection 4.3.2. In the first case, each entry of the channel vectors $h_D$, $h_U$, and $g_j$ follows the i.i.d zero mean and unit variance complex Gaussian distribution. The noise power is taken as $\sigma_n^2 = 1$ and the maximum transmit power at the BS and UL users is set to $P_{BS} = q_{uj} = 20$ dBW for all $U_j$. This setting resembles the case where the average signal-to-noise ratio (SNR) at transmitter sides is 20 dB. In the second case, the specific parameters are taken from Table 2 and the allowable transmit power at the BS and the users in the UL channel is fixed at $P_{BS} = q_{uj} = 10$ dBm.

An accurate model for the residual SI channel plays an important role in evaluating the SE performance of FD systems. Thus, theoretical studies and practical measurements on this issue are of significant importance and call for more research efforts. Herein, $H_{SI}$ is chosen as $CN_{N_hN_f}$ ($\frac{\sigma^2_{SI} K}{1+K} H_{SI} \otimes \frac{\sigma^2_{SI} I_{N_f} \otimes I_{N_f}}{}$), where $\otimes$ denotes the Kronecker product, $K$ is the Rician $K$-factor, $H_{SI}$ is a deterministic matrix, and $\sigma^2_{SI}$ is introduced to parameterize the capability of a certain SI cancellation design. We remark that according to [10] the Rician probability distribution with a small Rician $K$-factor should be used to characterize the residual SI channel after SI cancellation mechanisms. Without loss of generality, we set $K = 1$ and $H_{SI}$ to be the matrix of all ones for all experiments. Moreover, $\sigma^2_{SI}$ fixed at $-30$ dB for the first case and $-100$ dB for the second case.

Fig. 11 illustrates the convergence rate of Algorithms 5 and 6 for a given set of channel realizations generated randomly for the two cases. Each point on the curves of Fig. 11 is obtained by solving problems (89) and (99), respectively. Generally, we have observed that Algorithm 5 requires fewer iterations to converge than Algorithm 6. This observation is probably attributed to the fact that Algorithm 5 exploits the hidden convexity better since it searches for an improved solution over the whole feasible set in each iteration. We recall that SDPT3 is the dedicated solver for the type of problems in (89), and thus the choice of optimization software is limited for Algorithm 5. A recent work of [107] has reported that, among common general SDP solvers, SDPT3 is comparatively slow. The SDP formulation in Algorithm 6 allows for use of faster SDP
solvers such as SeDuMi or MOSEK. In return, the total time of Algorithm 6 to find a solution may be less than that of Algorithm 5 which is illustrated in Table 1.

In Table 1, we show the average runtime (in seconds) of Algorithms 5 and 6 for the two channel models mentioned above. The stopping criterion for the two algorithms is when the increase in the last 10 iterations is less than $10^{-5}$ [80]. All convex solvers considered in Table 1 are set to their default values. We observe that the per iteration solving time of Algorithm 6 is much less than that of Algorithm 5. Consequently, the total solving time of Algorithm 6 is smaller than that of Algorithm 5, especially when used with MOSEK solver.

### 4.3.2 Spectral efficiency performance

We now compare the achievable SE of the proposed beamformer designs for the FD system introduced in Section 4.1 with that of a traditional HD scheme having the relevant hardware configurations. In fact, as mentioned earlier, the application with the most potential for FD technology in cellular systems is in small-cells. To quantify the potential benefit of the FD transmission considered in this chapter, we evaluate the performance of the proposed algorithms under the 3GPP LTE specifications for small-cell deployments.

The general simulation parameters are taken from [2, 108] and listed in Table 2. Without loss of generality, per-user power constraints of users in the UL transmission are assumed to be equal, i.e., $\overline{q}_U = \overline{q}$. In particular, we consider two different settings of the transmit power constraints in both directions: (i) $(P_{BS}, \overline{q}) = (26 \text{ dBm}, 23 \text{ dBm})$ following the LTE 3GPP pico cell standard for outdoor [2] and (ii) $(P_{BS}, \overline{q}) = (10 \text{ dBm}, 10 \text{ dBm})$ according to the work of [10]. The number of antennas at the BS is set to six, of which four are used for transmitting and two for receiving, i.e., $N_T = 4$ and $N_R = 2$, respectively. All users in both directions are randomly dropped in a circular area of a radius $r = 100 \text{ m}$, centered at the FD capable BS in an outdoor small-cell scenario.

The channel vector from the BS to $D_i$ is given by $h_{D_i} = \sqrt{\kappa_{D_i}} \tilde{h}_{D_i}$ where $\tilde{h}_{D_i}$ follows $CN(0, I)$ that denotes the small scale fading, and $\kappa_{D_i} = 10^{-(PL_{\text{LOS}}/10)}$ represents the path loss, where $PL_{\text{LOS}}$ is calculated from a specific path loss model as shown in (100). The channel vector between the BS and $U_j$ is generated in the same way. For large scale fading, we adopt the path loss model presented in [2, 108]. More specifically, DL and UL channels are assumed to experience the path loss model for line-of-sight (LoS) communications as

$$PL_{\text{LOS}} = 103.8 + 20.9 \log_{10} d$$  \hspace{1cm} (100)
(a) Convergence rate for i.i.d channel realizations with $K_D = K_U = 4$ and $N_T = N_R = 4$.

(b) Convergence rate for channel realizations taken from the channel model in Section 4.3.2

Fig. 11. Convergence rate of Algorithms 5 and 6 for a set of random channel realizations. [38] ©2014 IEEE.
Table 1. Average run time (in seconds) for i.i.d and realistic channel models for various simulation setups. The proposed algorithms terminate if the gap of the objectives between the last 10 iterations is less than $\epsilon \leq 10^{-5}$. [38] ©2014 IEEE.

