UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

KHALED NAFEZ RAUF ARDAH

DECENTRALIZED ALGORITHMS FOR MULTICELL MULTIUSER MIMO WIRELESS MOBILE NETWORKS

FORTALEZA
2018

KHALED NAFEZ RAUF ARDAH

DECENTRALIZED ALGORITHMS FOR MULTICELL MULTIUSER MIMO WIRELESS MOBILE NETWORKS

Tese apresentada ao Curso de Doutorado em Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos para obtenção do Título de Doutor em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas

Orientador: Prof. Dr. Francisco Rodrigo Porto Cavalcanti

Coorientador: Prof. Dr. Yuri Carvalho Barbosa Silva

FORTALEZA
2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

A719d Ardah, Khaled Nafez Rauf.
Decentralized Algorithms for Multicell Multiuser MIMO Wireless Mobile Networks / Khaled Nafez Rauf Ardah. – 2018.
105 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2018.
Orientação: Prof. Dr. Francisco Rodrigo Porto Cavalcanti.
Coorientação: Prof. Dr. Yuri Carvalho Barbosa Silva.

1. Células pequenas. 2. maximização da taxa de soma ponderada. 3. minimização de potência de transmissão. 4. TDD dinâmico. 5. método de multiplicadores de direção alternada. I. Título.

CDD 621.38

KHALED NAFEZ RAUF ARDAH

DECENTRALIZED ALGORITHMS FOR MULTICELL MULTIUSER MIMO WIRELESS MOBILE NETWORKS

Presented Thesis for the Post-graduate Program in Teleinformatics Engineering of Federal University of Ceará as a partial requisite to obtain the Ph.D. degree in Teleinformatics Engineering. Concentration Area: Signals and Systems

Approved at: 09-03-2018.

EXAMINATION BOARD

Prof. Dr. Francisco Rodrigo Porto Cavalcanti (Orientador)
Universidade Federal do Ceará

Prof. Dr. Yuri Carvalho Barbosa Silva (Coorientador)
Universidade Federal do Ceará

Prof. Dr. JOSEF NOSSEK
Universidade Federal do Ceará

Prof. Dr. WALTER DA CRUZ FREITAS JUNIOR
Universidade Federal do Ceará

Prof. Dr. LUCIANO LEONEL MENDES
Instituto Nacional de Telecomunicações

Prof. Dr. Richard Demo
Universidade Federal do Paraná

To you: Palestine!

ACKNOWLEDGMENTS

While my name may be alone on the front cover of this thesis, I am by no means its sole contributor. Rather, there are a number of people behind this piece of work who deserve to be both acknowledged and thanked.

Firstly, I am forever indebted to my advisor Prof. Dr. Francisco. R. P. Cavalcanti and to my co-advisor Prof. Dr. Yuri C. B. Silva for receiving me at the Wireless Telecommunications Research Group (GTEL) and for the continuous support of my Ph.D study and research, for their patience, motivation, enthusiasm, and immense knowledge.

My sincere thanks also goes to Prof. Dr. Gabor Fodor, from Ericsson Research and Royal Institute of Technology, Stockholm-Sweden, for his patience, insightful comments, many Skype meetings sessions and uncountable e-mails discussing my results. Although, I have never met him in person, it has been a great pleasure working with him.

Also, I would like to use this opportunity to express my gratitude to my friends and colleagues at GTEL for their support and help throughout the course of this Ph.D journey. I also thank my family for supporting and encouraging me throughout my Ph.D journey and life. God bless you all.

Last but not the least, I acknowledge the technical and financial support from CAPES and Ericsson Innovation Center, Brazil, under EDB/UFC.35, EDB/UFC.42 and EDB/TIDE5G Technical Cooperation Contracts.

“Persistence is the shortest path to success!”
(Charles Chaplin)

RESUMO

Esta tese considera uma rede sem fio multicomponente multimídia MIMO e propõe algoritmos novos e descentralizados para resolver os seguintes problemas de pesquisa. Problema 1: como projetar os vetores de transmissão de feixe de transmissão que maximizam a taxa de soma ponderada do sistema (WSR), ao mesmo tempo em que satisfazem as restrições de energia nos transmissores, Problema 2: como projetar vetores robustos de transmissão de feixe de transmissão que minimizem a soma de potência de transmissão, satisfazendo Alvos de qualidade de serviço (QoS) dos usuários na presença de erros de canais e Problema 3: como selecionar de forma adaptável as direções de comunicação das células que maximizem o throughput dos usuários, considerando conjuntamente as condições de tráfego e os níveis de interferência. Em particular, são propostos três algoritmos diferentes e novos para resolver o Problema 1, que se baseiam na técnica de otimização alternada e garantidos para convergir para um ótimo local de WSR. Para facilitar a implementação dos algoritmos, um novo esquema de sinalização de OTA (over-the-air) é então proposto com base no modo de duplicação de divisão de tempo (TDD). Além disso, é proposto um novo algoritmo distribuído e robusto de formação de feixe coordenado (CBF) com base nas técnicas de programação semideterminada relaxada (SDP) e alternância de direção de multiplicadores (ADMM) para resolver o Problema 2, onde a forma de feixe robusta é abordada usando o pior caso criação de otimização. Além disso, propõe-se uma nova técnica de reconfiguração de células que maximiza o débito dos usuários, considerando conjuntamente tanto as condições de tráfego de IV como os níveis de interferência. As avaliações de algoritmos são realizadas usando simulador de computador, a partir do qual a eficácia dos algoritmos propostos é evidenciada, em comparação com os algoritmos de referência, em termos de eficiência espectral, eficiência de energia, taxa de convergência, sobrecarga de sinalização e complexidade. Os algoritmos propostos são descentralizados no sentido de que cada transmissor pode agir de forma independente, assim que tiver a informação necessária, o que torna os algoritmos propostos nesta tese especialmente adequados para redes sem fio atuais e futuras.

Palavras-chave: Células pequenas, maximização da taxa de soma ponderada, minimização de potência de transmissão, reconfiguração celular, TDD dinâmico, programação semideterminada, método de multiplicadores de direção alternada, otimização específica de enxame, programação linear inteira.

ABSTRACT

Small cells deployment is one of key technologies that is introduced to improve cellular communication systems' performance, since it provides a low-cost approach to reuse system resources. However, densifying cellular systems with small cells increases the inter-cell interference (ICI), which would degrade the system performance if not properly managed. Also, small cells are expected to have a burst-like traffic with strong fluctuation between uplink and downlink traffics, since the number of users served by small cells are expected to vary strongly with time and between adjacent cells. Complementing small cells with multiple-input multiple-output (MIMO) and dynamic TDD (DTDD) technologies can be seen as a key solution to cope with ICI effects and traffic fluctuations. While MIMO technology has great potential to achieve higher throughput, improve system capacity, and enhance spectral efficiency by serving multiple users and spatially eliminate/manage interference, DTDD technology allows each cell to adaptively reconfigure its communication direction based on the prevailing traffic demands and interference levels. This thesis considers a multicell multiuser MIMO wireless network and proposes novel and decentralized algorithms for solving the following research problems. Problem 1: how to design the transmit beamforming vectors that maximize the system weighted sum-rate (WSR), while satisfying the power constraints at transmitters, Problem 2: how to design a robust transmit beamforming vectors that minimize the sum transmit power, while satisfying the users' quality-of-service (QoS) targets in the presence of channel errors, and Problem 3: how to adaptively select the cells communication directions that maximize the users' throughput, while jointly considering their traffic conditions and interference levels. In particular, three different and novel algorithms are proposed for solving Problem 1, which are based on the alternating optimization technique and guaranteed to converge to a local WSR-optimum. Further, a novel distributed and robust coordinated beamforming (CBF) algorithm based on alternating direction method of multipliers (ADMM) technique is proposed for solving Problem 2, where the robust beamforming is tackled using a worst-case optimization criterion. For Problem 3, a novel cell reconfiguration technique is proposed that maximizes the users' throughput, while jointly considering both the prevailing traffic conditions and interference levels. Algorithms evaluations are carried out using computer simulation, from which the effectiveness of the proposed algorithms is evidenced, as compared to reference algorithms, in terms of spectral-efficiency, power-efficiency, convergence rate, signaling overhead, and complexity. The proposed algorithms are decentralized in the sense that each transmitter can act independently, as soon as it has the required information, which makes the proposed algorithms in this thesis especially suitable for current and future wireless networks.

Keywords: Sum rate maximization, coordinated beamforming, dynamic TDD, alternating direction method of multipliers, particle swarm optimization, integer linear programming.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

3GPP	3rd generation partnership project
ADMM	alternating direction method of multipliers
BC	broadcast channel
BS	base station
CBF	coordinated beamforming
CR	cognitive radio
CSI	channel state information
DoF	degree of freedom
DPC	dirty paper coding
DTDD	dynamic TDD
FDD	frequency division duplex
ICI	inter-cell interference
ILP	integer linear programming
KKT	Karush-Kuhn-Tucker
LMI	linear matrix inequality
LTE	long-term evolution
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
MMSE	minimum-mean-square error
mmWave	millimeter wavelength
MRT	maximum ratio transmission
MS	mobile station
NP	nondeterministic polynomial-time
OTA	over-the-air
PSO	particle swarm optimization
QoS	quality-of-service
SDMA	space-division multiple access
SDP	semidefinite programming
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
STDD	static TDD
TDD	time division duplexing
TTI	transmission time interval
WMMSE	weighted MMSE
WSR	weighted sum-rate
ZF	zero-forcing

NOTATIONS

In this thesis the following notations are used. Scalar variables are denoted by lower-case letters (a, b, ..., α, β, ...), vectors are written as boldface lower-case letters ($\mathbf{a}$, $\mathbf{b}$, ..., $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, ...), matrices correspond to boldface capitals ($\mathbf{A}$, $\mathbf{B}$, ...), and sets are written as calligraphic letters ($\mathcal{A}$, $\mathcal{B}$, ...). The meaning of the following symbols are listed below, unless otherwise explicitly stated:

$\mathbb{C}$	set of complex-valued numbers
$\mathbb{Z}$	set of binary (0-1) numbers
$\mathbb{R}$	set of integer real numbers
$(\mathbf{a})^{\mathrm{H}}$	complex conjugate transpose
$\|\mathbf{a}\|$	l_2 norm
$\|a\|$	the amplitude of a scaler
$\|\mathcal{A}\|$	cardinality of a set
$(\mathbf{A})^{-1}$	matrix inverse
$(\mathbf{A})^{\dagger}$	matrix pseudo-inverse
$\left\|\mathbf{A}\right\|$	determinant of a matrix
$\mathrm{Tr}(\mathbf{A})$	matrix trace operator
$\log(a)$	logarithm of base 2
$\mathbb{E}(\mathbf{a})$	statistical expectation
$\mathbf{A}_{[1:N]}$	the first N vectors of a matrix
$\mathbf{a}_{[i]}$	the i-th element of a vector
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product
$\mathbf{a} \odot \mathbf{b}$	dot-product between $\mathbf{a}$ and $\mathbf{b}$
bdiag{**a**/**A**}	denotes the block diagonalization of a given vectors/matrices

TABLE OF CONTENTS

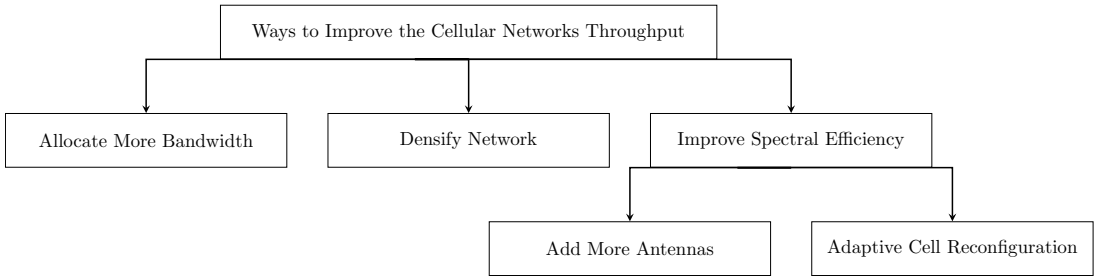
1 INTRODUCTION

Wireless communication cellular networks have drastically changed over the past few years. Originally designed for wireless voice communications, cellular networks of nowadays are dominated by wireless data communications. The amount of wireless voice and data traffic has grown at an exponential pace, and it will definitely continue to increase in the foreseeable future, where video-on-demand accounts for the majority of traffic. Therefore, it is important to evolve the current wireless communications technologies to meet the continuously increasing demand and to satisfy the rising expectations of service quality. The vision for future cellular networks is to achieve one-thousand folds (1000x) higher *area throughput* comparing to current cellular networks [1]. The area throughput is a performance metric that is measured in bits-per-second-per-square-kilometer (bit/s/km^2) and can be modeled using the following high-level formula:

$$\underbrace{\text{Area throughput}}_{\text{bit/s/km}^2} = \underbrace{\text{bandwidth}}_{\text{Hz}} \times \underbrace{\text{average cell density}}_{\text{cells/km}^2} \times \underbrace{\text{spectral efficiency per cell}}_{\text{bit/s/Hz/cell}}. \quad (1.1)$$

Consequently, there are three main ways to improve the cellular networks' throughput, as shown in Fig. 1.1, which are discussed in the following sections.

Figure 1.1 – Ways to improve the cellular networks throughput.



Source: Created by author.

1.1 Ways to improve the cellular networks throughput

1.1.1 Allocate more bandwidth

One potential solution to improve the cellular networks' throughput is to increase the system bandwidth. Current cellular networks use a frequency range below 6 GHz and utilize collectively more than 1 GHz of bandwidth [2]. Therefore, to achieve the 1000x vision in the future cellular networks, 1 THz of bandwidth or more is required, which entails using much higher frequency bands than 6 GHz. This is physically impractical since the frequency spectrum is a global resource that is shared among many different services, and the use of higher frequency bands limits the range and service reliability [2]. However, there are substantial bandwidths in

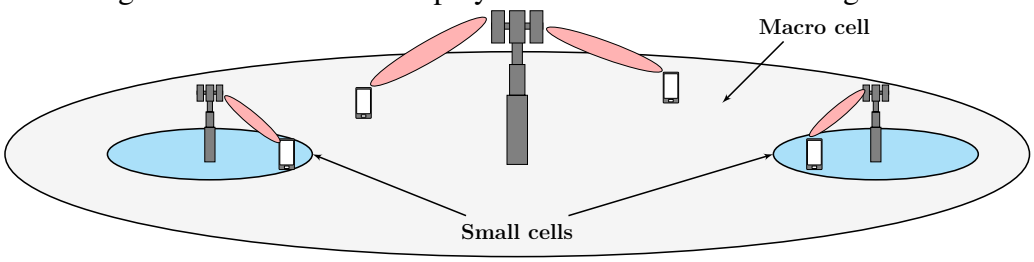
the millimeter wavelength (mmWave) bands (e.g., in the range 30–300 GHz) that are currently being investigated to be used in future cellular communications for short-range applications and small cells deployments [3]. The developed algorithms in this thesis assume use of low system bandwidth and frequency ranges, since the use of large system bandwidth and high-frequency bands adds some hardware constraints that should be considered in the practical implementations. In chapter 5 of this thesis, some future research directions are discussed including use of mmWave bands in the future cellular communications.

1.1.2 Densify network

Another potential solution to improve the cellular networks' throughput is to densify the cellular network by deploying more base stations (BSs) per area, which allows for reusing the system resources that can substantially improve the area's throughput. However, this approach is associated with high deployment costs, inter-cell interference issues, and is not suitable for high-speed mobile stations (MSs), as they would have to switch the serving BS very often. In this line, small cells deployment (also known as heterogeneous networks) were recently introduced by 3rd generation partnership project (3GPP) long-term evolution (LTE)-Advanced [4]. Small cells are meant to be deployed in hot-spots within macro cell coverage area (where users speed is low) to form a two-tier wireless communication network (see Fig. 1.2) to further increase the system coverage, capacity, and throughput [5]. It can also reduce the required transmit power, since the BSs and their MSs are closer to each other.

An interesting type of deployment contemplated in 3GPP LTE-Advanced networks is the non-co-channel small-cells deployment, where each network tier (small cells tier and macro cells tier) uses different carrier frequencies. Although the cross-tier interference can be eliminated in this case, the co-tier interference between small cells becomes a major problem that can significantly reduce the system throughput, especially if small cells are densely deployed. Therefore, interference management techniques become increasingly more important.

Figure 1.2 – Small cells deployed within macro cell coverage area.



Source: Created by author.

1.1.3 Improve spectral efficiency

Increasing cell density and using larger bandwidth have historically dominated the evolution of cellular networks, which explains why we are approaching a saturation point where

further improvements are increasingly complicated and expensive [2]. Therefore, the focus of this thesis is to propose methods and techniques to improve the system spectral efficiency.

Spectral efficiency measures the average number of bits that it can reliably be transmitted over channel under consideration per second per hertz (bits/s/Hz). The maximum spectral efficiency per user is determined by the channel capacity, which is defined by the Shannon formula that is in function of signal-to-interference-plus-noise ratio (SINR) as follows:

$$\text{Spectral efficiency} = \log\Big(1 + \text{SINR}\Big) = \log\Big(1 + \frac{\text{signal power}}{\text{interference power} + \text{noise powor}}\Big). \quad (1.2)$$

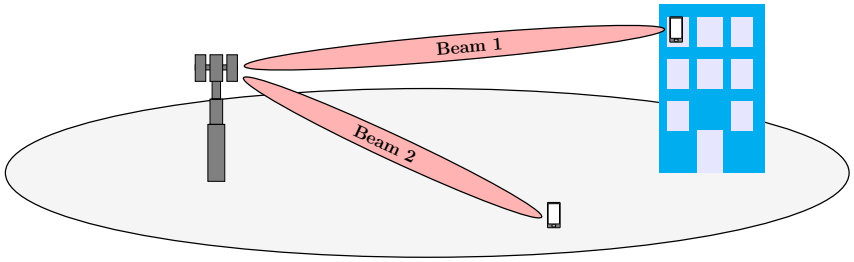
Therefore, to improve the system spectral efficiency, one should improve the SINR by increasing the signal power and/or reducing the interference power. There are different ways to improve the spectral efficiency in cellular networks. A direct approach is to use more transmit power. However, the positive effect quickly pushes the network into an interference-limited regime where no extraordinary spectral efficiency can be obtained. Furthermore, transmit power is a valuable resource that should be minimized and used efficiently. Therefore, different approaches to improve the cell spectral efficiency are more desired. In the following, the two approaches used to improve the system spectral efficiency in this thesis are discussed.

Add more antennas

One approach to improve the cell spectral efficiency is to deploy/add more antennas at transmitter and/or receiver to form what is called MIMO technology/transmission. In wireless communications, transmit beamforming technique is used to increase the signal power at the intended user and reduce interference to non-intended users [6]. A high signal power is achieved by transmitting the same data signal from all antennas, but with different amplitudes and phases, such that the signal components add coherently at the user. Low interference is accomplished by making the signal components add destructively at non-intended users [7]. This can be done by using linear or non-linear transmission techniques. Non-linear techniques have been shown to outperform linear techniques and achieve channel capacity. The capacity-achieving downlink strategy is non-linear and uses dirty paper coding (DPC) [8]. However, it is widely considered that DPC technique has limited practical applications, due to its high complexity. Therefore, linear transmission beamforming techniques have gained more interest and were proven to achieve the same sum rate scaling law as DPC [6], while maintaining lower complexity.

Since transmit beamforming focuses the signal energy at certain places, less energy arrives in other places. This allows for so-called space-division multiple access (SDMA), where a number of users spatially separated are served simultaneously using the same radio resource (e.g. frequency channel and/or time slot) [7] (as illustrated in Fig. 1.3). One beamforming vector is assigned to each user and can be matched to its channel. Unfortunately, the finite number of transmit antennas only provides a limited amount of spatial directivity. This means that there are energy leakages between the users, which act as interference. While it is fairly easy to design a beamforming vector that maximizes the signal power at the intended user, it is quite

Figure 1.3 – Multiuser MIMO beamforming (SDMA).



Source: Created by author.

difficult to strike a perfect balance between maximizing the signal power and minimizing the interference leakage. In fact, the optimization of multiuser transmit beamforming is generally a nondeterministic polynomial-time (NP) hard problem [7].

Nevertheless, the optimal transmit beamforming is known in special cases and has a closed-form solution [9]. For example, in the low signal-to-noise ratio (SNR) regime, where system can be regarded as noise-limited, the optimal transmit beamforming is given by the *egoistic* maximum ratio transmission (MRT) (also known as matched-filter) approach, where its main objective is to maximize the signal power at the intended user. In the high SNR regime, where system can be regarded as interference-limited, on the other hand, the optimal transmit beamforming is given by the *altruistic* zero-forcing (ZF) approach, where its main objective is to eliminate the interference-leakage to the non-intended users [7]. However, both approaches are far from optimal in the moderate SNR regime. Thus, different alternative and iterative approaches have been proposed, where BSs jointly optimize the transmit beamforming vectors to find a good balance between maximizing the signal power and minimizing the interference leakage [9].

To this end, different iterative algorithms to optimize the transmit beamforming were recently proposed for different system models and optimization criteria. For example, the SINR maximization (maxSINR) and the MMSE subject to maximum transmit power are well-known optimization problems in the literature [10]. It was shown in many papers, e.g. [10], that the MMSE algorithm has better sum rate performance, while the maxSINR algorithm has faster convergence rate. Therefore, the MMSE problem got more attention and many papers have extended it to consider solving the weighted MMSE (MMSE) problem [11]. Here, the weights are used to prioritize users and they can be adjusted to guarantee some fairness among them. Interestingly enough, it turns out that the WMMSE minimization problem and the WSR maximization problem can be made equivalent by adaptively adjusting the users' weights [12]. Therefore, the difficult WSR maximization problem is solved indirectly by solving the easier WMMSE minimization problem. In this thesis, chapter 2 elaborates more on this topic, where three novel algorithms are proposed for maximizing the users' WSR in multicell MIMO-broadcast channel (BC) systems.

Adaptive cell reconfiguration

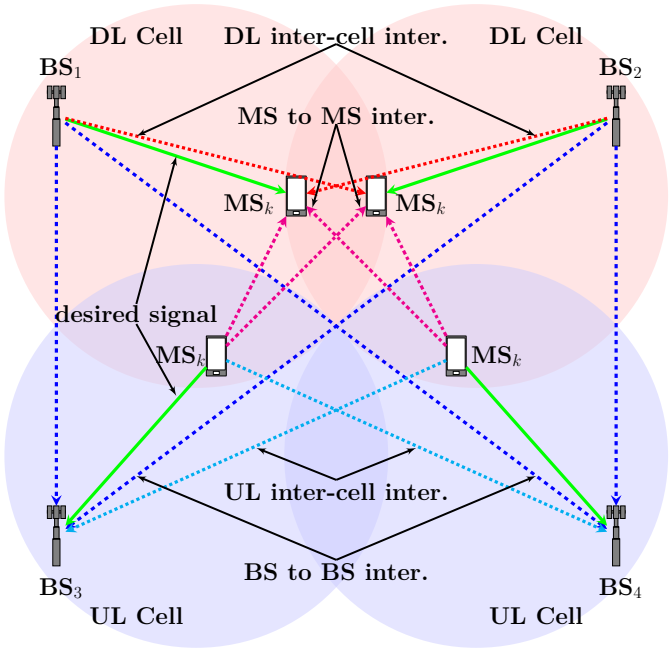
Another approach to improve the cell spectral efficiency is to adaptively select the cell's communication directions (downlink and uplink). It can be understood/shown that for a given channels realization, there is an optimal communication direction for each cell where it will have a higher signal power and a lower interference power that maximizes its spectral efficiency. This observation is rather old than new, since earlier cellular networks could have utilized it as well. However, the expected performance gain that can be achieved was low, since earlier cellular networks *almost* had symmetric traffic demands and co-channel interference on both communication directions. However, the introduction of small cells makes use of such approach more reasonably, since small cells are expected to have a burst-like traffic with strong fluctuation between uplink and downlink traffics. This is mainly due to the following observations: 1) the number of users served by small cells in hot-spots varies strongly with time and between adjacent cells and 2) the users of modern wireless networks typically demand a wide range of services, where each may have different traffic characteristics, in terms of packet size and maximum packet delay.

Among the two duplexing modes in LTE, time division duplexing (TDD) mode gains more importance for small cell deployments than frequency division duplex (FDD) mode [13], since it can be employed to provide unbalanced uplink-downlink data traffic. However, TDD in LTE is assumed to select a common TDD pattern for the whole network, which cannot be rapidly modified to match the instantaneous traffic demand [14]. Therefore, an adaptive cell reconfiguration technique called DTDD was recently introduced to cope with traffic fluctuations, which allows each cell to adaptively reconfigure its communication direction based on the prevailing traffic demands and interference levels. By means of simulation [15] and performance analysis [16], it was shown that DTDD technique enhances the system spectral and energy efficiency, especially in scenarios in which the offered traffic is time-varying and asymmetric in terms of uplink/downlink direction. However, allowing neighboring cells to have different transmission directions gives rise to BS-to-BS and MS-to-MS interference (see Fig. 1.4), among other impairments [17], which can severely degrade the system performance [18]. For that, several interference management techniques were investigated in [15], such as cell-clustering, power control, and interference suppressing and coordination techniques.

Clearly, the optimal performance in DTDD systems can be achieved by jointly optimizing the cells directions and transmit beamforming vectors. However, this will add a huge complexity on practical systems. Moreover, as specified in [15], the cell direction can be changed at the minimum every 10ms, in contrast to the transmit beamforming vectors that can be changed in every transmission time interval (TTI), i.e., every 1ms. Therefore, it is reasonable to separate both optimization problems; cells reconfiguration problem and transmit beamforming vectors design problem.

For the former problem i.e., cells reconfiguration problem, earlier works reconfigured each cell direction based only on the aggregate traffic in the cell [19]. While this approach is fairly

Figure 1.4 – Interference situations in dynamic TDD wireless networks.



Source: Created by author.

simple and inherently distributed, it cannot achieve the potential performance, as it disregards the interference effects that are particularly severe in DTDD systems. Recognizing this issue, cell reconfiguration schemes that account for the users' traffic demands and interference levels were recently proposed in [20, 21]. While these algorithms are shown to improve the system throughput, as compared to the conventional static TDD (STDD), they are scenario (traffic model) specific and do not consider the individual user's traffic characteristics [22], in terms of packet size and maximum packet delay. Therefore, it is important for the cell reconfiguration algorithm to support such different traffic characteristics. In this thesis, chapter 3 elaborates more on this topic, where a novel cell reconfiguration algorithm is proposed that takes into account both prevailing traffic conditions and multicell BS-to-BS and MS-to-MS interference levels.

For the latter problem i.e., transmit beamforming vectors design problem, the cells directions are generally assumed given and fixed, where the main objective is to optimize the transmit beamforming vectors that solve the considered optimization problem. To this end, one can optimize the transmit beamforming vectors that maximize the system WSR subject to a transmit power constraint or minimize the total transmit power subject to some users' QoS constraints. For the former problem, the proposed algorithms in chapter 2 can easily be extended to maximize the WSR in DTDD systems. We have done this extension in [23], but to avoid repeating a similar subject already treated in chapter 2, we turn our attention to the latter problem: minimize the total transmit power subject to some users' QoS constraints.

A notable technique for optimizing the transmit beamforming vectors is the CBF [24]. In CBF, each BS communicates with its own users, while minimizing the interference leakage to users in other cells. It has drawn significant attention recently due to its ability to handle

the interference problem using only channel state information (CSI), as compared to other interference management schemes that require data sharing as well, such as joint transmission [25]. It can be implemented in a centralized manner, where global CSI is made available to a central unit, or in a distributed manner, where each BS uses only local CSI. In the latter case, the coordination can be achieved by means of limited backhaul signaling between coordinated BSs. However, in practical scenarios, acquiring global CSI would drastically increase the backhaul signaling. Moreover, the BSs can never have perfect CSI, due to, for example, estimation errors and limited feedback channels [26]. Therefore, robust and distributed CBF solutions are much desired.

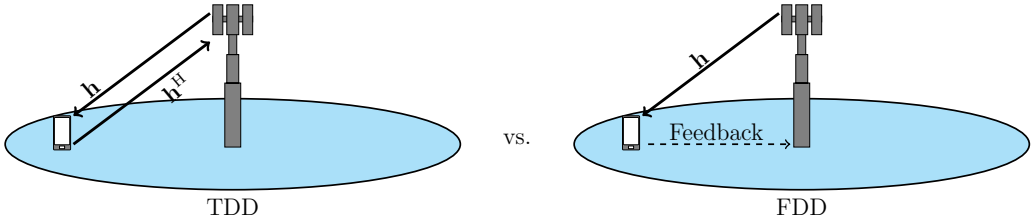
Earlier works on CBF were targeting the STDD systems, see [24, 27]. However, the interference situations in DTDD systems are more complicated, since the uplink and downlink users coexist at the same time among neighboring cells. Therefore, the interference management becomes more challenging and requires a special consideration from the optimization viewpoint. A possible solution is to formulate the optimization problem in DTDD systems as it is generally formulated in the cognitive radio (CR) networks [28], i.e., by assuming that the uplink cells are the *primary* cells and the downlink cells are the *secondary* cells and then include a threshold on the maximum BS-to-BS power from the downlink BSs to uplink BSs. In this case, not only the downlink performance targets can be guaranteed, but also the required uplink performance targets. In this thesis, chapter 4 elaborates more on this topic, where a novel distributed and robust CBF algorithm is proposed using relaxed semidefinite programming (SDP) [29] and ADMM [30] techniques for DTDD wireless networks.

1.2 Acquiring channel state information

The proposed algorithms in this thesis can be implemented in any *duplexing mode*, where the only difference would be on how the transceivers can acquire the CSI. There are two duplexing modes used in LTE cellular communication systems: TDD and FDD (see Fig. 1.5). The main difference between the two is on how the downlink and uplink communications are separated. In the TDD mode, the downlink and uplink are separated in time, where in the FDD mode, the downlink and uplink are instead separated in frequency. This means that the TDD mode uses a single frequency band for both downlink and uplink, while the TDD mode requires two separate frequency bands, one for downlink and another for uplink. Therefore, the channel responses in TDD mode are reciprocal, i.e., the downlink/uplink channel is only a transpose of uplink/downlink channel, thus, the downlink channel can be estimated from the uplink channel, and vice-versa. However, in FDD mode, the downlink and uplink channel responses are always different, which means that the estimates of the downlink/uplink channel responses need to be fed back to the BS/MS to enable transmit/receive beamforming computation.

The main method for CSI acquisition is pilot signaling, where a predefined pilot signal is transmitted from an antenna. Any other antenna in the network can simultaneously receive the transmission and compare it with the known pilot signal to estimate the channel from the

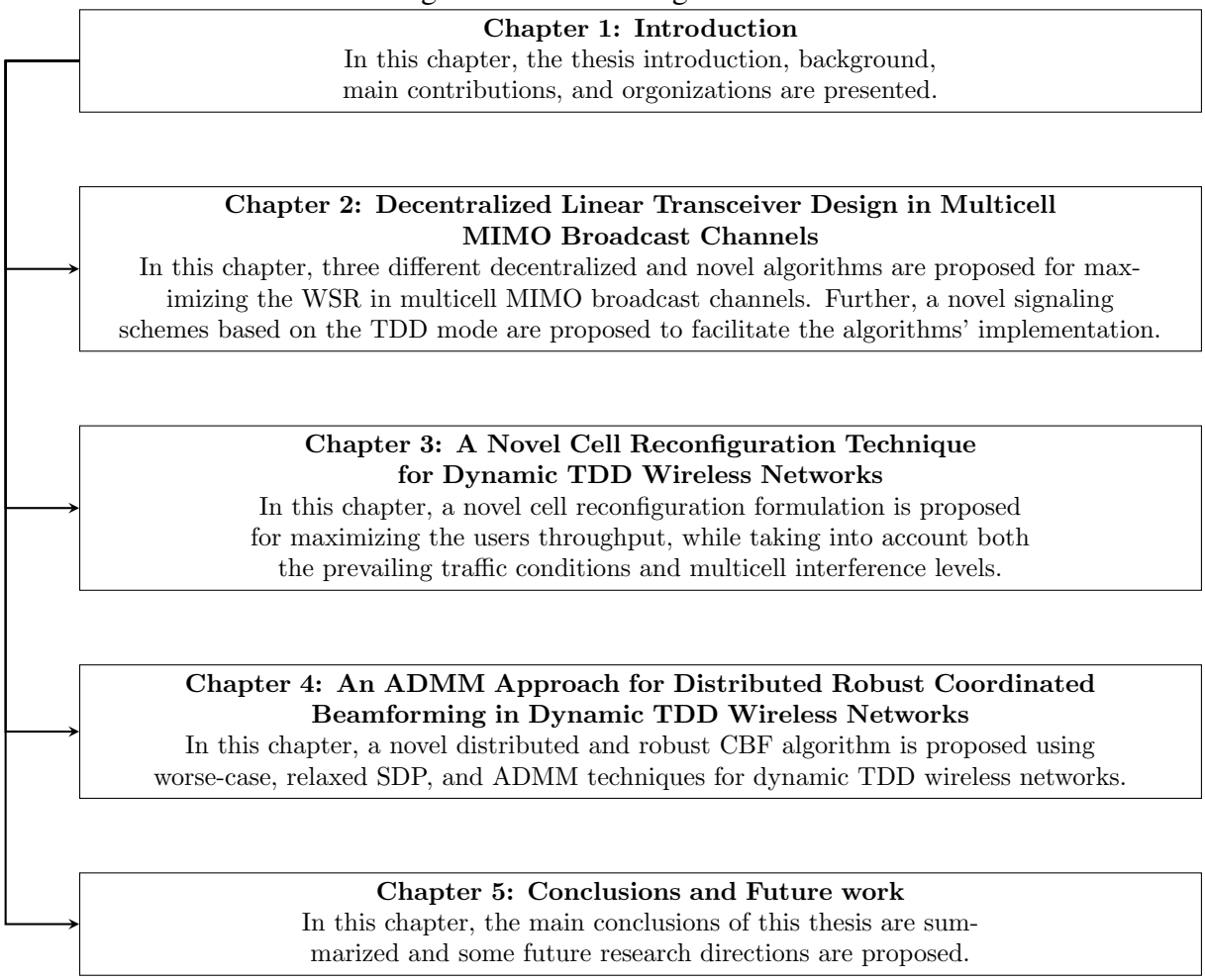
Figure 1.5 – Acquiring CSI: TDD vs. FDD.



Source: Created by author.

transmitting antenna. If we instead need to estimate the channel response from two transmitting antennas, two orthogonal pilot signals are generally required to separate the signals from the two antennas. The orthogonality is achieved by spending two samples on the transmission. The number of orthogonal pilot signals is proportional to the number of transmit antennas, while any number of receive antennas can "listen" to the pilots simultaneously and estimate their individual channels to the transmitters [2]. Note that, every pilot signal that is transmitted could have been a signal that carried payload data. Therefore, one should minimize the pilot signaling to improve the spectral efficiency.

Figure 1.6 – Thesis organization.



Source: Created by author.

1.3 Thesis organization and contributions

The content of this thesis is organized in five chapters, as shown in Fig. 1.6. Each chapter is made self-contained and one can read them in any sequence. The main contributions are summarized as follows:

- Chapter 2 considers a multicell MIMO-BC system model and propose three different decentralized and novel WSR maximization algorithms, which are based on the alternating optimization technique and are guaranteed to converge to a local WSR-optimum. The proposed algorithms in this chapter are summarized as follows:
  - The first algorithm uses an interference pricing approach, where each BS maximizes its own utility that is formed by the local users' WSR minus the priced-ICI leakage. The transmit beamforming matrices are obtained directly by investigating the Karush-Kuhn-Tucker (KKT) conditions of the formulated WSR maximization problem. Through computer simulations, it is shown that the proposed algorithm can achieve better sum rate performance than reference algorithms, while using fewer iterations.
  - The second algorithm designs the transmit beamforming that maximizes the network-wide WSR by generalizing the solution steps of the first algorithm. Interestingly, it is proven that the WSR maximization via interference pricing can be made equivalent to the network-wide WSR maximization whenever the MSs have single-antenna, i.e., in the multicell multiple-input single-output (MISO)-BC networks. However, the interference pricing approach is shown to have some performance loss when the MSs have multiple antenna, as compared to the network-wide approach.
  - The third algorithm is an implicit interference pricing approach, where each BS self-prices its ICI leakage to other cells. Through computer simulations, it is shown that the self-pricing approach has negligible performance loss, as compared to the network-wide approach when the BSs have enough degree of freedom (DoF). In this case, the self-pricing approach is more appealing for practical systems, since it does not require feedback of variables from other cells.

  The proposed algorithms are decentralized in the sense that each BS can solve for its transmit beamforming independently, as soon as it has the required information. Assuming that each BS can acquire the local CSI between itself and all the MSs in the system, as generally assumed in the literature, a novel over-the-air (OTA) signaling scheme is proposed based on the TDD mode to facilitate the algorithms' implementation. In contrast to some existing signaling schemes found in literature,

the proposed signaling scheme reduces the signaling overhead and requires no feedback of variables between BSs.

- Chapter 3 considers a multicell multiuser system model and proposes a novel cell reconfiguration formulation that takes into account both prevailing traffic conditions and multicell BS-to-BS and MS-to-MS interference levels. The proposed optimization problem is then solved optimally using the integer linear programming (ILP) algorithm [31]. However, due to its high computational complexity, a heuristic solution is then proposed based on the particle swarm optimization (PSO) algorithm [32], which is shown to achieve near optimal performance with much lower computational complexity. System level evaluations are carried out using gaming-network-traffic model from [22], from which the effectiveness of the proposed scheme is evidenced in terms of the packet throughput as compared to conventional STDD and other reference schemes, that disregard the DTDD specific inter-cell interference effects. The proposed scheme is general in the sense that it is not traffic model dependent, and therefore can easily be applied to any traffic model and/or deployment scenario. This feature makes the proposed scheme especially suitable for current and future wireless networks.

- Chapter 4 considers a multicell MISO-BC system model and proposes a novel distributed and robust CBF algorithm using relaxed SDP [29] and ADMM [30] techniques for DTDD wireless networks. The design objective is to minimize the total transmit power of downlink BSs, while satisfying the performance targets of downlink and uplink MSs. More precisely, it is assumed that each downlink MS has a predefined minimum SINR target and each uplink MS has a predefined maximum interference threshold. At first, the perfect CSI case is considered, for which a centralized algorithm is proposed to solve the aforementioned optimization problem using relaxed SDP technique [29]. To obtain the beamforming solution in a distributed way, a distributed algorithm is then proposed using relaxed SDP and ADMM techniques. Afterwards, both solutions (centralized and distributed) are extended to account for CSI errors based on worst-case optimization approach [29], where each infinitely nonconvex worst-case constraint is transformed to only one linear matrix inequality (LMI) constraint using the S-Lemma [33]. Using computer simulations, it is shown that the proposed algorithm has a better energy-efficiency than the centralized robust algorithm from [34] and a faster convergence rate than the primal decomposition technique used in [35].

- Chapter 5 draws the main conclusions along with some future research directions.

1.4 Scientific production

The content and contributions presented in this thesis were published with the following information:

1 **K. Ardah**, Y. C. B. Silva, and F. R. P. Cavalcanti, “Block diagonalization for multicell multiuser MIMO systems with other-cell interference,” in Proc. Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT), Sep 2016, **best paper award.**

2 **K. Ardah**, Y. C. B. Silva, and F. R. P. Cavalcanti, “Decentralized Linear Transceiver Design in Multicell MIMO Broadcast Channels,” Journal of Communication and Information Systems (JCIS), Vol.32, No.1, Oct 2017, **invited paper.**

3 **K. Ardah**, Y. C. B. Silva, W. C. Freitas Jr., F. R. P. Cavalcanti, and G. Fodor, “An ADMM Approach to Distributed Coordinated Beamforming in Dynamic TDD Networks,” in Proc. IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Dec, 2017.