<table>
<thead>
<tr>
<th>$N_T$</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_R = 2$</td>
<td>Algorithm 1 (SDPT3)</td>
<td>2.61</td>
<td>3.74</td>
<td>5.61</td>
<td>9.46</td>
<td>15.14</td>
</tr>
<tr>
<td>$K_D = 2$</td>
<td>Algorithm 2 (SeDuMi)</td>
<td>1.43</td>
<td>2.63</td>
<td>3.77</td>
<td>6.66</td>
<td>11.54</td>
</tr>
<tr>
<td>$K_U = 2$</td>
<td>Algorithm 2 (MOSEK)</td>
<td>0.089</td>
<td>0.26</td>
<td>0.45</td>
<td>1.09</td>
<td>2.68</td>
</tr>
<tr>
<td>i.i.d channel model</td>
<td>Algorithm 1 (SDPT3)</td>
<td>4.17</td>
<td>6.28</td>
<td>9.29</td>
<td>15.04</td>
<td>23.76</td>
</tr>
<tr>
<td>realistic channel model (given in Sec. 4.3.2)</td>
<td>Algorithm 2 (SeDuMi)</td>
<td>2.36</td>
<td>4.13</td>
<td>6.11</td>
<td>10.50</td>
<td>17.64</td>
</tr>
<tr>
<td></td>
<td>Algorithm 2 (MOSEK)</td>
<td>0.21</td>
<td>0.61</td>
<td>1.01</td>
<td>2.47</td>
<td>5.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$K_D$</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_T = 4$</td>
<td>Algorithm 1 (SDPT3)</td>
<td>3.74</td>
<td>9.64</td>
<td>13.01</td>
<td>16.27</td>
<td>18.76</td>
</tr>
<tr>
<td>$N_R = 2$</td>
<td>Algorithm 2 (SeDuMi)</td>
<td>2.63</td>
<td>6.25</td>
<td>8.12</td>
<td>9.98</td>
<td>12.77</td>
</tr>
<tr>
<td>$K_D = 2$</td>
<td>Algorithm 2 (MOSEK)</td>
<td>0.26</td>
<td>1.24</td>
<td>1.66</td>
<td>2.52</td>
<td>3.09</td>
</tr>
<tr>
<td>$K_U = 2$</td>
<td>Algorithm 1 (SDPT3)</td>
<td>6.28</td>
<td>17.33</td>
<td>22.84</td>
<td>27.55</td>
<td>31.58</td>
</tr>
<tr>
<td>realistic channel model (given in Sec. 4.3.2)</td>
<td>Algorithm 2 (SeDuMi)</td>
<td>4.13</td>
<td>10.59</td>
<td>14.19</td>
<td>17.05</td>
<td>22.54</td>
</tr>
<tr>
<td></td>
<td>Algorithm 2 (MOSEK)</td>
<td>0.61</td>
<td>2.24</td>
<td>2.90</td>
<td>4.34</td>
<td>5.24</td>
</tr>
</tbody>
</table>
Fig. 12. Location of users for the two specific simulation settings considered in the numerical results section. [38] ©2014 IEEE.
Table 2. Simulation parameters. [38] ©2014 IEEE.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier frequency</td>
<td>2 GHz</td>
</tr>
<tr>
<td>System bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Thermal noise</td>
<td>-174 dBm/Hz</td>
</tr>
<tr>
<td>Receiver noise figure (at DL users)</td>
<td>9 dB</td>
</tr>
<tr>
<td>Receiver noise figure (at BS)</td>
<td>5 dB</td>
</tr>
<tr>
<td>Maximum transmit power at BS ($P_{BS}$)</td>
<td>10 or 26 dBm</td>
</tr>
<tr>
<td>Maximum transmit power per user ($\bar{q}$)</td>
<td>10 or 23 dBm</td>
</tr>
</tbody>
</table>

where $PL_{LOS}$ is in dB, and $d$ is the distance (in kilometers) between the BS and a specific user. Similarly, the channel coefficient from $U_j$ to $D_i$ is modeled as $g_{ji} = \sqrt{\kappa_{ji}} \tilde{g}_{ji}$ where $\tilde{g}_{ji}$ follows $CN(0, 1)$ and $\kappa_{ji} = 10^{-PL_{NLOS}/10}$ denotes the large scale fading. Since there is a high possibility of obstructions between users deployed in an outdoor environment, we assume that the channel from $U_j$ to $D_i$ encounters the path loss model for non-line-of-sight (NLOS) transmission. That is, $PL_{NLOS}$ (in dB) is written as

$$PL_{NLOS} = 145.4 + 37.5 \log_{10} d_{CCI}$$

(101)

where $d_{CCI}$ is now the distance (in kilometers) from a user in the UL transmission to another user in the DL direction. The residual SI channel model is mentioned in Subsection 4.3.1.

To have a fair comparison between the FD and HD systems, we made the following assumptions. First, the BS of the HD counterpart is assumed to use all antennas in both DL and UL transmissions, i.e., $N_T + N_R$. For the HD case, since the DL and UL transmissions are separated, and thus the SEs of the DL and UL channels can be computed independently. Specifically, we use the iterative water-filling algorithm introduced in [97] to find the optimal SE of the UL channel. Note that the problem of SE maximization in the DL direction is NP-hard which requires extremely high computational complexity to find optimal solution [81]. Herein, we employ an efficient solution proposed in [70], which was shown to achieve near optimal performance, to calculate the SE of the DL transmission. Then, the resulting SEs of the DL and UL channels in the HD counterpart are divided by two since each of them is assumed to share 50% of the temporal resource [10]. For the FD case, the SEs of the DL and UL channels are simply calculated by (78) and (81), respectively, after achieving the solutions of the problem in (82).
Average spectral efficiency gain (%)

Self-interference strength, $\sigma_{SI}^2$ (dB)

(a) Average SE gain of DL channel.