4 **K. Ardah**, G. Fodor, Y. C. B. Silva, W. C. Freitas, and F. R. P. Cavalcanti, “A novel cell reconfiguration technique for dynamic TDD wireless networks”, IEEE Wireless Communications Letters, vol. PP, No. 99, pp. 1–1, 2017.

5 **K. Ardah**, G. Fodor, Y. C. B. Silva, W. C. Freitas Jr., "Efficient Signaling Mechanisms for MU MIMO Systems", regular patent filed, No. P71600US1, March, 2018.

2 DECENTRALIZED LINEAR TRANSCEIVER DESIGN IN MULTICELL MIMO BROADCAST CHANNELS

2.1 Introduction

MIMO technology has great potential to eliminate/manage interference, achieve higher throughput, and enhance system capacity [36]. Using multiple antennas, the BSs can transmit to multiple users simultaneously using linear or non-linear transmission techniques [6] to achieve a linear increase of system throughput in the number of BS antennas. In single-cell networks, the non-linear DPC technique [8] is known to achieve the channel capacity. However, it is widely considered that DPC has limited practical applications, due to its high complexity. Therefore, linear transmission techniques (also called beamforming) have gained more interest and were proven to achieve the same sum rate scaling law as DPC [6], while maintaining low complexity.

A notable scheme in this area is called block diagonalization (BD) [37]. In single cell networks, conventional Block Diagonalization (cBD) completely eliminates intra-cell interference by forcing each user to transmit on the null space of the other users. However, in multicell networks, cBD would ignore the ICI, which would affect the users' performance. For that purpose, the authors in [38] have proposed enhanced BD (eBD), which uses a whitening filter to reduce ICI effects. Nevertheless, both cBD and eBD algorithms have high dimensionality restrictions, since they both rely on transmit beamforming to eliminate intra-cell interference and ignore the receive beamforming. Motivated by the last observation, we have proposed iterative BD (iBD) in [39], which eliminates the intra-cell interference by jointly optimizing the transmit and receive beamforming matrices and also accounts for the ICI presence. We have shown that iBD has better sum rate performance than both cBD and eBD, while significantly reducing dimensionality restrictions. However, it is also shown in [39] that all the BD approaches become interference-limited in the presence of high ICI power. The main limitation of BD technique is that each user has an *altruistic* behavior with regard to other users in the same cell (since the intra-cell interference is completely eliminated) and an *egoistic* behavior with regard to users in adjacent cells (since nothing is done to reduce the ICI). Thus, the BD cannot achieve a good balance between users' beamforming behavior, which prevents it from achieving the optimum sum rate [7].

An alternative approach is to jointly design the transmit beamforming of all users in all cells. This approach is named as coordinated multi-point (CoMP) in the literature and can be classified into joint processing (JP) [25] and CBF techniques [24]. In contrast to JP, each user in a CBF system is served by a single BS and thus, the BSs do not need to share the users' data or to be time and phase synchronized. Therefore, CBF has gained a lot of attention and has been extensively studied in the literature with different optimization criteria. For example, in [24, 40, 41] for sum-power minimization, in [42] for SINR balancing, in [43] for sum mean-square

error (MSE) minimization, and in [11, 12, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53] for WSR maximization.

Among all, the WSR maximization problem received more attention. The problem is non-convex and NP-hard [12], for which only local optima can be guaranteed via practical methods. Nevertheless, it has some desirable proprieties such as 1) it can prioritize the users and achieve some fairness among them by adjusting the weights, 2) it has an implicit users and streams selection, since, the number of active streams at convergence is almost always less than or equal to the number of BS antennas, and 3) it is always feasible when only constrained by transmit power. In the literature, the authors in [11] considered the single-cell MIMO BC system model and reformulated the problem into an equivalent problem that incorporates a weighted sum-MSE and establish a weighted sum-MSE duality that is solved iteratively using a geometric program (GP) formulation. For multicell MISO BC system model, the problem was addressed in [45, 47, 48]. In [45], the authors derived the KKT conditions of the problem and then devised an iterative algorithm to solve them, without the need of resorting to convex optimization methods. In [47], an iterative pricing algorithm was proposed based on game-theory, which is guaranteed to converge to an interference equilibrium that corresponds to a KKT point for the original WSR maximization problem. Among all, the global optimum solution is guaranteed only in [48], where the problem was solved using a branch-reduce-and-bound algorithm. On the other hand, the authors in [46] considered the multicell MISO-interference channel (IC) system model and established a relationship between the WSR and the virtual signal-to-interference-plus-noise Ratio (VSINR) by applying the KKT conditions, which led to a distributed and iterative algorithm. Recently, the authors in [12] considered a single-cell MIMO-BC and established a relationship between the WSR maximization problem and the WMMSE minimization problem by applying the KKT conditions. As a result, an iterative algorithm called WSR-WMMSE was proposed, which is based on the alternating optimization technique [54] and solves the quite hard WSR maximization problem indirectly by solving the easier WMMSE minimization problem. This later relationship has inspired many extensions, such as to the multicell MIMO-IC [49] and to the multicell MIMO-BC [50, 51, 52, 53].

2.2 Chapter contributions

In this chapter, a multicell MIMO-BC system model is considered and three different decentralized and novel WSR maximization algorithms are proposed based on the alternating optimization technique [54] and are guaranteed to converge to a local WSR-optimum. The proposed algorithms are summarized as follows.

- The first algorithm uses an interference pricing approach, same as in [47], where each BS maximizes its own utility that is formed by the local users' WSR minus the priced-ICI leakage. In [47], the authors assumed single-antenna users and formulated the problem as a relaxed SDP, whose solution requires each BS to first obtain the

transmit covariance matrices, followed by an operation to guarantee and extract the rank-one transmit beamforming vectors. Different from [47], the proposed algorithm in this chapter consider multi-antenna users and the transmit beamforming matrices are obtained directly by investigating the KKT conditions of the problem. The main ingredient is given by Lemma 1, which makes it possible to solve the transmit beamforming directly from the problem cost function, in contrast to the WSR-WMMSE from [12, 49, 50, 51, 52, 53]. Through computer simulations, it is shown that the proposed algorithm can achieve a comparable sum rate performance to WSR-WMMSE, while using fewer iterations.

- The second algorithm designs the transmit beamforming matrices that maximizes the network-wide WSR by generalizing the solution steps of the first algorithm. Interestingly, it is proven that the WSR maximization via interference pricing can be made equivalent to the network-wide WSR maximization whenever the MSs have single-antenna, i.e., in the multicell MISO BC. However, the interference pricing approach is shown to have some performance loss when the MSs have multiple antennas, as compared to the network-wide approach.

- The third algorithm is an implicit interference pricing approach, where each BS self-prices its ICI leakage to other cells. Through computer simulations, it is shown that the self-pricing approach has negligible performance loss, as compared to the network-wide approach, when the BSs have enough Degrees of Freedom (Dof). In this case, the self-pricing approach is more appealing for practical systems, since it does not require variables feedback between cells.

The proposed algorithms are decentralized in the sense that each BS can solve for its transmit beamforming independently, as soon as it has the required information. Here, we assume that each BS can acquire the local CSI between itself and all the MSs in the system, same as in [12, 49, 50, 51, 52, 53]. An effective technique for obtaining this CSI is the TDD operation, where uplink training in conjunction with reciprocity simultaneously provides the BSs with downlink and uplink channel estimates [13, 55, 56]. In this chapter, a novel OTA signaling scheme based on TDD mode is proposed to facilitate the algorithms' implementation. In contrast to some existing signaling schemes in [51, 52, 53], the proposed signaling scheme reduces the signaling overhead and requires no feedback of variables between BSs.

2.3 Chapter organization

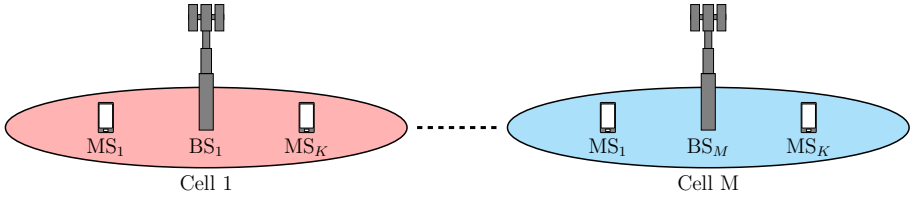
This chapter is organized as follows. Section 2.4 presents the system model. In section 2.5, the proposed block diagonalization (BD) algorithm is presented. The proposed algorithms and the over-the-air (OTA) signaling schemes for WSR maximization are presented

in sections 2.6 and 2.7, respectively. Finally, section 2.8 presents the numerical results and then section 2.9 conclude the chapter.

2.4 System model

Consider a multicell MIMO BC wireless network consisting of M cells, as in Fig.2.1. In each cell, there is one BS equipped with N_t antennas and K MSs, each equipped with N_r antennas. The BS of the n-th cell is denoted as BS_n and the k-th MS in each cell is denoted as MS_k. Let $\mathcal{M} \stackrel{\text{def}}{=} \{1, \ldots, M\}$ and $\mathcal{K} \stackrel{\text{def}}{=} \{\mathcal{K}_1, \ldots, \mathcal{K}_M\}$ denote the sets of all BSs and MSs, respectively, whereas $\mathcal{K}_n$ denotes the set of MSs associated with BS_n. The M BSs are assumed to operate over a common frequency channel and communicate with their K respective MSs using linear transmit beamforming. The scenario under consideration assumes that each MS is served by only one BS.

Figure 2.1 – Multicell MIMO-BC system diagram (M cells and K users per cell).



Source: Created by author.

The received signal at MS_k, $k \in \mathcal{K}_n$, is given as

$$\mathbf{y}_k = \underbrace{\mathbf{H}_{n,k}\mathbf{T}_k\mathbf{s}_k}_{\text{desired signal}} + \underbrace{\sum_{i \in \mathcal{K}_n \setminus k} \mathbf{H}_{n,k}\mathbf{T}_i\mathbf{s}_i}_{\text{intra-cell interference}} + \underbrace{\sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mathbf{H}_{m,k}\mathbf{T}_j\mathbf{s}_j}_{\text{inter-cell interference}} + \underbrace{\mathbf{z}_k}_{\text{noise}}, \tag{2.1}$$

where $\mathbf{H}_{m,k} \in \mathbb{C}^{N_r \times N_t}$ denotes the MIMO channel matrix from BS_m to MS_k, whose coefficients are independent and identically distributed (IID) complex Gaussian random variables, $\mathbf{T}_k \in \mathbb{C}^{N_t \times N_s}$ denotes the transmit beamforming, with N_s being the number of data streams, $\mathbf{s}_k \in \mathbb{C}^{N_s}$ denotes the transmitted data vector that is statistically independent with zero mean and $\mathbb{E}(\mathbf{s}_k\mathbf{s}_k^{\mathrm{H}}) = \mathbf{I}$, $\forall k \in \mathcal{K}$, and $\mathbf{z}_k \in \mathbb{C}^{N_r}$ denotes the IID complex Gaussian noise vector with zero mean and variance σ_k^2. To decode the desired signal, each MS_k multiplies its received signal vector $\mathbf{y}_k$ by the receive beamforming matrix $\mathbf{R}_k \in \mathbb{C}^{N_s \times N_r}$. Thus, the received data vector $\hat{\mathbf{s}}_k \in \mathbb{C}^{N_s}$ at MS_k is given as

$$\hat{\mathbf{s}}_k = \mathbf{R}_k\mathbf{y}_k. \tag{2.2}$$

2.5 Block diagonalization approach

Theoretically, the cBD algorithm from [37] can be interpreted as the equivalent ZF algorithm for the MIMO systems. The main objective is to completely eliminate the intra-cell

interference by forcing each user to transmit on the null space of all other users in the same cell. The cBD optimization problem of BS_n can be written as

$$\mathscr{P}_{\mathrm{cBD}}: \begin{cases} \max\limits_{\mathbf{T}_n} & r_n^{\mathrm{cBD}} = \log\left|\mathbf{I} + \frac{\mathbf{R}_n\mathbf{H}_n\mathbf{T}_n\mathbf{T}_n^{\mathrm{H}}\mathbf{H}_n^{\mathrm{H}}\mathbf{R}_n^{\mathrm{H}}}{\mathbf{R}_n\Upsilon_n\mathbf{R}_n^{\mathrm{H}}}\right|, \\ \text{s.t.} & \sum\limits_{j\in\mathcal{K}_n\setminus k}\mathbf{R}_k\mathbf{H}_{n,k}\mathbf{T}_j = 0,\ \forall k\in\mathcal{K}_n, \\ & \mathrm{Tr}[\mathbf{T}_n\mathbf{T}_n^{\mathrm{H}}] = p_n, \end{cases} \tag{2.3}$$

where p_n is the transmit power threshold, r_n^{cBD} is the BS_n achievable rate, and matrices $\mathbf{H}_n$, $\mathbf{T}_n$, $\mathbf{R}_n$ and Υ_n are defined as

$$\begin{cases} \mathbf{H}_n &= \left[\mathbf{H}_{n,1}^{\mathrm{T}}, \ldots, \mathbf{H}_{n,K}^{\mathrm{T}}\right]^{\mathrm{T}}, \\ \mathbf{T}_n &= [\mathbf{T}_1, \ldots, \mathbf{T}_K], \\ \mathbf{R}_n &= \mathrm{bdiag}\{\mathbf{R}_1, \ldots, \mathbf{R}_K\}, \\ \Upsilon_n &= \mathrm{bdiag}\{\Upsilon_1, \ldots, \Upsilon_K\}, \end{cases} \tag{2.4}$$

where Υ_k denotes the ICI plus noise covariance matrix of MS_k, $k\in\mathcal{K}_n$, which is given as

$$\Upsilon_k = \sum_{m\in\mathcal{M}\setminus n}\sum_{j\in\mathcal{K}_m}\mathbf{H}_{m,k}\mathbf{T}_j\mathbf{T}_j^{\mathrm{H}}\mathbf{H}_{m,k}^{\mathrm{H}} + \sigma_k^2\mathbf{I}_{N_r}. \tag{2.5}$$

As one can notice, problem $\mathscr{P}_{\mathrm{cBD}}$ does nothing to deal with the ICI that is being received from the other cells or is leaking to the other cells, since its main objective is to maximize each cell's achievable rate such that all intra-cell interference is eliminated. The main advantage, though, is that $\mathscr{P}_{\mathrm{cBD}}$ is completely distributed between M cells and has a closed-form solution as follows. The transmit beamforming matrix of MS_k, $k\in\mathcal{K}_n$ is given as

$$\mathbf{T}_k = \mathbf{G}_k\mathbf{F}_k\mathbf{P}_k^{\frac{1}{2}}, \tag{2.6}$$

where $\mathbf{P}_k$ holds on its diagonal the power allocation, $\mathbf{G}_k$ holds the orthogonal basis vectors of the null space of the *intra-cell users' channels* (i.e., all cell n users' channels except user k channel), and $\mathbf{F}_k$ holds the right singular vectors of the *effective channel* of MS_k. Let the *intra-cell users' channels* of MS_k, $k\in\mathcal{K}_n$, be given as

$$\mathbf{H}_k^{(-k)} = [\mathbf{H}_{n,j}^{\mathrm{T}},\ \forall j\in\mathcal{K}_n\setminus k]^{\mathrm{T}}. \tag{2.7}$$

Then, to calculate $\mathbf{G}_k$, let the singular value decomposition (SVD) of $\mathbf{H}_k^{(-k)}$ be given as

$$\mathbf{H}_k^{(-k)} = \mathbf{U}_k^{(-k)}\Sigma_k^{(-k)}[\mathbf{V}_k^{(-k)} \quad \mathbf{G}_k], \tag{2.8}$$

where $\mathbf{G}_k$ is the last $(N_t - l_k^{(-k)})$ right singular vectors, in which $l_k^{(-k)}$ denotes the rank of $\mathbf{H}_k^{(-k)}$. Further, let the *effective channel* of MS_k be given as

$$\mathbf{H}_k^e = \mathbf{H}_{n,k}\mathbf{G}_k. \tag{2.9}$$

Then, to calculate $\mathbf{F}_k$, let the SVD of $\mathbf{H}_k^e$ be given as

$$\mathbf{H}_k^e = \mathbf{U}_k^e \begin{bmatrix} \Sigma_k^e & 0 \\ 0 & 0 \end{bmatrix} [\mathbf{V}_k^{e(1)} \quad \mathbf{V}_k^{e(0)}], \tag{2.10}$$

where Σ_k^e is an $[l_k^e \times l_k^e]$ diagonal matrix, $\mathbf{V}_k^{e(1)}$ contains the first l_k^e singular vectors, in which l_k^e denotes the rank of $\mathbf{H}_k^e$. Therefore, assuming the values of Σ_k^e are in a decreasing order, we choose $\mathbf{F}_k$ and $\mathbf{R}_k$ to be the first N_s vectors of $\mathbf{V}_k^{e(1)}$ and $\mathbf{U}_k^e$, respectively, i.e.,

$$\mathbf{F}_k = \mathbf{V}_{k\,[1:N_s]}^{e(1)} \text{ and } \mathbf{R}_k = \mathbf{U}_{k[1:N_s]}^{e\,\mathrm{H}}. \tag{2.11}$$

With the transmit and receive beamforming matrices calculated as above, the BS_n rate function (optimization problem) is reduced to

$$r_n^{\mathrm{cBD}} = \max_{\mathbf{P}_n} \log \left| \mathbf{I} + \frac{\Sigma_n^2 \mathbf{P}_n}{\mathbf{R}_n \Upsilon_n \mathbf{R}_n^{\mathrm{H}}} \right|, \tag{2.12}$$

where $\Sigma_n = \mathrm{bdiag}(\Sigma_{n1}^e, \ldots, \Sigma_{nK}^e)$ and $\mathbf{P}_n$ is a diagonal matrix that holds the optimal power loading found using water-filling method [57] on the Σ_n diagonal elements. Note that the water-filling method is applied individually on each sub-matrix of Σ_n assuming that total transmit power of the BS is divided equally between its K users. Clearly, the equal power allocation is a suboptimal solution. However, this is done to make sure that all K users are allocated for transmission.

From above, it can be seen that one of the main issue with cBD is that it does nothing to reduce the effects of the ICI each user is receiving. This issue has been considered in [38], where the authors proposed the eBD algorithm to account for the ICI presence. The eBD algorithm is summarized as follows. First, to suppress the ICI effects, MS_k uses the whitening matrix $\mathbf{W}_k = \Upsilon_k^{-\frac{1}{2}}$ at the received signal. Let $\mathbf{W}_n = \mathrm{bdiag}[\mathbf{W}_{n1}, \ldots, \mathbf{W}_{nK}]$. Then, the BS_n rate function r_n^{cBD} can be written as

$$\begin{aligned} r_n^{\mathrm{eBD}} &= \max_{\widehat{\mathbf{T}}_n} \log \left| \mathbf{I} + \frac{\widehat{\mathbf{R}}_n \mathbf{W}_n \mathbf{H}_n \widehat{\mathbf{T}}_n \widehat{\mathbf{T}}_n^{\mathrm{H}} \mathbf{H}_n^{\mathrm{H}} \mathbf{W}_n^{\mathrm{H}} \widehat{\mathbf{R}}_n^{\mathrm{H}}}{\widehat{\mathbf{R}}_n \mathbf{W}_n \Upsilon_n \mathbf{W}_n^{\mathrm{H}} \widehat{\mathbf{R}}_n^{\mathrm{H}}} \right| \\ &= \max_{\widehat{\mathbf{T}}_n} \log \left| \mathbf{I} + \widehat{\mathbf{R}}_n \widehat{\mathbf{H}}_n \widehat{\mathbf{T}}_n \widehat{\mathbf{T}}_n^{\mathrm{H}} \widehat{\mathbf{H}}_n^{\mathrm{H}} \widehat{\mathbf{R}}_n^{\mathrm{H}} \right|, \end{aligned} \tag{2.13}$$

where $\widehat{\mathbf{H}}_n = \mathbf{W}_n \mathbf{H}_n$ (thus, $\widehat{\mathbf{H}}_k = \mathbf{W}_k \mathbf{H}_{n,k}$). As in the case of cBD, the transmit beamforming matrix of MS_k is given as $\widehat{\mathbf{T}}_k = \widehat{\mathbf{G}}_k \widehat{\mathbf{F}}_k \widehat{\mathbf{P}}_k^{\frac{1}{2}}$, where $\widehat{\mathbf{G}}_k$ is calculated similar to (2.8) from

$$\widehat{\mathbf{H}}_k^{(-k)} = [\widehat{\mathbf{H}}_j^{\mathrm{T}}, \forall j \in \mathcal{K}_n \backslash k]^{\mathrm{T}} = [(\mathbf{W}_j \mathbf{H}_{n,j})^{\mathrm{T}}, \forall j \in \mathcal{K}_n \backslash k]^{\mathrm{T}}. \tag{2.14}$$

The $\widehat{\mathbf{F}}_k$ and $\widehat{\mathbf{R}}_k$ matrices are calculated similar to (2.11) from the MS_k effective channel

$$\widehat{\mathbf{H}}_k^e = \widehat{\mathbf{H}}_k \widehat{\mathbf{G}}_k = \mathbf{W}_k \mathbf{H}_{n,k} \widehat{\mathbf{G}}_k. \tag{2.15}$$

Consequently, the BS_n rate function r_n^{eBD} given by (2.13) is reduced to

$$r_n^{\mathrm{eBD}} = \max_{\widehat{\mathbf{P}}_n} \log \left| \mathbf{I} + \widehat{\Sigma}_n^2 \widehat{\mathbf{P}}_n \right|. \tag{2.16}$$

From above, one can see that both BD approaches, cBD and eBD, have the same dimensionality restrictions. The expressions given by (2.7) and (2.14) have dimension of $[(K-1)N_r \times N_t]$ with rank of $[N_t - (K-1)N_r]$. Therefore, to have N_s columns in the null space, $[N_t - (K-1)N_r]$ should be larger than N_s, i.e., $[N_t - (K-1)N_r] \geq N_s$. Moreover, it is important to note that both approaches use only the transmit beamforming to eliminate the intra-cell interference, i.e., the receive beamforming is not utilized. Motivated by the last observation, one possible way to reduce the dimensionality restrictions is to utilize the receive beamforming matrix when calculating the transmit beamforming matrix [39]. To achieve this end, the receive beamforming matrix $\widetilde{\mathbf{R}}_k$ can be included in (2.14), then we have

$$\widetilde{\mathbf{H}}_k^{(-k)} = [(\widetilde{\mathbf{R}}_j \widehat{\mathbf{H}}_j)^{\mathrm{T}}, \forall j \in \mathcal{K}_n \backslash k]^{\mathrm{T}} = [(\widetilde{\mathbf{R}}_j \mathbf{W}_j \mathbf{H}_{n,j})^{\mathrm{T}}, \forall j \in \mathcal{K}_n \backslash k]^{\mathrm{T}}. \tag{2.17}$$

Note that $\widetilde{\mathbf{H}}_k^{(-k)}$ has dimension of $[(K-1)N_s \times N_t]$, which is no longer in function of N_r. Calculating the null space from $\widetilde{\mathbf{H}}_k^{(-k)}$ is always satisfied if, and only if, the number of data streams transmitted by a BS is less than or equal to its number of transmit antennas, i.e., the condition of $[N_t - (K-1)N_s] \geq N_s$ should be satisfied. The following steps are much similar to the ones above. The transmit beamforming is given as $\widetilde{\mathbf{T}}_k = \widetilde{\mathbf{G}}_k \widetilde{\mathbf{F}}_k \widetilde{\mathbf{P}}_k^{\frac{1}{2}}$, where $\widetilde{\mathbf{G}}_k$ is calculated similarly from $\widetilde{\mathbf{H}}_k^{(-k)}$. The $\widetilde{\mathbf{F}}_k$ and $\widetilde{\mathbf{R}}_k$ matrices are calculated from the MS_k effective channel

$$\widetilde{\mathbf{H}}_k^e = \widehat{\mathbf{H}}_k \widetilde{\mathbf{G}}_k = \mathbf{W}_k \mathbf{H}_{n,k} \widetilde{\mathbf{G}}_k. \tag{2.18}$$

Since the transmit and receive beamforming matrices are now coupled, the BS is required to conduct some iterations in order to achieve BD. Therefore, we refer to this approach as iterative BD (iBD) and summarize it in Algorithm 1.

Algorithm 1: Proposed iBD algorithm.

1: Initialize $\widetilde{\mathbf{R}}_k^{(1)}$, $\widetilde{\mathbf{T}}_k^{(1)}$, $\forall k \in \mathcal{K}$ and set $t = 1$.
2: BS_n, $\forall n$: Transmit data using $\widetilde{\mathbf{T}}_k^{(t)}$, $\forall k \in \mathcal{K}_n$.
3: MS_k, $\forall k$: Calculate $\Upsilon_k^{(t)}$ and feed it back to BS_n.
4: BS_n, $\forall n$: Calculate $\widetilde{\mathbf{R}}_k^{(t)}$ and $\widetilde{\mathbf{T}}_k^{(t)}$, $\forall k \in \mathcal{K}_n$ as:
5: - Construct $\widetilde{\mathbf{H}}_k^{(-k)(t)}$ using $\widetilde{\mathbf{R}}_k^{(t)}$.
6: - Calculate $\widetilde{\mathbf{G}}_k^{(t)}$ from $\widetilde{\mathbf{H}}_k^{(-k)(t)}$.
7: - Construct $\widetilde{\mathbf{H}}_k^{e(t)}$ using $\widetilde{\mathbf{G}}_k^{(t)}$.
8: - Calculate $\widetilde{\mathbf{F}}_k^{(t+1)}$ and $\widetilde{\mathbf{R}}_k^{(t+1)}$ from $\widetilde{\mathbf{H}}_k^{e(t)}$.
9: Repeat steps 2-4 (until convergence)

At the first step, Algorithm 1 initializes the transmit and receive beamforming matrices for all users. For instance, $\widetilde{\mathbf{T}}_k^{(1)}$ can be initialized using the MRT approach and $\widetilde{\mathbf{R}}_k^{(1)} = \mathbf{I}$. At the t-th iteration, each BS transmits pilot signals precoded with $\widetilde{\mathbf{T}}_k^{(t)}$ at step-2 so that each MS_k

can calculate the ICI covariance matrix, i.e., $\Upsilon_k^{(t)}$, and feed it back to its serving BS via feedback channels. After that, each BS updates the transmit and receive beamforming of its users at step-4. The aforementioned steps are repeated until convergence. Note that the transmit and receive beamforming matrices of all users are calculated at the BSs. Therefore, at convergence, each BS would forward the receive beamforming matrices to its users using the feedforward channels. With the transmit and receive beamforming matrices calculated as given by Algorithm 1, the BS_n rate function r_n^{eBD} given by (2.13) is reduced to

$$r_n^{\mathrm{iBD}} = \max_{\widetilde{\mathbf{P}}_n} \log \left| \mathbf{I} + \widetilde{\Sigma}_n^2 \widetilde{\mathbf{P}}_n \right|. \tag{2.19}$$

Note that both equations (2.15) and (2.18) have the same structure. The following theorem indicates their relation.

Theorem 1 *If the number of data streams transmitted to any user is equal to the number of its receive antennas, i.e., if $N_s = N_r$, then, both eBD and iBD are equivalent and have the same exact performance.*

Proof 1 *At first, one can note that $\widetilde{\mathbf{H}}_k^{-k}$ given by (2.17) can be written in function of $\widehat{\mathbf{H}}_k^{-k}$ given by (2.14). To show this, let us define $\widetilde{\mathbf{R}}_n^{-k} = \mathrm{bdiag}[\widetilde{\mathbf{R}}_j, \forall j \in \mathcal{K}_n \backslash k]$. Then we can write $\widetilde{\mathbf{H}}_k^{-k} = \widetilde{\mathbf{R}}_n^{-k} \widehat{\mathbf{H}}_k^{-k}$. Note that $\widetilde{\mathbf{R}}_n^{-k}$ is an orthogonal unitary matrix. Therefore, if $N_s = N_r$, then we have $\widetilde{\mathbf{R}}_n^{H-k} \widetilde{\mathbf{R}}_n^{-k} = \mathbf{I}$, otherwise, if $N_s < N_r$, then $\widetilde{\mathbf{R}}_n^{H-k} \widetilde{\mathbf{R}}_n^{-k} \neq \mathbf{I}$. Assuming $N_s = N_r$, we have*

$$\widetilde{\mathbf{H}}_k^{H-k} \widetilde{\mathbf{H}}_k^{-k} = \widehat{\mathbf{H}}_k^{H-k} \widetilde{\mathbf{R}}_n^{H-k} \widetilde{\mathbf{R}}_n^{-k} \widehat{\mathbf{H}}_k^{-k} = \widehat{\mathbf{H}}_k^{H-k} \widehat{\mathbf{H}}_k^{-k}. \tag{2.20}$$

Therefore, we have $\widetilde{\mathbf{H}}_k^{-i} \propto \widehat{\mathbf{H}}_k^{-i}$. This end result proves that both matrices are proportional to each other. Consequently, their individual null spaces are also proportional to each other, i.e., $\widetilde{\mathbf{G}}_k \propto \widehat{\mathbf{G}}_k$. Therefore, the singular values calculated using (2.10) assuming $\widetilde{\mathbf{H}}_k^e$ given by (2.18) are exactly equal to the ones calculated assuming $\widehat{\mathbf{H}}_k^e$ given by (2.15), which completes the proof.

It's worth noting that if the system has only one cell, then the eBD algorithm is equivalent to cBD. In this case, the iBD algorithm is also equivalent to cBD, only if $N_s = N_r$, which is a straightforward result of Theorem 1.

2.6 Weighted sum rate maximization approach

This section uses the WSR as the transceiver design criterion. It is assumed that each MS employs single-user detection by treating the interference as additive noise. Therefore, the achievable rate of MS_k, $k \in \mathcal{K}_n$, can be written as

$$r_k = \log \left| \mathbf{I}_{N_s} + \mathbf{T}_k^{\mathrm{H}} \mathbf{H}_{n,k}^{\mathrm{H}} \Phi_k^{-1} \mathbf{H}_{n,k} \mathbf{T}_k \right|, \tag{2.21}$$

where Φ_k denotes the received interference plus-noise covariance matrix for MS_k, $k \in \mathcal{K}_n$, which is given as

$$\Phi_k = \sum_{i \in \mathcal{K}_n \backslash k} \mathbf{H}_{n,k} \mathbf{T}_i \mathbf{T}_i^{\mathrm{H}} \mathbf{H}_{n,k}^{\mathrm{H}} + \Upsilon_k, \tag{2.22}$$

whereas Υ_k is given by (2.5), which denotes the ICI plus noise covariance matrix of MS_k. Here, we assume that each MS_k uses MMSE receive beamforming, which is given as [12]

$$\begin{aligned} \mathbf{R}_k^{\mathrm{MMSE}} &= \arg\min_{\mathbf{R}_k} \mathbb{E}[\|\mathbf{R}_k \mathbf{y}_k - \mathbf{s}_k\|^2] \\ &= \mathbf{T}_k^{\mathrm{H}} \mathbf{H}_{n,k}^{\mathrm{H}} \Omega_k^{-1}, \end{aligned} \tag{2.23}$$

where $\Omega_k = \mathbf{H}_{n,k} \mathbf{T}_k \mathbf{T}_k^{\mathrm{H}} \mathbf{H}_{n,k}^{\mathrm{H}} + \Phi_k$. Using (2.23), the MSE-matrix of MS_k is given as [51]

$$\begin{aligned} \mathbf{E}_k &= \mathbb{E}[\|\mathbf{R}_k^{\mathrm{MMSE}} \mathbf{y}_k - \mathbf{s}_k\|^2] \\ &= \mathbf{I}_{N_s} - \mathbf{R}_k^{\mathrm{MMSE}} \mathbf{H}_{n,k} \mathbf{T}_k, \end{aligned} \tag{2.24}$$

which can be equivalently expressed as [12]

$$\mathbf{E}_k = (\mathbf{I}_{N_s} + \mathbf{T}_k^{\mathrm{H}} \mathbf{H}_{n,k}^{\mathrm{H}} \Phi_k^{-1} \mathbf{H}_{n,k} \mathbf{T}_k)^{-1}. \tag{2.25}$$

The latter form of $\mathbf{E}_k$ in (2.25) shows that the rate function given by (2.21) can be equivalently expressed as

$$r_k = \log \left| \mathbf{E}_k^{-1} \right|. \tag{2.26}$$

Note that $\mathbf{E}_k$ must be Hermitian, since from (2.25), $\mathbf{E}_k$ equals a quantity (right-hand side) that is Hermitian, which means that $\mathbf{E}_k = \mathbf{E}_k^{\mathrm{H}}$. Furthermore, the following lemma is needed throughout the rest of this chapter.

Lemma 1 *Given the MSE-matrix $\mathbf{E}_k$ as in (2.24), or equivalently as in (2.25), the receive beamforming matrix $\mathbf{R}_k$ can be written as*

$$\mathbf{R}_k = \mathbf{E}_k \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1}. \tag{2.27}$$

Proof 2 *In the lemma, we claim that $\mathbf{R}_k = \mathbf{E}_k \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1}$. To prove this, assume $N_s = N_r$ and solve for $\mathbf{R}_k$ from (2.24) as $\mathbf{R}_k = (\mathbf{I} - \mathbf{E}_k)(\mathbf{H}_{n,k} \mathbf{T}_k)^{-1}$. Then, our claim is that*

$$\mathbf{E}_k \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1} \stackrel{\text{def}}{=} (\mathbf{I} - \mathbf{E}_k)(\mathbf{H}_{n,k} \mathbf{T}_k)^{-1} \tag{2.28}$$

$$\mathbf{E}_k \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1} = (\mathbf{H}_{n,k} \mathbf{T}_k)^{-1} - \mathbf{E}_k (\mathbf{H}_k \mathbf{T}_k)^{-1} \tag{2.29}$$

$$\mathbf{T}_k^H \mathbf{H}_k^H \Phi_k^{-1} \stackrel{(a)}{=} \mathbf{E}_k^{-1} [(\mathbf{H}_{n,k} \mathbf{T}_k)^{-1} - \mathbf{E}_k (\mathbf{H}_{n,k} \mathbf{T}_k)^{-1}] \tag{2.30}$$

$$\mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1} \stackrel{(b)}{=} \mathbf{E}_k^{-1} (\mathbf{H}_{n,k} \mathbf{T}_k)^{-1} - (\mathbf{H}_{n,k} \mathbf{T}_k)^{-1}, \tag{2.31}$$

where (a) is obtained by left-multiplying both sides by $\mathbf{E}_k^{-1}$ *and (b) is obtained by simplifying (a). From (2.25),* $\mathbf{E}_k^{-1} = (\mathbf{I}_{N_r} + \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1} \mathbf{H}_{n,k} \mathbf{T}_k)$*. Substitute* $\mathbf{E}_k^{-1}$ *into (b) and when simplifying the resulting expression we have*

$$\mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1} = \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1}, \tag{2.32}$$

which completes the proof.

2.6.1 Per-Cell WSR maximization via interference pricing

This section considers an interference pricing approach for designing the transmit beamforming. The main idea is to manage the ICI received by a user by pricing the interfering BSs. Similar to [47, 58], we define the interference price as the marginal decrease in the user rate due to a marginal increase in the received interference. Mathematically, the MS_k interference price is given as

$$\pi_k = \nabla_{\Phi_k} r_k. \tag{2.33}$$

Using the result of $\nabla \log \left|\mathbf{X}\right| = \text{Tr}(\mathbf{X}^{-1} \nabla \mathbf{X})$, where $\mathbf{X}$ is a matrix [59], then π_k is given as

$$\pi_k = \text{Tr}(\mathbf{E}_k \mathbf{T}_k^{\text{H}} \mathbf{H}_{n,k}^{\text{H}} \Phi_k^{-2} \mathbf{H}_{n,k} \mathbf{T}_k). \tag{2.34}$$

By observing (2.34) and the Lemma 1 result, we have the following corollary.

Corollary 1 *The* MS_k *interference price* π_k *given by (2.34) can be equivalently written as*

$$\pi_k = Tr(\mathbf{R}_k^H \mathbf{E}_k^{-1} \mathbf{R}_k). \tag{2.35}$$

Proof 3 *According to Lemma 1, the receive beamforming can be written as* $\mathbf{R}_k = \mathbf{E}_k \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-1}$*. Then, the* MS_k *interference price* π_k *given by (2.34) is reduced to*

$$\begin{aligned} \pi_k &= Tr(\mathbf{E}_k \mathbf{T}_k^H \mathbf{H}_{n,k}^H \Phi_k^{-2} \mathbf{H}_{n,k} \mathbf{T}_k) \\ &\overset{(a)}{=} Tr(\mathbf{R}_k \mathbf{R}_k^H \mathbf{E}_k^{-1}) \\ &\overset{(b)}{=} Tr(\mathbf{R}_k^H \mathbf{E}_k^{-1} \mathbf{R}_k) \end{aligned}$$

where (a) is obtained by substituting $\mathbf{R}_k$ *into the first equality and (b) is obtained by using the results of* $Tr(\mathbf{XYZ}) = Tr(\mathbf{YZX}) = Tr(\mathbf{ZXY})$ *[59], which completes the proof.*

Let $\boldsymbol{\pi}_n = \{\pi_j, \forall j \in \mathcal{K}_m, \forall m \in \mathcal{M} \backslash n\}$ denote the vector that collects all interference prices of all users in the system except BS_n users. Then, define the following MS-specific function

$$f_k(\boldsymbol{\pi}_n) = \mu_k r_k - \text{Tr}(\mathbf{L}_k), \tag{2.36}$$

where $\mu_k > 0$ denotes the weight associated to MS_k and $\mathbf{L}_k$ defines the priced-ICI caused by the MS_k beamforming $\mathbf{T}_k$, which is given as

$$\mathbf{L}_k = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \pi_j \mathbf{H}_{n,j} \mathbf{T}_k \mathbf{T}_k^{\mathrm{H}} \mathbf{H}_{n,j}^{\mathrm{H}}. \tag{2.37}$$

Afterwards, each BS_n, $\forall n \in \mathcal{M}$, updates its transmit beamforming $\mathbf{T}_k$, $\forall k \in \mathcal{K}_n$, as the solution to the following interference-priced WSR maximization problem

$$\mathscr{P}_{\mathrm{WSRP}} = \begin{cases} \max\limits_{\substack{\mathbf{T}_k, \\ \forall k \in \mathcal{K}_n}} & \sum\limits_{k \in \mathcal{K}_n} f_k(\boldsymbol{\pi}_n) \\ \text{s.t.} & \sum\limits_{k \in \mathcal{K}_n} \mathrm{Tr}(\mathbf{T}_k \mathbf{T}_k^{\mathrm{H}}) = p_n, \end{cases} \tag{2.38}$$

where we have used an equality power constraint rather than the often used $\sum_{k \in \mathcal{K}_n} \mathrm{Tr}(\mathbf{T}_k \mathbf{T}_k^{\mathrm{H}}) \leq p_n$, since the WSR optimum is reached at maximum transmit power [12]. From problem $\mathscr{P}_{\mathrm{WSRP}}$, one can see that this approach is different from the BD approach in the sense that the ICI received by a user is being managed by the interfering BSs and not the serving BS.

In [47], the authors addressed problem $\mathscr{P}_{\mathrm{WSRP}}$ from a game-theoretic view-point assuming single-antenna users, where the objective function $f_k(\boldsymbol{\pi}_n)$ is interpreted as a user utility function that penalizes the user rate by the ICI that he is leaking. The problem in [47], however, was formulated as a relaxed SDP and its solution would require each BS to obtain first the transmit covariance matrices, i.e., $\mathbf{Q}_k \overset{\mathrm{def}}{=} \mathbf{T}_k \mathbf{T}_k^{\mathrm{H}} \succeq 0$, followed by operations to guarantee and extract the rank-one transmit beamforming vectors, i.e., $\mathbf{T}_k$. It was proven in [47] that problem $\mathscr{P}_{\mathrm{WSRP}}$ is guaranteed to converge to an equilibrium point that corresponds to a KKT point for the original problem $\mathscr{P}_{\mathrm{WSRP}}$. In the following, we present a different solution to problem $\mathscr{P}_{\mathrm{WSRP}}$. The solution is obtained by investigating the KKT conditions of problem $\mathscr{P}_{\mathrm{WSRP}}$ and with the help of the Lemma 1 result. The solution of $\mathscr{P}_{\mathrm{WSRP}}$ w.r.t. transmit beamforming for MS_k, $\forall k \in \mathcal{K}$, is given by Proposition 1.

Proposition 1 *Let the receive beamforming $\mathbf{R}_k$ and MSE matrix $\mathbf{E}_k$ for MS_k be given by (2.23) and (2.24), respectively, and by utilizing the Lemma 1 result, then the solution of problem $\mathscr{P}_{WSRP}$ w.r.t. transmit beamforming $\mathbf{T}_k$ for MS_k, $\forall k \in \mathcal{K}_n$, is given as*

$$\mathbf{T}_k^{WSRP} = (\mathbf{A}_k + \mathbf{B}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{n,k}^H \mathbf{R}_k^H \mu_k, \tag{2.39}$$

where λ_n, $\forall n \in \mathcal{M}$, are the Lagrange multipliers associated with the $\mathscr{P}_{WSRP}$ constraint functions, $\mathbf{A}_k$ and $\mathbf{B}_k$ are given as

$$\mathbf{A}_k = \sum_{i \in \mathcal{K}_n \setminus k} \mu_i \mathbf{H}_{n,i}^H \mathbf{R}_i^H \mathbf{E}_i^{-1} \mathbf{R}_k \mathbf{H}_{n,i}, \tag{2.40}$$

$$\mathbf{B}_n = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \pi_j \mathbf{H}_{n,j}^H \mathbf{H}_{n,j}. \tag{2.41}$$

Proof 4 *From the KKT conditions, a local optimum must satisfy* $\nabla_{\mathbf{T}_k}\mathscr{L} = 0, \forall k \in \mathcal{K}_n$, *where* $\nabla_{\mathbf{T}_k}\mathscr{L}$ *defines the complex gradient operator of* $\mathscr{L}$ *with respect to* $\mathbf{T}_k$ *and* $\mathscr{L}$ *defines the Lagrangian function of problem* $\mathscr{P}_{WSRP}$, *which is given as*

$$\begin{aligned}\mathscr{L} &= \sum_{k\in\mathcal{K}_n} f_k(\boldsymbol{\pi}_n) - \lambda_n\Big(\sum_{k\in\mathcal{K}_n} Tr(\mathbf{T}_k\mathbf{T}_k^H) - p_n\Big)\\ &= \sum_{k\in\mathcal{K}_n}(\mu_k r_k - \mathbf{L}_k) - \lambda_n\Big(\sum_{k\in\mathcal{K}_n} Tr(\mathbf{T}_k\mathbf{T}_k^H) - p_n\Big).\end{aligned} \tag{2.42}$$

The gradient is a matrix with the $[p,q]$*-th element defined as* $[\nabla_{\mathbf{T}_k}\mathscr{L}]_{[p,q]} = \nabla_{[\mathbf{T}_k]_{[p,q]}}\mathscr{L}$. *In order to calculate* $\nabla_{\mathbf{T}_k}\mathscr{L}$, *we need to calculate first* $\nabla_{\mathbf{T}_k} r_k$, $\nabla_{\mathbf{T}_k} r_j, \forall j \in \mathcal{K}_n\backslash k$, *and* $\nabla_{\mathbf{T}_k}\mathbf{L}_k$ *by utilizing the following results from [59]:* $\nabla \log|\mathbf{X}| = Tr(\mathbf{X}^{-1}\nabla\mathbf{X})$ *and* $\nabla(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}(\nabla\mathbf{X})\mathbf{X}^{-1}$, *where* $\mathbf{X}$ *is a matrix.*

First, $\nabla_{\mathbf{T}_{k[p,q]}} r_k = Tr(\mathbf{E}_k\nabla_{\mathbf{T}_{k[p,q]}}\mathbf{E}_k^{-1})$. *Here,* $\nabla_{\mathbf{T}_{k[p,q]}}\mathbf{E}_k^{-1} = \mathbf{e}_q\mathbf{e}_p^H\mathbf{H}_{n,k}^H\Phi_k^{-1}\mathbf{H}_{n,k}\mathbf{T}_k$, *where* $\mathbf{e}_p$ ($\mathbf{e}_q$) *is a vector of* N_t (N_r) *dimension with one at the* p (q) *-th element and zeros elsewhere. Then, we have*

$$\begin{aligned}\nabla_{\mathbf{T}_{k[p,q]}} r_k &= Tr(\mathbf{E}_k\mathbf{e}_q\mathbf{e}_p^H\mathbf{H}_{n,k}^H\Phi_k^{-1}\mathbf{H}_{n,k}\mathbf{T}_k)\\ &= \mathbf{e}_p^H\mathbf{H}_{n,k}^H\Phi_k^{-1}\mathbf{H}_{n,k}\mathbf{T}_k\mathbf{E}_k\mathbf{e}_q.\end{aligned} \tag{2.43}$$

Since $\nabla_{\mathbf{T}_{k[p,q]}} r_k = [\nabla_{\mathbf{T}_k} r_k]_{[p,q]}$, *then we have* $\nabla_{\mathbf{T}_k} r_k = \mathbf{H}_{n,k}^H\Phi_k^{-1}\mathbf{H}_{n,k}\mathbf{T}_k\mathbf{E}_k$. *Furthermore,* $\nabla_{\mathbf{T}_{k[p,q]}} r_j, \forall j \in \mathcal{K}_n\backslash k$, *is given as* $\nabla_{\mathbf{T}_{k[p,q]}} r_j = Tr(\mathbf{E}_j\nabla_{\mathbf{T}_{k[p,q]}}\mathbf{E}_j^{-1})$. *First,* $\nabla_{\mathbf{T}_{k[p,q]}}\mathbf{E}_j^{-1}$ *is given as* $\nabla_{\mathbf{T}_{k[p,q]}}\mathbf{E}_j^{-1} = \mathbf{T}_j^H\mathbf{H}_{n,j}^H[-\Phi_j^{-1}[\nabla_{\mathbf{T}_{k[p,q]}}\Phi_j]\Phi_j^{-1}]\mathbf{H}_{n,j}\mathbf{T}_j$, *whereas* $\nabla_{\mathbf{T}_{k[p,q]}}\Phi_j = \mathbf{H}_{n,j}\mathbf{T}_k\mathbf{e}_q\mathbf{e}_p^H\mathbf{H}_{n,j}^H$. *Combining all results together, we have*

$$\begin{aligned}\nabla_{\mathbf{T}_{k[p,q]}} r_j &= Tr(\mathbf{E}_j\mathbf{T}_j^H\mathbf{H}_{n,j}^H[-\Phi_j^{-1}\mathbf{H}_{n,j}\mathbf{T}_k\mathbf{e}_q\mathbf{e}_p^H\mathbf{H}_{n,j}^H\Phi_j^{-1}]\mathbf{H}_{n,j}\mathbf{T}_j)\\ &= -\mathbf{e}_p^H\mathbf{H}_{n,j}^H\Phi_j^{-1}\mathbf{H}_{n,j}\mathbf{T}_j\mathbf{E}_j\mathbf{T}_j^H\mathbf{H}_{n,j}^H\Phi_j^{-1}\mathbf{H}_{n,j}\mathbf{T}_k\mathbf{e}_q.\end{aligned} \tag{2.44}$$

Therefore, $\nabla_{\mathbf{T}_k} r_j, \forall j \in \mathcal{K}_n\backslash k$, *is given as*

$$\nabla_{\mathbf{T}_k} r_j = -\mathbf{H}_{n,j}^H\Phi_j^{-1}\mathbf{H}_{n,j}\mathbf{T}_j\mathbf{E}_j\mathbf{T}_j^H\mathbf{H}_{n,j}^H\Phi_j^{-1}\mathbf{H}_{n,j}\mathbf{T}_k. \tag{2.45}$$

Finally, $\nabla_{\mathbf{T}_k}\mathbf{L}_k$ *is given as* $\nabla_{\mathbf{T}_k}\mathbf{L}_k = \sum_{m\in\mathcal{M}\backslash n}\sum_{j\in\mathcal{K}_m}\mu_j\mathbf{H}_{n,j}^H\mathbf{H}_{n,j}\mathbf{T}_k$. *From above,* $\nabla_{\mathbf{T}_k}\mathscr{L}$ *is given as*

$$\nabla_{\mathbf{T}_k}\mathscr{L} = \mu_k\mathbf{H}_{n,k}^H\Phi_k^{-1}\mathbf{H}_{n,k}\mathbf{T}_k\mathbf{E}_k - \tilde{\mathbf{A}}_k\mathbf{T}_k - \mathbf{B}_n\mathbf{T}_k - \lambda_n\mathbf{T}_k, \tag{2.46}$$

where

$$\tilde{\mathbf{A}}_k = \sum_{i\in\mathcal{K}_n\backslash k}\mu_i\mathbf{H}_{n,i}^H\Phi_i^{-1}\mathbf{H}_{n,i}\mathbf{T}_i\mathbf{E}_i\mathbf{T}_i^H\mathbf{H}_{n,i}^H\Phi_i^{-1}\mathbf{H}_{n,i}, \tag{2.47}$$

$$\mathbf{B}_n = \sum_{m\in\mathcal{M}\backslash n}\sum_{j\in\mathcal{K}_m}\mu_j\mathbf{H}_{n,j}^H\mathbf{H}_{n,j}. \tag{2.48}$$

From (2.46), it can be seen that it is not possible to solve for $\mathbf{T}_k$ *directly. However, according to Lemma 1, we can write* $\mathbf{R}_k^H = \Phi_k^{-1}\mathbf{H}_{n,k}\mathbf{T}_k\mathbf{E}_k$ *(note that* $\mathbf{E}_k = \mathbf{E}_k^H$*). Then, the gradient function (2.46) can be written as*

$$\nabla_{\mathbf{T}_k}\mathscr{L} = \mu_k\mathbf{H}_{n,k}^H\mathbf{R}_k^H - \mathbf{A}_k\mathbf{T}_k - \mathbf{B}_n\mathbf{T}_k - \lambda_n\mathbf{T}_k, \tag{2.49}$$

where

$$\mathbf{A}_k = \sum_{i \in \mathcal{K}_n \setminus k} \mu_i \mathbf{H}_{n,i}^H \mathbf{R}_i^H \mathbf{E}_i^{-1} \mathbf{R}_i \mathbf{H}_{n,i}. \tag{2.50}$$

From (2.49), we can solve for $\mathbf{T}_k$ *directly as*

$$\mathbf{T}_k = (\mathbf{A}_k + \mathbf{B}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{n,k}^H \mathbf{R}_k^H \mu_k. \tag{2.51}$$

Thus, we have the result given in proposition 1.

In (2.39), the $\lambda_n, \forall n \in \mathcal{M}$, are calculated to satisfy the power constraint at $\mathrm{BS}_n, \forall n \in \mathcal{M}$, by using the KKT condition $\lambda_n \big(\sum_{k \in \mathcal{K}_n} \mathrm{Tr}(\mathbf{T}_k \mathbf{T}_k^{\mathrm{H}}) - p_n \big) = 0$ and by utilizing the fact that the transmit power is monotonically decreasing with respect to increasing λ_n [49]. The closed-form solution can be obtained by readapting the approach shown in [12] as

$$\mathbf{T}_k^{\mathrm{WSRP}} = \beta_n \bar{\mathbf{T}}_k, \text{ where } \begin{cases} \bar{\mathbf{T}}_k = (\mathbf{A}_k + \mathbf{B}_n + \alpha_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{n,k}^{\mathrm{H}} \mathbf{R}_k^{\mathrm{H}} \mu_k, \\ \alpha_n = \sum\limits_{i \in \mathcal{K}_n} \mathrm{Tr}(\mu_i \mathbf{E}_i^{-1} \mathbf{R}_i \mathbf{R}_i^{\mathrm{H}}) / p_n, \\ \beta_n = \sqrt{\dfrac{p_n}{\sum\limits_{i \in \mathcal{K}_n} \mathrm{Tr}(\bar{\mathbf{T}}_i \bar{\mathbf{T}}_i^{\mathrm{H}})}}. \end{cases} \tag{2.52}$$

2.6.2 Network-wide WSR maximization

In this section, we consider the general network-wide WSR maximization problem. Mathematically, the WSR maximization problem can be written as [49]

$$\mathscr{P}_{\mathrm{WSRM}} = \begin{cases} \max\limits_{\substack{\mathbf{T}_k, \\ \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n}} & \sum\limits_{n \in \mathcal{M}} \sum\limits_{k \in \mathcal{K}_n} \mu_k r_k \\ \text{s.t.} & \sum\limits_{k \in \mathcal{K}_n} \mathrm{Tr}(\mathbf{T}_k \mathbf{T}_k^{\mathrm{H}}) = p_n, \forall n \in \mathcal{M}. \end{cases} \tag{2.53}$$

Problem $\mathscr{P}_{\mathrm{WSRM}}$ has been addressed in [12, 49, 50, 51, 52, 53]. For all the algorithms presented in these references, $\mathscr{P}_{\mathrm{WSRM}}$ was solved by exploiting its relationship to the WMMSE minimization problem, which was initially shown in [12]. Different from all, in the following, we propose a novel solution that directly solves $\mathscr{P}_{\mathrm{WSRM}}$. Similar to $\mathscr{P}_{\mathrm{WSRP}}$, the solution is obtained by investigating the KKT conditions of problem $\mathscr{P}_{\mathrm{WSRM}}$ and with the help of the Lemma 1 result. The solution of $\mathscr{P}_{\mathrm{WSRM}}$ w.r.t. transmit beamforming for $\mathrm{MS}_k, \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$, is given by Proposition 2.

Proposition 2 *Let the receive beamforming* $\mathbf{R}_k$ *and the MSE matrix* $\mathbf{E}_k$ *for* MS_k *be given by (2.23) and (2.24), respectively, and by utilizing the Lemma 1 result, then the solution to* $\mathscr{P}_{WSRM}$ *w.r.t. transmit beamforming for* MS_k, $\forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$, *is given as*

$$\mathbf{T}_k^{WSRM} = (\mathbf{A}_k + \mathbf{C}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{n,k}^H \mathbf{R}_k^H \mu_k, \tag{2.54}$$

where $\lambda_n, \forall n \in \mathcal{M}$, are the Lagrange multipliers associated with the $\mathscr{P}_{WSRM}$ constraint functions, $\mathbf{A}_k$ is given by (2.40), and $\mathbf{C}_n$ is given as

$$\mathbf{C}_n = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mu_j \mathbf{H}_{n,j}^H \mathbf{R}_j^H \mathbf{E}_j^{-1} \mathbf{R}_j \mathbf{H}_{n,j}. \tag{2.55}$$

Proof 5 *The proof can be shown by generalizing the derivation steps shown in proof of Proposition 1 and thus omitted here for brevity.*

By observing (2.54), we can see that it is closely related to (2.39). Theorem 2 shows the connection between both equations in a special case.

Theorem 2 *Both equations $\mathbf{T}_k^{WSRP}$ and $\mathbf{T}_k^{WSRM}$ given by (2.39) and (2.54), respectively, are equal if $N_r = 1$ and the MS_k interference price π_k given by (2.35) is replaced by $\tilde{\pi}_k$ that is given as*

$$\tilde{\pi}_k = \mu_k \pi_k. \tag{2.56}$$

Proof 6 *When $N_r = 1$, the interference price π_k in (2.35) reduces to $\pi_k = \mathbf{R}_k^H \mathbf{E}_k^{-1} \mathbf{R}_k$, since both terms $\mathbf{R}_k$ and $\mathbf{E}_k$ are scalars. By substituting π_k into the $\mathbf{A}_k$ and $\mathbf{C}_n$ terms, we have*

$$\mathbf{A}_k = \sum_{i \in \mathcal{K}_n \setminus k} \mu_i \pi_i \mathbf{H}_{n,i}^H \mathbf{H}_{n,i}.$$
$$\mathbf{C}_n = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mu_j \pi_j \mathbf{H}_{n,j}^H \mathbf{H}_{n,j}.$$

Since $\mathbf{A}_k$ is common in both, the only difference is between the $\mathbf{B}_n$ and $\mathbf{C}_n$ terms. Now, comparing $\mathbf{B}_n$ to $\mathbf{C}_n$, we can see that both terms are equal if each interference price in $\mathbf{B}_n$ is replaced by $\tilde{\pi}_j = \mu_j \pi_j$, which completes the proof.

The result of Theorem 2 establishes a relationship between problems $\mathscr{P}_{\text{WSRM}}$ and $\mathscr{P}_{\text{WSRP}}$. When $N_r = 1$, the problems $\mathscr{P}_{\text{WSRP}}$ and $\mathscr{P}_{\text{WSRM}}$ are exactly equivalent. In this case, the receive beamforming and MSE terms are scalars and directly specify the interference prices of the MSs. However, when $N_r > 1$, $\mathscr{P}_{\text{WSRP}}$ would provide a suboptimal solution to $\mathscr{P}_{\text{WSRM}}$. In this latter case, the interference prices cannot exploit the spatial dimension that the receive beamforming brings, since $\mathbf{R}_k^{\text{H}} \mathbf{E}_k^{-1} \mathbf{R}_k$ has a dimension of $N_r \times N_r$, irrespective of the number of data streams N_s, whereas the interference price $\pi_k = \text{Tr}(\mathbf{R}_k^{\text{H}} \mathbf{E}_k^{-1} \mathbf{R}_k)$ is represented by a scalar. Therefore, when $N_r > 1$, the $\mathbf{C}_n$ term given by (2.55) contains extra information, as compared to $\mathbf{B}_n$ given by (2.41), which can be exploited by the BSs to reshape the interference.

Similar to (2.39), the $\lambda_n, \forall n \in \mathcal{M}$, in (2.54) are calculated to satisfy the power constraint at $\text{BS}_n, \forall n \in \mathcal{M}$. The closed-form solution of (2.54) can be obtained similar to (2.52), by replacing the $\mathbf{B}_n$ term with $\mathbf{C}_n$.

2.6.3 WSR maximization based on self-pricing

In this approach, we consider a different strategy for maximizing the WSR in the sense that each BS would self-price the ICI it is leaking to other cells. In this regard, when compared to $\mathscr{P}_{\text{WSRP}}$, the BSs do not need to collect the interference prices from the users when calculating the transmit beamforming.

In order to show this, let us assume for a moment that the rate function of a given user, say MS_j, $j \in \mathcal{K}_m$, is mostly degraded by the transmit beamforming from a single interfering BS, say BS_n, $n \neq m$. This can be translated to many scenarios, such as the BSs other than BS_n have a mutual interference that is negligible, or they are using a transmit beamforming strategy that eliminates ICI by any means. Therefore, the interference plus noise covariance matrix of MS_j can be approximated as

$$\Psi_j \approx \sum_{k \in \mathcal{K}_n} \mathbf{H}_{n,j} \mathbf{T}_k \mathbf{T}_k^{\text{H}} \mathbf{H}_{n,j}^{\text{H}} + \mathbf{I}_{N_r}, \tag{2.57}$$

where the approximation is used due to the assumption that the intra-cell interference as well as the ICI from other interfering BSs than BS_n are negligible. By (2.57), the achievable rate of MS_j is given as

$$r_j = \log \left| \mathbf{I}_{N_s} + \mathbf{T}_j^{\text{H}} \mathbf{H}_{m,j}^{\text{H}} \Psi_j^{-1} \mathbf{H}_{m,j} \mathbf{T}_j \right|. \tag{2.58}$$

Considering the high SNR regime, the rate function of MS_j (2.58) can be approximated as

$$\begin{aligned} r_j &\approx \log \left| \mathbf{T}_j^{\text{H}} \mathbf{H}_{m,j}^{\text{H}} \Psi_j^{-1} \mathbf{H}_{m,j} \mathbf{T}_j \right| \\ &= \log \left| \mathbf{T}_j^{\text{H}} \mathbf{H}_{m,j}^{\text{H}} \mathbf{H}_{m,j} \mathbf{T}_j \right| - \log \left| \Psi_j \right|. \end{aligned} \tag{2.59}$$

From (2.59), it can be noticed that the second term in the right-hand side ($\log \left| \Psi_j \right|$) represents the rate loss at MS_j due to the beamforming at BS_n. An important point to observe is that this rate loss at MS_j is already known to BS_n, as it denotes the interference leakage from BS_n to MS_j. Therefore, BS_n can consider an implicit interference pricing approach to reduce the interference leakage towards the MSs of other cells.

Let $\Psi_n = \{\Psi_j, \forall j \in \mathcal{K}_m, \forall m \in \mathcal{M} \backslash n\}$ and define the following BS_n specific function

$$g_n(\Psi_n) = \sum_{k \in \mathcal{K}_n} \mu_k r_k - \sum_{m \in \mathcal{M} \backslash n} \sum_{j \in \mathcal{K}_m} \log \left| \Psi_j \right|. \tag{2.60}$$

Using (2.60), each BS_n, $\forall n \in \mathcal{M}$, updates the transmit beamforming $\mathbf{T}_k$, $\forall k \in \mathcal{K}_n$, as solution to the following optimization problem

$$\mathscr{P}_{\text{WSRH}} = \begin{cases} \max\limits_{\mathbf{T}_k, \forall k \in \mathcal{K}_n} & g_n(\Psi_n) \\ \text{s.t.} & \sum\limits_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_k \mathbf{T}_k^{\text{H}}) = p_n. \end{cases} \tag{2.61}$$

The solution of $\mathscr{P}_{\text{WSRH}}$ w.r.t. transmit beamforming for MS_k, $\forall n \in \mathcal{M}$, $\forall k \in \mathcal{K}_n$, is given by Proposition 3.

Proposition 3 *Let the receive beamforming $\mathbf{R}_k$ and MSE matrix $\mathbf{E}_k$ for MS_k be given by (2.23) and (2.24), respectively, and by utilizing the Lemma 1 result, then the solution to $\mathscr{P}_{WSRH}$ w.r.t. transmit beamforming for MS_k, $\forall n \in \mathcal{M}$, $\forall k \in \mathcal{K}_n$, is given as*

$$\mathbf{T}_k^{WSRH} = (\mathbf{A}_k + \mathbf{D}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{n,k}^H \mathbf{R}_k^H \mu_k, \tag{2.62}$$

where $\mathbf{A}_k$ is given by (2.40) and

$$\mathbf{D}_n = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mathbf{H}_{n,j}^H \Psi_j^{-1} \mathbf{H}_{n,j}. \tag{2.63}$$

Proof 7 *The derivation steps are similar to the ones shown in proof of Proposition 1 and thus omitted here for brevity.*

Similar to (2.39), λ_n, $\forall n \in \mathcal{M}$, in (2.62) are calculated to satisfy the power constraint at BS_n, $\forall n \in \mathcal{M}$. The closed-form solutions of (2.62) can be obtained similar to (2.52), by replacing the $\mathbf{B}_n$ term with $\mathbf{D}_n$ given by (2.63).

2.6.4 Algorithm design and convergence analysis

To solve either problem of $\mathscr{P}_{\text{WSRP}}$, $\mathscr{P}_{\text{WSRM}}$, or $\mathscr{P}_{\text{WSRH}}$, an algorithm based on alternating optimization can be used [49, 50, 51, 52, 53]. The basic idea is to optimize each problem with respect to one variable at a time, while keeping the rest of the variables fixed. The proposed algorithm to solve either optimization problem is summarized in Algorithm 2. We refer to this algorithm as WSRP when solving $\mathscr{P}_{\text{WSRP}}$, as WSRM when solving $\mathscr{P}_{\text{WSRM}}$, and as WSRH when solving $\mathscr{P}_{\text{WSRH}}$.

Algorithm 2: WSR Maximization via Alternate Optimization.

1: Initialize $\mathbf{T}_k^{(1)}$, $\forall k \in \mathcal{K}$.
2: MS_k, $\forall k$: Calculates $\mathbf{R}_k^{(t)}$ using (2.23) for given $\mathbf{T}_k^{(t)}$, $\forall k \in \mathcal{K}$.
3: MS_k, $\forall k$: Calculates $\mathbf{E}_k^{(t)}$ using (2.24) for given $\mathbf{T}_k^{(t)}$, $\mathbf{R}_k^{(t)}$, $\forall k \in \mathcal{K}$.
4: **if** $\mathscr{P}_{\text{WSRP}}$ **then**
5: $\quad \text{BS}_n$, $\forall n$: Updates $\mathbf{T}_k^{(t+1)}$, $\forall k \in \mathcal{K}_n$, using (2.39) for given $\mathbf{R}_k^{(t)}$, $\mathbf{E}_k^{(t)}$, $\forall k \in \mathcal{K}$.
6: **end if**
7: **if** $\mathscr{P}_{\text{WSRM}}$ **then**
8: $\quad \text{BS}_n$, $\forall n$: Updates $\mathbf{T}_k^{(t+1)}$, $\forall k \in \mathcal{K}_n$, using (2.54) for given $\mathbf{R}_k^{(t)}$, $\mathbf{E}_k^{(t)}$, $\forall k \in \mathcal{K}$.
9: **end if**
10: **if** $\mathscr{P}_{\text{WSRH}}$ **then**
11: $\quad \text{BS}_n$, $\forall n$: Updates $\mathbf{T}_k^{(t+1)}$, $\forall k \in \mathcal{K}_n$, using (2.62) for given $\mathbf{R}_k^{(t)}$, $\mathbf{E}_k^{(t)}$, $\forall k \in \mathcal{K}$.
12: **end if**
13: Repeat steps 2-12 (until convergence)

In step 1, Algorithm 2 initializes the transmit beamforming matrices for all users in the system by any means, e.g. MRT approach. Afterwards, the algorithm alternates between the

following three steps. In steps 2 and 3, all MSs calculate, in parallel, their receive beamforming and MSE matrices, respectively, for the given transmit beamforming. Next, all BSs calculate, in parallel, their transmit beamforming, using either approach, for the given receive beamforming and MSE matrices. If this iterative process converges, it converges to a fixed point that is a stationary point of the WSR-objective function [12]. It is worth noting that in a single-cell scenario, all proposed algorithms coincide, where each algorithm differs from the other two in the ICI handling.

For alternating optimization, monotonic convergence of the objective to a stationary (locally optimal) point is guaranteed, if each step has a unique optimum [60, Proposition 2.7.1]. The requirement for the transmit beamforming optimization to be unique is that the matrix to be inverted in (2.39), (2.54) or (2.62) is invertible, i.e., $(\mathbf{A}_k + \mathbf{B}_n/\mathbf{C}_n/\mathbf{D}_n + \lambda_n \mathbf{I})^{-1}$ does exist. One sufficient condition for invertibility is that all the power constraints are active, i.e., $\sum_{k\in\mathcal{K}_n} \mathrm{Tr}(\mathbf{T}_k \mathbf{T}_k^{\mathrm{H}}) = p_n$ so that we always have $\lambda_n > 0$, which is the case in the formulation of our problems, since the WSR optimum is reached at maximum transmit power [12]. Another condition is that there are at least N_t active vectors whose effective channels are linearly independent, i.e., $\mathrm{rank}(\mathbf{A}_k + \mathbf{B}_n/\mathbf{C}_n/\mathbf{D}_n) \geq N_t$. In practice, the cases when the matrix is non-invertible, and the optimal beamforming solution is not unique, are very rare [53]. Nevertheless, if the matrix is not invertible, the pseudo-inverse may be used to get a solution. However, since the original problems $\mathscr{P}_{\mathrm{WSRP}}$, $\mathscr{P}_{\mathrm{WSRM}}$, and $\mathscr{P}_{\mathrm{WSRH}}$ are non-convex, a globally optimal point cannot be found, in general, via alternating optimization. Moreover, different initializations and iteration orders may converge to different local WSR-optima [49, 50, 51, 52, 53].

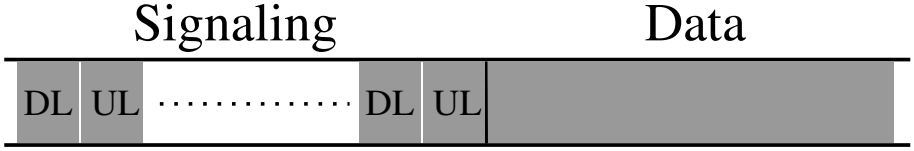
2.7 WSR maximization signaling schemes

The proposed algorithms above are decentralized, where each BS can calculate its transmit beamforming locally once it has the required information. Here, we assume that each BS_n has access to the local CSI, i.e., $\mathbf{H}_{n,j}$, $\forall j \in \mathcal{K}$, as in [51, 52, 53]. TDD operation is an effective technique for obtaining this CSI, where uplink training in conjunction with reciprocity provides the BSs with downlink and uplink CSI simultaneously [13, 55].

In the following, a novel OTA signaling schemes are proposed based on TDD mode to facilitate the algorithm implementation. It is assumed that 1) each BS and MS has orthogonal pilot symbols (training) in the downlink and uplink direction, respectively, for the OTA signaling, 2) each TDD frame is divided into two parts; signaling and data parts, as shown in Fig. 2.2, where the signaling part is further divided into downlink and uplink sub-parts to facilitate the variables' exchange between BSs and MSs, and 3) all exchanged variables are perfectly estimated at each iteration.

At the downlink, it is assumed that each BS_n transmits pilot signals that are precoded with the transmit beamforming $\mathbf{T}_k$, $\forall k \in \mathcal{K}_n$. Thus, each MS_k can estimate the downlink equivalent channels $\mathbf{H}_{m,k}\mathbf{T}_j$, $\forall m \in \mathcal{M}$, $\forall j \in \mathcal{K}_m$, and update its receive beamforming $\mathbf{R}_k$ and MSE-matrix $\mathbf{E}_k$. On the other hand, to update the transmit beamforming, each algorithm has

Figure 2.2 – TDD frame structure.



Source: Created by author.

different signaling needs. Therefore, we propose the following two signaling schemes.

2.7.1 Signaling scheme A

In this scheme, it is assumed that each MS_k transmits an uplink pilot signal that is precoded with the receive beamforming $\mathbf{R}_k$. Thus, each BS_n can estimate the uplink equivalent channels $\mathbf{R}_j\mathbf{H}_{n,j}$, $\forall j \in \mathcal{K}$, and calculate $\mathbf{E}_k$, $\forall k \in \mathcal{K}_n$, using (2.24). This information is sufficient to calculate $\mathbf{A}_k$, $\forall k \in \mathcal{K}_n$, which is common in the three algorithms.

For WSRH, $\mathbf{A}_k$, $\forall k \in \mathcal{K}_n$, is all that is needed to update the transmit beamforming $\mathbf{T}_k$, $\forall k \in \mathcal{K}_n$, where the second term $\mathbf{D}_n$ can be calculated locally. However, for WSRP, each BS_n would require vector $\boldsymbol{\pi}_n$ to calculate $\mathbf{B}_n$, which collects all interference prices from the users of other cells. The direct approach, as assumed in [47], is to let each MS_k, $k \in \mathcal{K}_n$, calculate its interference price π_k and feed it back to its serving BS, i.e., BS_n. Then, all BSs perform broadcast-and-gather operation of their interference prices using the backhaul. Different from [47], we assume that each BS_n first recalculates the receive beamforming as $\mathbf{R}_k = (\mathbf{I}-\mathbf{E}_k)(\mathbf{H}_{n,k}\mathbf{T}_k)^{-1}$, $\forall k \in \mathcal{K}_n$, using local information (see Appendix **??**), and then calculates the interference prices $\pi_k = \text{Tr}[\mathbf{R}_k^{\text{H}}\mathbf{E}_k^{-1}\mathbf{R}_k]$, $\forall k \in \mathcal{K}_n$, and exchanges them with the other BSs using the backhaul. Thus, we do not need the feedback from the MSs. On the other hand, for WSRM, each BS_n would require $\mu_j\mathbf{E}_j$, $\forall j \in \mathcal{K}\backslash\mathcal{K}_n$, to calculate $\mathbf{C}_n$. Using this signaling scheme, one possible way, as assumed in [53], is to let the BSs exchange them using the backhaul.

From above, we can see that signaling Scheme A is best applicable to WSRH, since no further variables feedback is required. However, for WSRM (WSRP), the feedback of matrices (scalars) between BSs is required. To reduce the signaling overhead of WSRM and WSRP, we further propose the following signaling scheme.

2.7.2 Signaling scheme B

In this scheme, we assume that each MS_k transmits a pilot signal that is precoded with $\sqrt{\mu_k}\mathbf{E}_k^{-\frac{1}{2}}\mathbf{R}_k$ [51, 53]. Thus, each BS_n can estimate the uplink equivalent channels $\sqrt{\mu_j}\mathbf{E}_j^{-\frac{1}{2}}\mathbf{R}_j\mathbf{H}_{n,j}$, $\forall j \in \mathcal{K}$, which are sufficient to calculate $\mathbf{A}_k$, $\forall k \in \mathcal{K}_n$, and $\mathbf{B}_n$ or $\mathbf{C}_n$. However, with WSRM and WSRP, each BS_n still requires $\mathbf{R}_k$, $\forall k \in \mathcal{K}_n$, to calculate the uplink equivalent channels $\mu_k\mathbf{R}_k\mathbf{H}_{n,k}$, $\forall k \in \mathcal{K}_n$. One possible way, as proposed in [51], is to let each MS_k transmit two consecutive uplink pilot signals; one precoded with $\sqrt{\mu_k}\mathbf{E}_k^{-\frac{1}{2}}\mathbf{R}_k$ and another precoded with $\mathbf{R}_k$. However, this approach would unnecessarily increase the signaling overhead. In the

following, an alternative approach is proposed, where the main idea is to let each BS_n recalculate $\mathbf{R}_k$, $\forall k \in \mathcal{K}_n$, using only local information and thus reduce the signaling overhead.

Let $\mathbf{X}_k = \sqrt{\mu_k}\mathbf{E}_k^{-\frac{1}{2}}\mathbf{R}_k\mathbf{H}_{n,k}$ denote the uplink equivalent channel with MS_k estimated at BS_n. Substitute $\mathbf{R}_k = (\mathbf{I}-\mathbf{E}_k)(\mathbf{H}_{n,k}\mathbf{T}_k)^{-1}$ into $\mathbf{X}_k$ and simplify the resulting expression, then we have

$$\mathbf{X}_k = (\sqrt{\mu_k}\mathbf{E}_k^{-\frac{1}{2}} - \sqrt{\mu_k}\mathbf{E}_k^{\frac{1}{2}})\mathbf{Y}_k, \tag{2.64}$$

where $\mathbf{Y}_k = (\mathbf{H}_{n,k}\mathbf{T}_k)^{-1}\mathbf{H}_{n,k}$. Right multiply both sides of the latter equation by $\mathbf{Y}_k^{\dagger}$ and again simplify the resulting expression, then we have

$$\mathbf{E}_k^{-\frac{1}{2}} = \frac{1}{\sqrt{\mu_k}}\mathbf{M}_k + \mathbf{E}_k^{\frac{1}{2}}, \tag{2.65}$$

where $\mathbf{M}_k = \mathbf{X}_k\mathbf{Y}_k^{\dagger}$ (where $\mathbf{Y}_k^{\dagger} = \mathbf{Y}_k^{\mathrm{H}}(\mathbf{Y}_k\mathbf{Y}_k^{\mathrm{H}})^{-1}$), which is formed using local information. Then, equation (2.65) can be solved for $\mathbf{E}_k$ iteratively as given by Algorithm 3.

Algorithm 3: Proposed signaling scheme B.

1: Construct $\mathbf{M}_k$ and initialize $\mathbf{E}_k^{(1)}$ randomly.
2: Set $\bar{\mathbf{E}}_k^{(t)} = \frac{1}{\sqrt{\mu_k}}\mathbf{M}_k + \mathbf{E}_k^{\frac{1}{2}(t)}$.
3: Set $\mathbf{E}_k^{(t+1)} = \bar{\mathbf{E}}_k^{-2(t)}$.
4: Repeat steps 2-3 (until convergence).
5: At convergence, solve $\mathbf{R}_k = (\mathbf{I}-\mathbf{E}_k)(\mathbf{H}_{n,k}\mathbf{T}_k)^{-1}$.

In step 1, Algorithm 3 constructs the local matrix $\mathbf{M}_k$ and randomly initializes $\mathbf{E}_k^{(1)}$. Given those initial matrices, the algorithm alternates between steps 2 and 3 at each iteration. At the t-th iteration, the algorithm solves for $\mathbf{E}_k^{-\frac{1}{2}(t)}$ given $\mathbf{E}_k^{\frac{1}{2}(t)}$ at step 2 (denoted as $\bar{\mathbf{E}}_k^{(t)}$). Then, at step 3, the algorithm solves for $\mathbf{E}_k^{(t+1)}$ given $\bar{\mathbf{E}}_k^{(t)}$. Those two steps are repeated until convergence. If matrix $\mathbf{M}_k$ is assumed perfect, the algorithm is able to recalculate the MSE-matrix $\mathbf{E}_k$ perfectly. Then, using $\mathbf{E}_k$ at step 5 we can calculate for the receive beamforming matrix $\mathbf{R}_k$. Algorithm 3 convergence behavior is shown numerically in the next section. A proof of convergence is open and we leave it for a future work.

2.8 Numerical results

This section evaluates the performance of the proposed algorithms by means of simulation. We consider a flat Rayleigh fading scenario with uncorrelated channels between antennas, i.e., each element of $\mathbf{H}_{n,k}$, $\forall n \in \mathcal{M}$, $\forall k \in \mathcal{K}_n$, is an i.i.d. complex Gaussian random variable with zero mean and unit variance. For each simulated algorithm, we initialize the transmit beamforming matrices using the MRT approach, i.e.,

$$\mathbf{T}_k^{(1)} = \mathbf{V}_{k[1:Ns]}\sqrt{\frac{p_n}{KN_s}}, \forall k \in \mathcal{K}, \tag{2.66}$$

where $\mathbf{V}_k$ denotes the matrix holding in its columns the right singular vectors of $\mathbf{H}_{n,k}$ arranged in a decreasing order w.r.t their singular values. Moreover, we assume that the noise variance $\sigma_k^2 = \sigma^2 = 1$. The transmit power of BS_n is represented by SNR as $\text{SNR} = p_n/\sigma^2$. For comparison purpose, we show simulation results of the WSR-WMMSE algorithm from [49, 50, 51, 52, 53], and the MRT approach.

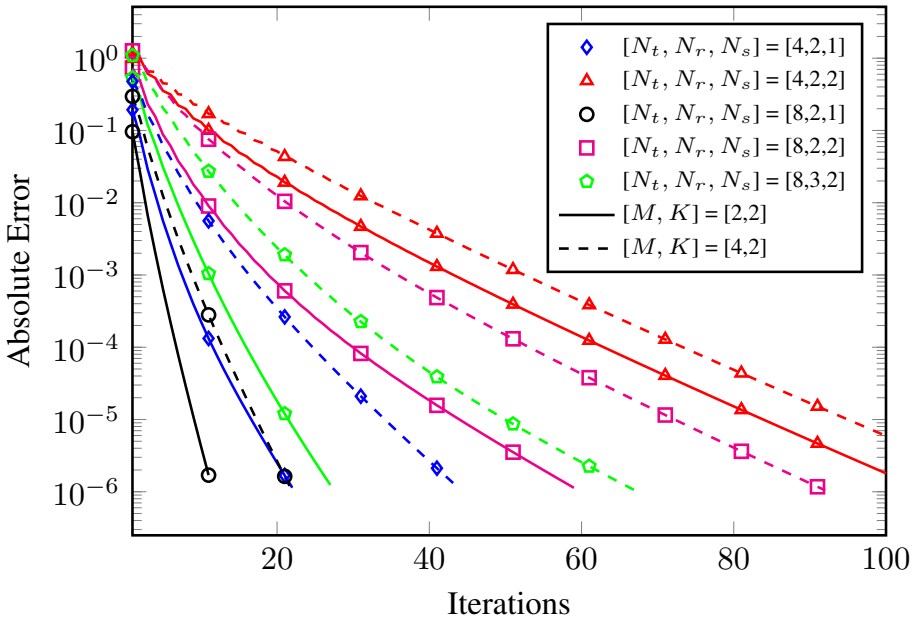
Example 1: algorithm 3 convergence

This example shows simulation results to evaluate the convergence behavior of Algorithm 3. Fig. 2.3 shows the log-scale convergence results of Algorithm 3 for the first user, i.e., MS_1, in terms of the Absolute Error that is defined as

$$\text{Absolute Error} = \left| \sum_{i,j} \left(\mathbf{E}_1^{(t)} - \mathbf{E}_1^{\star} \right) \right|, \tag{2.67}$$

where $\mathbf{E}_1^{\star}$ is the perfect MSE-matrix of MS_1 and $\mathbf{E}_1^{(t)}$ is the obtained MSE-matrix at the t-th iteration. We assume system of $M = 2$ and $K = 2$ and each simulated point is averaged over 1,000 channel realizations.