(b) Average SE gain of UL channel.
Fig. 13 depicts the SE gains in percentage of the FD system over the HD one as a function of $\sigma_{SI}^2$ for the scenario as shown in Fig. 12a. A general observation is that FD transmission can significantly improve the SE of the HD one when the SI is substantially suppressed. Specifically, as shown in Fig. 13c, the total SE gain of the FD system is 45.6% and 55% for the cases $(P_{BS}, \bar{q}_j) = (26 \text{ dBm}, 23 \text{ dBm})$ and $(P_{BS}, \bar{q}_j) = (10 \text{ dBm}, 10 \text{ dBm})$ at $\sigma_{SI}^2 = -130 \text{ dB}$, respectively. However, when $\sigma_{SI}^2 = -55 \text{ dB}$, the HD system performs better than the FD one for both cases of transmit power constraint. This observation simply means that the SI cancellation mechanism should be efficient enough for the FD system to compete against the HD counterpart. In addition, the simulation results also indicate that the SI needs to be canceled at least 75 dB (i.e., $\sigma_{SI}^2 < -75 \text{ dB}$) for the case $(P_{BS}, \bar{q}) = (10 \text{ dBm}, 10 \text{ dBm})$ and at least 83 dB (i.e., $\sigma_{SI}^2 < -83 \text{ dB}$) for the case $(P_{BS}, \bar{q}) = (26 \text{ dBm}, 23 \text{ dBm})$ for the FD system to attain better SE in both DL and UL transmissions, compared to the HD one. These requirements can be achieved by a recent advanced SI cancellation technique reported in [7–10, 12–14].
To obtain more insights into the performance of the FD system, we also study the gains of the DL and UL channels separately in Figs. 13a and 13b, respectively. We can see that, while the SE of the UL transmission of the FD system is always deteriorated as $\sigma^2_{SI}$ increases, that of the DL channel decreases until a certain value of $\sigma^2_{SI}$ ($-100$ dB and $-90$dB for $(P_{BS}, \eta) = (26$ dBm, $23$ dBm) and $(P_{BS}, \eta) = (10$ dBm, $10$ dBm), respectively) and increases after that. The degradation on the SE of the UL channel is obvious and due to the fact that a large value of $\sigma^2_{SI}$ results in a greater amount of residual SI power being added to the background noise at the FD BS. To explain different trends in the SE of the DL channel, we first recall that the main goal of the proposed designs is to maximize the total SE of the FD system, i.e., jointly optimizing both UL and DL transmissions. When the residual SI power is quite small, the joint optimization schemes slightly reduce the actual transmit power of the DL channel to maintain the SE of the UL channel. For a large value of $\sigma^2_{SI}$, the SI is comparable or even dominates the desired signals of the users in the UL channel. Hence, data detection for UL users becomes more erroneous, incredibly deteriorating the UL performance. For such a case, the total SE of the FD system is mostly determined by the DL transmission since the SE of the UL channel is extremely low. Thus, it is better to reduce the transmit power in the UL channel and concentrate on maximizing the SE of the DL channel. As a result, the SE of the UL channel greatly declined. Specifically, the SE of the UL direction of the FD system is significantly smaller than that of the HD one as $\sigma^2_{SI} \geq -80$ and $\sigma^2_{SI} \geq -70$ for $(P_{BS}, \eta) = (26$ dBm, $23$ dBm) and $(P_{BS}, \eta) = (10$ dBm, $10$ dBm), respectively. It is worth noting that a reduction in the transmit power of users in UL channel results in a decrease in the CCI. This explains the increment of the SE gain in the DL transmission as $\sigma^2_{SI}$ is greater than a certain threshold. An interesting observation from Fig. 13c is that the SE gain of the FD system is higher when the maximum transmit power is smaller. This is due to the fact that smaller maximum transmit powers create a smaller amount of residual SI as well as CCI.

In Fig. 14, we show empirical cumulative distribution function (CDF) of the total SE gain of the FD for the scenario in Fig. 12a. Obviously, for the same power setting, a smaller value of $\sigma^2_{SI}$ results in better SE gain. On the other hand, for the same $\sigma^2_{SI}$, a lower transmit power yields better SE gain. These observations are consistent with the observation in Fig. 13c.
Fig. 14. CDF of total SE gains for 5000 random channel realizations for the scenario shown in Fig. 12a. [38] ©2014 IEEE.
Average spectral efficiency gain CDF

$\sigma^2_{SI} = -90, (P_{BS}, q_{U_j}) = (26, 23)$

$\sigma^2_{SI} = -90, (P_{BS}, q_{U_j}) = (10, 10)$

$\sigma^2_{SI} = -80, (P_{BS}, q_{U_j}) = (26, 23)$

$\sigma^2_{SI} = -80, (P_{BS}, q_{U_j}) = (10, 10)$

(a) CDF of average SE gain of DL channel.