Figure 2.3 – Log-scale convergence behavior of Algorithm 3.



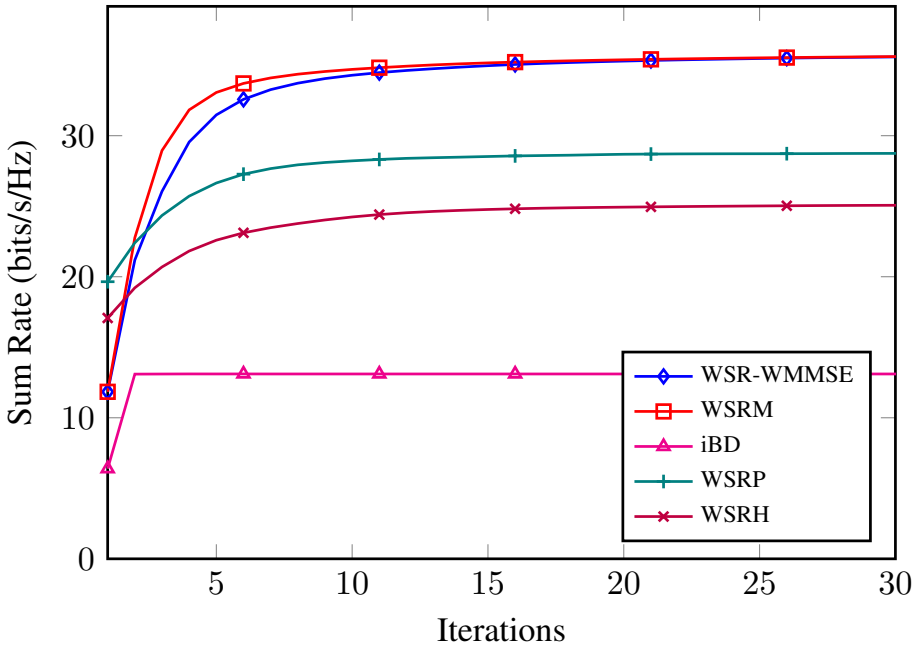
Source: Created by author.

From Fig. 2.3, it can be seen that Algorithm 3 has a fast convergence rate, where it is able to obtain the perfect MSE-matrix using a few iterations. Note that, when N_t increases, N_s decreases, or MK decreases, the algorithm has faster convergence rate. We note that all simulated channel realizations have converged to the perfect MSE-matrix, although the convergence of some channel realizations is not necessarily monotonic.

Example 2: convergence behavior of algorithms 1 and 2

This example shows simulation results to evaluate the convergence behavior of Algorithms 1 and 2. Fig. 2.4 shows the averaged sum rate convergence results assuming SNR = 10 dB and $[M, K, N_t, N_r, N_s] = [3, 3, 9, 2, 2]$.

Figure 2.4 – Convergence behavior of algorithms 1 and 2.



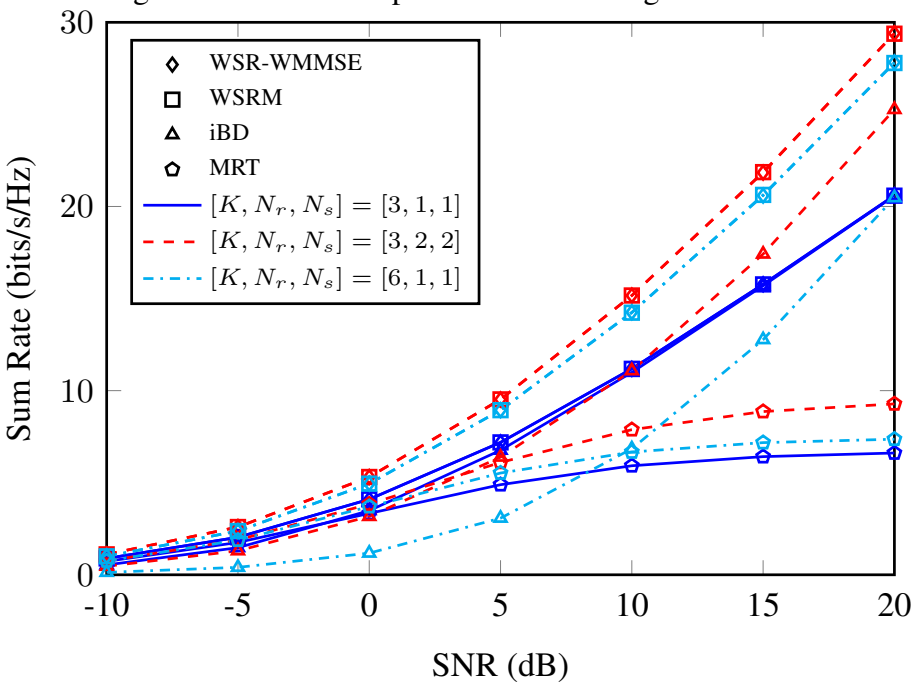
Source: Created by author.

From Fig. 2.4, we can see that all iterative algorithms have a fast convergence rate, within 1-to-2 iterations for iBD and within 10-to-15 iterations for the other algorithms. It's worth noting that WSRM has a slightly faster convergence rate than WSR-WMMSE, although both algorithms seem to converge almost to the same point. However, the convergence speed of either algorithm varies for the individual channel realizations. For some channel realizations, WSRM appears to converge slightly faster and to a higher sum rate than WSR-WMMSE and vice-versa for the other channel realizations.

Example 3: sum rate performance for single-cell case

This example shows simulation results to evaluate the sum rate performance of Algorithms 1 and 2 in the single cell case, i.e., $M = 1$. Fig. 2.5 shows the average sum rate results for a range of SNR values, assuming $\mu_k = \mu = 1$ and $N_t = 6$. Note that, when $M = 1$, algorithms WSRM, WSRP, and WSRH are all equivalent, since the $\mathbf{B}_n$, $\mathbf{C}_n$ and $\mathbf{D}_n$ terms are all identity matrices and all algorithms share the $\mathbf{A}_k$ calculation. Therefore, for this example, we only show WSRM results.

Figure 2.5 – Sum rate performance for single cell case.

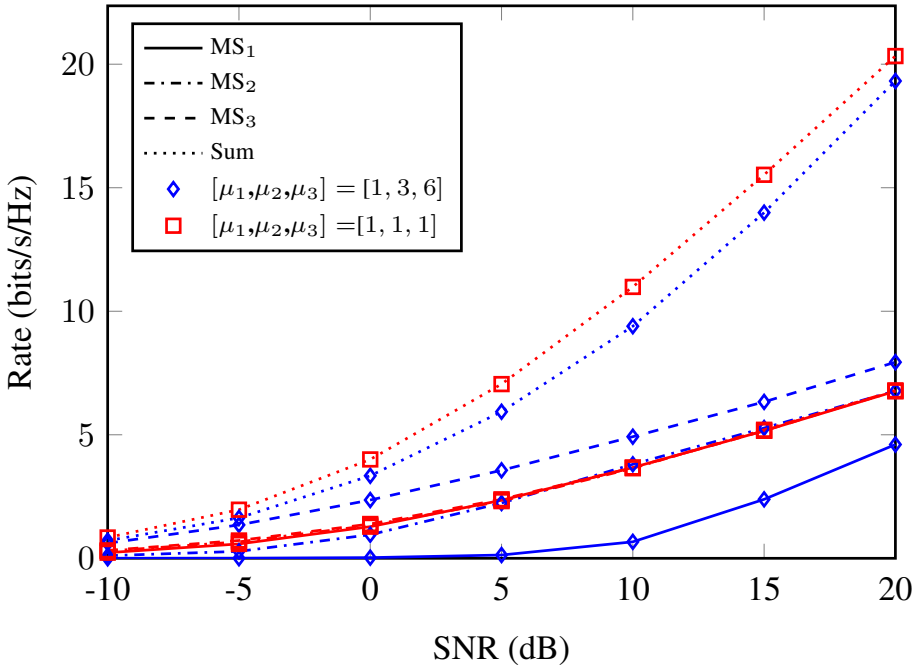


Source: Created by author.

From Fig. 2.5, it can be seen that when $[K, N_r, N_s] = [3, 1, 1]$ (solid-lines), iBD has very close performance to both WSR-WMMSE and WSRM, which are seem to have the same sum rate performance. However, for the other simulated scenarios, when K and/or N_s increases while keeping N_t fixed, iBD has large performance loss as compared to WSR-WMMSE and WSRM, although it maintain the same multiplexing gain. None of the iBD and MRT algorithms can achieve a good balance between the altruistic and egoistic behaviors of users. While iBD has complete altruistic behavior, MRT has complete egoistic behavior. Consequently, they have performance loss, as compared to WSR-WMMSE and WSRM.

To examine the impact of user-weights, Fig. 2.6 shows the simulation results for a system with equal and unequal user-weights and $[M, K, N_t, N_r, N_s] = [1, 3, 6, 1, 1]$. From Fig. 2.6, it can be seen that when all users have equal weights, they achieve equal performance, in average. However, when a user has a larger weight than others, MS_3 in this case, the algorithm favors him and, thus, achieves better performance. In terms of sum rate performance, the system with equal users' weights has better performance than otherwise. The reason behind this is that when the algorithm favors one user over the others, the user(s) with lower weight would have a degradation in his(their) performance, MS_1 in this case. In general, the increase of one user's rate does not compensate the loss of the other users' rate. Thus, the algorithm would lose in terms of sum rate.

Figure 2.6 – Sum rate performance with equal and different user-weights.

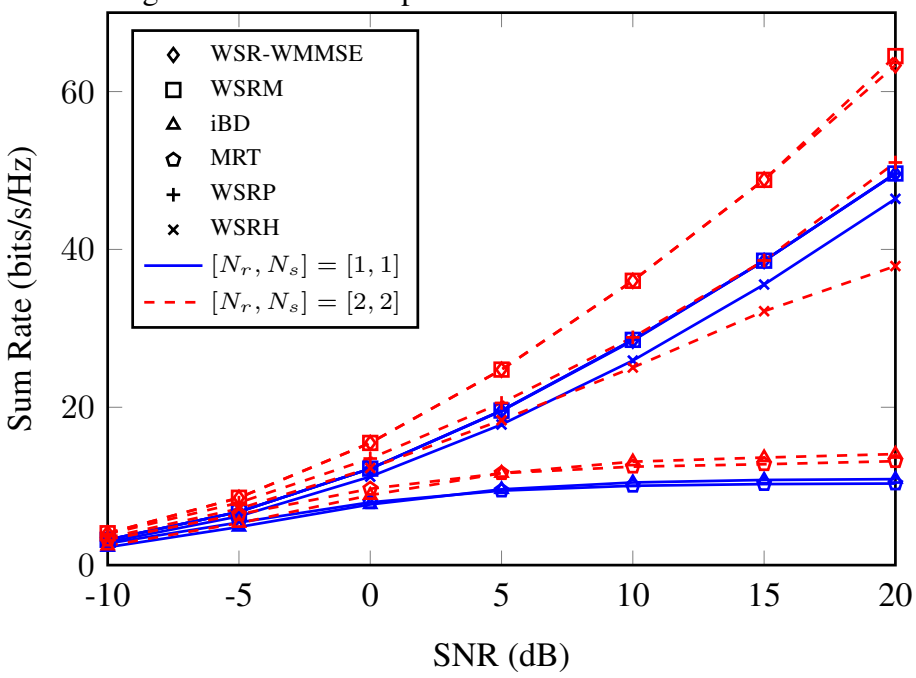


Source: Created by author.

Example 4: sum rate performance for multicell case

This example shows simulation results to evaluate the sum rate performance of Algorithms 1 and 2 in the multicell case, i.e., $M > 1$, assuming $\mu_k = \mu = 1$. Fig. 2.7 shows the average sum rate results for a range of SNR values, where $[M, K, N_t] = [3, 3, 9]$.

Figure 2.7 – Sum rate performance for multicell case.

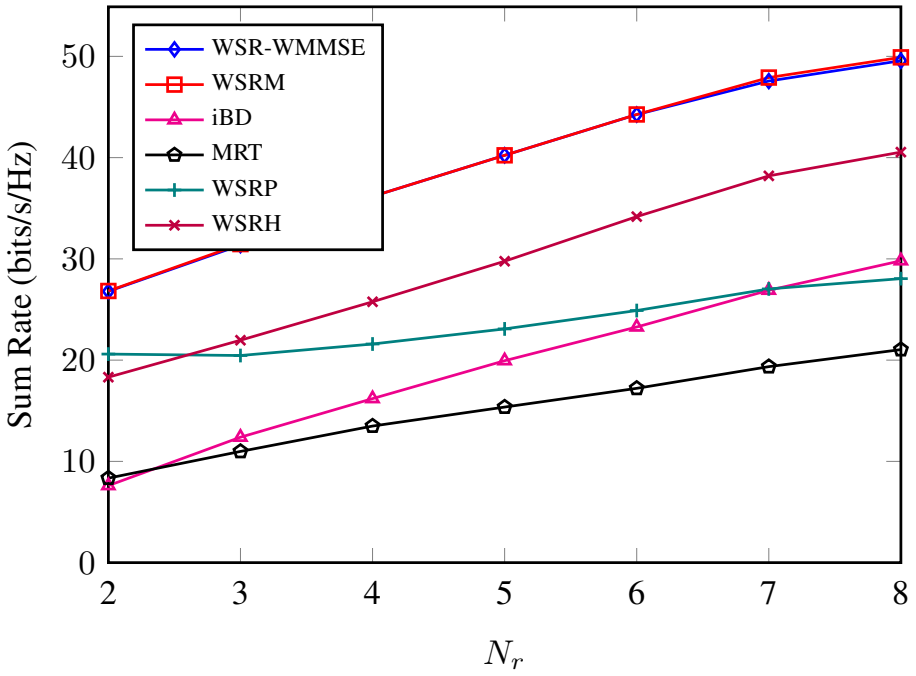


Source: Created by author.

From Fig. 2.7, we can see that WSRM and WSRP have the exact sum rate performance when $N_r = N_s = 1$ (solid-lines), since both algorithms are equivalent as it is shown in Theorem 2. However, when $N_r = N_s = 2$, we can see that WSRP has some performance loss, as compared to WSRM, and the performance loss increases as the SNR value increases. Furthermore, both iBD and MRT have a flat performance as the SNR increases, due to severe ICI. For WSRH, we can see that it has close performance to WSRM for the entire SNR range with a small performance loss when $N_r = N_s = 1$. However, increasing N_s, the performance loss increases, as well, since N_t is fixed. WSRH is a self-pricing algorithm that is distributed between cells. Thus, the algorithm has less transmit coordination than WSRM and WSR-WMMSE.

In Fig. 2.8, we show sum rate performance while varying the number of MS antennas N_r and fixing the other parameters. We assume SNR = 10 dB and $[M, K, N_t, N_r, N_s] = [3, 3, 6, N_r, 2]$. From Fig. 2.8, we can see that all algorithms have better sum rate performance as N_r increases. However, WSRP has a much slower increase rate than others, which is translated to a higher rate loss, as compared to WSRM. On the other hand, as N_r increases, iBD starts to have better sum rate performance than MRT, as compared to results from Fig. 2.7. The reason behind this is that when N_r increases, the interference whitening method has better impact on reducing ICI effects, and thus, better performance.

Figure 2.8 – Sum rate performance for multicell case while varying number of MS antennas.

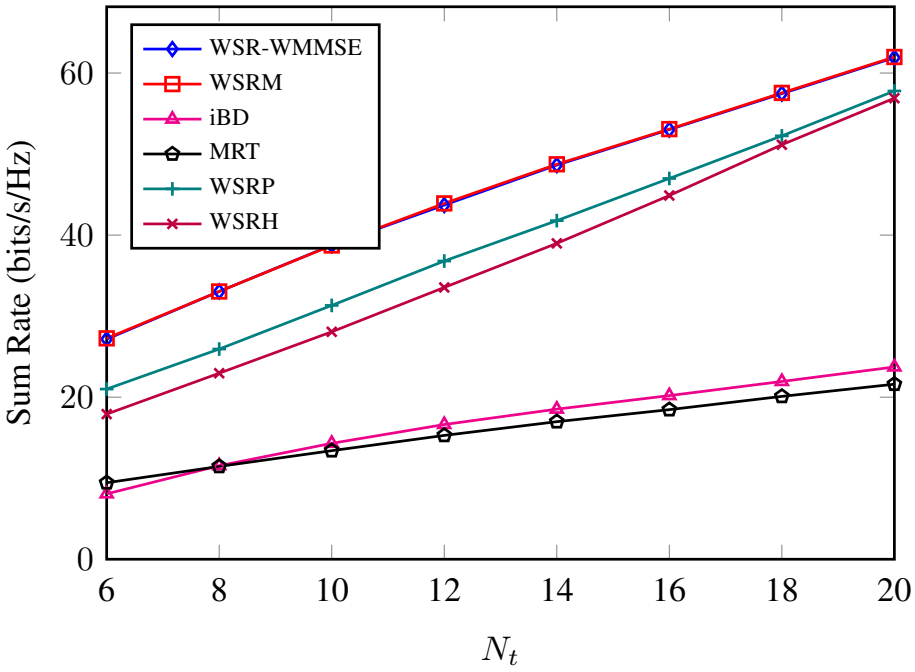


Source: Created by author.

In Fig. 2.9, we show sum rate performance while varying the number of BS antennas and fixing the other parameters. We assume SNR = 10 dB and $[M, K, N_t, N_r, N_s] = [3, 3, N_t, 2, 2]$. As discussed above, we can see from Fig. 2.9 that all algorithms have better sum rate performance as N_t increases. Different from Fig. 2.8, WSRP and WSRH sum rate increases as N_t increases,

thus reducing the rate loss as compared to WSRM. Meanwhile, iBD has a slower sum rate increase as compared to results from Fig. 2.8.

Figure 2.9 – Sum rate performance for multicell case while varying number of BS antennas.



Source: Created by author.

2.9 Chapter conclusions

This chapter has considered the WSR maximization problem in multicell MIMO BC and proposed three different algorithms, which are based on the alternating optimization technique and are guaranteed to converge to a local WSR-optimum. For all algorithms, the transmit beamforming matrices are obtained by investigating the KKT conditions of the problems with the help of Lemma 1. In contrast to the WSR-WMMSE algorithm from [49, 50, 51, 52, 53], which solves the WSR maximization problem indirectly by solving the WMMSE minimization problem, the proposed algorithms in this chapter provide a direct solution to the WSR maximization problem. Using computer simulations, it was shown that the proposed algorithms achieve comparable sum rate performance to the WSR-WMMSE algorithm, while using fewer iterations. Further, it was shown that the network-wide WSR maximization can be equivalently solved using an interference pricing approach if 1) each MS is equipped with a single-antenna and 2) the users' weights are included in the interference prices. Furthermore, two different signaling schemes based on TDD mode were also proposed to facilitate the implementation of the algorithms. Different from existing schemes, the proposed signaling schemes reduce the signaling overhead and require no variables feedback between BSs.

3 A NOVEL CELL RECONFIGURATION TECHNIQUE FOR DYNAMIC TDD WIRELESS NETWORKS

3.1 Introduction

Recently, DTDD was introduced to cope with traffic fluctuations by allowing each cell to adaptively reconfigure its communication direction based on the prevailing traffic demands and interference levels [15]. By means of simulation [15] and performance analysis [16], it was shown that the adaptive cell reconfiguration technique enhances the system spectral and energy efficiency, especially in scenarios in which the offered traffic is time-varying and asymmetric in terms of uplink/downlink direction. However, allowing neighboring cells to have different transmission directions gives rise to BS-to-BS and MS-to-MS interference (see Fig. 1.4), which can severely degrade the system performance [18].

Earlier works for DTDD systems either assumed that the UL/DL configurations are given [51] or reconfigured each cell direction based only on the aggregate traffic in the cell [19]. While this latter approach is fairly simple and inherently distributed, it cannot achieve the potential performance, as it disregards the interference effects that are particularly severe in DTDD systems. Recognizing this issue, cell reconfiguration schemes that account for the users' traffic demands and interference levels were recently proposed in [20, 21]. Specifically, the authors in [20] formulated the cell reconfiguration problem as a non-cooperative game with the objective of minimizing the total traffic flow delays. The authors in [21] used a cluster-based approach [61] and proposed a cell reconfiguration algorithm, where the highly interfering cells are first grouped into a cluster, in which each cell selects a direction that matches its own traffic demands and receives the least interference. While the proposed algorithms in [20, 21] improve the throughput as compared to the conventional STDD, they are scenario (traffic model) specific and do not consider the users' traffic characteristics. Specifically, users of modern wireless networks typically demand a wide range of services, where each may have different traffic characteristics, in terms of packet size and maximum packet delay [22]. Therefore, it is important for the cell reconfiguration algorithm to support such different traffic characteristics.

3.2 Chapter contributions

This chapter proposes a novel cell reconfiguration formulation that takes into account both prevailing traffic conditions and multicell BS-to-BS and UE-to-UE interference levels. The proposed optimization problem is then solved optimally using the ILP algorithm [31]. However, due to its high computational complexity, a heuristic solution is then proposed based on the PSO algorithm [32], which is shown to achieve near optimal performance with much lower computational complexity. System level evaluations are carried out, from which the effectiveness of our proposed scheme is evidenced in terms of the packet throughput as compared

to conventional STDD and other reference schemes, that disregard the DTDD specific inter-cell interference effects. Although the system simulator uses the gaming network traffic model from [22], the proposed scheme is general in the sense that it is not traffic model dependent, and therefore can easily be applied to any traffic model and/or deployment scenario. This feature makes the proposed scheme especially suitable for current and future wireless networks.

3.3 Chapter organization

This chapter is organized as follows. Section 3.4 presents the system model and problem formulation. In section 3.5, the proposed ILP approach is presented. The proposed algorithm based on the PSO technique is presented in sections 3.6. Finally, section 3.7 presents the numerical results and then section 3.8 concludes the chapter.

3.4 System Model and Problem Formulation

Consider a multicell multiuser system model comprising M cells, where each cell has one BS that is communicating with/serving K MSs, where each MS is served by only one BS. Let $\mathcal{M} \stackrel{\text{def}}{=} \{1, \ldots, M\}$ and $\mathcal{K} \stackrel{\text{def}}{=} \{\mathcal{K}_1, \ldots, \mathcal{K}_M\}$ denote the sets of all BSs and MSs, respectively, where $\mathcal{K}_n$ denotes the set of MSs associated with the n-th BS. The BS of the n-th cell is denoted as BS_n and the k-th MS in each cell is denoted as MS_k. In each TTI, it is assumed that each MS_k has predefined downlink and uplink weights denoted as $\alpha_{k,1} \geq 1$ and $\alpha_{k,2} \geq 1$ respectively, where $\alpha_{k,1} > \alpha_{k,2}$ means that MS_k prefers the downlink to uplink direction, and vice-versa. Throughout this chapter, subscript 1 is used for downlink and 2 for uplink. The link-weights can be optimized in a way that reflects the user traffic main characteristics, such as the packet size and/or packet maximum delay. A possible link-weights optimization criterion is shown later in Section 3.7.

Let $\boldsymbol{\beta} \in \mathbb{Z}^{M \times 2}$ be a binary matrix of two columns, where each column represents one communication direction, i.e., $\boldsymbol{\beta}_{n,1} = 1$ if cell n is at downlink and $\boldsymbol{\beta}_{n,2} = 1$ if cell n is at uplink. It is assumed that each cell is allowed to be in one direction in each TTI. Therefore, the summation of the two columns of $\boldsymbol{\beta}$ should be a vector of all ones, i.e., $\boldsymbol{\beta}_{n,1} + \boldsymbol{\beta}_{n,2} = 1, \forall n$. Further, let $\gamma_{k,1}$ and $\gamma_{k,2}$ define the two functions of MS_k, $k \in \mathcal{K}_n$, which are given as

$$\gamma_{k,1} = \frac{\alpha_{k,1} p_k g_{n,k}}{I_{k,1} + \sigma_k^2}, \tag{3.1}$$

$$\gamma_{k,2} = \frac{\alpha_{k,2} q_k g_{n,k}}{I_{k,2} + \sigma_k^2}, \tag{3.2}$$

where p_k (q_k) is the downlink (uplink) transmit power of the MS_k, σ_k^2 is the noise variance, $I_{k,1}$ and $I_{k,2}$ are the out-of-cell interference (OCI) power received at downlink and uplink, respectively,

which are given as

$$I_{k,1} = \underbrace{\sum_{m\in\mathcal{M}\setminus n} \boldsymbol{\beta}_{m,1} \sum_{j\in\mathcal{K}_m} \alpha_{j,1}^{-1} p_j g_{m,k}}_{\text{DL inter-cell interference}} + \underbrace{\sum_{m\in\mathcal{M}\setminus n} \boldsymbol{\beta}_{m,2} \sum_{j\in\mathcal{K}_m} \alpha_{j,2}^{-1} q_j \tilde{g}_{j,k}}_{\text{MS to MS interference}}, \tag{3.3}$$

$$I_{k,2} = \underbrace{\sum_{m\in\mathcal{M}\setminus n} \boldsymbol{\beta}_{m,2} \sum_{j\in\mathcal{K}_m} \alpha_{j,2}^{-1} q_j g_{n,j}}_{\text{UL inter-cell interference}} + \underbrace{\sum_{m\in\mathcal{M}\setminus n} \boldsymbol{\beta}_{m,1} \sum_{j\in\mathcal{K}_m} \alpha_{j,1}^{-1} p_j \hat{g}_{m,n}}_{\text{BS to BS interference}}, \tag{3.4}$$

where $g_{n,k}$ is the channel gain from BS_n to MS_k, $\hat{g}_{n,m}$ is the channel gain from BS_n to BS_m, and $\tilde{g}_{j,k}$ is the channel gain from MS_j to MS_k. Throughout this chapter, letters n and m are used to index cells/BSs, while letters k, i and j are used to index users/MSs.

Note that in functions $\gamma_{k,1}$ and $\gamma_{k,2}$, the signal links are weighted directly, while the interfering links are weighted inversely. This makes each user have a stronger signal link and weaker interfering links in the preferred direction, which increases the probability of being optimized in the preferred direction. Further, the functions $\gamma_{k,1}$ and $\gamma_{k,2}$ are SINR-like functions if $\alpha_{k,d} = 1, \forall k \in \mathcal{K}, \forall d \in \{1, 2\}$. Therefore, the rate-like function of MS_k in direction d is given as

$$r_{k,d} = \log(1 + \gamma_{k,d}). \tag{3.5}$$

From above, the proposed cell reconfiguration optimization problem can be written as

$$\mathscr{P}_1 = \begin{cases} \max\limits_{\boldsymbol{\beta}\in\mathbb{Z}^{M\times 2}} & \sum\limits_{n\in\mathcal{M}} \sum\limits_{k\in\mathcal{K}_n} \sum\limits_{d\in\{1,2\}} \boldsymbol{\beta}_{n,d} r_{k,d} \\ \text{s.t.} & \sum\limits_{d\in\{1,2\}} \boldsymbol{\beta}_{n,d} = 1, \forall n. \end{cases} \tag{3.6}$$

Problem $\mathscr{P}_1$ is a mixed integer linear programming (MILP) problem, which is non-convex and very hard to solve even for small scale problems [62]. Here, similar to [31], we consider the low-SINR regime. In this case, the rate maximization problem is simplified by approximating the Shannon capacity with its first order Taylor series around zero SINR. That is, one can approximate MS_k, $k \in \mathcal{K}_n$, objective function as $\boldsymbol{\beta}_{n,d} r_{k,d} \approx \boldsymbol{\beta}_{n,d} \gamma_{k,d}$. Then, problem $\mathscr{P}_1$ can be written as

$$\mathscr{P}_2 = \begin{cases} \max\limits_{\boldsymbol{\beta}\in\mathbb{Z}^{M\times 2}} & \sum\limits_{n\in\mathcal{M}} \sum\limits_{k\in\mathcal{K}_n} \sum\limits_{d\in\{1,2\}} \boldsymbol{\beta}_{n,d} \gamma_{k,d} \\ \text{s.t.} & \sum\limits_{d\in\{1,2\}} \boldsymbol{\beta}_{n,d} = 1, \forall n. \end{cases} \tag{3.7}$$

Although $\mathscr{P}_2$ is still an MILP problem, it can be transformed into an ILP from, as it will be shown in the next section, at the expense of additional variables and constraints.

3.5 Centralized cell reconfiguration - method 1

This section proposes the first solution method to solve the optimization problem $\mathscr{P}_2$ based on the ILP technique [31]. To begin with, let $\psi_{k,1}$ and $\psi_{k,2}$ be the two functions of

MS_k, $k \in \mathcal{K}_n$, that are given as

$$\psi_{k,1} = \frac{\boldsymbol{\beta}_{n,1}}{I_{k,1} + \sigma_k^2}, \tag{3.8}$$

$$\psi_{k,2} = \frac{\boldsymbol{\beta}_{n,2}}{I_{k,2} + \sigma_k^2}. \tag{3.9}$$

Without loss of generality, we assume $\sigma_k^2 = 1$. Thus, $\psi_{k,d} \leq 1$, $\forall k \in \mathcal{K}$, $\forall d \in \{1, 2\}$. Using the cross multiplication, we obtain

$$\psi_{k,d} I_{k,d} + \psi_{k,d} \sigma_k^2 = \boldsymbol{\beta}_{n,d},\ d \in \{1, 2\}. \tag{3.10}$$

Substituting $I_{k,d}$, $\forall d \in \{1, 2\}$, that are given by (3.3) and (3.4) into the latter equation, we have the two following functions:

$$\sum_{m \in \mathcal{M} \setminus n} \psi_{k,1} \boldsymbol{\beta}_{m,1} \sum_{j \in \mathcal{K}_m} \alpha_{j,1}^{-1} p_j g_{m,k} + \sum_{m \in \mathcal{M} \setminus n} \psi_{k,1} \boldsymbol{\beta}_{m,2} \sum_{j \in \mathcal{K}_m} \alpha_{j,2}^{-1} q_j \tilde{g}_{j,k} + \psi_{k,1} \sigma_k^2 = \boldsymbol{\beta}_{n,1}, \tag{3.11}$$

$$\sum_{m \in \mathcal{M} \setminus n} \psi_{k,2} \boldsymbol{\beta}_{m,2} \sum_{j \in \mathcal{K}_m} \alpha_{j,2}^{-1} q_j g_{n,j} + \sum_{m \in \mathcal{M} \setminus n} \psi_{k,2} \boldsymbol{\beta}_{m,1} \sum_{j \in \mathcal{K}_m} \alpha_{j,1}^{-1} p_j \hat{g}_{m,n} + \psi_{k,2} \sigma_k^2 = \boldsymbol{\beta}_{n,2}. \tag{3.12}$$

Note that (3.11) and (3.12) are non-linear functions due to $\psi_{k,d} \boldsymbol{\beta}_{m,d}$. However, we can write both functions in a linear form with some added constraints as follows.

Let $\mathbf{X} \in \mathbb{R}_+^{MK \times 2 \times M \times 2}$ define a non-negative matrix such that $\mathbf{X}_{k,d,m,u} = \psi_{k,d} \boldsymbol{\beta}_{m,u}$. Note that, $\boldsymbol{\beta}_{m,d} \in \{0, 1\}$ and $\sum_{d \in \{1,2\}} \boldsymbol{\beta}_{m,d} = 1$, imply that if $\boldsymbol{\beta}_{m,1} = 1$, then $\boldsymbol{\beta}_{m,2} = 0$, $\mathbf{X}_{k,d,m,1} = \psi_{k,d}$, and $\mathbf{X}_{k,d,m,2} = 0$. However, if $\boldsymbol{\beta}_{m,1} = 0$, then $\boldsymbol{\beta}_{m,2} = 1$, $\mathbf{X}_{k,d,m,1} = 0$, and $\mathbf{X}_{k,d,m,2} = \psi_{k,d}$. Therefore, the following constraints are included:

$$\mathbf{X}_{k,d,m,u} \geq 0, \tag{3.13}$$

$$\mathbf{X}_{k,d,m,u} \geq \boldsymbol{\beta}_{m,u} + \psi_{k,d} - 1, \tag{3.14}$$

$\forall m \in \mathcal{M}$, $\forall k \in \mathcal{K}$, $\forall d, u \in \{1, 2\}$. Furthermore, note that for every $\mathbf{X}_{k,d,m,u}$, if $\mathbf{X}_{k,1,m,u} = \psi_{k,1}$, then $\mathbf{X}_{k,2,m,u} = 0$, and vice-versa. Therefore, the following constraints are included:

$$\min\{\psi_{k,1}, \psi_{k,2}\} \leq \sum_{d \in \{1,2\}} \mathbf{X}_{k,d,m,u}, \tag{3.15}$$

$$\sum_{d \in \{1,2\}} \mathbf{X}_{k,d,m,u} \leq \max\{\psi_{k,1}, \psi_{k,2}\}. \tag{3.16}$$

The above constraints (3.15) and (3.16) are non-convex. However, they can be replaced with the following inequality constraints after introducing the auxiliary variables

$\mathbf{z} \in \mathbb{R}_{+}^{MK}$ and $\mathbf{w} \in \mathbb{R}_{+}^{MK}$, where constraints (3.15) and (3.16) can be written respectively as:

$$\begin{cases} w_k \geq \psi_{k,1}, \\ w_k \geq \psi_{k,2}, \\ \sum\limits_{d\in\{1,2\}} \mathbf{X}_{k,d,m,u} \geq w_k - (\psi_{k,1} + \psi_{k,2}). \end{cases} \tag{3.17}$$

$$\begin{cases} z_k \leq \psi_{k,1}, \\ z_k \leq \psi_{k,2}, \\ \sum\limits_{d\in\{1,2\}} \mathbf{X}_{k,d,m,u} \leq (\psi_{k,1} + \psi_{k,2}) - z_k, \end{cases} \tag{3.18}$$

Using the above constraints, the functions (3.11) and (3.12) can be written equivalently in a linear form as

$$\sum_{m\in\mathcal{M}\setminus n} \mathbf{X}_{k,1,m,1} \sum_{j\in\mathcal{K}_m} \alpha_{j,1}^{-1} p_j g_{m,k} + \sum_{m\in\mathcal{M}\setminus n} \mathbf{X}_{k,1,m,2} \sum_{j\in\mathcal{K}_m} \alpha_{j,2}^{-1} q_j \tilde{g}_{j,k} + \psi_{k,1}\sigma_k^2 = \boldsymbol{\beta}_{n,1}. \tag{3.19}$$

$$\sum_{m\in\mathcal{M}\setminus n} \mathbf{X}_{k,2,m,1} \sum_{j\in\mathcal{K}_m} \alpha_{j,1}^{-1} p_j \hat{g}_{m,n} + \sum_{m\in\mathcal{M}\setminus n} \mathbf{X}_{k,2,m,2} \sum_{j\in\mathcal{K}_m} \alpha_{j,2}^{-1} q_j g_{n,j} + \psi_{k,2}\sigma_k^2 = \boldsymbol{\beta}_{n,2}. \tag{3.20}$$

Finally, the optimization problem $\mathscr{P}_2$ can be written in a linear form as

$$\mathscr{P}_3 = \begin{cases} \max\limits_{\boldsymbol{\psi},\boldsymbol{\beta},\mathbf{X},\mathbf{z},\mathbf{w}} & \sum\limits_{n\in\mathcal{M}} \sum\limits_{k\in\mathcal{K}_n} \left(\psi_{k,1}\alpha_{k,1}p_k g_{n,k} + \psi_{k,2}\alpha_{k,2}q_k g_{n,k}\right) \\ \text{s.t.} & \sum\limits_{d\in\{1,2\}} \boldsymbol{\beta}_{n,d} = 1, \forall n, \text{(3.13), (3.14), (3.17)} - \text{(3.20)} \end{cases} \tag{3.21}$$

Problem $\mathscr{P}_3$ is an ILP problem, for which there exists a multitude of powerful algorithms and solvers (such as CPLEX, GUROBI, and MOSEK), which deal with large-scale problems (involving hundreds or even thousands of variables/constraints), in very fast time-scales, and provide solutions with global optimality guarantees. Note that we did not use any approximation to transform the non-linear problem $\mathscr{P}_2$ to a linear one, problem $\mathscr{P}_3$, but the cost that we pay for this transformation is an increase in the number of variables. For an arbitrary number of integer variables v, the number of linear programming subproblems to be solved is at least $(\sqrt{2})^v$ [63]. Meanwhile, the number of iterations needed to solve one linear programming problem with c constraints and v variables is approximately $2(v+c)$, and each iteration encompasses $(vc-c)$ multiplications, $(vc-c)$ summations, and $(v-c)$ comparisons [63, 64]. Thus, the required total number of operations is $(\sqrt{2})^v[2(v+c)(2vc+v-3c)]$. As in problem $\mathscr{P}_3$ there are $v = 2M$ integer variables and $c = 10MK + 2M$ constraints, its complexity can be approximated as $O\left((\sqrt{2})^{2M} M^3 K^2\right)$.

3.6 Centralized cell reconfiguration - method 2

This section proposes an alternative solution method to solve problem $\mathscr{P}_2$ based on the PSO algorithm [32]. PSO is a meta-heuristic global optimization method, which belongs

to the family of algorithms that are based on the concept of swarm intelligence [65]. It was developed originally by Kennedy and Eberhart [32] in analogy to the behavior of bird flocks and fish schools. Due to its metaheuristic nature, which allows obtaining solutions also for non-differentiable problems which may be irregular, noisy or dynamically changing with time, PSO algorithm has found a wide range of application in many domains of computer science and applied mathematics, such as for the calculation of neural network weights [66, 67], time series analysis [68], business optimization [69] and many others.