(b) CDF of average SE gain of UL channel.
The performance of the FD is further explored in the next numerical experiment, in which we study the CDF of the average SE gain of the FD system for a number of random scenarios. The results in Fig. 15 are plotted for 1000 scenarios, where all users are uniformly distributed in a circle area of a radius \( r = 100 \) meters centered at the BS. As can be seen in Fig. 15c, the total average SE of FD systems is higher than that of the HD one for most of the scenarios. For example, the SE gains are larger than 20% and 28% for the power settings \((P_{\text{BS}}, q_{U_j}) = (26 \text{ dBm}, 23 \text{ dBm})\) and \((P_{\text{BS}}, q_{U_j}) = (10 \text{ dBm}, 10 \text{ dBm})\), respectively for a half of the simulated scenarios at \(\sigma_{\text{SI}}^2 = -80 \text{ dB}\). Not surprisingly, the SE gain of the DL channel is rather sensitive to scenarios which determine the degree of CCI. On the other hand, positions of users have a small impact on the SE of the UL transmission when \(\sigma_{\text{SI}}^2 = -90 \text{ dB}\). The reason is that the residual SI in this case is relatively lower than the received signal strength for most of the scenarios. However, the situation dramatically changes as \(\sigma_{\text{SI}}^2\) increases to \(-80 \text{ dB}\), where more dependency between topology and SE gain is observed. Thus, the number of scenarios that can yield a received signal strength higher than the SI power is reduced for a larger value of \(\sigma_{\text{SI}}^2\).
Distance from the uplink user to the downlink user, $d_{CCI}$ (m)

(a) Average SE of DL channel.

(b) Average SE of UL channel.
Next, we study the impact of CCI on the SE of the FD system. For this purpose, we fix $\sigma_S^2$ at −100 dB, and consider a setting shown in Fig. 12b. In this simulation setup, we vary the distance between $U_1$ and $D_1$, denoted by $d_{CCI}$, and plot the resulting SEs of the FD system in Fig. 16. Each value of $d_{CCI}$ on the x-axis of Fig. 16 corresponds to a position of $U_1$, while $D_1$ is held fixed. We observe that the SE of $D_1$ increases as $U_1$ moves far away from $D_1$. Especially when the two users are close (e.g., $d_{CCI} < 64.82$ m), the performance of the FD DL transmission can be worse than that of the HD one. The reason is straightforward since decreasing $d_{CCI}$ leads to an increase in CCI which then degrades the SE of the DL channel. On the other hand, the location of $U_1$ has a small impact on the SE of the UL transmission for a fixed small value of $\sigma_S^2$. The results in Fig. 16 indicate that the CCI is a critical factor that needs to be controlled for successful deployment of FD systems.

In the final numerical experiment we plot the CDF of the average total SE of the FD system with and without accounting for the CCI. The problem of beamformer design without taking CCI into account was studied in [37]. It is obvious that the proposed designs in this chapter outperform the one with no CCI in [37] as expected. For instance, the proposed designs attain 2 bits/s/Hz of total SE higher than the scheme in [37] for
Fig. 17. CDF of average total SE of the proposed design and the design with no consideration of CCI in [37] for 1000 random scenarios. For each topology, the average total SE is calculated over 500 random channel realizations. [38] ©2014 IEEE.

approximately 60% of the simulated scenarios when $K_D = 3$ and $K_U = 2$. As the total number of users is reduced, the SE becomes smaller due to a decrease in the available multiuser diversity gain.

4.4 Summary and discussion

In this chapter we have attempted to explore the potential gains of the FD communications for small-cell networks. The focus of this chapter is driven by the fact that small-cell networks are envisioned to be one of importance use cases for FD communications. To do so we have adopted beamforming technique and considered the problem of joint SE maximization of DL and UL transmissions under some power constraints. First, the design problem is formulated as a rank constrained optimization one, and then the rank relaxation technique is employed to attain a relaxed problem, which is still nonconvex. We note that the rank relaxation technique, commonly known as semidefinite relaxation method under various contexts, is widely used to solve the
problem of linear precoder design in MIMO DL channels, e.g., in [27, 71, 99, 109–111]. Very often, the relaxed problems in those cases are convex and general convex program solvers can be called upon to find the solutions. Moreover, in some special cases, the rank relaxation is proved to be tight [71, 111, 112]. The same property unfortunately does not carry over into our case.

To solve the nonconvex relaxed problem we have proposed two iterative algorithms, one based on the concept of the Frank-Wolfe algorithm and the other based on the framework of SPCA method. The idea of both proposed methods is to approximate the nonconvex problem by a convex formulation in each iteration. While the first approach needs to solve a sequence of MAXDET programs, the second one relies on solving a series of SDPs. Using the two proposed solutions we have carried out several numerical experiments under 3GPP LTE small-cell setups to evaluate the SE performance of the FD scheme. It has been shown that the SE of the FD system is remarkably larger than that of the HD one as the capability of current SI cancellation schemes is efficient. Our designs have proved that the FD transmission is a promising technique to improve the SE of small-cell wireless communications systems.
5 Conclusion and future work

5.1 Conclusion

This thesis has proposed transmission schemes to maximize the SE and EE of a FD system, in which a FD capable BS communicates with multiple HD users in the DL and UL channels at the same time on the same frequency. For this problem, we have proposed a joint optimization approach that simultaneously optimizes the DL and UL transmissions due to the natural coupling of both channels of the FD operation. Since the joint design problems are non-convex in nature, it is hard to find the globally optimal solutions. Thus, we have applied a rank relaxation method to the non-convex problems of interest and then presented low-complexity iterative algorithms to find locally optimal solutions. The proposed iterative methods are summarized as follows.

– For the optimization problem with the assumption of no CCI in Chapter 3, the iterative algorithms of the linear precoder design for maximizing SE and EE have been devised in the light of SPCA. In this way, the non-convex relaxed problems are approximated by the convex programs in each iteration. Then a solution of the resulting convex program, which has an analytical structure, can be found by numerical methods such as the BCA and dual decomposition.

– For the general design taking the residual SI and CCI into account in Chapter 4, two iterative algorithms of the SE maximization problem have been developed. The first one is based on the concept of the Frank-Wolfe algorithm and the other based on the framework of SPCA method. The idea of both proposed methods is to approximate the non-convex problem by a convex formulation in each iteration. While the first approach needs to solve a sequence of MAXDET programs, the second one relies on solving a series of SDPs.

We have carried out extensive numerical experiments to evaluate the SE and EE performance of the FD systems of interests. It has been indicated that the SE of the considered FD system is better than that of the HD counterpart as the residual SI power is efficiently canceled. The proposed EE design is superior to the proposed SE design with regard to EE, and yields lower EE than the HD system in the low transmit power regime. However, in the high transmit power region, the EE design outperforms the HD one as the residual SI power is small.
5.2 Future work

The work considered in this thesis also opens several possibilities for future research as follows.

- The solutions for the linear processing design problems in this thesis can be slightly modified to arrive at a centralized joint design for a multi-cell deployment scenario. Specifically, if all the CSI can be timely forwarded to the centralized processing unit, a joint design can be derived easily by including the multicell interference term when appropriate. Moreover, distributed solutions and related matters such as synchronization are of more interest from a practical perspective.

- The proposed designs in Chapters 3 and 4 have been developed under the assumption of the knowledge of residual SI channel and perfect CSI of wireless links known at the FD BS and users. Thus, the achievable SE and EE of the considered FD system with the proposed transmission schemes can be considered as the upper bounds. However, there are several reasons such as transmitter/receiver dynamic range, noisy channel estimate, channel feedback delay, and processing delay, etc., resulting in imperfect/outdated CSI of communications links as well as the SI channel in real systems. Therefore, this leads to the development of a new class of transmission strategies for such cases, e.g., using the framework of robust optimization as done in [113, 114].

- As mentioned in Chapter 3, when a number of users exceeds an acceptable number of users which can be served in the wireless system, a set of users must be selected in advance by a scheduler. Note that many existing user scheduling algorithms [82–84, 115, 116] in literature can be applied in combination with our proposed designs for the FD MU-MIMO system but they do not probably achieve the best performance. The reason is that these user schedulers were designed for one direction transmission only (e.g., for DL channel [82–84] or for UL channel [115, 116]), and did not consider the impact of the residual SI and CCI incurred by the nature of the FD operation. Therefore, new user scheduling schemes for the FD MU-MIMO network need to be developed to exploit the multiuser diversity gain in both DL and UL channels.

- Another issue of the FD system is that depending on network deployment (user locations), residual SI caused by FD transmission and typical propagation characteristics of wireless channels (time variation, frequency fluctuation), some HD scenarios are superior to the FD ones. To obtain the best trade-off of FD and HD networks,
the conditions for which FD or HD performs surpassingly should be investigated, and from this the adaptive mode switching schemes between the FD and HD need to be proposed. The future research can also include new necessary protocols and algorithms to support coexistence between the two technologies due to the majority of devices in existing networks operate in HD mode.

– A promising research topic is on exploring the potential gains of the combination of massive MIMO with FD communications. Massive MIMO uses a very large number of antennas with low power each. Thus it may be easier to implementation of FD communications on massive MIMO platform. In this regard a new class of solutions methods for large-scale radio resource management problems are a rich research area.
References


  Microwave Journal 6(9): 63–71.
  antenna design and interference cancellation algorithms for in-band full duplex relays. IEEE 
42. Riihonen T, Werner S & Wichman R (2011) Mitigation of loopback self-interference in 
  658–662.
44. Ju H, Oh E & Hong D (2009) Improving efficiency of resource usage in two-hop full 
  duplex relay systems based on resource sharing and interference cancellation. IEEE Trans. Wireless 
  Commun. 8(8): 3933–3938.
  4381–4393.
47. Duarte M, Sabharwal A, Aggarwal V, Jana R, Ramakrishnan K, Rice C & Shankaranarayanan 
  capacity with common-carrier full duplex radios. In: Proc. IEEE International Conference 
  on Communications (ICC 2014), pp. 4987–4993.
  63(12): 5250–5266.
  perspective of PHY and MAC layers. IEEE Communications Surveys Tutorials PP(99).
  full-duplex wireless: Challenges and opportunities. IEEE J. Sel. Areas Commun. 32(9): 
  1637–1652.
  Signals, Systems and Computers (ASILOMAR), pp. 29–33.
  pp. 11–16.


Appendix 1: On the rank of optimal solutions to the problem in (41). [37] ©2013 IEEE.

As mentioned earlier, the proposed precoder design offers optimal solutions that meet the rank constraint automatically. In this appendix, we show that the iterative algorithm in which a generic method is used to compute the optimal precoders in each iteration can also yield the same result for a specific condition. Explicitly, let $\lambda^*$ be the optimal dual variable associated with the sum power constraint $\sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) \leq P_{BS}$ and $Q_{D_i}^*$'s be the optimal solutions to (43) after the iterative procedure converges. Then, we have the following claim.

Claim 1. If $\lambda^* > 0$, then it follows that rank$(Q_{D_i}^*) \leq N_{D_i}$, for all $i = 1, \ldots, K_D$.

The proof of Claim 1 follows similar arguments as in [117, 118]. We begin by forming the Lagrangian function of (43) for step $n+1$ of the iterative procedure, which is given by

$$
\mathcal{L}((Q_{D_i}), (Q_{U_j}), \lambda, \mu, (\bar{Q}_{D_i}), (\bar{Q}_{U_j})) =
\left(\text{R}_{1B}^{(n)}((Q_{D_i}), (Q_{U_j}))-\lambda \left( \sum_{i=1}^{K_D} \text{Tr}(Q_{D_i}) - P_{BS} \right) \right)
- \sum_{j=1}^{K_U} \mu_j \left( \text{Tr}(Q_{U_j}) - q_{U_j} \right) + \sum_{i=1}^{K_D} \text{Tr}(Q_{D_i} Q_{D_i}) + \sum_{i=1}^{K_D} \text{Tr}(\bar{Q}_{D_i} Q_{D_i})
\tag{102}
$$