3.6.1 PSO algorithm review

The PSO algorithm process is as follows: Given an optimization (objective) function $f(\mathbf{x})$, our task is to optimize (maximize or minimize) this function by finding the variable vector $\mathbf{x}$ that is in D-dimension, where D is the number of elements in $\mathbf{x}$. At first, the algorithm generates and initializes a swarm comprising S particles, where each particle i, $i \in \{1, \ldots, S\}$ has its own position $\mathbf{x}_i$ and direction $\mathbf{v}_i$ vectors, both are in D-dimension as well. Here, each particle position $\mathbf{x}_i$ represents a possible solution to the optimization function $f(\mathbf{x})$. Then, the algorithm searches the optimal solution by iteratively evolving the particles. In each iteration, each particle i position, $\mathbf{x}_i$, and direction, $\mathbf{v}_i$, vectors are updated considering two extreme values. The first one is the personal best position $\bar{\mathbf{x}}_i$, which corresponds to the personal best cost function $f_i(\bar{\mathbf{x}}_i)$, and the other one is the global best position $\bar{\mathbf{x}}$ found by the whole swarm, which corresponds to the global best cost function $f(\bar{\mathbf{x}})$ [70]. The direction and position vectors are updated in the PSO algorithm using the following equations [71]:

$$\mathbf{v}_i^{(t+1)} = w\mathbf{v}_i^{(t)} + c_1\mathbf{r}_1^{(t)} \odot (\bar{\mathbf{x}}_i - \mathbf{x}_i^{(t)}) + c_2\mathbf{r}_2^{(t)} \odot (\bar{\mathbf{x}} - \mathbf{x}_i^{(t)}), \tag{3.22}$$

$$\mathbf{x}_i^{(t+1)} = \mathbf{x}_i^{(t)} + \mathbf{v}_i^{(t+1)}, \tag{3.23}$$

where $\odot$ denotes the dot-product operator. In function (3.22), $\mathbf{v}_i^{(t)}$ and $\mathbf{x}_i^{(t)}$ are the direction and position vectors for particle i in the t-th iteration, respectively. The direction vectors in function (3.22) govern the way particles move across the search space and are made of the contribution of three terms: the first one, defined the inertia or momentum, prevents the particle from drastically changing direction by keeping track of the previous flow direction; the second term, called the cognitive component, accounts for the tendency of particles to return to their own previously found best positions; the last one, named the social component, identifies the propensity of a particle to move towards the best position of the whole swarm. The parameter w is called the inertia weight, whose role is to balance the global and local search. The parameters c_1 and c_2 are called cognition learning and social learning rates, respectively, which are constants and they respectively regulate the maximal step size towards the personal best particle and the global best particle. In addition, $\mathbf{r}_1^{(t)}$ and $\mathbf{r}_2^{(t)}$ are random real number vectors drawn from uniform distribution in [0,1] in each iteration. Accordingly, the trajectories drawn by the particles are semi-random in nature, as they derive from the contribution of systematic attraction towards the personal and

global best solutions and stochastic weighting of these two acceleration terms. To improve the readability of this section, the defined variables above are summarized in Table 3.1.

Table 3.1 – PSO variables

Parameter	Meaning
$f(\mathbf{x})$	fitness function
S	number of swarm particles
$\mathbf{x}_i$	position vector of particle i
$\mathbf{v}_i$	direction vector of particle i
$\bar{\mathbf{x}}_i$	personal best position vector of particle i
$f(\bar{\mathbf{x}}_i)$	personal best cost function of particle i
$\bar{\mathbf{x}}$	global best position vector of all particles
$f(\bar{\mathbf{x}})$	global best cost function of all particles
w, c_1, c_2	constant real-valued numbers
$\mathbf{r}_1$, $\mathbf{r}_2$	random vectors drawn uniformly from distribution [0,1]

Source: Created by author.

However, for some optimization problems, the optimization vector $\mathbf{x}$ has only integer values in [0, 1]. For that, Kennedy and Eberhart proposed a reworking of the PSO algorithm in 1997 called binary particle swarm optimization (BPSO) [72]. In the BPSO algorithm, function (3.22) remains the same, except that all position vectors are integers in [0, 1]. On the other hand, the position vectors are updated using the following rule:

$$\begin{aligned} if \quad & \vartheta_{i[j]}^{(t+1)} < \mathrm{Sig}(\mathbf{v}_{i[j]}^{(t+1)}) \\ then \quad & \mathbf{x}_{i[j]}^{(t+1)} = 1 \\ else \quad & \mathbf{x}_{i[j]}^{(t+1)} = 0, \end{aligned} \tag{3.24}$$

where $\mathbf{v}_{i[j]}^{(t+1)}$ and $\mathbf{x}_{i[j]}^{(t+1)}$ represent the j-th element of particle i direction and position vectors, respectively, $\vartheta_{i[j]}^{(t+1)}$ is a quasi-random number selected from a uniform distribution in [0,1], and $\mathrm{Sig}(\mathbf{v}_{i[j]}^{(t+1)})$ is a Sigmoid limiting transformation function defined as

$$\mathrm{Sig}(\mathbf{v}_{i[j]}^{(t+1)}) = \frac{1}{1+\exp(-\mathbf{v}_{i[j]}^{(t+1)})}. \tag{3.25}$$

A schematic description of the basic PSO algorithm is summarized in Algorithm 4. The iterative process described in Algorithm 4 is repeated until a stopping criterion is met. This could be, e.g., a pre-specified total number of iterations, a maximum number of iterations since the last update of global best or a predefined target value of the fitness. Here it must be stressed that, in PSO, as in all stochastic evolutionary algorithms, the term convergence may refer to two different scenarios: convergence as the limit of a series of solution (where, for instance, all or most of the particles, reach the same point in the search space, which may not necessarily be the optimum), indicating the algorithm stability, and convergence to a local (or the global) optimum of the problem, which may be achieved by one or more particles (through personal bests or global best), irrespective of the overall behavior of the swarm.

Algorithm 4: Classical PSO algorithm

1: **For each of the S particles:**
2: a) Initialize the direction and position vectors $\mathbf{v}_i^{(0)}$, $\mathbf{x}_i^{(0)}$, $\forall i \in \{1, \ldots, S\}$.
3: b) Initialize the particle's best position to its initial position $\bar{\mathbf{x}}_i = \mathbf{x}_i^{(0)}$.
4: c) Calculate the fitness of each particle $f(\mathbf{x}_j^{(0)})$.
5: d) Initialize the global best as $\bar{\mathbf{x}} = \mathbf{x}_j^{(0)}$, where $f(\mathbf{x}_j^{(0)}) \geq f(\mathbf{x}_i^{(0)})$, $\forall i \neq j$.
6: **Repeat the following steps until a stopping criterion is met:**
7: a) Update particle direction using (3.22).
8: b) Update particle position using (3.23)/(3.24) for binary/non-binary variables.
9: c) Evaluate the fitness of particle $f(\mathbf{x}_i^{(t+1)})$.
10: d) If $f(\mathbf{x}_i^{(t+1)}) \geq f(\bar{\mathbf{x}}_i)$, update personal best as $\bar{\mathbf{x}}_i = \mathbf{x}_i^{(t+1)}$.
11: e) If $f(\mathbf{x}_i^{(t+1)}) \geq f(\bar{\mathbf{x}})$, update global best as $\bar{\mathbf{x}} = \mathbf{x}_i^{(t+1)}$.
12: **At end of the iterative process, the best solution is given by $\bar{\mathbf{x}}$.**

In the following we discus how to use Algorithm 4 to solve our optimization problem that is given by $\mathscr{P}_2$. Note that we can consider the un-relaxed (original) optimization problem given by $\mathscr{P}_1$ to obtain the cells' directions (thus giving the following proposed solution one more advantage). However, to keep consistency with the first proposed solution, ILP method, problem $\mathscr{P}_2$ is considered.

3.6.2 Fitness function and coding rule

The fitness function is used as the performance evaluation of particles in the swarm. The fitness function is defined to be equal to the objective function of problem $\mathscr{P}_2$, i.e.,

$$f(\beta) = \sum_{n \in \mathcal{M}} \sum_{k \in \mathcal{K}_n} \sum_{d \in \{1,2\}} \boldsymbol{\beta}_{n,d} \gamma_{k,d}, \tag{3.26}$$

where the main objective is to find β such that $f(\beta)$ is maximized. In Algorithm 4, the dimension of particles' position and direction vectors are set in the M-dimension, i.e., $\mathbf{x}_i \in \mathbb{Z}^M$ and $\mathbf{v}_i \in \mathbb{R}^M$, $\forall i \in \{1, \ldots, S\}$. However, as β in the fitness function (3.26) is a matrix, i.e., $\beta \in \mathbb{Z}^{M \times 2}$, the following transformation rule is considered to achieve $\mathbf{x} \in \mathbb{Z}^M \rightarrow \beta \in \mathbb{Z}^{M \times 2}$, where

$$\begin{aligned} \textit{if} \quad & \mathbf{x}_{i[j]} = 1 \\ \textit{then} \quad & \beta_{i,1} = 1 \text{ and } \beta_{i,2} = 0 \\ \textit{else} \quad & \beta_{i,1} = 0 \text{ and } \beta_{i,2} = 1. \end{aligned} \tag{3.27}$$

3.6.3 Position and direction initialization

As shown in Algorithm 4, being an iterative algorithm, PSO requires an initial estimate of the particles' positions and directions. The choice of the way these two entities are initialized plays a determinant role in defining what is the probability that particles travel outside the boundaries of the search space and, as a consequence, in affecting the convergence properties of the solution. In particular, there is a general agreement in the literature – and it is also the

author's opinion – that initializing the particles' positions so that they cover as uniformly as possible the search space is the best option [71].

In the simulator, the particles positions $\mathbf{x}_i$, $\forall i \in \{1, \ldots, S\}$ are initialized as follows. Let C denote the set containing the 2^M possible solution vectors, i.e., the number of possible solutions using exhaustive search method. For example, if $M = 2$, then the set C is defined as

$$C = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}, \tag{3.28}$$

where each column vector of C represents a possible solution. If the number of particles $S \leq 2^M$, which is the general case, the position vectors $\mathbf{x}_i$, $\forall i \in \{1, \ldots, S\}$, are initialized with subset of C, while making sure that no two particles have the same initial vector/solution. Otherwise, if $S \geq 2^M$, one or more particles can share the same initial vector/solution. On the other hand, the direction vectors $\mathbf{v}_i$, $\forall i \in \{1, \ldots, S\}$, are initialized randomly from the uniform distribution [0,1].

3.6.4 Choice of the acceleration constant parameters

As evident from function (3.22), the values of the acceleration constants w, c_1 and c_2 govern the extent to which the particles move towards the individual and global best particle, modulating the relative contributions of the social and cognitive terms. Different authors (see, e.g. [65] or [73]) have investigated the effect of these coefficients on the particles' trajectories and on the convergence properties of the algorithm, showing that as the acceleration constants are increased, the frequency of oscillation of the particle around the optimum increases while smaller values result in sinusoidal patterns. In the simulator, the constants are chosen as [74]:

$$\begin{cases} w = \xi, \\ c_1 = \xi\phi_1, \\ c_2 = \xi\phi_2, \\ \xi = \frac{2\kappa}{|2-\phi-\sqrt{\phi^2-4\phi}|}, \text{ where } \phi = \phi_1 + \phi_2 \text{ and } \kappa \text{ is a constant number.} \end{cases} \tag{3.29}$$

It was shown in [74] that if we set $\kappa = 1$ and $\phi_1 = \phi_2 = 2.05$, the PSO algorithm is guaranteed to converge, but not necessarily to the global optimal.

3.6.5 Complexity analysis

Algorithm 4 complexity can be measured by the number of mathematical operations required by each particle in each iteration. Note that the computational complexity of Algorithm 4 is dominated by the calculation of the fitness function (3.26). To calculate (3.26), each particle requires $2MK(6K(M-1)+2)$ multiplications, $2MK(2K(M-1))$ summations, and $2MK$ divisions, all of scalar numbers. Thus, the complexity of Algorithm 4 is given as

$$2MKSt_{max}[(6K(M-1)+2)+(2K(M-1))+1] \approx O(ST(M^2K^2)), \tag{3.30}$$

where t_{max} denotes the total number of iterations performed by Algorithm 4. Note that, as compared to the first method, which has complexity of $O\left((\sqrt{2})^{2M}M^3K^2\right)$, the second method using the PSO Algorithm 4 has much lower complexity, especially with small numbers of particles S and iterations t_{max}. However, as Algorithm 4 may not achieve the optimal solution as in the first method, it can be seen as a trade-off solution between the complexity and system performance.

3.7 Numerical results

This section evaluates the proposed cell reconfiguration scheme using a system-level simulator. In particular, the time-correlated one-ring channel model from [75] is used, which has frequency-nonselective Rayleigh distribution. It is assumed that each BS has $N_t = 2$ antennas and each MS has a single antenna. The channel gains at each TTI are given as $g_{n,k} = \rho_{n,k}\|\mathbf{h}_{n,k}\|^2$, $\hat{g}_{n,m} = \rho_{n,m}\|\mathbf{H}_{n,m}\|_F^2$, and $\tilde{g}_{j,k} = \rho_{j,k}|h_{j,k}|^2$, where $\rho_{n,k}$, $\rho_{n,m}$ and $\rho_{j,k}$ are path-loss parameters implemented as proposed by the 3GPP TR [15, Table 6.3-1,p.59], which are shown in Table 3.2. Accordingly, the channel gains follow exponential distribution. Each cell has a radius of 50 meters, while users are uniformly distributed within the serving area, and the minimum distance between a MS and its serving BS is 10 meters. In the simulation, the channel gains are normalized with noise power σ_k^2 to always satisfy $\psi_{kd} < 1$ in (3.8).

Table 3.2 – Pathloss model

Communication Link	Pathloss Model
BS to BS	$98.4 + 20\log_{10}(d_{n,m})$, where $d_{n,m}$ in kilo meters
BS to MS	$103.8 + 20.9\log_{10}(d_{n,k})$, where $d_{n,k}$ in kilo meters
MS to MS	$55.78 + 40\log_{10}(d_{j,k})$, where $d_{j,k}$ in meters

Source: Created by author.

The traffic model follows the gaming-traffic model from [22], where the downlink and uplink packet sizes and the downlink packet arrival values are generated as $x = \lfloor a - b\ln(-\ln y)\rfloor$, in which $\lfloor\cdot\rfloor$ represents the floor operator, y is drawn uniformly from the range [0, 1], while a and b are given in Table 3.3. In addition, we add an extra 2 Bytes to each generated packet to account for the header size.

Table 3.3 – Gaming network traffic parameters

Parameter	Downlink	Uplink
Initial packet	1st TTI	1st TTI
Packet size	a = 120 Bytes, b = 36	a = 55 Bytes, b = 5.7
Packet arrival	a = 55 ms, b = 6	every 40 ms

Source: Created by author.

Let $\Upsilon_{k,1}$ and $\Upsilon_{k,2}$ denote the user k downlink and uplink buffer sizes. At the t-th TTI, the users' weights $\alpha_{k,1}$ and $\alpha_{k,2}$ are updated as

$$\alpha_{k,d} = 1 + \frac{\Upsilon_{k,d}}{\Upsilon_k}, \tag{3.31}$$

where $\Upsilon_k = \Upsilon_{k,1} + \Upsilon_{k,2}$, which assigns larger weight for the direction that has larger buffer size and the addition of 1 is used to guarantee that we always have $\alpha_{kd} \geq 1$, as we have assumed above. At the end of the t-th TTI, assuming the user k has been scheduled to transmit on the d direction, the corresponding buffer size is updated as $\Upsilon_{k,d} = \Upsilon_{k,d} - r^t_{k,d}$, where $r^t_{k,d}$ is the transmission rate of user k at the t-th TTI, that is determined by the modulation order and code rate as given in [76, Table I]. Further, we assume that the downlink and uplink transmit power $p_k = q_k = 23$ dBm, noise variance $\sigma^2_k = -100$ dBm, and system bandwidth 180 kHz. We use PT as the performance metric, which is defined as the ratio of successfully transmitted bits over the required transmission time.

For comparison purposes, simulation results for the following algorithms are also shown:

1. STDD algorithm, where all cells transmit on the same direction and switch the communication direction every t_{switch} TTIs. In our simulator, we assume $t_{switch} = 10$ TTIs.

2. Distributed cell reconfiguration algorithm, where it simply neglect the OCI parameters in problem $\mathscr{P}_3$, thus decoupling the optimization problem between the M cells. Therefore, the cell reconfiguration problem of the n-th cell can be written as

$$\mathscr{P}_4 = \begin{cases} \max\limits_{\boldsymbol{\beta}_n} & \sum\limits_{k\in\mathcal{K}_n} \left(\boldsymbol{\beta}_{n,1}\alpha_{k,1}p_k g_{n,k} + \boldsymbol{\beta}_{n,2}\alpha_{k,2}q_k g_{n,k}\right) \\ \text{s.t.} & \sum\limits_{d\in\{1,2\}} \boldsymbol{\beta}_{n,d} = 1, \end{cases} \tag{3.32}$$

for which the optimal solution is given as

$$\boldsymbol{\beta}_n = u_n\Big(\sum_{k\in\mathcal{K}_n} \alpha_{k,1}p_k g_{n,k}, \sum_{k\in\mathcal{K}_n} \alpha_{k,2}q_k g_{n,k}\Big), \tag{3.33}$$

where $u_n(A, B)$ is a function defined as

$$u_n(A, B) = \begin{cases} \beta_{n,1} = 1, \beta_{n,2} = 0, & \text{if } A \geq B, \\ \beta_{n,1} = 0, \beta_{n,2} = 1, & \text{if } A < B. \end{cases} \tag{3.34}$$

3. Traffic Only cell reconfiguration algorithm, where the n-th cell chooses its communication direction as

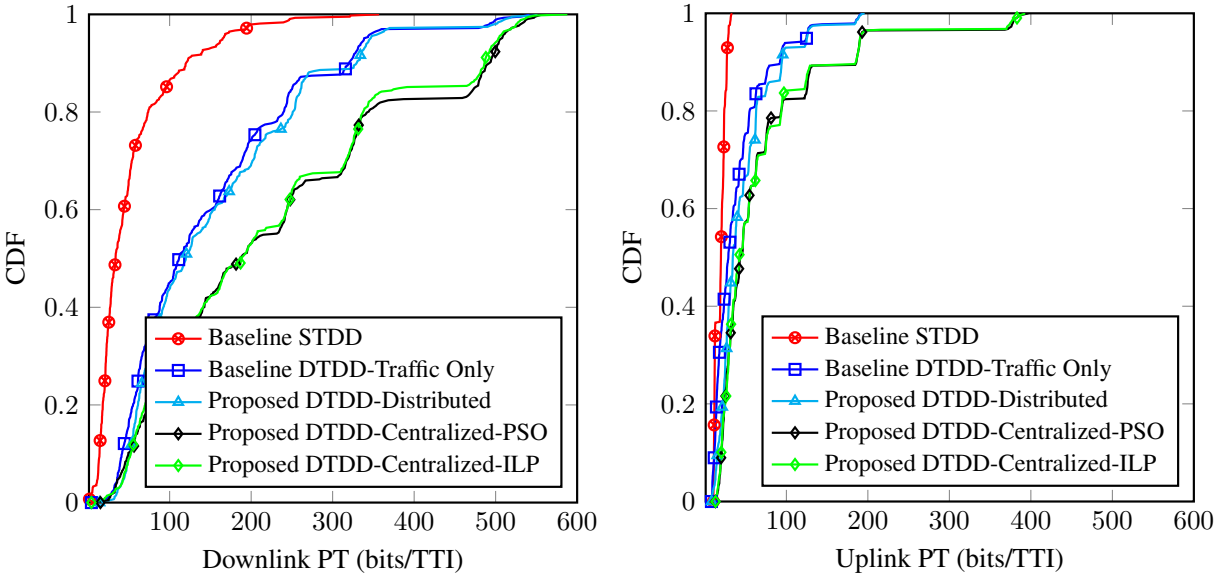
$$\boldsymbol{\beta}_n = u_n\Big(\sum_{k\in\mathcal{K}_n} \Upsilon_{k,1}, \sum_{k\in\mathcal{K}_n} \Upsilon_{k,2}\Big). \tag{3.35}$$

Example 1: Users' UL and DL packets throughput

Fig. 3.1 shows the cumulative distribution function (CDF) plots of users' downlink and uplink PT, while assuming $M = 5$, $K = 5$, and $S = 3$. Fig. 3.1 indicates that the DTDD cell reconfiguration methods have higher spectral efficiency than the STDD method, whereas the

proposed DTDD cell reconfiguration methods (centralized and distributed) have higher spectral efficiency than the baseline Traffic Only cell reconfiguration method. More specifically, the proposed DTDD Centralized-ILP (distributed) cell reconfiguration method achieves 4.2 (2.8) times, as compared to the STDD method, and 1.6 (1.1) times, as compared to the baseline Traffic Only method. The underlying reason for this improved performance is that the proposed method jointly considers the users' traffic demands and the interference levels when choosing the cells' directions. In contrast, the Traffic Only method considers only the traffic demands and neglects the interference effects. Note that the centralized method clearly has better performance than the distributed method as it considers the OCI, while the distributed method considers only the intra-cell interference. Furthermore, we can see that the heuristic centralized method based on the PSO algorithm, Proposed DTDD-Centralized-PSO, has very close performance, around 98%, to that achieved by the centralized method based on the ILP technique, Proposed DTDD-Centralized-ILP.

Figure 3.1 – CDF plots of users' packets throughput.



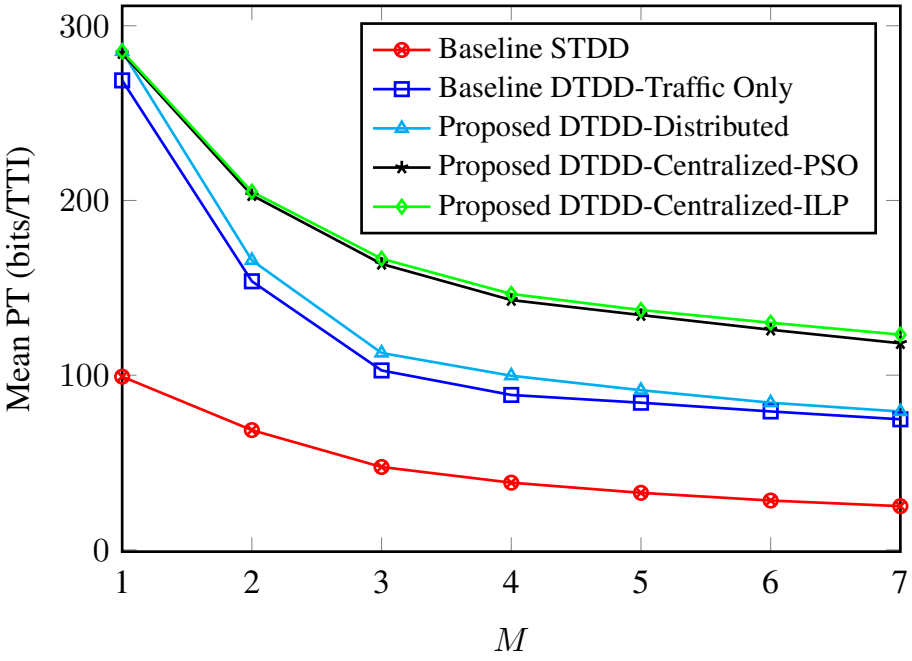
Source: Created by author.

Example 2: Users' mean packets throughput

Figs. 3.2 and 3.3 show the users' mean PT while varying the number of cells M and the number of users K, respectively, assuming $S = 3$. In Fig. 3.2, it is assumed that $K = 5$, while in Fig. 3.3, it is assumed that $M = 5$. From Figs. 3.2 and 3.3, we can see that the proposed cell-reconfiguration algorithm (centralized and distributed) achieves higher performance than the reference algorithms, Traffic Only and STDD. Note that the proposed distributed and centralized methods are equal whenever $M = 1$, while the proposed distributed and Traffic Only methods are equal whenever $K = 1$. The Proposed DTDD-Centralized-PSO has negligible performance loss, as compared to the Proposed DTDD-Centralized-ILP, with all simulation scenarios. Further, all algorithms have degraded performance as M and/or K increases. However, the proposed scheme

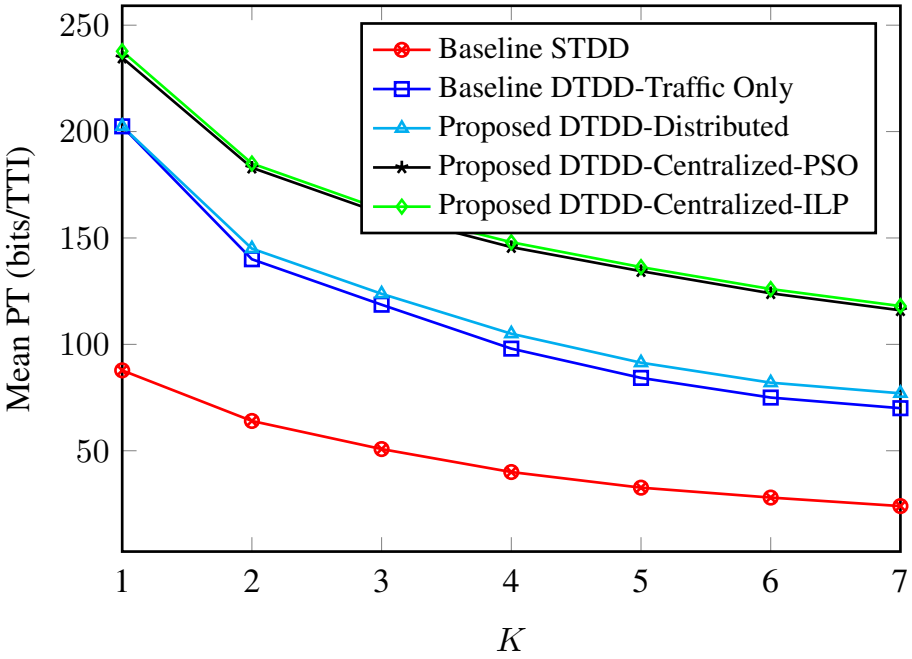
maintains its gain over the reference algorithms.

Figure 3.2 – Users' mean PT while varying the number of cells.



Source: Created by author.

Figure 3.3 – Users' mean PT while varying the number of local users.



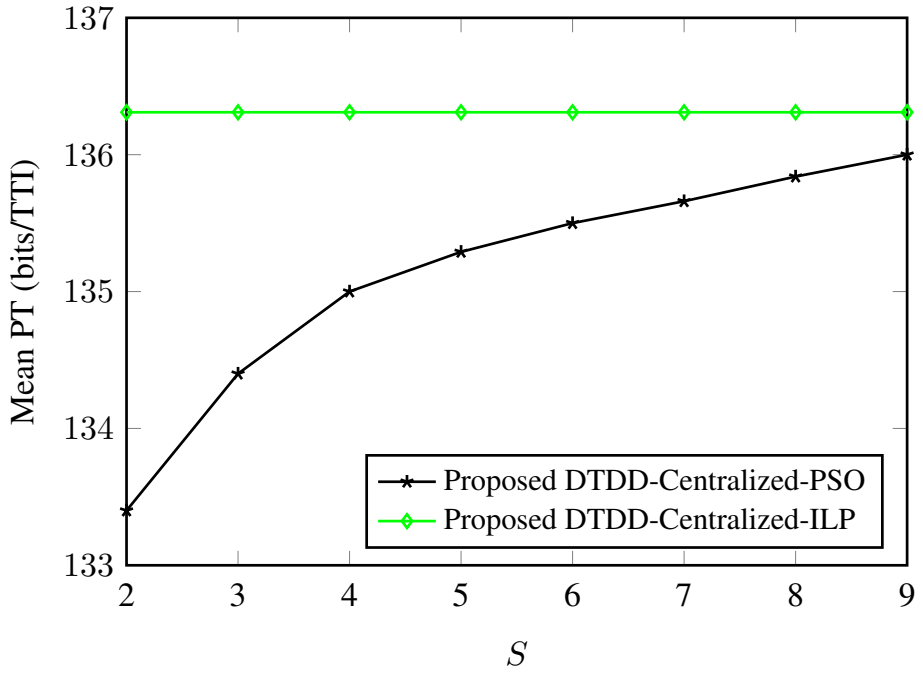
Source: Created by author.

Example 3: Impact of swarm population on system performance

Next, the impact of number of swarm particles S on system performance is investigated. Fig. 3.4 shows the users' mean PT while varying the number of swarm particles S, while

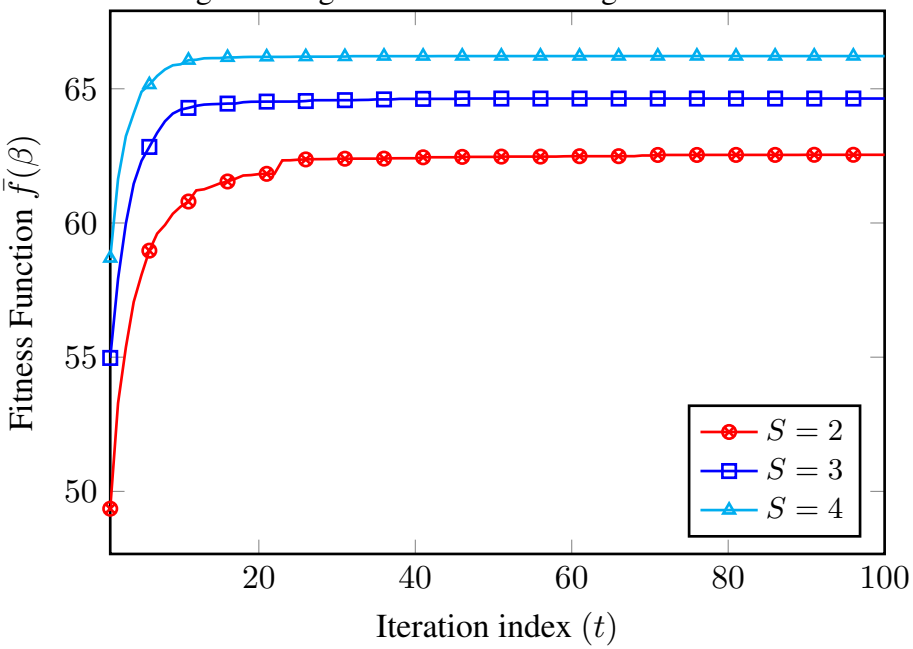
in Fig. 3.5, we show the average convergence behavior of the global best fitness function $\bar{f}(\beta)$. For both Figs. 3.4 and 3.5, we assume $M = 5$ and $K = 5$. From Fig. 3.4, we can see that increasing the number of swarm particles S, the performance of Proposed DTDD-Centralized-PSO gets closer to that achieved by Proposed DTDD-Centralized-ILP. Another important note is that increasing number of particles S, the PSO algorithm achieves a faster convergence rate and higher performance, as can be observed in Fig. 3.5. Here, increasing the number of particles S increases the diversity of the swarm and its exploration ability. However, increasing S will lead to more computational complexity. Fortunately, we can see from Fig. 3.4 that even when $S = 2$, the Proposed DTDD-Centralized-PSO achieves approximately 98% of the performance achieved by the Proposed DTDD-Centralized-ILP.

Figure 3.4 – Users' mean PT while varying the number of swarm particles.



Source: Created by author.

Figure 3.5 – Average convergence behavior of the global best fitness function.



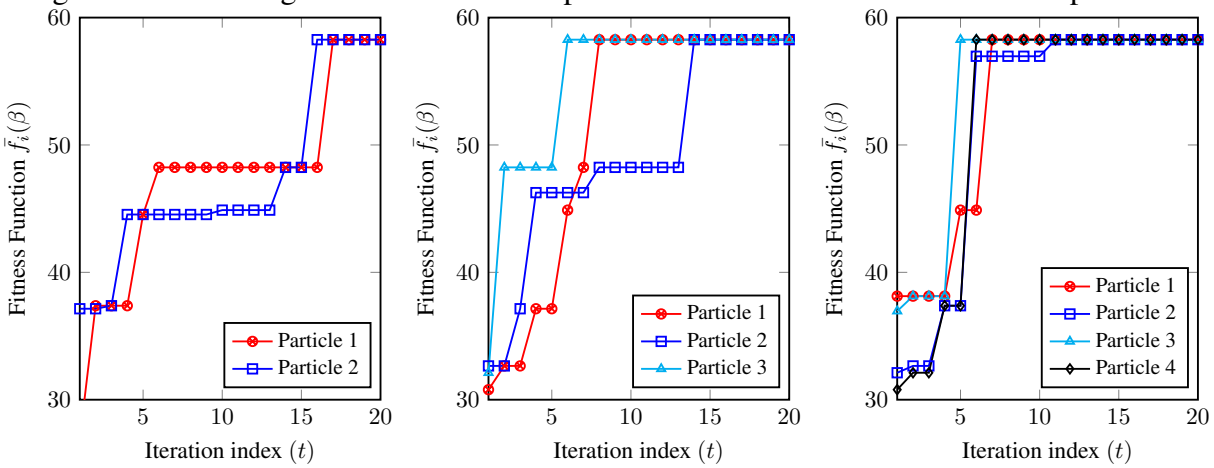
Source: Created by author.

Example 4: Proposed PSO algorithm convergence

Finally, in Fig. 3.6, we show the convergence behavior of the personal best fitness function $\bar{f}_i(\beta), i \in \{1, \ldots, S\}$ with one channel realization, assuming $M = 5, K = 5$, and $S \in \{2, 3, 4\}$. For this particular channel realization, the PSO algorithm converges to the same solution vector with all considered scenarios. However, we can see that when $S = 2$, left figure, the best solution vector is achieved by Particle 2 after 15 iterations. Note that, in this case, Particle 1 also converged to the same solution vector in the next iteration. This behavior holds true as well when $S = 3$ and $S = 4$, although with a lower number of iterations.

Nevertheless, we have noticed that for some channel realizations, the PSO algorithm has faster convergence rate and higher performance with smaller S than with larger S. The reason behind this is that it might happen that with smaller S, one or more particles is actually initialized with the optimal solution vector. Thus, the algorithm would have the best solution from the very first iteration. However, as S increases, the probability that at least one of the particles is initialized with the optimal solution vector increases as well. Therefore, on average, with larger S, the PSO algorithm has faster convergence rate and better performance than with smaller S.

Figure 3.6 – Convergence behavior of the personal best fitness function of swarm particles.



Source: Created by author.

3.8 Chapter conclusions

This chapter proposed a novel low-complexity cell reconfiguration scheme for DTDD systems that jointly considers the interference-levels and the user traffic characteristics when determining the cells' directions to maximize the system throughput. The proposed optimization problem is solved optimally using the ILP algorithm. However, due to its high computational complexity, a heuristic solution is proposed based on the PSO algorithm, which is shown to achieve near optimal performance with much lower computational complexity. In particular, we have shown using a system-level evaluations that when $M = 5$ and $K = 5$, the proposed centralized (distributed) scheme improves the users' throughput by 4.2 (2.8) times, as compared to the conventional static TDD, and by 1.6 (1.1) times, as compared to Traffic Only cell reconfiguration schemes. The proposed scheme can be easily applied to a broad range of traffic models and/or deployment scenarios, which makes it suitable for current and future wireless networks.

4 AN ADMM APPROACH FOR DISTRIBUTED ROBUST COORDINATED BEAMFORMING IN DYNAMIC TDD WIRELESS NETWORKS

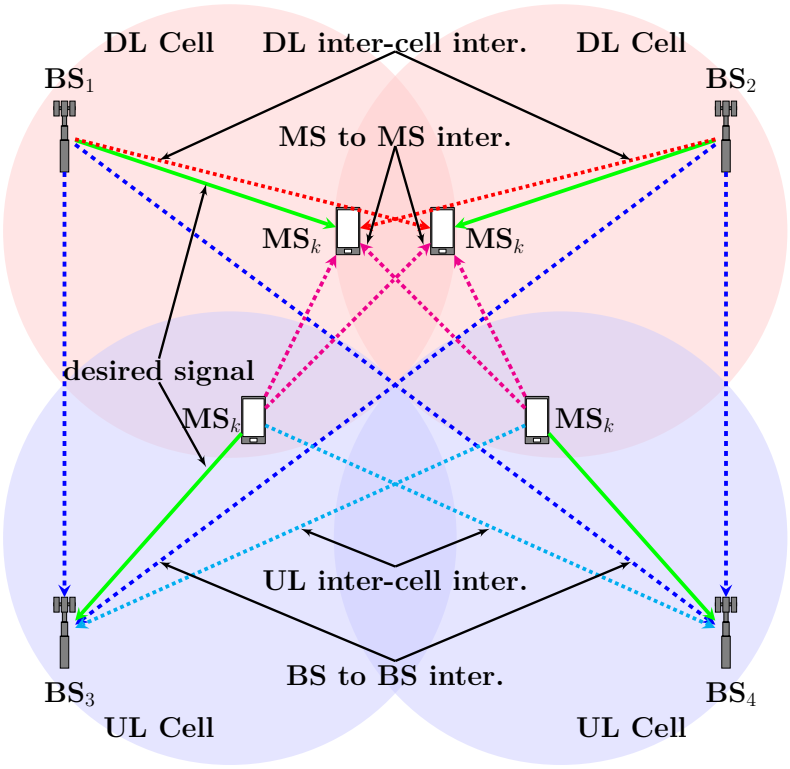
4.1 Introduction

Recently, 3GPP LTE has introduced small cell deployment [5] as one of the key techniques to meet the exponential growth of traffic demands. Small cells using low-power nodes are meant to be deployed in hot-spots, where the number of users vary strongly with time and between adjacent cells. As a result, small cells are expected to have a burst-like traffic with strong fluctuation between uplink and downlink traffics [17]. For that, TDD mode gains more importance for small cell deployments than FDD mode [13], since it can be employed to provide unbalanced uplink-downlink data traffic. However, TDD in LTE is assumed to select a common TDD pattern for the whole network, which cannot be rapidly modified to match the instantaneous traffic demand [14]. As a solution, DTDD technique has been recently introduced [15], where each BS is allowed to dynamically reconfigure its TDD pattern based on its instantaneous traffic demand and/or interference status. In [15, 77], DTDD system performance was evaluated with different performance metrics and found that the DTDD system provides a significant improvement in throughput as compared to the static TDD.

However, cross-link interference arises in DTDD, as illustrated in Fig. 4.1, among other impairments [17], and requires a proper management to realize the DTDD advantages. Earlier studies [21, 78] have often considered interference avoidance schemes. In [21], a cell-clustering scheme was proposed by grouping a number of cells into a cluster according to some metric(s), where cells in the same cluster adopt the same TDD configuration. Authors in [78] proposed dynamic time slot allocation for an adaptive and flexible interference avoidance scheme. Adaptive power control techniques have been used as well, such as in [79], to reduce and compensate the cross-link interference.

In this line, MIMO techniques have been used as well, such as in [80], to spatially suppress the interference. A notable scheme in this area is the CBF technique [81], which has drawn significant attention recently due to its ability to handle the interference problem using only CSI, as compared to other interference management schemes that require data sharing as well, such as joint transmission [25]. In CBF, each BS communicates with its own users, while minimizing the interference leakage to users in other cells. It can be implemented in a centralized manner, where global CSI is made available to a central unit, or in a distributed manner, where each BS uses only local CSI. In the latter case, the coordination can be achieved by means of limited backhaul signaling. However, in practical scenarios, acquiring global CSI would drastically increase the backhaul signaling, thus limiting the centralized approach practical applications. Moreover, the BSs can never have perfect CSI, due to, for example, estimation errors and limited feedback channels [26]. Therefore, robust and distributed CBF solutions are much desired.

Figure 4.1 – Cross-link interference in DTDD wireless networks.



Source: Created by author.

Robust optimization is usually addressed by either a probabilistic or non-probabilistic approach [82]. With the probabilistic approach, the CSI errors are often modeled as Gaussian random variables, where the robustness can then be achieved in a statistical sense [83]. With the non-probabilistic approach, the CSI errors are often modeled such that they fall inside a bounded uncertainty set. The system is then optimized to operate under the CSI worst-case condition [29]. On the other hand, distributed implementation can be achieved by a variety of techniques, for instance, by utilizing the downlink-uplink duality [24], dual and primal decomposition techniques [84], or ADMM technique [30]. In the literature, a wide range of CBF algorithms were recently proposed using different assumptions. For example, centralized and robust CBF algorithms based on worst-case criterion were proposed in [34] for the sum power minimization problem and in [85] for the weighted sum rate maximization problem. Distributed and non-robust CBF algorithms were proposed in [9, 24] for sum-power minimization by utilizing the downlink-uplink duality and in [41] by using the ADMM technique. A distributed and robust CBF algorithm based on worst-case criterion was proposed in [27] by using the ADMM technique for the sum power minimization problem. However, all these algorithms consider static TDD systems. In DTDD systems, the interference situations are more complicated since the uplink and downlink users coexist at the same time among neighboring cells. Therefore, the interference management becomes more challenging and requires a special consideration from the optimization view point.