where $\mu_j$ is the dual variables associated with the power constraint $\text{Tr}(Q_{U_j}) \leq q_{U_j}$ in the UL channel, $Q_{D_i} \succeq 0$ and $Q_{U_j} \succeq 0$ are the dual variables for the positive semidefinite constraints $Q_{D_i} \succeq 0$ and $Q_{U_j} \succeq 0$, respectively. We can then rewrite (102) as

$$
\mathcal{L}((Q_{D_i}), (Q_{U_j}), \lambda, \mu, (\bar{Q}_{D_i}), (\bar{Q}_{U_j})) =
\sum_{i=1}^{K_S} \log |I + \bar{H}_{S_i} Q_{D_i} \bar{H}_{S_i}^H| - \lambda \left( \sum_{i=1}^{K_S} \text{Tr}(Q_{D_i}) - P_{BS} \right)
- \sum_{j=1}^{K_U} \mu_j \left( \text{Tr}(Q_{U_j}) - q_{U_j} \right) - \text{Tr}\left( (\mathcal{T}^{(n)})^{-1} \frac{1}{N_0} \sum_{i=1}^{K_S} \bar{G}_{S_i} Q_{D_i} \bar{G}_{S_i}^H \right) + \sum_{i=1}^{K_S} \text{Tr}(\bar{Q}_{D_i} Q_{D_i})
+ \sum_{i=1}^{K_S} \text{Tr}(\bar{Q}_{U_j} Q_{U_j}) + \log |I + \frac{1}{N_0} \sum_{i=1}^{K_S} \bar{G}_{S_i} Q_{D_i} \bar{G}_{S_i}^H + \frac{1}{N_0} \sum_{j=1}^{K_U} \bar{H}_{U_j} Q_{U_j} \bar{H}_{U_j}^H|.
\tag{103}
$$
Taking the first derivative of $L(\{Q_{D_i}\}, \{Q_{U_j}\}, \lambda, \mu, \{\bar{Q}_{D_i}\}, \{\bar{Q}_{U_j}\})$ with respect to $Q_{D_i}$ and set it to zero, we have

$$\tilde{H}_D^H (I + \tilde{H}_D Q_{D_i}^* \tilde{H}_D^{-1}) - \tilde{H}_D + \frac{1}{N_0} \tilde{G}_S^H (\Xi^{-1} - (\Upsilon^{(n)})^{-1}) \tilde{G}_S - \lambda I + Q_{D_i}^* = 0$$  \hspace{1cm} (104)

where $\Xi \triangleq I + \frac{1}{N_0} \sum_{i=1}^{K_D} \tilde{G}_S Q_{D_i}^* \tilde{G}_S^H + \frac{1}{N_0} \sum_{j=1}^{K_U} H_{U_j} Q_{U_j}^* H_{U_j}^H$. In fact, (104) is the stationary property of the optimal solutions for (43) and we have used the fact that $\nabla X \log |I + HXH^H| = HXH$. Next, using the complementary slackness property $\bar{Q}_{D_i} Q_{D_i}^* = 0$, we obtain

$$\tilde{H}_D^H (I + \tilde{H}_D Q_{D_i}^* \tilde{H}_D^{-1}) - \tilde{H}_D Q_{D_i}^* = (\lambda^* I + \tilde{G}_S^H (N_0 \Upsilon^{(n)})^{-1} \tilde{G}_S - \tilde{G}_S^H (N_0 \Xi)^{-1} \tilde{G}_S) Q_{D_i}^*.$$  \hspace{1cm} (105)

From the definition of $\Upsilon^{(n)}$, the following inequality is obvious

$$\Xi = I + \frac{1}{N_0} \sum_{i=1}^{K_D} \tilde{G}_S Q_{D_i}^* \tilde{G}_S^H + \frac{1}{N_0} \sum_{j=1}^{K_U} H_{U_j} Q_{U_j}^* H_{U_j}^H \geq I + \frac{1}{N_0} \sum_{i=1}^{K_D} \tilde{G}_S Q_{D_i}^* \tilde{G}_S^H = \Upsilon^{(n)}.$$  \hspace{1cm} (106a)

Note that the equality in (106b) holds true due to the fact that the iterative process has converged. Since $\lambda^* > 0$, we can conclude that $\lambda^* I + \tilde{G}_S^H (N_0 \Upsilon^{(n)})^{-1} \tilde{G}_S - \tilde{G}_S^H (N_0 \Xi)^{-1} \tilde{G}_S \geq \lambda^* I > 0$. Then, it follows from (105) that $\text{rank}(Q_{D_i}) \leq \text{rank}(\tilde{H}_D) = \min(K_D, K^*_D)$, which completes the proof. Interestingly, we observe that $\lambda^* > 0$ always holds true in the simulation setups considered in Chapter 3. and in various setups as well.
Appendix 2: Proof of convergence of algorithm 5 and 6.

[38] ©2014 IEEE.

In this appendix we adopt the techniques from [80] to prove the convergence of Algorithms 5 and 6 (i.e., the iterative MAXDET-based algorithm and the iterative SDP-based algorithm, respectively) to a KKT point. Let us start with the convergence proof of Algorithm 5. First, we note that the affine majorization in (86) has the following two important properties which are the key to show the convergence to a KKT point of Algorithm 5.

\[
\begin{align*}
g^{(n)}(Q^{(n)}, q^{(n)}) &= g(Q^{(n)}, q^{(n)}), \quad (107) \\
\nabla g^{(n)}(Q^{(n)}, q^{(n)}) &= \nabla g(Q^{(n)}, q^{(n)}), \quad (108)
\end{align*}
\]

where property (107) means that the inequality in (86) is tight when \((Q, q) = (Q^{(n)}, q^{(n)})\) and property (108) is obvious due to the first order approximation. Note that the gradient in (108) is with respect to \(Q\) and \(q\). To proceed further, let \(S\) denote the feasible set of (89), i.e., the set of \(Q\) and \(q\) that satisfy the constraint (82b), (82c) and (82d). We note that \(S\) is a compact convex set. Further let \(u^{(n+1)}\) be the obtained optimal objective of (89) at iteration \(n+1\). According to the updating rule in Algorithm 5, we can derive the following inequalities

\[
\begin{align*}
u^{(n+1)} &= h(Q^{(n+1)}, q^{(n+1)}) - g^{(n)}(Q^{(n+1)}, q^{(n+1)}) \quad (109) \\
&= \max_{(Q, q) \in S} h(Q, q) - g^{(n)}(Q, q) \quad (110) \\
&\geq h(Q^{(n)}, q^{(n)}) - g^{(n)}(Q^{(n)}, q^{(n)}) \quad (111) \\
&= h(Q^{(n)}, q^{(n)}) - g(Q^{(n)}, q^{(n)}) \quad (112) \\
&\geq h(Q^{(n)}, q^{(n)}) - g^{(n-1)}(Q^{(n)}, q^{(n)}) = u^{(n)} \quad (113)
\end{align*}
\]

where (111) follows from the fact that the objective at the optimal solution is greater than the one at any feasible solution, i.e., \(f(x^*) = \max_{x \in \mathcal{X}} f(x) \geq f(x_0)\) where \(x^*\) and \(x_0\) are an optimal solution and any feasible solution, respectively, (112) is due to (107), (113) is due to the affine majorization in (86). In fact, we have shown that the sequence \(\{u^{(n)}\}\) is nondecreasing. Furthermore, the value of \(\{u^{(n)}\}\) is bounded above due to the limited transmit power, and thus it is guaranteed to converge. We note that the function \(f(X) = \log \det(X)\) is differentiable and strictly concave on \(X > 0\) [85, Section
we first introduce the set of dual variables for the constraints in (89) which is listed in Table 3. Thus, it is straightforward to see that the set of equations in (116)-(120) are actually the KKT conditions of the optimal value at iteration $n$ due to [80, Lemma 3.1]. As a result, the sequence $(Q^{(n)}, q^{(n)})$ converges to an accumulation point denoted by $(Q^*, q^*)$. To establish the convergence to a KKT point, we first introduce the set of dual variables for the constraints in (89) which is listed in Table 3.

It is easy to check that the Slater’s condition holds for the convex program at all iterations of Algorithm 5. Thus, the KKT conditions are necessary and sufficient for optimality [85, Section 5.5]. With the dual variables introduced in Table 3, the KKT conditions of the optimal value at iteration $n$ (see [85] for more details) are given as

$$ \nabla_{Q_i} h(Q^{(n)}, q^{(n)}) - \nabla_{Q_i} g^{(n)}(Q^{(n)}, q^{(n)}) - \mu I + Z_{B_i} = 0, \quad \forall i = 1, \ldots, K_D. \quad (114) $$

$$ \partial_{q_{ij}} h(Q^{(n)}, q^{(n)}) - \partial_{q_{ij}} g^{(n)}(Q^{(n)}, q^{(n)}) + \lambda_{ij} - \tilde{\lambda}_{ij} = 0, \quad \forall j = 1, \ldots, K_U. \quad (115) $$

$$ \lambda_{ij}, q_{ij}^{(n)} = 0; \tilde{\lambda}_{ij}, (q_{ij}^{(n)} - \tilde{q}_{ij}) = 0, \quad \forall j = 1, \ldots, K_U. \quad (116) $$

$$ \text{Tr}(Q_{B_i}^{(n)} Z_{B_i}) = 0, \quad \forall i = 1, \ldots, K_D. \quad (117) $$

$$ \mu \left( \sum_{i=1}^{K_D} \text{Tr}(Q_{B_i}^{(n)}) - P_{BS} \right) = 0. \quad (118) $$

Due to property (108), we can replace $\nabla_{Q_i} g^{(n)}(Q^{(n)}, q^{(n)})$ and $\partial_{q_{ij}} g^{(n)}(Q^{(n)}, q^{(n)})$ by $\nabla_{Q_i} g(Q^{(n)}, q^{(n)})$ and $\partial_{q_{ij}} g(Q^{(n)}, q^{(n)})$ on convergence (i.e., as $n \to \infty$), respectively. Thus,

$$ \nabla_{Q_i} h(Q^{(n)}, q^{(n)}) - \nabla_{Q_i} g(Q^{(n)}, q^{(n)}) - \mu I + Z_{B_i} = 0, \quad \forall i = 1, \ldots, K_D. \quad (119) $$

$$ \partial_{q_{ij}} h(Q^{(n)}, q^{(n)}) - \partial_{q_{ij}} g(Q^{(n)}, q^{(n)}) + \lambda_{ij} - \tilde{\lambda}_{ij} = 0, \quad \forall j = 1, \ldots, K_U. \quad (120) $$

It is straightforward to see that the set of equations in (116)-(120) are actually the KKT conditions for the problem (85) and thus completes the proof. We note that the KKT

### Table 3. Constraints and their corresponding dual variables.

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Dual variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 \leq q_{ij}$</td>
<td>$\lambda_{ij}$</td>
</tr>
<tr>
<td>$\frac{\text{sh)cine } q_{ij}}{\sum_{k=1}^{K_D} \text{Tr}(Q_{B_i}) \leq P_{BS}}$</td>
<td>$\lambda_{ij}$</td>
</tr>
<tr>
<td>$\sum_{k=1}^{K_D} \text{Tr}(Q_{B_i}) \leq P_{BS}$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>$Q_{B_i} \geq 0$</td>
<td>$Z_{B_i}$</td>
</tr>
</tbody>
</table>