For instance, a decentralized and non-robust algorithm was proposed in [51] for the weighted sum rate maximization problem in DTDD systems. The problem was formulated similar

to the one proposed in [45] for the static TDD systems, by treating the cross-link interference as an inter-cell interference. Although the algorithm is able to achieve a local-optimal solution, the uplink performance can be highly degraded especially if the downlink and uplink transmit powers highly differ, which is the case in general. In this case, the algorithm naturally disregards the uplink connections (cells). A possible solution is to formulate the optimization problem in DTDD systems as it is generally formulated in the CR networks [28], i.e., by assuming that the uplink cells are the *primary* cells and the downlink cells are the *secondary* cells and then include a threshold on the maximum cross-link interference power from the downlink to uplink cells. In this case, not only the downlink performance can be guaranteed, but also the uplink performance. Considering CR networks, a centralized and robust algorithm based on worst-cast criterion was proposed in [28] for the max-min SINR problem and in [86] for the sum MSE minimization problem. In [35], a distributed and non-robust algorithm was proposed for sum power minimization based on the primal decomposition.

4.2 Chapter contributions

Motivated by the above algorithms, this chapter proposes a novel distributed and robust CBF algorithm using relaxed SDP and ADMM techniques for DTDD wireless networks. The design objective is to minimize the total transmit power of downlink BSs, while satisfying the performance targets of downlink and uplink MSs. More precisely, it is assumed that each downlink MS has a predefined minimum SINR target and each uplink MS has a predefined maximum interference threshold that can it tolerate. At first, a centralized algorithm using relaxed SDP technique is proposed considering the perfect CSI case. To obtain the beamforming solution in a distributed way, a distributed algorithm is then proposed using relaxed SDP and ADMM techniques. Afterwards, both solutions (centralized and distributed) are extended to account for the CSI errors based on worst-case optimization approach [29], where each infinitely nonconvex worst-case constraint is transformed to only one linear matrix inequality (LMI) constraint using the S-Lemma [33]. Using computer simulations, it is shown that the proposed algorithm has a better energy-efficiency than the centralized robust algorithm from [34] and a faster convergence rate than the primal decomposition technique used in [35].

4.3 Chapter organization

This chapter is organized as follows. Section 4.4 presents the system model, while section 4.5 presents the problem formulation. In sections 4.6 and 4.7, the proposed centralized and distributed algorithms assuming perfect CSI are presented, respectively. Then, sections 4.8 and 4.9 extend both distributed and centralized algorithms for robust optimization considering imperfect CSI, respectively. Finally, section 4.10 presents the numerical results and then section 4.11 concludes the chapter.

4.4 System model

Consider a DTDD system model comprising M cells, where in each cell there is one BS equipped with N antennas and K local MSs, each equipped with a single-antenna. The n-th BS is denoted as BS_n and the k-th MS in each cell is denoted as MS_k. Let $\mathcal{M} \overset{\text{def}}{=} \{1, \ldots, M\}$ and $\mathcal{K} \overset{\text{def}}{=} \{\mathcal{K}_1, \ldots, \mathcal{K}_M\}$ denote the sets of all BSs and MSs, respectively, where $\mathcal{K}_n$ denotes the set of MSs associated with BS_n. At each time instant, it is assumed that there is a set of BSs $\mathcal{M}^{dl} \subset \mathcal{M}$ in the downlink direction and a set of BSs $\mathcal{M}^{ul} \subset \mathcal{M}$ in the uplink direction, where $\mathcal{M}^{dl} \bigcap \mathcal{M}^{ul} = \{\emptyset\}$ and $\mathcal{M}^{dl} \cup \mathcal{M}^{ul} = \mathcal{M}$. Similarly, it is assumed that $\mathcal{K}^{dl} \subset \mathcal{K}$ denotes the set of MSs in the downlink direction and $\mathcal{K}^{ul} \subset \mathcal{K}$ denotes the set of MSs in the uplink direction, where $\mathcal{K}^{dl} \bigcap \mathcal{K}^{ul} = \{\emptyset\}$ and $\mathcal{K}^{dl} \cup \mathcal{K}^{ul} = \mathcal{K}$. The received signal at MS_k, $k \in \mathcal{K}_n$, $n \in \mathcal{M}^{dl}$ (at downlink), is given as

$$y_k = \sum_{m \in \mathcal{M}^{dl}} \sum_{j \in \mathcal{K}_m} \mathbf{h}_{m,k}^{\text{H}} \mathbf{t}_j s_j + \sum_{m \in \mathcal{M}^{ul}} \sum_{j \in \mathcal{K}_m} \sqrt{q_j} h_{j,k} s_j + n_k, \tag{4.1}$$

where $\mathbf{h}_{m,k} \in \mathbb{C}^N$ denotes the channel vector from BS_m to MS_k, $h_{j,k} \in \mathbb{C}$ denotes the channel from MS_j to MS_k, $\mathbf{t}_k \in \mathbb{C}^N$ denotes the MS_k transmit beamforming vector, q_j denotes the MS_j uplink transmit power, $s_k \in \mathbb{C}$ denotes the MS_k data symbol, where $\mathbb{E}(|s_k|^2) = 1$, and $n_k \in \mathbb{C}$ denotes the additive white Gaussian noise with zero mean and variance σ_k^2. On the other hand, the received signal at BS_n after combining relative to MS_k, $k \in \mathcal{K}_n$, $n \in \mathcal{M}^{ul}$ (at uplink), is given as

$$y_k = \mathbf{r}_k^{\text{H}} \left(\sum_{m \in \mathcal{M}^{ul}} \sum_{j \in \mathcal{K}_m} \mathbf{h}_{n,j} \sqrt{q_j} s_j + \sum_{m \in \mathcal{M}^{dl}} \sum_{j \in \mathcal{K}_m} \mathbf{H}_{m,n} \mathbf{t}_j s_j + \mathbf{n}_k \right), \tag{4.2}$$

where $\mathbf{r}_k \in \mathbb{C}^N$ denotes the MS_k unit-norm uplink receive beamforming vector, $\mathbf{H}_{m,n} \in \mathbb{C}^{N \times N}$ denotes the channel matrix from BS_m to BS_n, and $\mathbf{n}_k \in \mathbb{C}^N$ denotes the additive white Gaussian noise with zero mean and variance σ_k^2.

4.5 Problem formulation

This chapter considers minimizing the sum power of downlink BSs, while satisfying (i) the downlink MSs minimum SINR targets $\gamma_k, \forall k \in \mathcal{K}^{dl}$, and (ii) the maximum BS to BS interference power thresholds $\omega_k, \forall k \in \mathcal{K}^{ul}$, tolerated by the uplink MSs. Mathematically, the optimization problem can be written as

$$\mathscr{P}_A = \begin{cases} \min\limits_{\{\mathbf{r}_k\}, \{\mathbf{t}_k\}} & \sum\limits_{k \in \mathcal{K}^{dl}} \|\mathbf{t}_k\|^2 \\ \text{s.t.} & \text{A1: } \Gamma_k \geq \gamma_k, \forall k \in \mathcal{K}^{dl}, \\ & \text{A2: } \varpi_j \leq \omega_j, \forall j \in \mathcal{K}^{ul}, \end{cases} \tag{4.3}$$

where Γ_k denotes the SINR of MS_k, $k \in \mathcal{K}_n$, $n \in \mathcal{M}^{dl}$, which is given as

$$\Gamma_k = \frac{|\mathbf{h}_{n,k}^{\text{H}}\mathbf{t}_k|^2}{\underbrace{\sum_{i \in \mathcal{K}_n \setminus k} |\mathbf{h}_{n,k}^{\text{H}}\mathbf{t}_i|^2}_{\text{intra-cell interference}} + \underbrace{\sum_{m \in \mathcal{M}^{dl} \setminus n} \sum_{j \in \mathcal{K}_m} |\mathbf{h}_{m,k}^{\text{H}}\mathbf{t}_j|^2}_{\text{inter-cell interference}} + \varphi_k}, \tag{4.4}$$

in which φ_k denotes the MS to MS interference plus noise power that is given as

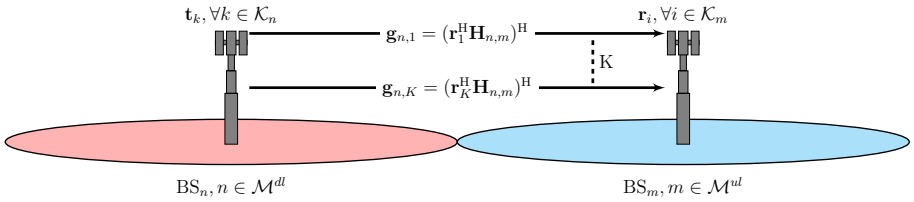
$$\varphi_k = \sum_{j \in \mathcal{K}^{ul}} |h_{j,k}|^2 q_j + \sigma_k^2. \tag{4.5}$$

Furthermore, ϖ_j denotes the BS to BS interference power that affects the MS_j, $j \in \mathcal{K}_m$, $m \in \mathcal{M}^{ul}$, transmission, for which the following two scenarios are considered:

- Scenario 1: In this scenario, denoted hereafter as S1, it is assumed that each BS_n, $n \in \mathcal{M}^{dl}$, knows the *equivalent-channel* vectors relative to MS_j, $\forall j \in \mathcal{K}_m$, $\forall m \in \mathcal{M}^{ul}$, (see Fig. 4.2) which are given as

$$\mathbf{g}_{n,j} = \left(\mathbf{r}_j^{\text{H}}\mathbf{H}_{n,m}\right)^{\text{H}}, \forall j \in \mathcal{K}_m, \forall m \in \mathcal{M}^{ul}. \tag{4.6}$$

Figure 4.2 – Scenario 1.



Source: Created by author.

Thus, the BS to BS interference power relative to MS_j, $j \in \mathcal{K}_m$, $m \in \mathcal{M}^{ul}$, is given as

$$\varpi_j^{\text{S1}} = \sum_{n \in \mathcal{M}^{dl}} \sum_{k \in \mathcal{K}_n} \mathbf{g}_{n,j}^{\text{H}}\mathbf{t}_k\mathbf{t}_k^{\text{H}}\mathbf{g}_{n,j}. \tag{4.7}$$

For this scenario, problem $\mathscr{P}_A$ requires a joint optimization over transmit and receive beamforming vectors, which is non-convex and NP-hard problem [35]. Alternatively, one could use an iterative approach, i.e., for given transmit beamforming vectors, optimize the receive beamforming vectors, and vice-versa. While the transmit beamforming vectors for solving problem $\mathscr{P}_A$ are, in general, not known, the optimal receive beamforming vectors are known to be those maximizing the individual SINRs, which have a closed form-solution given by the minimum variance distortionless response (MVDR) beamforming approach [87]. More precisely, for any given and fixed transmit beamforming vectors, the receive beamforming vector of MS_k, $k \in \mathcal{K}_n$, $n \in \mathcal{M}^{ul}$, can be updated as [87]

$$\mathbf{r}_k = \frac{\Omega_k^{-1}\mathbf{h}_{n,k}q_k}{\|\Omega_k^{-1}\mathbf{h}_{n,k}q_k\|}, \tag{4.8}$$

where Ω_k represents the interference plus noise covariance matrix, which is given as

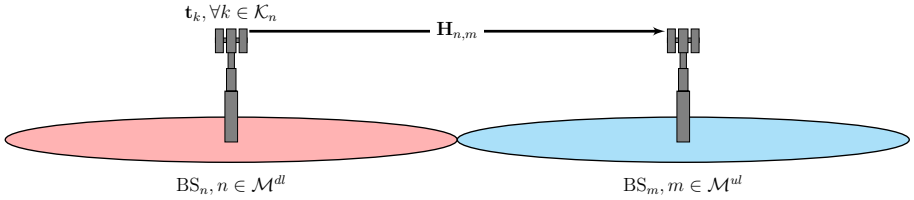
$$\Omega_k = \sum_{j \in \mathcal{K}^{ul} \setminus k} \mathbf{h}_{n,j} \mathbf{h}_{n,j}^{\mathrm{H}} q_j + \sum_{m \in \mathcal{M}^{dl}} \sum_{j \in \mathcal{K}_m} \mathbf{H}_{m,n} \mathbf{t}_j \mathbf{t}_j^{\mathrm{H}} \mathbf{H}_{m,n}^{\mathrm{H}} + \sigma_k^2 \mathbf{I}. \tag{4.9}$$

Before leaving this section, it should be noted that for fixed receive beamforming vectors, problem $\mathscr{P}_A$ is similar to the optimization problem considered recently in [35]. The distributed algorithm in [35] uses the primal decomposition technique, where the problem is decoupled by introducing auxiliary variables that are updated later using the sub-gradient technique. However, the update of the variables problem is unbounded, which often leads to infeasible solutions. Our preliminarily tests show that this undesired situation often happens, especially with large scale problems. To resolve this issue and further extend it for robust beamforming, this chapter addresses problem $\mathscr{P}_A$ and proposes a robust and distributed algorithm based on SDP and ADMM techniques.

- Scenario 2: In this scenario, denoted hereafter as S2, it is assumed that each downlink BS_n, $n \in \mathcal{M}^{dl}$, knows only the BS to BS channel matrices $\mathbf{H}_{n,m}$, $\forall m \in \mathcal{M}^{ul}$, but not the receive beamforming vectors $\mathbf{r}_j$, $\forall j \in \mathcal{K}_m$, $\forall m \in \mathcal{M}^{ul}$ (see Fig. 4.3). Therefore, for this scenario, the BS to BS interference power received at uplink BS_m, $m \in \mathcal{M}^{ul}$, is given as

$$\varpi_m^{\mathrm{S2}} = \sum_{n \in \mathcal{M}^{dl}} \sum_{k \in \mathcal{K}_n} \mathrm{tr}\left[\mathbf{H}_{n,m} \mathbf{t}_k \mathbf{t}_k^{\mathrm{H}} \mathbf{H}_{n,m}^{\mathrm{H}}\right]. \tag{4.10}$$

Figure 4.3 – Scenario 2.



Source: Created by author.

Note that the MSs of uplink cell m, i.e., MS_i, $\forall i \in \mathcal{K}_m^{ul}$, will experiences the same BS to BS interference power received at BS_m before combining. Therefore, to satisfy all BS to BS interference thresholds ω_i, $\forall i \in \mathcal{K}_m$, the downlink BSs should consider the minimum one as $\omega_m = \arg\min_i \{\omega_i, \forall i \in \mathcal{K}_m\}$. Thus, for this scenario, we replace the constraints A2 of problem $\mathscr{P}_A$ by

$$\varpi_m^{\mathrm{S2}} \leq \omega_m, \forall m \in \mathcal{M}^{ul}. \tag{4.11}$$

Note that for some values of ω_m, the BS to BS constraints A2 might be inactive. In this case, removing constraints A2 would not change the obtained solutions (transmit beamforming vectors) and thus both scenarios (S1 and S2) become equivalent. In the rest of this chapter, it is assumed that the BS to BS constraints A2 are active for all or some of the uplink MSs.

4.6 Perfect CSI case: centralized algorithm

This section presents a centralized algorithm to solve problem $\mathscr{P}_A$ that is based on the relaxed-SDP technique (i.e., by dropping the rank-one constraints) [29]. At first, problem $\mathscr{P}_A$ can be readily written in a relaxed-SDP form as

$$\mathscr{P}_B = \begin{cases} \min\limits_{\{\mathbf{T}_k, \forall k \in \mathcal{K}^{dl}\}} & \sum\limits_{n \in \mathcal{M}^{dl}} \mathrm{tr}[\widetilde{\mathbf{T}}_n] \\ \text{s.t.} & \text{B1:} \quad \mathbf{h}_{n,k}^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_{n,k} \geq \mathbf{J}_k + \varphi_k, \ \forall k \in \mathcal{K}^{dl}, \\ & \text{B2:} \begin{cases} \sum\limits_{n \in \mathcal{M}^{dl}} \mathbf{g}_{n,i}^{\mathrm{H}} \widetilde{\mathbf{T}}_n \mathbf{g}_{n,i} \leq \omega_i, \ \forall i \in \mathcal{K}^{ul}, & \text{for S1} \\ \sum\limits_{n \in \mathcal{M}^{dl}} \mathrm{tr}\left[\mathbf{H}_{n,m} \widetilde{\mathbf{T}}_n \mathbf{H}_{n,m}^{\mathrm{H}}\right] \leq \omega_m, \ \forall m \in \mathcal{M}^{ul}, & \text{for S2} \end{cases} \end{cases} \tag{4.12}$$

where $\mathbf{T}_k$, $\widetilde{\mathbf{T}}_m$, $\mathbf{D}_k$ and $\mathbf{J}_k$ matrices are defined as

$$\begin{cases} \mathbf{T}_k &= \mathbf{t}_k \mathbf{t}_k^{\mathrm{H}} \geq 0, \\ \widetilde{\mathbf{T}}_m &= \sum\limits_{i \in \mathcal{K}_m} \mathbf{T}_i \geq 0, \\ \mathbf{D}_k &= \frac{1}{\gamma_k} \mathbf{T}_k - \sum\limits_{j \neq k \in \mathcal{K}_n} \mathbf{T}_j \geq 0, \\ \mathbf{J}_k &= \sum\limits_{m \in \mathcal{M}^{dl} \setminus n} \mathbf{h}_{m,k}^{\mathrm{H}} \widetilde{\mathbf{T}}_m \mathbf{h}_{m,k}. \end{cases} \tag{4.13}$$

Problem $\mathscr{P}_B$ is a convex relaxed SDP problem [29], which can be efficiently solved using the interior point methods [82]. Several convex solvers are freely available, e.g. SeDuMi [88] or CVX [89]. The centralized algorithm for solving problem $\mathscr{P}_B$, with either scenario, for the transmit beamforming vectors is summarized in Algorithm 5. We refer to this algorithm as S1-Cent-CSI for Scenario 1 and as S2-Cent-CSI for Scenario 2.

Algorithm 5: Centralized CBF Algorithm.

1: Initialization: $\mathbf{r}_k(0), \forall k \in \mathcal{K}^{ul}$.
2: Solve problem $\mathscr{P}_B$ for $\{\mathbf{T}_k(r)\}, \forall k \in \mathcal{K}^{dl}$.
3: Update $\mathbf{r}_k(r+1), \forall k \in \mathcal{K}^{ul}$, using (4.8).
4: For S1, set $r = r + 1$ and go back to step (2). For S2, break.

The optimal solution of S1-Cent-CSI, in each iteration, or S2-Cent-CSI, is said to be achieved if each of the obtained matrices $\mathbf{T}_k$, $\forall k \in \mathcal{K}^{dl}$, has rank equal to one [29]. From the optimization theory view point, this result means that the strong duality holds for problem $\mathscr{P}_A$, as it has been shown in [35]. Therefore, the ostensible relaxation of problem $\mathscr{P}_B$ is not a relaxation and it is exactly equivalent to the original problem $\mathscr{P}_A$. In this case, the optimal rank-one solutions can be found using the eigenvalue decomposition [90]. In particular, the rank-one solution of $\mathbf{T}_k$ is given as

$$\mathbf{t}_k = \sqrt{\lambda_k^{max}} \hat{\mathbf{t}}_k, \tag{4.14}$$

where λ_k^{max} is the maximum eigenvalue of $\mathbf{T}_k$ (in this case it is the only non-zero one according to Sylvester's inequality) and $\hat{\mathbf{t}}_k$ is the eigenvector associated with λ_k^{max}. However, there may be situations where the solution is not unique [82]. In this case, there might be one or more of the obtained matrices $\mathbf{T}_k$, $\forall k \in \mathcal{K}^{dl}$, to have a rank greater than one. Although this situation almost never appears in practice [82], the optimal rank-one solutions can still be extracted in polynomial time according to Theorem 3, proved in [91].

Theorem 3 *Suppose that* $\mathbf{X}$ *is a Hermitian positive semidefinite matrix of rank* R*, and* $\mathbf{C}_1$ *and* $\mathbf{C}_2$ *are two given Hermitian matrices. Then there is a rank-one decomposition of* $\mathbf{X}$*, namely,* $\mathbf{X} = \sum_{r=1}^{R} \mathbf{x}_r \mathbf{x}_r^H$ *such that* $\mathbf{x}_r^H \mathbf{C}_1 \mathbf{x}_r = Tr[\mathbf{C}_1\mathbf{X}]/R$ *and* $\mathbf{x}_r^H \mathbf{C}_2 \mathbf{x}_r = Tr[\mathbf{C}_2\mathbf{X}]/R$ *for all* $r = 1, \ldots, R$*. Moreover, such a decomposition can be found in polynomial time.*

Note that, for each set of receive beamforming vectors, S1-Cent-CSI would obtain the optimal transmit beamforming vectors, if the problem is feasible. However, as the receive beamforming vectors are updated in each iteration, the global optimal solution of problem $\mathscr{P}_B$ with S1 can not be guaranteed.

In the rest of this chapter, our effort will be spent on solving problem $\mathscr{P}_A$ for downlink transmit beamforming vectors in a distributed way assuming the uplink receive beamforming vectors are given by (4.8).

4.7 Perfect CSI case: distributed algorithm

In the previous section, it was assumed that there exist a central-unit that has access to the global CSI to jointly optimize the downlink transmit beamforming vectors. This section relaxes this assumption and assumes that each downlink BS has access only to the local CSI (i.e., each BS_n, $\forall n \in \mathcal{M}^{dl}$, knows $\mathbf{h}_{n,i}$, $\forall i \in \mathcal{K}^{dl}$ and $\mathbf{g}_{n,i}$, $\forall i \in \mathcal{K}^{ul}$ for S1 or $\mathbf{H}_{n,m}$, $\forall m \in \mathcal{M}^{ul}$ for S2) and calculates its transmit beamforming vectors with limited signaling with other downlink BSs. The main idea is to decouple the optimization problem between the downlink cells, i.e., downlink BSs. In particular, each downlink BS is coupled with (i) each MS in the other downlink cells by inter-cell interference and (ii) each MS in the uplink cells by BS to BS interference power. Thus, to decouple the optimization problem, each downlink BS will first assume that those coupling terms are fixed, as auxiliary variables, and optimize its MSs transmit beamforming accordingly. After that, the auxiliary variables are gradually updated. These steps are detailed in the next sections.

4.7.1 Problem decoupling

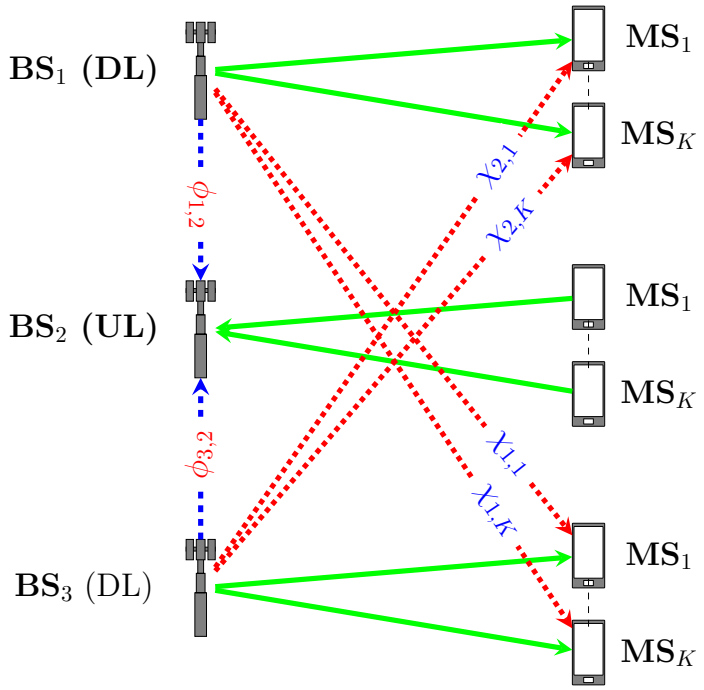
To begin with, let $\chi_{m,k}$ denotes the inter-cell interference power from downlink BS_m to downlink MS_k and $\phi_{m,k}/\phi_{m,n}$ denotes the BS to BS interference power from downlink BS_m to

uplink MS_k/BS_n (see Fig. 4.4), i.e.,

$$\chi_{m,k} \overset{\text{def}}{=} \mathbf{h}_{m,k}^{\mathrm{H}} \widetilde{\mathbf{T}}_m \mathbf{h}_{m,k}, \tag{4.15}$$

$$\begin{cases} \phi_{m,k} \overset{\text{def}}{=} \mathbf{g}_{m,k}^{\mathrm{H}} \widetilde{\mathbf{T}}_m \mathbf{g}_{m,k}, & \text{for S1,} \\ \phi_{m,n} \overset{\text{def}}{=} \mathrm{tr}\left[\mathbf{H}_{m,n} \widetilde{\mathbf{T}}_m \mathbf{H}_{m,n}^{\mathrm{H}}\right], & \text{for S2.} \end{cases} \tag{4.16}$$

Figure 4.4 – DTDD network diagram.



Source: Created by author.

With defined variables above, the SINR function of MS_k, $k \in \mathcal{K}_n$, $n \in \mathcal{M}^{dl}$, can be written as

$$\Gamma_k = \frac{|\mathbf{h}_{n,k}^{\mathrm{H}} \mathbf{t}_k|^2}{\sum\limits_{i \in \mathcal{K}_n \backslash k} |\mathbf{h}_{n,k}^{\mathrm{H}} \mathbf{t}_i|^2 + \sum\limits_{m \in \mathcal{M}^{dl} \backslash n} \chi_{m,k} + \varphi_k}, \tag{4.17}$$

while, the BS to BS interference power for both scenarios can be written as

$$\varpi_i^{\mathrm{S1}} = \sum_{n \in \mathcal{M}^{dl}} \phi_{n,k}, \forall i \in \mathcal{K}_m, \forall m \in \mathcal{M}^{ul}, \tag{4.18}$$

$$\varpi_m^{\mathrm{S2}} = \sum_{n \in \mathcal{M}^{dl}} \phi_{n,m}, \forall m \in \mathcal{M}^{ul}. \tag{4.19}$$

Further, let $\boldsymbol{\chi}$ and $\boldsymbol{\phi}$ denote the two vectors that collect all the inter-cell and BS to BS interference power variables, respectively, i.e.,

$$\boldsymbol{\chi} \overset{\text{def}}{=} [\{\chi_{n,k}\}, \forall n \in \mathcal{M}^{dl}, \forall k \in \mathcal{K}^{dl} \backslash \mathcal{K}_n]^{\mathrm{T}}, \tag{4.20}$$

$$\boldsymbol{\phi} \overset{\text{def}}{=} \begin{cases} [\{\phi_{n,k}\}, \forall n \in \mathcal{M}^{dl}, \forall k \in \mathcal{K}^{ul}]^{\mathrm{T}}, & \text{for S1,} \\ [\{\phi_{n,m}\}, \forall n \in \mathcal{M}^{dl}, \forall m \in \mathcal{M}^{ul}]^{\mathrm{T}}, & \text{for S2.} \end{cases} \tag{4.21}$$

Furthermore, let $\boldsymbol{\chi}_n$ and $\boldsymbol{\phi}_n$ denote the two vectors that collect all the inter-cell and BS to BS interference variables, respectively, that are relevant only to BS_n, i.e.,

$$\boldsymbol{\chi}_n \overset{\text{def}}{=} [\{\chi_k\}, \forall k \in \mathcal{K}_n^{dl}, \quad \{\chi_{n,k}\}, \forall k \in \mathcal{K}^{dl} \backslash \mathcal{K}_n]^{\mathrm{T}} \tag{4.22}$$

$$\boldsymbol{\phi}_n \overset{\text{def}}{=} \begin{cases} [\{\phi_{n,k}\}, \forall k \in \mathcal{K}^{ul}]^{\mathrm{T}}, & \text{for S1} \\ [\{\phi_{n,m}\}, \forall m \in \mathcal{M}^{ul}]^{\mathrm{T}}, & \text{for S2} \end{cases} \tag{4.23}$$

where $\chi_k \overset{\text{def}}{=} \sum_{m \in \mathcal{M}^{dl} \backslash n} \chi_{m,k}$. From above, it can be shown that there exist permutation matrices Ξ_n and Π_n such that

$$\boldsymbol{\chi}_n = \Xi_n \boldsymbol{\chi}, \forall n \in \mathcal{M}^{dl}, \tag{4.24}$$

$$\boldsymbol{\phi}_n = \Pi_n \boldsymbol{\phi}, \forall n \in \mathcal{M}^{dl}. \tag{4.25}$$

Let $\mathsf{P}_n = \text{tr}[\widetilde{\mathbf{T}}_n]$. Then, using the above defined vectors, problem $\mathscr{P}_B$ can be written as

$$\mathscr{P}_C = \begin{cases} \min\limits_{\substack{\{\boldsymbol{\chi}_n\},\{\boldsymbol{\phi}_n\} \\ \{\mathsf{P}_n\},\{\mathbf{T}_k\}}} & \sum\limits_{n \in \mathcal{M}^{dl}} \mathsf{P}_n \\ \text{s.t.} & \text{C1:} \quad \{\boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \mathsf{P}_n, \{\mathbf{T}_k, \forall k \in \mathcal{K}_n\}\} \in \mathcal{S}_n, \forall n \in \mathcal{M}^{dl}, \\ & \text{C2:} \quad \boldsymbol{\chi}_n = \Xi_n \boldsymbol{\chi}, \forall n \in \mathcal{M}^{dl}, \\ & \text{C3:} \quad \boldsymbol{\phi}_n = \Pi_n \boldsymbol{\phi}, \forall n \in \mathcal{M}^{dl}, \\ & \text{C4:} \quad \widetilde{\Pi} \boldsymbol{\phi}_n \leq \boldsymbol{\omega}, \end{cases} \tag{4.26}$$

where $\widetilde{\Pi} = \sum_{n \in \mathcal{M}^{dl}} \Pi_n$, $\boldsymbol{\omega}$ denotes the vector that collects all the BS to BS interference power thresholds, which is given as

$$\boldsymbol{\omega} = \begin{cases} [\{\omega_i, \forall i \in \mathcal{K}^{ul}\}]^{\mathrm{T}}, & \text{for S1,} \\ [\{\omega_m, \forall m \in \mathcal{M}^{ul}\}]^{\mathrm{T}}, & \text{for S2,} \end{cases} \tag{4.27}$$

and $\mathcal{S}_n$ is the convex set of BS_n, $n \in \mathcal{M}^{dl}$, which is defined as

$$\mathcal{S}_n = \begin{cases} \text{C1-1:} & \mathbf{h}_{n,k}^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_{n,k} \geq \chi_k + \varphi_k, \forall k \in \mathcal{K}_n, \\ \text{C1-2:} & \mathbf{h}_{n,i}^{\mathrm{H}} \widetilde{\mathbf{T}}_n \mathbf{h}_{n,i} \leq \chi_{n,i}, \forall i \in \mathcal{K}^{dl} \backslash \mathcal{K}_n, \\ \text{C1-3:} & \begin{cases} \mathbf{g}_{n,i}^{\mathrm{H}} \widetilde{\mathbf{T}}_n \mathbf{g}_{n,i} \leq \phi_{n,i}, \forall i \in \mathcal{K}^{ul}, & \text{for S1} \\ \text{tr}\left[\mathbf{H}_{n,m} \widetilde{\mathbf{T}}_n \mathbf{H}_{n,m}^{\mathrm{H}}\right] \leq \phi_{n,m}, \forall m \in \mathcal{M}^{ul}, & \text{for S2} \end{cases} \\ \text{C1-4:} & \text{tr}[\widetilde{\mathbf{T}}_n] = \mathsf{P}_n \end{cases} \tag{4.28}$$

for all $n \in \mathcal{M}^{dl}$. Problem $\mathscr{P}_C$ is equivalent to problem $\mathscr{P}_B$, since all the inequality constraints hold with equality at the optimal solution.

4.7.2 Distributed algorithm via ADMM

This section applies the ADMM concepts [30] to solve problem $\mathscr{P}_C$ distributively. ADMM is a powerful dual decomposition technique that is perfectly suited for distributed constrained optimization problems. It blends the superior convergence properties of dual decomposition and the numerical robustness of augmented Lagrangian methods [30]. Therefore, ADMM is proved to be more numerically stable, faster in convergence, and can converge under more general conditions than dual decomposition can, e.g., without the requirements of strict convexity or finiteness of the objective function [30]. The ADMM algorithm solves problems in the form [30]

$$\min_{\mathbf{x}\in\mathbf{R}^n,\mathbf{z}\in\mathbf{R}^m} \quad f(\mathbf{x})+g(\mathbf{z}), \quad \text{s.t. } \mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{z}=\mathbf{c}, \tag{4.29}$$

where f and g are convex, $\mathbf{A}\in\mathbf{R}^{p\times n}$, $\mathbf{B}\in\mathbf{R}^{p\times m}$, and $\mathbf{c}\in\mathbf{R}^p$. The augmented Lagrangian of the above problem is given as

$$\mathscr{J}(\mathbf{x},\mathbf{z},\mathbf{y}) = f(\mathbf{x})+g(\mathbf{z})+\mathbf{y}^{\mathrm{T}}(\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{z}-\mathbf{c})+\frac{\rho}{2}\|\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{z}-\mathbf{c}\|_2^2. \tag{4.30}$$

where $\rho>0$ is the penalty parameter and $\mathbf{y}$ is the dual vector. Then, the ADMM consists of the following iterations

$$\mathbf{x}^{(t+1)} = \min_{\mathbf{x}} \quad \mathscr{J}(\mathbf{x},\mathbf{z}^{(t)},\mathbf{y}^{(t)}) \tag{4.31}$$

$$\mathbf{z}^{(t+1)} = \min_{\mathbf{z}} \quad \mathscr{J}(\mathbf{x}^{(t+1)},\mathbf{z},\mathbf{y}^{(t)}) \tag{4.32}$$

$$\mathbf{y}^{(t+1)} = \mathbf{y}^{t}+\rho(\mathbf{A}\mathbf{x}^{(t+1)}+\mathbf{B}\mathbf{z}^{(t+1)}-\mathbf{c}), \tag{4.33}$$

where t here denotes the iteration index. Notice that $\mathbf{x}$ and $\mathbf{z}$ are updated in an alternating fashion, which accounts for the term alternating direction. This is in contrast to the original method of multipliers, that considers updating both variables jointly. Separating the minimization over $\mathbf{x}$ and $\mathbf{z}$ into two steps is precisely what allows for decomposition when f or g are separable. According to [27, Lemma 2], a sufficient condition for convergence of the above steps is that $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ and $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ need to be invertible. If so, then the sequence generated by the above steps is bounded, and every limit point is an optimal solution of the original problem.

Applying the above concepts to problem $\mathscr{P}_C$, one can see that the solution is mainly composed of three iterative steps. The first step is a BS-wise step to update the local primal variables. The second step is a network-wise step to update the global variables. The third step is again a BS-wise step to update the local dual variables. The three steps are detailed as follows. To begin with, we consider excluding constraints C4 at the first step, as it couples all the downlink BSs together, and include it only at the second step. The augmented-form of problem $\mathscr{P}_C$ (after

excluding constraints C4) is given as

$$\mathscr{P}_D = \begin{cases} \min\limits_{\substack{\{\boldsymbol{\chi}_n\},\{\boldsymbol{\phi}_n\}\\ \{\mathsf{P}_n\},\{\mathbf{T}_k\}}} & \sum\limits_{n\in\mathcal{M}^{dl}} \left(\mathsf{P}_n + \frac{\rho}{2}\|\Xi_n\boldsymbol{\chi} - \boldsymbol{\chi}_n\|^2 + \frac{\rho}{2}\|\Pi_n\boldsymbol{\phi} - \boldsymbol{\phi}_n\|^2 + \frac{\rho}{2}(\tilde{\mathsf{P}}_n - \mathsf{P}_n)^2\right) \\ \text{s.t.} & \text{D1:} \quad \{\boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \mathsf{P}_n, \{\mathbf{T}_k, \forall k \in \mathcal{K}_n\}\} \in \mathcal{S}_n, \forall n \in \mathcal{M}^{dl}, \\ & \text{D2:} \quad \boldsymbol{\chi}_n = \Xi_n\boldsymbol{\chi}, \forall n \in \mathcal{M}^{dl}, \\ & \text{D3:} \quad \boldsymbol{\phi}_n = \Pi_n\boldsymbol{\phi}, \forall n \in \mathcal{M}^{dl}, \\ & \text{D4:} \quad \tilde{\mathsf{P}}_n = \mathsf{P}_n, \forall n \in \mathcal{M}^{dl}, \end{cases} \tag{4.34}$$

where $\tilde{\mathsf{P}}_n \geq 0, \forall n \in \mathcal{M}^{dl}$, are slack variables. Note that problem $\mathscr{P}_D$ is clearly equivalent to problem $\mathscr{P}_C$ (after excluding constraint C4), since for any feasible solution, the added terms to the objective are zero. However, the added penalty terms bring numerical stability and faster convergence [27, 30]. The augmented Lagrangian of problem $\mathscr{P}_D$ is given by

$$\begin{aligned} \mathscr{L}(\boldsymbol{\chi}, \boldsymbol{\phi}, \{\tilde{\mathsf{P}}_n\}, \{\boldsymbol{\eta}_n\}, \{\boldsymbol{\xi}_n\}, \{\mu_n\}) = \sum_{n\in\mathcal{M}^{dl}} \Big(&\mathsf{P}_n + \frac{\rho}{2}\|\Xi_n\boldsymbol{\chi} - \boldsymbol{\chi}_n\|^2 + \frac{\rho}{2}\|\Pi_n\boldsymbol{\phi} - \boldsymbol{\phi}_n\|^2 + \\ &\frac{\rho}{2}(\tilde{\mathsf{P}}_n - \mathsf{P}_n)^2 + \boldsymbol{\eta}_n^{\mathrm{T}}(\Xi_n\boldsymbol{\chi} - \boldsymbol{\chi}_n) + \boldsymbol{\xi}_n^{\mathrm{T}}(\Pi_n\boldsymbol{\phi} - \boldsymbol{\phi}_n) + \\ &\mu_n(\tilde{\mathsf{P}}_n - \mathsf{P}_n)\Big), \end{aligned} \tag{4.35}$$

where $\boldsymbol{\eta}_n, \boldsymbol{\xi}_n$, and $\mu_n, \forall n \in \mathcal{M}^{dl}$, are the dual vectors associated with constraints D2, D3 and D4, respectively. Note that the Lagrangian (4.35) is separable between the downlink BSs. Thus, function (4.35) can be written as

$$\mathscr{L}(\boldsymbol{\chi}, \boldsymbol{\phi}, \{\tilde{\mathsf{P}}_n\}, \{\boldsymbol{\eta}_n\}, \{\boldsymbol{\xi}_n\}, \{\mu_n\}) = \sum_{n\in\mathcal{M}^{dl}} \mathscr{L}_n(\boldsymbol{\chi}, \boldsymbol{\phi}, \tilde{\mathsf{P}}_n, \boldsymbol{\eta}_n, \boldsymbol{\xi}_n, \mu_n). \tag{4.36}$$

From above, the $\mathrm{BS}_n, \forall n \in \mathcal{M}^{dl}$, augmented dual problem based ADMM is given as

$$\mathscr{P}_E = \begin{cases} \min\limits_{\substack{\boldsymbol{\chi}_n,\boldsymbol{\phi}_n\\ \mathsf{P}_n,\{\mathbf{T}_k\}}} & \mathscr{L}_n(\boldsymbol{\chi}, \boldsymbol{\phi}, \tilde{\mathsf{P}}_n, \boldsymbol{\eta}_n, \boldsymbol{\xi}_n, \mu_n) \\ \text{s.t.} & \{\boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \mathsf{P}_n, \{\mathbf{T}_k, \forall k \in \mathcal{K}_n\}\} \in \mathcal{S}_n. \end{cases} \tag{4.37}$$

Note that problem $\mathscr{P}_E$ is a BS-wise step that updates the local variables, $\boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \mathsf{P}_n, \mathbf{T}_k, \forall k \in \mathcal{K}_n$, for given global and local variables $\boldsymbol{\chi}, \boldsymbol{\phi}, \tilde{\mathsf{P}}_n, \boldsymbol{\eta}_n, \boldsymbol{\xi}_n, \mu_n$. Thus, at the $(t+1)$ iteration, each $\mathrm{BS}_n, \forall n \in \mathcal{M}^{dl}$, solves problem $\mathscr{P}_E$ to obtain $\boldsymbol{\chi}_n(t+1), \boldsymbol{\phi}_n(t+1), \mathsf{P}_n(t+1)$ and $\mathbf{T}_k(t+1), \forall k \in \mathcal{K}_n$. After that, each downlink $\mathrm{BS}_n, \forall n \in \mathcal{M}^{dl}$, performs a *broadcast and gather* operation of the obtained local parameters $\boldsymbol{\chi}_n(t+1), \boldsymbol{\phi}_n(t+1), \boldsymbol{\eta}_n(t)$, and $\boldsymbol{\xi}_n(t)$ with other downlink BSs. After gathering all the variables from other downlink BSs, each $\mathrm{BS}_n, \forall n \in \mathcal{M}^{dl}$, updates the global variables $\boldsymbol{\chi}(t+1)$ and $\boldsymbol{\phi}(t+1)$ and the local variable $\tilde{\mathsf{P}}_n(t+1)$ as a solution to the following

problems:

$$\begin{aligned}
\boldsymbol{\chi}(t+1) &= \min_{\boldsymbol{\chi}} \mathscr{L}(\boldsymbol{\chi}, \boldsymbol{\phi}, \{\tilde{\mathsf{P}}_n\}, \{\boldsymbol{\eta}_n\}, \{\boldsymbol{\xi}_n\}, \{\mu_n\}) \\
&\overset{(a)}{=} \min_{\boldsymbol{\chi}} \frac{\rho}{2} \sum_{n \in \mathcal{M}^{dl}} \|\Xi_n \boldsymbol{\chi} - \boldsymbol{\chi}_n(t+1)\|^2 + \sum_{n \in \mathcal{M}^{dl}} \boldsymbol{\eta}_n^{\mathrm{T}}(t)\Big(\Xi_n \boldsymbol{\chi} - \boldsymbol{\chi}_n(t+1)\Big), \qquad (4.38) \\
\boldsymbol{\phi}(t+1) &= \min_{\substack{\boldsymbol{\phi} \\ \text{s.t. } \widetilde{\Pi}\boldsymbol{\phi}=\boldsymbol{\omega}}} \mathscr{L}(\boldsymbol{\chi}, \boldsymbol{\phi}, \{\tilde{\mathsf{P}}_n\}, \{\boldsymbol{\eta}_n\}, \{\boldsymbol{\xi}_n\}, \{\mu_n\}) \\
&\overset{(b)}{=} \min_{\substack{\boldsymbol{\phi} \\ \text{s.t. } \widetilde{\Pi}\boldsymbol{\phi}=\boldsymbol{\omega}}} \frac{\rho}{2} \sum_{n \in \mathcal{M}^{dl}} \|\Pi_n \boldsymbol{\phi} - \boldsymbol{\phi}_n(t+1)\|^2 + \sum_{n \in \mathcal{M}^{dl}} \boldsymbol{\xi}_n^{\mathrm{T}}(t)\Big(\Pi_n \boldsymbol{\phi} - \boldsymbol{\phi}_n(t+1)\Big), \qquad (4.39) \\
\tilde{\mathsf{P}}_n(t+1) &= \min_{\tilde{\mathsf{P}}_n} \mathscr{L}_n(\boldsymbol{\chi}, \boldsymbol{\phi}, \{\tilde{\mathsf{P}}_n\}, \{\boldsymbol{\eta}_n\}, \{\boldsymbol{\xi}_n\}, \{\mu_n\}) \\
&\overset{(c)}{=} \min_{\tilde{\mathsf{P}}_n} \frac{\rho}{2}(\tilde{\mathsf{P}}_n - \mathsf{P}_n(t+1))^2 + \mu_n(t)\Big(\tilde{\mathsf{P}}_n - \mathsf{P}_n(t+1)\Big), \qquad (4.40)
\end{aligned}$$

where (a), (b) and (c) are obtained by removing the irrelevant terms. Note that the constraint C4, that was excluded in problem $\mathscr{P}_D$, is included in the global variables update in problem (4.39). Problems (4.38), (4.39) and (4.40) are convex quadratic problems. The closed-form solutions are given, respectively, as

$$\boldsymbol{\chi}(t+1) = \Xi^{\dagger}\Big(\tilde{\boldsymbol{\chi}}(t+1) - \frac{1}{\rho}\tilde{\boldsymbol{\eta}}(t)\Big), \qquad (4.41)$$

$$\boldsymbol{\phi}(t+1) = \Pi^{\dagger}\Big(\tilde{\boldsymbol{\phi}}(t+1) - \frac{1}{\rho}\tilde{\boldsymbol{\xi}}(t)\Big) + \frac{1}{\rho}(\Pi^{\mathrm{T}}\Pi)^{-1}\widetilde{\Pi}^{\mathrm{T}}\boldsymbol{\zeta}, \qquad (4.42)$$

$$\tilde{\mathsf{P}}_n(t+1) = \mathsf{P}_n(t+1) - \frac{1}{\rho}\mu_n(t), \qquad (4.43)$$

where $\boldsymbol{\zeta}$ denotes the dual vector associated with problem (4.39) constraint, which has a closed-form solution given as

$$\boldsymbol{\zeta} = \Big(\frac{1}{\rho}\widetilde{\Pi}(\Pi^{\mathrm{T}}\Pi)^{-1}\widetilde{\Pi}^{\mathrm{T}}\Big)^{-1}\Big(\boldsymbol{\omega} - \widetilde{\Pi}\Pi^{\dagger}\Big(\tilde{\boldsymbol{\phi}}(t+1) - \frac{1}{\rho}\tilde{\boldsymbol{\xi}}(t)\Big)\Big). \qquad (4.44)$$

Moreover, matrix Ξ is defined as

$$\Xi = [\Xi_{\mathcal{M}^{dl}(1)}^{\mathrm{T}} \cdots \Xi_{\mathcal{M}^{dl}(|\mathcal{M}^{dl}|)}^{\mathrm{T}}]^{\mathrm{T}}, \qquad (4.45)$$

where Π, $\tilde{\boldsymbol{\chi}}(t+1)$, $\tilde{\boldsymbol{\eta}}(t)$, $\tilde{\boldsymbol{\phi}}(t+1)$, and $\tilde{\boldsymbol{\xi}}(t)$ are defined in a similar way. Note that $\Xi^{\mathrm{T}}\Xi > 0$ and $\Pi^{\mathrm{T}}\Pi > 0$. Thus, both matrices are invertible, which is a sufficient condition for convergence of the above steps according to [27, Lemma 2]. Finally, each BS_n, $n \in \mathcal{M}^{dl}$, updates the dual vectors $\boldsymbol{\eta}_n(t+1)$, $\boldsymbol{\xi}_n(t+1)$ and $\mu_n(t+1)$ as

$$\boldsymbol{\eta}_n(t+1) = \boldsymbol{\eta}_n(t) + \rho\Big(\Xi_n \boldsymbol{\chi}(t+1) - \boldsymbol{\chi}_n(t+1)\Big), \qquad (4.46)$$

$$\boldsymbol{\xi}_n(t+1) = \boldsymbol{\xi}_n(t) + \rho\Big(\Pi_n \boldsymbol{\phi}(t+1) - \boldsymbol{\phi}_n(t+1)\Big), \qquad (4.47)$$

$$\mu_n(t+1) = \mu_n(t) + \rho\Big(\tilde{\mathsf{P}}_n(t+1) - \mathsf{P}_n(t+1)\Big). \qquad (4.48)$$

Note that the above dual variables update are a subgradient-like update with the penalty parameter ρ representing the step-size. It is known that the subgradient algorithms are guaranteed to converge as long as the update step-size is carefully chosen [35]. Unfortunately, the optimal value of the step-size, i.e., ρ, is not known. In general, it is dependent of the system parameters.

The distributed algorithm for solving problem $\mathscr{P}_A$, with either scenario, is summarized in Algorithm 6. We refer to this algorithm as S1-Dist-CSI for Scenario 1 and as S2-Dist-CSI for Scenario 2. In Algorithm 6 inner-iterations, each BS_n, $\forall n \in \mathcal{M}^{dl}$, is required to perform *broad-*

Algorithm 6: Distributed CBF Algorithm based ADMM Technique.

Inputs:
The global and local parameters $\boldsymbol{\chi}(0)$, $\boldsymbol{\phi}(0)$,$\tilde{\mathsf{P}}_n(0)$,$\boldsymbol{\eta}_n(0)$,$\boldsymbol{\xi}_n(0)$,$\mu_n(0)$, $\forall n \in \mathcal{M}^{dl}$.
Receive beamforming vectors $\mathbf{r}_k(0)$, $\forall k \in \mathcal{K}^{ul}$.
Penalty parameter ρ.
Set $r = 1$ outer-iterations index for updating receive beamforming of uplink users.
Set $t = 1$ inner-iterations index for updating transmit beamforming of downlink users.
while not converged **do**
  For fixed $\mathbf{r}_k(r)$, $\forall k \in \mathcal{K}^{ul}$, update $\mathbf{t}_k(t)$, $\forall k \in \mathcal{K}^{dl}$:
  while not converged **do**
    Solve problem $\mathscr{P}_E$ to obtain $\boldsymbol{\chi}_n(t+1)$, $\boldsymbol{\phi}_n(t+1)$ and $\tilde{\mathsf{P}}_n(t+1)$, $\forall n \in \mathcal{M}^{dl}$.
    Share $\boldsymbol{\chi}_n(t+1)$, $\boldsymbol{\phi}_n(t+1)$, $\boldsymbol{\eta}_n(t)$, and $\boldsymbol{\xi}_n(t)$, $\forall n \in \mathcal{M}^{dl}$, between downlink BSs.
    Update $\boldsymbol{\chi}(t+1)/\boldsymbol{\phi}(t+1)$ using (4.41)/(4.42).
    Update $\tilde{\mathsf{P}}_n(t+1)/\boldsymbol{\eta}_n(t+1)/\boldsymbol{\xi}_n(t+1)/\mu_n(t+1)$ using (4.43)/(4.46)/(4.47)/(4.48).
    Set $t = t+1$,
  end while
  For fixed $\mathbf{t}_k(t)$, $\forall k \in \mathcal{K}^{dl}$, update $\mathbf{r}_k(r+1)$, $\forall k \in \mathcal{K}^{ul}$ using (4.8).
  For S1, set $r = r+1$ and go back to step 8.
  For S2, break.
end while

cast and gather operation of the obtained local variables $\boldsymbol{\chi}_n(t+1)$ and $\boldsymbol{\phi}_n(t+1)$ at each iteration index (t). It can be checked that the total backhaul signaling at each iteration is

$$\begin{cases} 2|\mathcal{M}^{dl}|(|\mathcal{M}^{dl}|K + |\mathcal{M}^{ul}|K), & \text{for S1,} \\ 2|\mathcal{M}^{dl}|(|\mathcal{M}^{dl}|K + |\mathcal{M}^{ul}|), & \text{for S2.} \end{cases} \tag{4.49}$$

Clearly, the total backhaul signaling for S2 is less than S1, as it is in function of the number of uplink BSs rather than the number of uplink MSs. This means that the two scenarios have equal signaling overhead, per iteration, when there is only one user per uplink cell, i.e., $K = 1$. Furthermore, in practice, it might be required to stop the algorithm after some finite number of iterations, say $t^{\max}$. However, since Algorithm 6 inner-iterations operate in the dual domain, the minimum SINR and maximum BS to BS interference power thresholds might be violated. In this case, each BS_n, $\forall n \in \mathcal{M}^{dl}$, is required to perform one more iteration in the primal domain to guarantee that both constraints are satisfied. With this respect, each BS_n, $\forall n \in \mathcal{M}^{dl}$,

solves the following problem

$$\mathscr{P}_F = \begin{cases} \min\limits_{\{\mathbf{T}_k, \forall k \in \mathcal{K}_n\}} & \mathsf{P}_n \\ \text{s.t.} & \{\boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \mathsf{P}_n, \{\mathbf{T}_k, \forall k \in \mathcal{K}_n\}\} \in \mathcal{S}_n, \end{cases}$$

for given $\boldsymbol{\chi}_n = \Xi_n \boldsymbol{\chi}(t^{\max})$, $\boldsymbol{\phi}_n = \Pi_n \boldsymbol{\phi}(t^{\max})$ (where $\widetilde{\Pi}\boldsymbol{\phi}(t^{\max}) = \boldsymbol{\omega}$), and $\{\mathbf{r}_i(r)\}$.

4.8 Imperfect CSI case: distributed algorithm

In the previous section, we have assumed that the downlink BSs have perfect knowledge of local CSI. This simplified the presentation of the distributed algorithm, but it is clearly an ideal model as, in real implementation, the BSs can only have imperfect CSI due to a variety of reasons; e.g., imperfect channel estimation, feedback quantization, inadequate channel reciprocity [26]. As a result, the MSs' target values can no longer be guaranteed. To overcome this, a robust optimization of the transmit beamforming vectors should be considered, i.e., Algorithms 5 and 6 should be modified such that they would account for the CSI errors.

In this chapter, it assumed that the CSI errors fall inside a bounded uncertainty set. Thus, the main task is to modify the proposed algorithms to operate under the CSI worst-case condition. At the end of this section, it will be clear that the only change to be made to Algorithm 6 is only on step (a-1), whereas the remaining steps are similar. To start with, let the true/perfect channels be given as

$$\mathbf{h}_{n,k} = \hat{\mathbf{h}}_{n,k} + \mathbf{e}_{n,k}, \tag{4.50}$$

$$\mathbf{H}_{n,m} = \hat{\mathbf{H}}_{n,m} + \mathbf{E}_{n,m}, \tag{4.51}$$

where $\hat{\mathbf{h}}_{n,k} \in \mathbb{C}^N$ and $\hat{\mathbf{H}}_{n,m} \in \mathbb{C}^{N\times N}$ denote the estimated channels available to the downlink BSs. The CSI errors are denoted by $\mathbf{e}_{n,k} \in \mathbb{C}^N$ and $\mathbf{E}_{n,m} \in \mathbb{C}^{N\times N}$, which are assumed to be bounded and take values from the set

$$\mathcal{H} \stackrel{\text{def}}{=} \begin{cases} \{\mathbf{e}_{n,k} \big| \|\mathbf{e}_{n,k}\|^2 = \mathbf{e}_{n,k}^{\mathrm{H}} \mathbf{I}_N \mathbf{e}_{n,k} \le \epsilon_{n,k}\}, \forall n \in \mathcal{M}^{dl}, \forall k \in \mathcal{K}, \\ \{\mathbf{E}_{n,m} \big| \mathrm{tr}(\mathbf{E}_{n,m}\mathbf{E}_{n,m}^{\mathrm{H}}) = \mathrm{vec}(\mathbf{E}_{n,m})^{\mathrm{H}} \mathbf{I}_{N^2} \mathrm{vec}(\mathbf{E}_{n,m}) \le \acute{\epsilon}_{n,m}\}, \forall n, m \in \mathcal{M}, \end{cases} \tag{4.52}$$

where $\epsilon_{n,k} > 0$ and $\acute{\epsilon}_{n,m} > 0$ control the degree of errors associated with the $\mathbf{h}_{n,k}$ and $\mathbf{H}_{n,m}$ channels, respectively.

To account for CSI errors, the constraint functions of convex set $\mathcal{S}_n$, $n \in \mathcal{M}_n$, given by (4.28) should be modified to operate under the worst-case condition. The worst-case optimization for constraints C1-1 and C1-2 of BS_n, $n \in \mathcal{M}^{dl}$ can be written, respectively, as

$$\min_{\forall \mathbf{e}_{n,k} \in \mathcal{H}} (\hat{\mathbf{h}}_{n,k} + \mathbf{e}_{n,k})^{\mathrm{H}} \mathbf{D}_k (\hat{\mathbf{h}}_{n,k} + \mathbf{e}_{n,k}) \ge \chi_k + \varphi_k, \forall k \in \mathcal{K}_n, \tag{4.53}$$

$$\max_{\forall \mathbf{e}_{n,i} \in \mathcal{H}} (\hat{\mathbf{h}}_{n,i} + \mathbf{e}_{n,i})^{\mathrm{H}} \widetilde{\mathbf{T}}_n (\hat{\mathbf{h}}_{n,i} + \mathbf{e}_{n,i}) \le \chi_{n,i}, \forall i \in \mathcal{K}^{dl} \backslash \mathcal{K}_n. \tag{4.54}$$

For S1, the worst-case optimization for constraint C1-3 of BS_n, $n \in \mathcal{M}^{dl}$, can be written as

$$\max_{\forall \mathbf{e}_{n,i} \in \mathcal{H}} (\hat{\mathbf{g}}_{n,i} + \mathbf{e}_{n,i})^{\text{H}} \widetilde{\mathbf{T}}_n (\hat{\mathbf{g}}_{n,i} + \mathbf{e}_{n,i}) \leq \phi_{n,i}, \forall i \in \mathcal{K}^{ul}. \tag{4.55}$$

For S2, the worst-case optimization for constraint C1-3 of BS_n, $n \in \mathcal{M}^{dl}$ can be written as

$$\max_{\forall \mathbf{E}_{n,m} \in \mathcal{H}} \text{tr}\left[(\hat{\mathbf{H}}_{n,m} + \mathbf{E}_{n,m}) \widetilde{\mathbf{T}}_n (\hat{\mathbf{H}}_{n,m} + \mathbf{E}_{n,m})^{\text{H}}\right] \leq \phi_{n,m}, \forall m \in \mathcal{M}^{ul}. \tag{4.56}$$

Note that, due to the worst-case design criterion, each problem of (4.53), (4.54), (4.55) and (4.56) contains infinitely many constraints. Fortunately, each of these problems can be written in linear matrix inequality (LMI) form using the S-Lemma [33, p.655].

Lemma 2 (S-Lemma) *Let $f(\mathbf{x}) = \mathbf{x}^H \mathbf{A} \mathbf{x} + 2\mathbf{a}^H \mathbf{x} + c$ and $h(\mathbf{x}) = \mathbf{x}^H \mathbf{B} \mathbf{x} + 2\mathbf{b}^H \mathbf{x} + d$ be two quadratic functions having symmetric matrices $\mathbf{A}$ and $\mathbf{B}$. Given any pair of quadratic functions (f, h), if $h(\mathbf{x}) \leq 0$ satisfies Slater's condition, namely, there is an $\mathbf{x} \in \mathbf{C}^N$ such that $h(\mathbf{x}) < 0$, the following two statements are always equivalent:*

1. *$\forall \mathbf{x}$ satisfying $h(\mathbf{x}) \leq 0 \Rightarrow f(\mathbf{x}) \geq 0$.*
2. *There exists a $\lambda \geq 0$ such that $f(\mathbf{x}) + \lambda h(\mathbf{x}) \geq 0$, i.e,*

$$\begin{bmatrix} \mathbf{A} & \mathbf{a} \\ \mathbf{a}^H & c \end{bmatrix} + \lambda \begin{bmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^H & d \end{bmatrix} \geq 0$$

To make use of Lemma 2, consider, for example, the minimization problem (4.53). It can be noticed that the objective function and the channel uncertainty constraint of problem (4.53) can be expressed as $f(\mathbf{e}_{n,k})$ and $h(\mathbf{e}_{n,k})$, respectively, as

$$f(\mathbf{e}_{n,k}) = \mathbf{e}_{n,k}^{\text{H}} \mathbf{D}_k \mathbf{e}_{n,k} + 2\mathbf{e}_{n,k}^{\text{H}} \mathbf{D}_k \hat{\mathbf{h}}_{n,k} + \hat{\mathbf{h}}_{n,k}^{\text{H}} \mathbf{D}_k \hat{\mathbf{h}}_{n,k} - \chi_k - \varphi_k, \tag{4.57}$$

$$h(\mathbf{e}_{n,k}) = \mathbf{e}_{n,k}^{\text{H}} \mathbf{I}_N \mathbf{e}_{n,k} - \epsilon_{n,k}. \tag{4.58}$$

Then, according to Lemma 2, problem (4.53) can be written equivalently in linear matrix inequality (LMI) form as

$$\mathbf{F}_k \overset{\text{def}}{=} \begin{bmatrix} \mathbf{D}_k + \lambda_{n,k} \mathbf{I}_N & \mathbf{D}_k \hat{\mathbf{h}}_{n,k} \\ (\mathbf{D}_k \hat{\mathbf{h}}_{n,k})^{\text{H}} & \hat{\mathbf{h}}_{n,k}^{\text{H}} \mathbf{D}_k \hat{\mathbf{h}}_{n,k} - f_k \end{bmatrix} \geq 0, \forall k \in \mathcal{K}_n \tag{4.59}$$

where $f_k = \epsilon_{n,k} \lambda_{n,k} + \chi_k + \varphi_k$. Using similar steps, one can write the maximization problems (4.54), (4.55) and (4.56) in LMI form, respectively, as

$$\mathbf{J}_{n,i} \overset{\text{def}}{=} \begin{bmatrix} \lambda_{n,i} \mathbf{I}_N - \widetilde{\mathbf{T}}_n & -\widetilde{\mathbf{T}}_n \hat{\mathbf{h}}_{n,i} \\ -(\widetilde{\mathbf{T}}_n \hat{\mathbf{h}}_{n,i})^{\text{H}} & j_{n,i} - \hat{\mathbf{h}}_{n,i}^{\text{H}} \widetilde{\mathbf{T}}_n \hat{\mathbf{h}}_{n,i} \end{bmatrix} \geq 0, \forall i \in \mathcal{K}^{dl} \backslash \mathcal{K}_n, \tag{4.60}$$

$$\mathbf{L}_{n,i} \overset{\text{def}}{=} \begin{bmatrix} \lambda_{n,i} \mathbf{I}_N - \widetilde{\mathbf{T}}_n & -\widetilde{\mathbf{T}}_n \hat{\mathbf{g}}_{n,i} \\ -(\widetilde{\mathbf{T}}_n \hat{\mathbf{g}}_{n,i})^{\text{H}} & l_{n,i} - \hat{\mathbf{g}}_{n,i}^{\text{H}} \widetilde{\mathbf{T}}_n \hat{\mathbf{g}}_{n,i} \end{bmatrix} \geq 0, \forall i \in \mathcal{K}^{ul}, \tag{4.61}$$

$$\mathbf{U}_{n,m} \overset{\text{def}}{=} \begin{bmatrix} \acute{\lambda}_{n,m} \mathbf{I}_{N^2} - (\mathbf{I}_N \otimes \widetilde{\mathbf{T}}_n) & -\text{vec}(\hat{\mathbf{H}}_{n,m} \widetilde{\mathbf{T}}_n) \\ -\text{vec}(\hat{\mathbf{H}}_{n,m} \widetilde{\mathbf{T}}_n)^{\text{H}} & u_{n,m} - \text{tr}[\hat{\mathbf{H}}_{n,m} \widetilde{\mathbf{T}}_n \hat{\mathbf{H}}_{n,m}^{\text{H}}] \end{bmatrix} \geq 0, \forall m \in \mathcal{M}^{ul}, \tag{4.62}$$

where $j_{n,i} = \chi_{n,i} - \epsilon_{nn_i}\lambda_{n,i}$, $l_{n,i} = \phi_{n,i} - \epsilon_{n,i}\lambda_{n,i}$ and $u_{n,m} = \phi_{n,m} - \acute{\epsilon}_{n,m}\acute{\lambda}_{n,m}$. Note that in formulating the functions (4.62), the following trace properties are utilized: for give matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, we have $\mathrm{tr}(\mathbf{A}+\mathbf{B}+\mathbf{C}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B}) + \mathrm{tr}(\mathbf{C})$ and that $\mathrm{tr}(\mathbf{ABC}) = \mathrm{vec}(\mathbf{A})^{\mathrm{H}}(\mathbf{I}\otimes\mathbf{B})\mathrm{vec}(\mathbf{C})$.

From above, it can be seen that using the S-Lemma, each problem with infinite number of constraints is converted into just one LMI constraint. This comes with the cost of adding extra variables $\{\lambda_{n,k}, \acute{\lambda}_{n,m}\}$ that indirectly represent the worst channel conditions in the uncertainty set. Thus, if we can find $\{\lambda_{n,k} \geq 0, \acute{\lambda}_{n,m} \geq 0\}$, then all the constraints (4.53), (4.54), (4.55) and (4.56) are satisfied for all $\{\mathbf{e}_{n,k}, \mathbf{E}_{n,m}\} \in \mathcal{H}$.

Therefore, to account for CSI errors, each BS_n, $n \in \mathcal{M}^{dl}$, should modify its non-robust convex set $\mathcal{S}_n$ with the worst-case based robust convex set $\widehat{\mathcal{S}}_n$, which is given as

$$\widehat{\mathcal{S}}_n = \left\{ \begin{array}{l} \mathbf{F}_k \succeq 0, \forall k \in \mathcal{K}_n, \\ \mathbf{J}_{n,i} \succeq 0, \forall i \in \mathcal{K}^{dl} \backslash \mathcal{K}_n, \\ \begin{cases} \mathbf{L}_{n,i} \succeq 0, \forall i \in \mathcal{K}^{ul}, & \text{for S1} \\ \mathbf{U}_{n,m} \succeq 0, \forall m \in \mathcal{M}^{ul}, & \text{for S2} \end{cases} \\ \quad \mathrm{tr}[\widetilde{\mathbf{T}}_n] = \mathsf{P}_n \end{array} \right\}, \forall n \in \mathcal{M}^{dl}. \tag{4.63}$$

Finally, problem $\mathscr{P}_E$ can be rewritten considering the worst-case beamforming design as

$$\mathscr{P}_J = \left\{ \begin{array}{ll} \underset{\substack{\boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \mathsf{P}_n \\ \{\mathbf{T}_k\} \\ \{\lambda_{n,k}\}, \{\acute{\lambda}_{n,m}\}}}{\min} & \mathscr{L}_n(\boldsymbol{\chi}(t), \boldsymbol{\phi}(t), \mathsf{P}_n(t), \boldsymbol{\eta}_n(t), \boldsymbol{\xi}_n(t), \mu_n(t)) \\ \text{s.t.} & \text{J1: } \{\boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \mathsf{P}_n, \{\mathbf{T}_k\}, \{\lambda_{n,k}\}, \{\acute{\lambda}_{n,m}\}\} \in \widehat{\mathcal{S}}_n. \end{array} \right. \tag{4.64}$$

Hence, the distributed robust algorithm of either scenario can be obtained by updating Algorithm 6 step (a-1) to consider solving problem $\mathscr{P}_J$ instead of $\mathscr{P}_E$. We refer to this algorithm as S1-Dist-Robust for Scenario 1 and as S2-Dist-Robust for Scenario 2.

4.9 Imperfect CSI case: centralized algorithm

The centralized robust optimization problem of either scenario can be written similar to problem $\mathscr{P}_C$, by replacing the constraint C1 with the constraint of problem $\mathscr{P}_J$, i.e., constraint J1. The resulting centralized problem can then be solved by an algorithm similar to Algorithm 5 using off-the-shelf convex solvers, e.g. SeDuMi [88]. We refer to this algorithm as S1-Cent-Robust for Scenario 1 and as S2-Cent-Robust for Scenario 2.

4.10 Numerical results

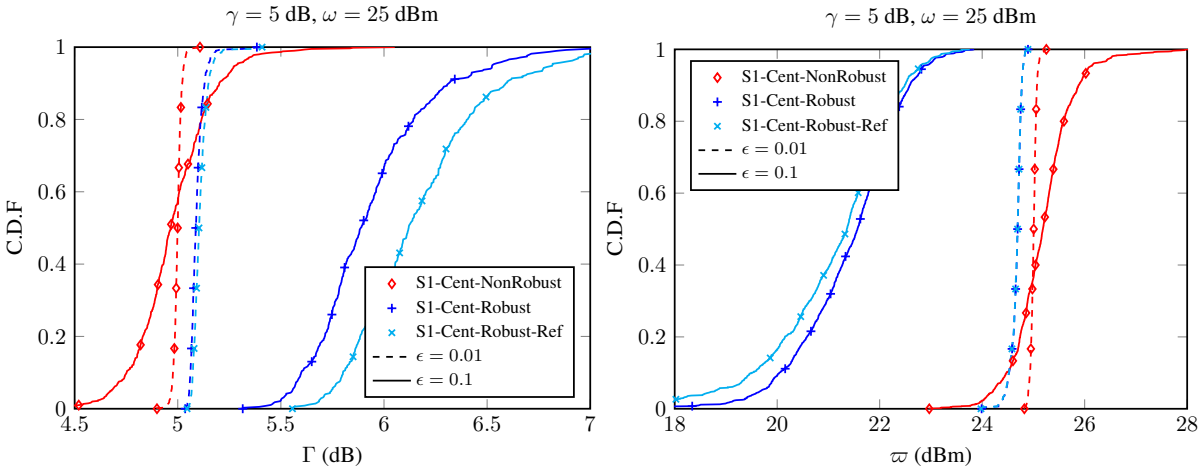
We consider a flat Rayleigh fading scenario with uncorrelated channels between antennas, i.e., each element of $\mathbf{h}_{n,k}$, $\mathbf{e}_{n,k}$, $\mathbf{H}_{n,m}$, and $\mathbf{E}_{n,m}$ is an i.i.d. complex Gaussian random

variable with zero mean and unit variance. To simplify the exposition, we assume equal input design parameters for all MSs, i.e., $\gamma_k = \gamma, \forall k \in \mathcal{K}^{dl}$, $\omega_k = \omega, \forall k \in \mathcal{K}^{ul}$, and $\epsilon_{n,k} = \acute{\epsilon}_{n,m} = \epsilon, \forall n, m \in \mathcal{M}, \forall k \in \mathcal{K}$. We initialize the receive beamforming vectors of uplink MSs using (4.8) assuming the initial transmit beamforming vectors of downlink MSs are given by the MRT approach, i.e., $\mathbf{t}_k(0) = \mathbf{h}_{n,k}/\|\mathbf{h}_{n,k}\|$. Moreover, we assume that the uplink transmit power $q_k = 23$ dBm, $\forall k \in \mathcal{K}^{ul}$, and the noise power $\sigma^2 = 1$. We consider a system of $[M, K, N] = [3, 3, 8]$, where $|\mathcal{M}^{dl}| = 2$ and $|\mathcal{M}^{ul}| = 1$.

Example 1: impact of channel errors

In this example, we show simulation results to examine the impact of CSI errors on the satisfaction of design targets. Fig. 4.5 shows the CDF plots for the achievable SINR at the downlink MSs (Γ) and the BS to BS interference power at the uplink MSs (ϖ) for two error-bound (ϵ) values. For comparison, we show simulation results of the centralized robust algorithm shown in [34]. We refer to this reference algorithm as S1-Cent-Robust-Ref.

Figure 4.5 – CDF plots of SINR at downlink MSs and BS-to-BS interference power at uplink MSs.



Source: Created by author.

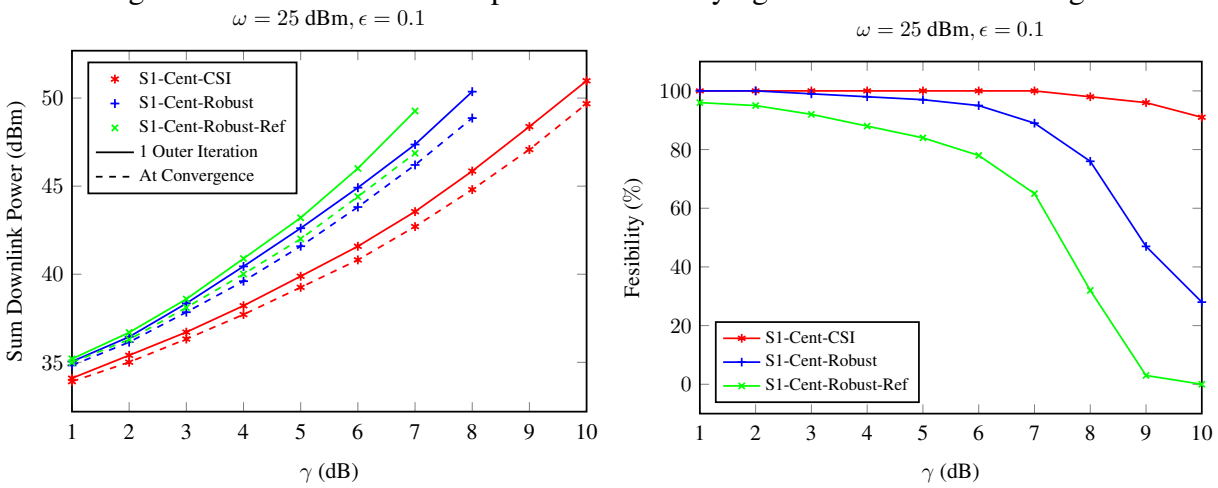
At first, we note note that in Fig. 4.5, S1-Cent-NonRobust algorithm is equivalent to S1-Cent-CSI algorithm with estimated channels given by (4.50) and (4.51) being used as the input channels. From Fig. 4.5, it can be observed that, with S1-Cent-NonRobust, the design targets γ and ω are unsatisfied with high probability and the unsatisfaction level increases as the CSI error-bound increases. On the other hand, with both robust algorithms, S1-Cent-Robust and S1-Cent-Robust-Ref, the design targets are always satisfied and the satisfaction level increases as the CSI error-bound increases. Note that, when $\epsilon = 0.01$, both robust algorithms almost have the same satisfaction level. However, when $\epsilon = 0.1$, the satisfaction level of S1-Cent-Robust-Ref is higher than S1-Cent-Robust. This shows that S1-Cent-Robust-Ref is overreacting to the CSI error-bounds, as compared to S1-Cent-Robust. In the next example, it will be shown that

this overreaction of S1-Cent-Robust-Ref would result on using a higher transmit power than S1-Cent-Robust.

Example 2: impact of input parameters

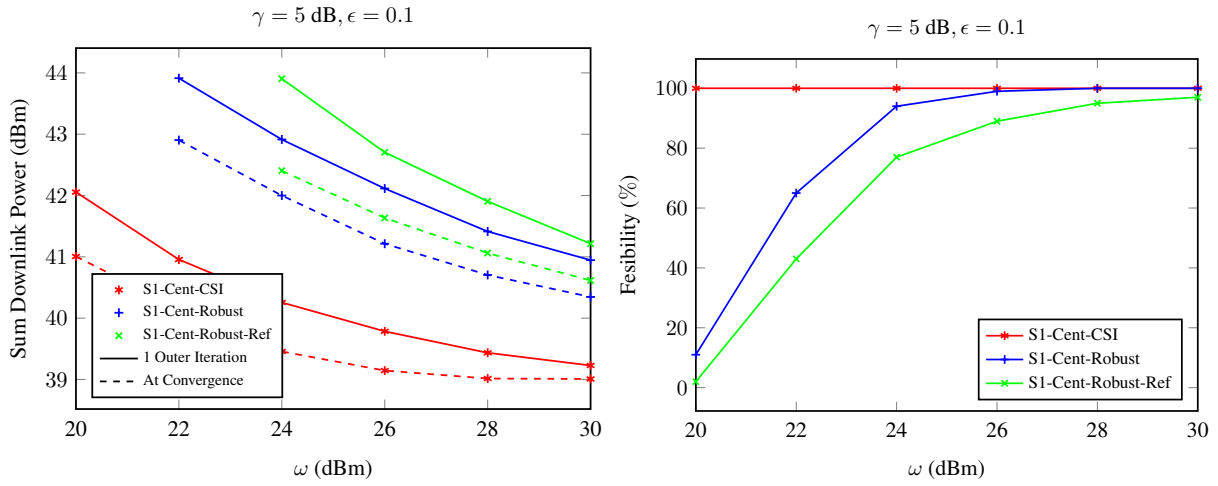
In this example, we show simulation results to evaluate the impact of input parameters γ, ω and ϵ on system performance (sum downlink power and feasibility). Figs. 4.6, 4.7 and 4.8 show the averaged simulation results, while varying the minimum SINR target γ, the maximum BS to BS interference power threshold ω, and the channel uncertainty upper-bound ϵ, respectively. For each simulation scenario, we vary one of the input parameters assuming the others are fixed.

Figure 4.6 – Performance comparison while varying the minimum SINR target.



Source: Created by author.

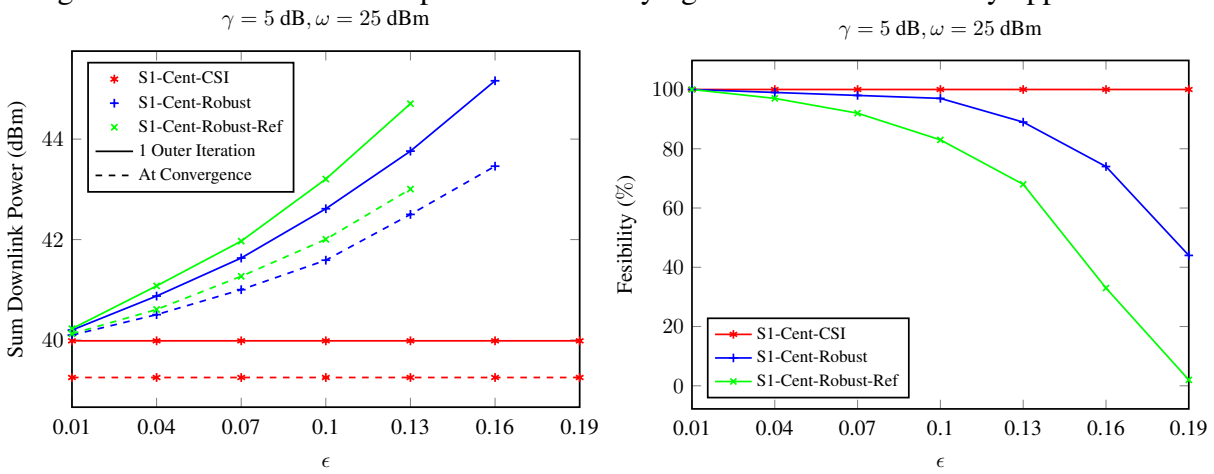
Figure 4.7 – Performance comparison while varying the maximum BS-to-BS interference power threshold.



Source: Created by author.

At first, we note that in Figs. 4.6, 4.7 and 4.8, the feasibility plots are a measure of the percentage of the feasible channel-realizations to the total number of tested ones. The results

Figure 4.8 – Performance comparison while varying the channel uncertainty upper-bound.



Source: Created by author.

with less than 50% feasibility rate are not shown on the performance plots. Furthermore, the results of *1 Outer Iteration* are obtained using the first initialization of the receive beamforming vectors, i.e., $\{\mathbf{r}_k(0)\}$, while the *At Convergence* results are obtained with updating the receive beamforming vectors until convergence.