112
conditions for the convex program after convergence are also the necessary ones for local optimality of the problem (85). Indeed since \((\mathbf{Q'}, \mathbf{q'})\) is an optimal solution to the convex program at convergence, it satisfies [75, Section 2.1]

\[
\langle \nabla g(\mathbf{q'}) - \nabla h(\mathbf{Q'}, \mathbf{q'}), (\mathbf{Q'}, \mathbf{q'}) - (\mathbf{Q'}, \mathbf{q'}) \rangle \geq 0 \quad \text{for all } (\mathbf{Q'}, \mathbf{q'}) \in \mathcal{S} \quad (121)
\]

where \(\langle \cdot, \cdot \rangle\) stands for the inner product of the arguments, i.e., \(\langle \mathbf{X}, \mathbf{Y} \rangle = \text{Tr}(\mathbf{X}^\top \mathbf{Y})\), the subtraction in (121) is element-wise, and the gradient is with respect to \(\mathbf{Q}\) and \(\mathbf{q}\). As mentioned previously, we can replace \(\nabla g(\mathbf{q'})\) by \(\nabla g(\mathbf{Q'}, \mathbf{q'})\), and thus (121) becomes

\[
\langle \nabla g(\mathbf{Q'}, \mathbf{q'}) - \nabla h(\mathbf{Q'}, \mathbf{q'}), (\mathbf{Q'}, \mathbf{q'}) - (\mathbf{Q'}, \mathbf{q'}) \rangle \geq 0 \quad \text{for all } (\mathbf{Q'}, \mathbf{q'}) \in \mathcal{S} \quad (122)
\]

which is the first order necessary condition for local optimality of the problem (85) [75, Section 2.1].

The proof of Algorithm 6 follows in the same spirit. As mentioned earlier for the convex approximation in (95), \(\tilde{F}(\mathbf{t}_i, \beta_{D_i}, \psi_{D_{i_j}}^{(n)}) = f(\mathbf{t}_i, \beta_{D_i})\) when \(\psi_{D_{i_j}}^{(n)} = \mathbf{t}_i / \beta_{D_i}\), that is

\[
F(\mathbf{t}_i, \beta_{D_i}, \psi_{D_{i_j}}^{(n)})|_{\psi_{D_{i_j}}^{(n)} = \mathbf{t}_i / \beta_{D_i}} = \mathbf{t}_i \beta_{D_i} = f(\mathbf{t}_i, \beta_{D_i}). \quad (123)
\]

Furthermore, we also have

\[
\frac{\partial F(\mathbf{t}_i, \beta_{D_i}, \psi_{D_{i_j}}^{(n)})}{\partial \mathbf{t}_i} |_{\psi_{D_{i_j}}^{(n)} = \mathbf{t}_i / \beta_{D_i}} = \frac{1}{\psi_{D_{i_j}}^{(n)}} \bigg|_{\psi_{D_{i_j}}^{(n)} = \mathbf{t}_i / \beta_{D_i}} = \frac{\partial f(\mathbf{t}_i, \beta_{D_i})}{\partial \mathbf{t}_i} \quad (124)
\]

and

\[
\frac{\partial F(\mathbf{t}_i, \beta_{D_i}, \psi_{D_{i_j}}^{(n)})}{\partial \beta_{D_i}} |_{\psi_{D_{i_j}}^{(n)} = \mathbf{t}_i / \beta_{D_i}} = \frac{\partial f(\mathbf{t}_i, \beta_{D_i})}{\partial \beta_{D_i}}. \quad (125)
\]

Let \(\mathcal{S}(n)\) be the feasible set of the convex program solved at iteration \(n\). Due to the updating rule in Algorithm 6 (i.e., \(\psi_{D_{i_j}}^{(n+1)} = \mathbf{t}_i / \beta_{D_i}^{(n+1)}\)), follows that \(F(\mathbf{t}_i, \beta_{D_i}^{(n+1)}, \psi_{D_{i_j}}^{(n+1)}) = f(\mathbf{t}_i, \beta_{D_i})\). Similarly, we have \(G(x_{D_{i_j}}, \mathbf{Q}, q_{D_{i_j}}^{(n)}, \mathbf{Q}^{(n)}, \mathbf{q}^{(n)}) = -g(x_{D_{i_j}}, \mathbf{Q}, \mathbf{q}^{(n)})\). This means that \((x_{D_{i_j}}^{(n)}, \mathbf{Q}^{(n)}, q_{D_{i_j}}^{(n)}) \in \mathcal{S}(n+1)\) and thus \(u^{(n+1)} \geq u^{(n)}\) where \(u^{(n)}\) is the objective of (91) at iteration \(n\). The convergence proof to a solution that satisfies KKT conditions follows the same steps from (109) to (120) presented above.
547. Pääkkönen, Tiina (2015) Improving the energy efficiency of processes: reduction of the crystallization fouling of heat exchangers
549. Skön, Jukka-Pekka (2015) Intelligent information processing in building monitoring systems and applications
550. Irannezhad, Masoud (2015) Spatio-temporal climate variability and snow resource changes in Finland
551. Pekkinen, Leena (2015) Information processing view on collaborative risk management practices in project networks
553. Nelo, Mikko (2015) Inks based on inorganic nanomaterials for printed electronics applications
558. Aula, Matti (2016) Optical emission from electric arc furnaces
559. Ferdinand, Nuwan Suresh (2016) Low complexity lattice codes for communication networks
560. Xue, Qiang (2016) Analysis of near-optimal relaying schemes for wireless tandem and multicast relay networks

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/
Vu Thuy Dan Nguyen

TRANSMISSION STRATEGIES FOR FULL-DUPLEX MULTIUSER MIMO COMMUNICATIONS SYSTEMS