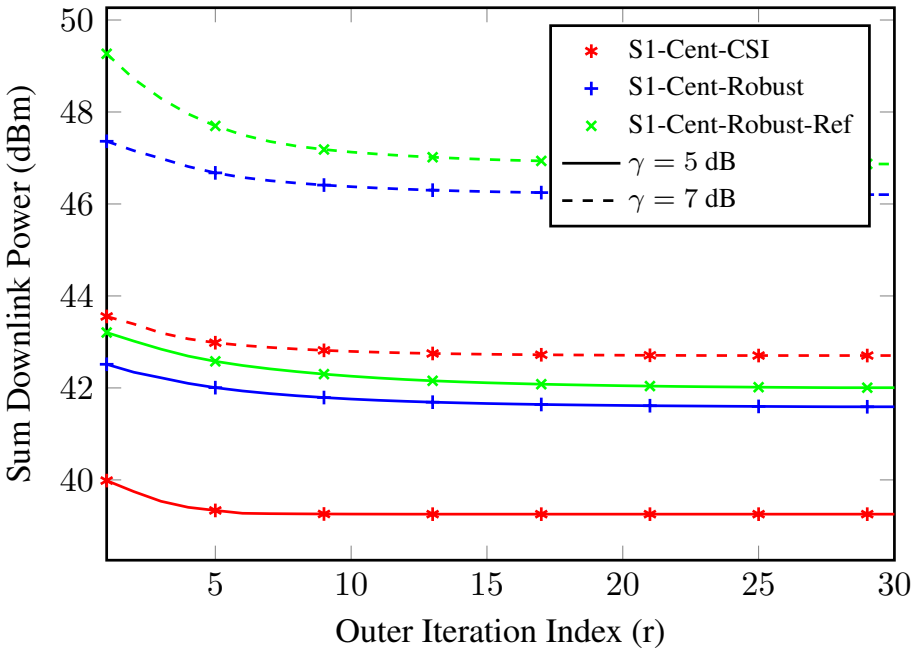
From Figs. 4.6, 4.7 and 4.8, it can be seen that when the input parameters γ and ϵ increase, the sum of the downlink power increases and the feasibility rate decreases for both robust algorithms. These results hold true, as well, when the input parameter ω decreases. From all results, it can be seen that the proposed algorithm S1-Cent-Robust has better energy-efficiency and wider feasibility range, as compared to the reference algorithm S1-Cent-Robust-Ref. Furthermore, we can see that updating the receive uplink beamforming vectors of the uplink MSs is critical on improving the energy-efficiency of the downlink BSs. Fig. 4.9 demonstrates the average convergence behavior outer-iterations (r) for updating the uplink MSs receive beamforming for two γ values. From Fig. 4.9, we can see that all the considered algorithms converge approximately within 10 iterations.

Example 3: comparing distributed to centralized algorithms

It was pointed out earlier that for any given receive beamforming vectors for uplink MSs, there are an optimal transmit beamforming vectors for the downlink MSs. This example shows simulation results comparing S1-Cent-Robust to S1-Dist-Robust for calculating the transmit beamforming vectors when using the first initialization of the receive beamforming vectors, i.e., $\{\mathbf{r}_k(0)\}$. Fig. 4.10 shows the sum downlink power of 25 channel realizations.

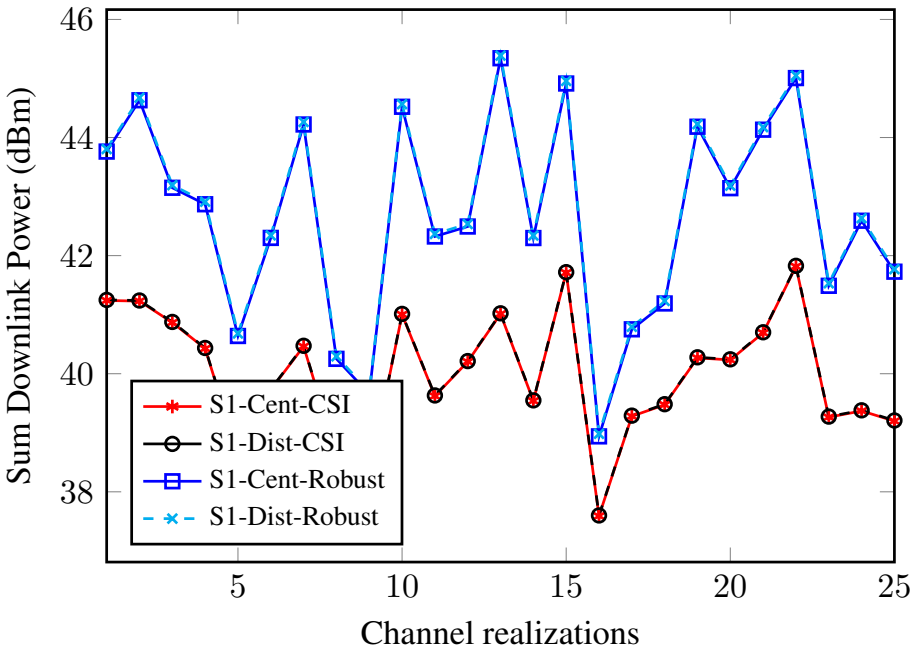
From Fig. 4.10, it can be observed that S1-Dist-Robust achieves *near-optimal* solutions as compared to S1-Cent-Robust. Similar results can be obtained using the distributed algorithm shown in [35] for non-robust beamforming, i.e., assuming perfect CSI. The algorithm uses two iterative steps: in the first step, the local variables are updated using the uplink-downlink duality [24], while in the second step, the global variables are updated using the sub-gradient

Figure 4.9 – Convergence behavior of centralized algorithms while updating the uplink receive beamforming.
$\omega = 25$ dBm, $\epsilon = 0.1$



Source: Created by author.

Figure 4.10 – Comparison between centralized and distributed algorithms with 25 channel realizations.
$\gamma = 5$ dB, $\omega = 25$ dBm, $\epsilon = 0.1$



Source: Created by author.

technique with a step-size α. We refer to this algorithm as S1-Dist-CSI-Ref.

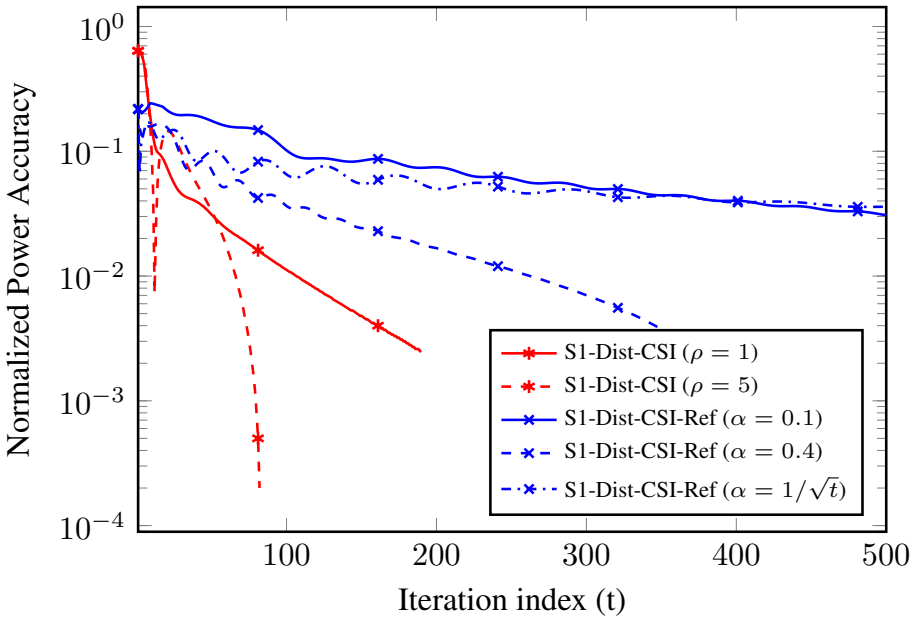
The main difference between S1-Dist-CSI and S1-Dist-CSI-Ref is on the convergence

behavior. In Fig. 4.11, the convergence behavior of both algorithms is shown for one channel realization (rather than the average to expose the convergence behavior of each algorithm) in terms of normalized power accuracy, which is defined as

$$\text{Normalized power accuracy} = \frac{|\mathsf{P}(t) - \mathsf{P}^{\star}|}{\mathsf{P}^{\star}}, \tag{4.65}$$

where $\mathsf{P}(t) = \sum_{n \in \mathcal{M}^{dl}} \mathsf{P}_n(t)$ is the sum of downlink power at iteration (t) with the distributed algorithm and $\mathsf{P}^{\star}$ is the sum of downlink power with the centralized algorithm.

Figure 4.11 – Convergence behavior of distributed algorithms with one channel realization. $\gamma = 5$ dB, $\omega = 25$ dBm



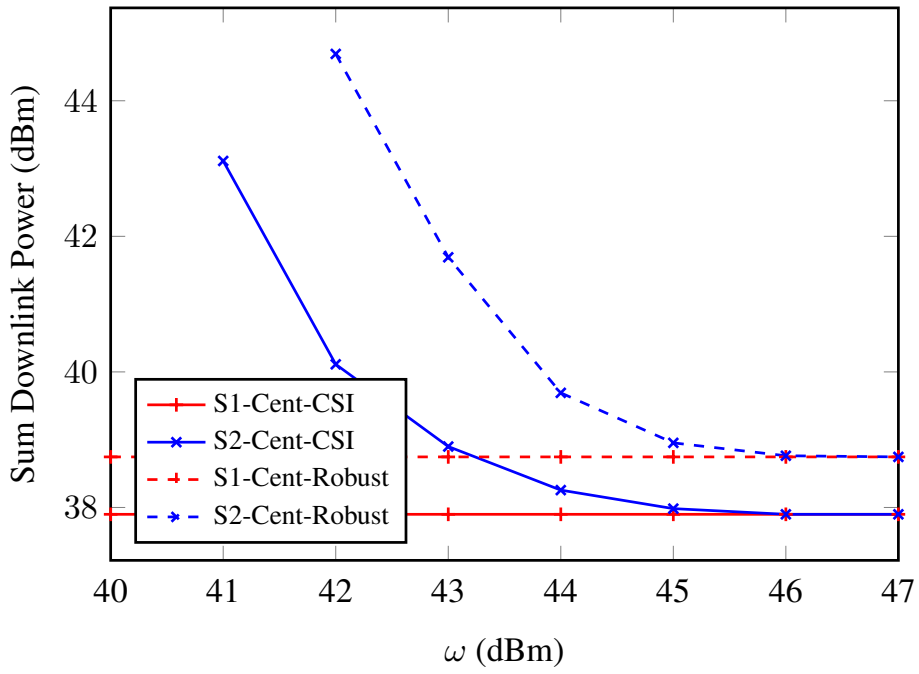
Source: Created by author.

From Fig. 4.11, it can be observed that both distributed approaches converge to *near-optimal* solutions after few iterations for low-to-moderate convergence tolerance value. Nevertheless, it can be seen that S1-Dist-CSI has a faster convergence rate, as compared to S1-Dist-CSI-Ref. We note that the convergence behavior of either algorithm depends on the selection of ρ and α parameters. Unfortunately, the optimal value of either parameter is not known, and, in general, it is dependent on the system-scale and input parameter values.

Example 4: comparing scenario 1 to scenario 2

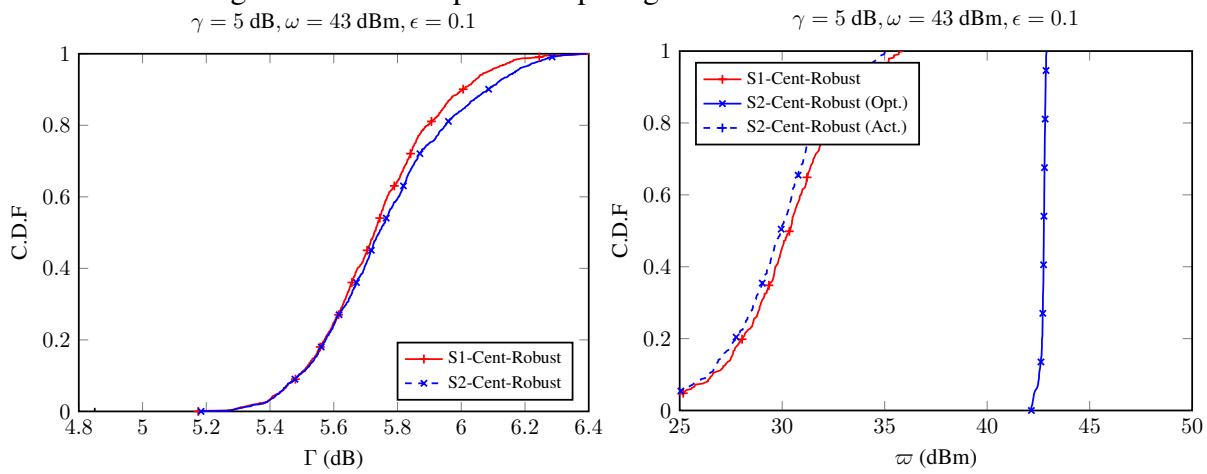
In this example, we show simulation results comparing Scenario 1 to Scenario 2. So far, we have shown that Scenario 2 has advantage over Scenario 1 in terms of signaling overhead (as pointed out in (4.49)), whereas Scenario 1 has advantage over Scenario 2 in terms of energy-efficiency, since it can exploit the receive beamforming vectors (as shown in the previous simulation results). In addition, Fig. 4.12 shows simulation results comparing both scenarios for a range of ω values.

Figure 4.12 – Performance comparison between scenario 1 and scenario 2.
$\gamma = 5$ dB, $\epsilon = 0.1$



Source: Created by author.

Figure 4.13 – CDF plots comparing scenario 1 with scenario 2.



Source: Created by author.

From Fig. 4.12, we can see that Scenario 2 requires much higher ω values to be feasible than Scenario 1. However, after a certain ω value (after 46 dBm for the considered system), both scenarios become equivalent, since the BS to BS interference power thresholds become inactive. Nevertheless, in Scenario 2, each uplink MS would have its BS apply the receive beamforming vector in the actual data reception phase. Fig. 4.13 shows the CDF plots comparing both scenarios for the achievable SINR of downlink MSs and the BS to BS interference power of uplink MSs.

In Fig. 4.13, the Opt/Act result refers to the BS to BS interference power of uplink

MSs before/after applying the receive beamforming vectors (i.e., at the optimization phase and at the actual reception phase). From the figure, it can be seen that while the optimized value of the BS to BS interference power is quite high with Scenario 2, the actual BS to BS interference power received by uplink MSs, after applying the receive beamforming, is much smaller and even closer to that with Scenario 1.

4.11 Chapter conclusions

This chapter considered DTDD wireless networks and proposed a distributed and robust CBF algorithm based on the relaxed semidefinite programming (SDP) and ADMM techniques. The design objective was to minimize the sum power of downlink BSs, while satisfying the worst-case of minimum SINR targets for the downlink MSs and the worst-case maximum interference thresholds for the uplink MSsm where each infinitely nonconvex worst-case constraint is transformed into only one LMI constraint by using the S-Lemma. Detailed simulation results are presented, with a wide range of input parameters and system scale, to investigate the effectiveness of the proposed robust algorithm for interference mitigation in DTDD networks. It is shown that the proposed algorithm outperforms the reference algorithms, where it has a better energy-efficiency and a faster convergence rate.

5 CONCLUSIONS AND FUTURE WORKS

5.1 Algorithms summary and conclusions

This thesis has dealt with the design of decentralized algorithms for multicell multiuser wireless networks. The contributions of this thesis have addressed the following main research topics/problems:

- Problem 1: How to design decentralized transmit beamforming algorithms that maximize the system WSR, while satisfying the power constraints at transmitters.
- Problem 2: How to adaptively select the cells communication directions in order to maximize the users' throughput, while jointly considering their traffic conditions and interference levels.
- Problem 3: How to design a decentralized and robust transmit beamforming algorithm that minimizes the sum transmit power, while satisfying the users' QoS targets in the presence of channel errors.

Problem 1 is an interesting beamforming design that is more in line with the future wireless networks, where users will demand different kinds of applications and services, each with different traffic characteristics (e.g., packet size and maximum packet delay). Therefore, maximizing the system sum rate while considering the users' traffic characteristics is a very desired problem formulation. However, the problem is shown to be non-convex and NP hard, where most of the proposed solutions in the literature approach Problem 1 indirectly by solving the WMMSE minimization problem. The algorithm using this latter approach is often called WSR-WMMSE. Differently, the proposed algorithms in this thesis approached Problem 1 directly by investigating its KKT conditions. In particular, this thesis proposed three different novel algorithms for solving Problem 1 based on the alternating optimization technique, which are guaranteed to converge to a local WSR-optimum. The first algorithm is an interference pricing approach, where each cell maximizes its own utility that is formed by the local users' WSR minus the priced ICI leakage. The second algorithm designs transceivers that maximize the network-wide WSR. The third algorithm is an implicit interference pricing approach, where each cell self-prices its ICI leakage and, thus, does not require variables feedback between cells. The performance of the three proposed algorithms was investigated using numerical examples, where it was shown that the proposed algorithms have better sum rate and faster convergence rate, as compared to the WSR-WMMSE and some other state-of-art algorithms. Furthermore, a novel OTA signaling scheme based on TDD mode was also proposed to facilitate the algorithms' implementation, which reduces the signaling overhead and requires no backhaul feedback, as compared to some existing signaling schemes. The main conclusions that can be drawn from this chapter are summarized as follows:

- Problem 1 is an important beamforming design that is more applicable for future cellular networks, since it has the ability to prioritize users according to, e.g., their traffic characteristics, and has implicit users and streams selection mechanism. Further, it is shown to outperform the state-of-art beamforming design approaches, like MRT and ZF/BD.

- There exists a direct solution to Problem 1, which can be achieved by investigating its KKT conditions with help of Lemma 1. This is in contrast to the generally used approach in the literature that indirectly solves Problem 1 by solving the WMMSE minimization problem, which can be made equivalent to the original WSR maximization problem by adaptively adjusting the users' weights.

- Also, the WSR maximization via interference pricing approach can be made equivalent to the network-wide WSR maximization, whenever the mobile stations are equipped with single-antenna.

Problem 2 targets a question that is most probably the first question raised since the introduction of DTDD technique. Earlier works targeting Problem 3 reconfigured each cell direction based only on the aggregate traffic in the cell. This approach is fairly simple and inherently distributed, but it cannot achieve the potential performance as it disregards the interference effects that are particularly severe in DTDD systems. Other solutions were proposed that account for the users' traffic demands and interference levels, as well. However, most of those solutions are scenario (traffic model) specific and do not consider the individual user's traffic characteristics. Therefore, it is important for the cell reconfiguration algorithm to support such different traffic characteristics. This thesis proposed a novel cell reconfiguration formulation to maximizes the users' throughput, while jointly considering both the prevailing traffic conditions and multicell interference levels. Realistic system level simulations indicate that the proposed scheme outperforms not only the static TDD system but also other reference schemes, that disregard the DTDD specific interference effects. The main conclusions that can be drawn from this chapter are summarized as follows:

- DTDD technique enhances the system spectral and energy efficiencies, as compared to STDD, especially in scenarios in which the offered traffic is time-varying and asymmetric in terms of uplink/downlink direction.

- It is important for the cell reconfiguration algorithm to jointly consider the users' traffic characteristics and interference levels when selecting the cells directions.

Problem 3 approaches the transmit beamforming design from a different perspective than Problem 1. The future wireless cellular networks are expected to be much more densified with small cells as compared to current cellular networks. Therefore, minimizing the transmit power while satisfying the users' QoS targets will be a critical optimization problem for cellular

networks design. The problem itself is very well known in the literature, where many solutions can be found. However, their extension to DTDD wireless networks is not direct, due to more complicated interference situations in DTDD networks, since the uplink and downlink users coexist at the same time among neighboring cells. Therefore, Problem 2 in DTDD networks requires a special consideration from the optimization viewpoint. A possible solution is to formulate the optimization problem in DTDD as it is generally formulated in the CR networks, i.e., by assuming the uplink cells are the *primary* cells and the downlink cells are the *secondary* cells and then include a threshold on the maximum BS-to-BS interference power from the downlink to uplink cells. In this case, not only the required downlink performance can be guaranteed, but also the required uplink performance. However, due to the fact that the transmitters can never have perfect CSI (due to, e.g., estimation errors and limited feedback channels), the performance targets can no longer be guaranteed if the CSI errors are not taken into account. Therefore, robust and fully distributed transmit beamforming solutions are much desired. To this end, this thesis proposed a novel distributed and robust CBF algorithm based on the relaxed SDP and ADMM techniques for solving Problem 2, where the robust beamforming was tackled using a worst-case optimization criterion. From simulation results, it was shown that the proposed algorithm outperforms some reference algorithms by achieving a better energy-efficiency and a faster convergence rate. The main conclusions that can be drawn from this chapter are summarized as follows:

- Cross-link interference imposes great challenge to DTDD systems, which can severely degrade the system performance, especially the uplink cells' performance, if not properly managed.
- Robust optimization techniques are critical for beamforming designs to guarantee the users' QoS targets, due to unavoidable CSI estimation errors.
- There is a trade-off between system energy-efficiency and the beamforming design algorithms' complexity and signaling overhead. Iteratively updating the transmit and receive beamforming vectors provides more system energy-efficiency but increases the algorithms' complexity and signaling overhead.
- ADMM is a powerful dual decomposition technique that is perfectly suited for distributed constrained optimization problems, which has superior convergence properties and numerical robustness as compared to many other decomposition techniques, thus making it a favorable choice for practical implementations.

5.2 Future works

The contents of this thesis can be extended in many directions. In the following, some of the possible future work directions are pointed out.

5.2.1 Effects of CSI estimation errors

The proposed algorithms in chapter 2 assumed perfect knowledge of CSI, which is impossible to have in the practical implementation due to unavoidable CSI estimation errors. CSI estimation has two main effects on system performance. First, it has affected on resources utilization, since some resources will be used for the CSI estimation, thus, reducing the data transmission resources. Second, the CSI estimation errors have a great impact on the beamforming design quality and system performance. Although those two aspects will not change the main conclusions of the chapter, since all the proposed and reference algorithms will face the same CSI quality, it is still interesting to investigate the effects of CSI estimation errors on system performance.

5.2.2 Maximize WSR in DTDD systems

The proposed algorithms in chapter 2 for the WSR maximization can be readily implemented in DTDD systems by treating the cross-link interference (BS to BS and MS to MS interference) as an inter-cell interference. However, we have noticed that if the downlink and uplink transmit powers highly differ, the uplink performance can be highly degraded. A possible solution is to include an interference-power threshold from downlink cells to uplink cells, similar to the considered approach in chapter 4.

5.2.3 Massive MIMO and millimeter waves

The proposed algorithms in chapters 2 and 4 can be extended to consider use of massive MIMO setup, i.e., very large number of antennas at transmitter and/or receiver. The use of massive MIMO has become a key enabler to meet the data rate demands of the future cellular systems [1]. Unfortunately, the bottleneck that limits the successful incorporation of massive MIMO into cellular networks is the large physical size of the antenna arrays at currently used cellular frequencies (below 6 GHz). For this reason, massive MIMO is being considered in conjunction with mmWave frequencies [3], where antenna arrays of reasonable physical sizes are feasible. From a signal processing perspective, transmission with massive arrays increases the processing complexity derived from the computation of the required beamformers when digital schemes are applied to hundreds of antennas. Multi-antenna digital beamforming is carried out at baseband and, thus, it requires an architecture with as many radio-frequency (RF) chains as antenna ports to do the digital to analog data conversion and subsequent upmixing to RF. However, the use of one RF chain per antenna is not only a very costly option for massive MIMO systems, but it also leads to an extremely high power consumption. Motivated by this, there is now a growing interest in alternative transmission schemes based on analog beamforming and hybrid beamforming architectures, where all or part of the processing is based on RF beamforming. However, RF beamforming adds hardware constraints (see [92] for more details), which should be considered when designing the beamforming solutions.

5.2.4 LTE-based system evaluations

The simulation results in chapter 3 assumed that each cell can change its communication direction in every TTI. This, however, is not realistic, as cells are allowed to change their communication directions at minimum every 10 TTIs according to 3GPP TR in [15]. Further, the simulation results can also be extended to consider an LTE-TDD-like simulator, which comprises 7 subframes configurations [13], as shown in Table 5.1. For instance, we can select one configuration for each cell that better matches the output of the proposed algorithms. Furthermore, different and more optimal users' weights optimization approaches can also be investigated.

Table 5.1 – TDD frames configurations.

Configuration Number	Subframe Number										UL-DL
-	0	1	2	3	4	5	6	7	8	9	-
0	D	S	U	U	U	D	S	U	U	U	6-2
1	D	S	U	U	D	D	S	U	U	D	4-4
2	D	S	U	D	D	D	S	U	D	D	2-6
3	D	S	U	U	U	D	D	D	D	D	3-6
4	D	S	U	U	D	D	D	D	D	D	2-7
5	D	S	U	D	D	D	D	D	D	D	1-8
6	D	S	U	U	U	D	S	U	U	D	5-3

Source: Created by author.

BIBLIOGRAPHY

1 ANDREWS, J. G. et al. What Will 5G Be? **IEEE J. Sel. Areas Commun.**, v. 32, n. 6, p. 1065–1082, June 2014.

2 BJÖRNSON, E.; HOYDIS, J.; SANGUINETTI, L. Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency. **Foundations and Trends® in Signal Processing**, v. 11, n. 3-4, p. 154–655, 2017.

3 PI, Z.; KHAN, F. An introduction to millimeter-wave mobile broadband systems. **IEEE Communications Magazine**, v. 49, n. 6, p. 101–107, June 2011.

4 KHANDEKAR, A. et al. **LTE-Advanced: Heterogeneous networks**. In: PROC. European Wireless Conference (EW). [S.l.: s.n.], Apr. 2010. p. 978–982.

5 STANZE, O.; WEBER, A. Heterogeneous Networks With LTE-Advanced Technologies. **Bell Labs Technical Journal**, v. 18, n. 1, p. 41–58, 2013.

6 AR, E. T.; TELATAR, I. E. Capacity of Multi-antenna Gaussian Channels. **European Transactions on Telecommunications**, v. 10, p. 585–595, June 1999.

7 BJÖRNSON, E.; BENGTSSON, M.; OTTERSTEN, B. Optimal Multiuser Transmit Beamforming: A Difficult Problem with a Simple Solution Structure [Lecture Notes]. **IEEE Signal Process. Mag.**, v. 31, n. 4, p. 142–148, July 2014.

8 WEINGARTEN, H.; STEINBERG, Y.; SHAMAI, S. S. The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel. **IEEE Trans. Inf. Theory**, v. 52, n. 9, p. 3936–3964, Sept. 2006. ISSN 0018-9448.

9 LI, J. et al. Coordinated beamforming design using duality theory with dynamic cooperation clusters. **IET Communications**, v. 6, n. 12, p. 1662–1669, Aug. 2012. ISSN 1751-8628.

10 SCHMIDT, D. A. et al. Comparison of Distributed Beamforming Algorithms for MIMO Interference Networks. **IEEE Transactions on Signal Processing**, v. 61, n. 13, p. 3476–3489, July 2013.

11 BOGALE, T. E.; VANDENDORPE, L. **Weighted sum rate optimization for downlink multiuser MIMO systems with per antenna power constraint: Downlink-uplink duality approach**. In: PROC. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). [S.l.: s.n.], Mar. 2012. p. 3245–3248.

12 CHRISTENSEN, S. S. et al. **Weighted Sum-Rate Maximization Using Weighted MMSE for MIMO-BC Beamforming Design**. In: PROC. IEEE International Conference on Communications. [S.l.: s.n.], June 2009. p. 1–6.

13 CHEN, S. et al. A comprehensive survey of TDD-based mobile communication systems from TD-SCDMA 3G to TD-LTE(A) 4G and 5G directions. **China Communications**, v. 12, n. 2, p. 40–60, Feb. 2015. ISSN 1673-5447.

14 SHEN, Z. et al. Dynamic uplink-downlink configuration and interference management in TD-LTE. **IEEE Communications Magazine**, v. 50, n. 11, p. 51–59, Nov. 2012. ISSN 0163-6804.

15 3GPP. **Evolved Universal Terrestrial Radio Access (E-UTRA): Further enhancements to LTE Time Division Duplex (TDD) for Downlink-Uplink (DL-UL) interference management and traffic adaptation**: Technical Specification Group Radio Access Network. [S.l.], June 2012. (Release 11, 36.828).

16 GUPTA, A. K. et al. **Rate analysis and feasibility of dynamic TDD in 5G cellular systems**. In: PROC. IEEE Int. Conf. on Commun. (ICC). [S.l.: s.n.], May 2016. p. 1–6.

17 WANG, C. et al. **HARQ Signalling Design for Dynamic TDD System.** In: PROC. IEEE Vehicular Technology Conference (VTC). [S.l.: s.n.], Sept. 2014. p. 1–5.

18 ERIKSSON, E. Dynamic TDD for Enhanced Traffic Adaptation. **IEEE Commun. Mag.**, v. 56, n. 4, p. 45–53, July 2013.

19 SUN, F.; ZHAO, Y.; SUN, H. **Centralized Cell Cluster Interference Mitigation for Dynamic TDD DL/UL Configuration with Traffic Adaptation for HTN Networks**. In: PROC. IEEE Veh. Technol. Conf. (VTC). [S.l.: s.n.], Sept. 2015. p. 1–5.

20 ELBAMBY, M. S. et al. **Dynamic uplink-downlink optimization in TDD-based small cell networks**. In: PROC. Int. Symposium on Wireless Commun. Systems (ISWCS). [S.l.: s.n.], Aug. 2014. p. 939–944.

21 LIN, Y.-T.; CHAO, C.-C.; WEI, H.-Y. **Dynamic TDD interference mitigation by using Soft Reconfiguration**. In: PROC. Int. Conf. on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QSHINE). [S.l.: s.n.], Aug. 2015. p. 352–357.

22 NGMN. NGMN Radio Access Performance Evaluation Methodology. Version 1.3: **NGMN White Paper**, 2007.

23 UFC.42 TEAM. **Fourth Report: Interference Management for Super-Dense Scenarios**. [S.l.], Oct. 2016.

24 DAHROUJ, H.; YU, W. Coordinated beamforming for the multicell multi-antenna wireless system. **IEEE Trans. Wireless Commun.**, v. 9, n. 5, p. 1748–1759, May 2010. ISSN 1536-1276.

25 GESBERT, D. et al. Multi-Cell MIMO Cooperative Networks: A New Look at Interference. **IEEE J. Sel. Areas Commun.**, v. 28, n. 9, p. 1380–1408, Dec. 2010.

26 LOVE, D. J. et al. An overview of limited feedback in wireless communication systems. **IEEE J. Sel. Areas Commun.**, v. 26, n. 8, p. 1341–1365, Oct. 2008. ISSN 0733-8716.

27 SHEN, C. et al. Distributed Robust Multicell Coordinated Beamforming With Imperfect CSI: An ADMM Approach. **IEEE Trans. Signal Process.**, v. 60, n. 6, p. 2988–3003, June 2012.

28 ZHENG, G.; WONG, K. K.; OTTERSTEN, B. Robust Cognitive Beamforming With Bounded Channel Uncertainties. **IEEE Transactions on Signal Processing**, v. 57, n. 12, p. 4871–4881, Dec. 2009.

29 GERSHMAN, A. B. et al. Convex optimization-based beamforming: From receive to transmit and network designs. **IEEE Signal Processing Magazine**, v. 27, n. 3, p. 62–75, 2010.

30 BOYD, S. et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. **Foundations and Trends in Machine Learning**, Now Publishers Inc., v. 3, n. 1, p. 1–122, 2011.

31 SABERI, N. Throughput Maximizing Frequency and Power Scheduling for Wireless Ad-Hoc Networks in the Low-SINR Regime. **Wireless Engineering and Technology**, v. 3, n. 3, p. 106–112, 2012.

32 EBERHART, R.; KENNEDY, J. **A new optimizer using particle swarm theory**. In: IN Proc. of the Sixth International Symposium on Micro Machine and Human Science. [S.l.: s.n.], Oct. 1995. p. 39–43.

33 BOYD, S.; VANDENBERGHE, L. **Convex Optimization**. [S.l.]: Camebridge University Press, 2004.

34 TAJER, A.; PRASAD, N.; WANG, X. Robust Linear Precoder Design for Multi-Cell Downlink Transmission. **IEEE Transactions on Signal Processing**, v. 59, n. 1, p. 235–251, Jan. 2011.

35 PENNANEN, H.; TÖLLI, A.; LATVA-AHO, M. Multi-Cell Beamforming With Decentralized Coordination in Cognitive and Cellular Networks. **IEEE Transactions on Signal Processing**, v. 62, n. 2, p. 295–308, Jan. 2014.

36 ZHENG, L.; TSE, D. N. C. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. **IEEE Trans. Inf. Theory**, v. 49, n. 5, p. 1073–1096, May 2003.

37 SPENCER, Q. H.; SWINDLEHURST, A. L.; HAARDT, M. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. **IEEE Trans. Signal Process.**, v. 52, n. 2, p. 461–471, Feb. 2004. ISSN 1053-587X.

38 SHIM, S. et al. Block diagonalization for multi-user MIMO with other-cell interference. **IEEE Trans. Wireless Commun.**, v. 7, n. 7, p. 2671–2681, July 2008. ISSN 1536-1276.

39 ARDAH, K.; SILVA, Y. C. B.; CAVALCANTI, F. R. P. **Block Diagonalization For Multicell Multiuser MIMO Systems with Other-Cell Interference**. In: PROC. Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT). [S.l.: s.n.], Sept. 2016. p. 1–5.

40 SONG, B.; CRUZ, R. L.; RAO, B. D. Network Duality for Multiuser MIMO Beamforming Networks and Applications. **IEEE Trans. Commun.**, v. 55, n. 3, p. 618–630, Mar. 2007. ISSN 0090-6778.

41 JOSHI, S.; CODREANU, M.; LATVA-AHO, M. **Distributed resource allocation for MISO downlink systems via the alternating direction method of multipliers**. In: PROC. Asilomar Conference on Signals, Systems and Computers (ASILOMAR). [S.l.: s.n.], Nov. 2012. p. 488–493.

42 _____. **Distributed SINR balancing for MISO downlink systems via the alternating direction method of multipliers**. In: PROC. 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt). [S.l.: s.n.], May 2013. p. 318–325.

43 JOSE, J. et al. Pilot Contamination and Precoding in Multi-Cell TDD Systems. **IEEE Trans. Wireless Commun.**, v. 10, n. 8, p. 2640–2651, Aug. 2011. ISSN 1536–1276.

44 CODREANU, M. et al. **MIMO Downlink Weighted Sum Rate Maximization with Power Constraints per Antenna Groups**. In: PROC. IEEE Vehicular Technology Conference - Spring. [S.l.: s.n.], Apr. 2007. p. 2048–2052.

45 VENTURINO, L.; PRASAD, N.; WANG, X. Coordinated linear beamforming in downlink multi-cell wireless networks. **IEEE Trans. Wireless Commun.**, v. 9, n. 4, p. 1451–1461, Apr. 2010. ISSN 1536–1276.

46 PARK, S. H. et al. New Beamforming Techniques Based on Virtual SINR Maximization for Coordinated Multi-Cell Transmission. **IEEE Trans. Wireless Commun.**, v. 11, n. 3, p. 1034–1044, Mar. 2012. ISSN 1536-1276.

47 GARZÁS, J. J. E. et al. Interference Pricing Mechanism for Downlink Multicell Coordinated Beamforming. **IEEE Trans. Commun.**, v. 62, n. 6, p. 1871–1883, June 2014. ISSN 0090-6778.

48 BJÖRNSON, E. et al. Robust Monotonic Optimization Framework for Multicell MISO Systems. **IEEE Trans. Signal Process.**, v. 60, n. 5, p. 2508–2523, May 2012. ISSN 1053-587X.

49 SHIN, J.; MOON, J. Weighted-Sum-Rate-Maximizing Linear Transceiver Filters for the K-User MIMO Interference Channel. **IEEE Trans. Commun.**, v. 60, n. 10, p. 2776–2783, Oct. 2012. ISSN 0090-6778.

50 SHI, Q. et al. An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel. **IEEE Trans. Signal Process.**, v. 59, n. 9, p. 4331–4340, Sept. 2011. ISSN 1053-587X.

51 JAYASINGHE, P. et al. **Bi-directional signaling for dynamic TDD with decentralized beamforming**. In: IEEE International Conference on Communication Workshop (ICCW). [S.l.: s.n.], June 2015. p. 185–190.

52 LAGEN, S.; AGUSTIN, A.; VIDAL, J. Decentralized Coordinated Precoding for Dense TDD Small Cell Networks. **IEEE Trans. Wireless Commun.**, v. 14, n. 8, p. 4546–4561, Aug. 2015. ISSN 1536-1276.

53 KOMULAINEN, P.; TÖLLI, A.; JUNTTI, M. Effective CSI Signaling and Decentralized Beam Coordination in TDD Multi-Cell MIMO Systems. **IEEE Trans. Signal Process.**, v. 61, n. 9, p. 2204–2218, May 2013. ISSN 1053-587X.

54 HUNGER, R. et al. **Alternating Optimization for MMSE Broadcast Precoding**. In: IEEE International Conference on Acoustics Speech and Signal Processing Proceedings. [S.l.: s.n.], May 2006. v. 4, p. 757–760.

55 KOBAYASHI, M.; JINDAL, N.; CAIRE, G. Training and Feedback Optimization for Multiuser MIMO Downlink. **IEEE Trans. Commun.**, v. 59, n. 8, p. 2228–2240, Aug. 2011. ISSN 0090–6778.

56 BIGUESH, M.; GERSHMAN, A. B. Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals. **IEEE Trans. Signal Process.**, v. 54, n. 3, p. 884–893, Mar. 2006.

57 PALOMAR, D. P.; FONOLLOSA, J. R. Practical algorithms for a family of waterfilling solutions. **IEEE Trans. Signal Process.**, v. 53, n. 2, p. 686–695, Feb. 2005. ISSN 1053-587X.

58 SHI, C. et al. **Distributed Interference Pricing for the MIMO Interference Channel**. In: PROC. IEEE International Conference on Communications. [S.l.: s.n.], June 2009. p. 1–5.

59 PETERSEN, K. B.; PEDERSEN, M. S. **The Matrix Cookbook**. [S.l.]: Technical University of Denmark, Nov. 2012. Available from: <http://www2.imm.dtu.dk/pubdb/p.php?3274>.

60 BERTSEKAS, D. **Nonlinear Programming**. [S.l.]: Athena Scientific, 1999.

61 LIANG, L. et al. A Cluster-Based Energy-Efficient Resource Management Scheme for Ultra-Dense Networks. **IEEE Access**, v. 4, p. 6823–6832, 2016. ISSN 2169-3536.

62 GOTSIS, A. G.; ALEXIOU, A. **Global network coordination in densified wireless access networks through integer linear programming**. In: IEEE Annual Int. Symposium on Personal, Indoor, and Mobile Radio Commun. [S.l.: s.n.], 2013. p. 1548–1553.

63 SIERKSMA, G. **Linear and Integer Programming: Theory and Practice**. [S.l.]: Taylor & Francis, 2001.

64 ZHANG, Y. J.; LETAIEF, K. B. Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems. **IEEE Trans. Wireless Commun.**, v. 3, n. 5, p. 1566–1575, Sept. 2004. ISSN 1536-1276.

65 EBERHART, R.; Y.SHI; KENNEDY, J. **Swarm Intelligence**. [S.l.]: Morgan Kaufmann, 2001.

66 MEISSNER, M.; SCHMUKER, M.; SCHNEIDER, G. Optimized Particle Swarm Optimization (OPSO) and its application to artificial neural network training. **BMC Bioinformatics**, v. 7, n. 1, p. 125, Mar. 2006.

67 MOHAMMADI, N.; MIRABEDINI, S. Comparison of particle swarm optimization and back propagation algorithms for training feed forward neural network. **J. Math. Computer Sci.**, v. 12, n. 2, p. 113–123, 2014.

68 HADAVANDI, E.; GHANBARI, A.; ABBASIAN-NAGHNEH, S. **Developing a Time Series Model Based on Particle Swarm Optimization for Gold Price Forecasting**. In: THIRD International Conference on Business Intelligence and Financial Engineering. [S.l.: s.n.], Aug. 2010. p. 337–340.

69 YANG, X.-S.; DEB, S.; FONG, S. Accelerated Particle Swarm Optimization and Support Vector Machine for Business Optimization and Applications. In: **In Proc. Networked Digital Technologies: Third International Conference**. Ed. by Simon Fong. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 53–66. ISBN 978-3-642-22185-9.

70 BERGH, F. van den; ENGELBRECHT, A. P. A Cooperative approach to particle swarm optimization. **IEEE Transactions on Evolutionary Computation**, v. 8, n. 3, p. 225–239, June 2004.

71 MARINI, F.; WALCZAK, B. Particle swarm optimization (PSO). A tutorial. **Chemometrics and Intelligent Laboratory Systems**, v. 149, Part B, p. 153–165, 2015. ISSN 0169-7439.

72 KENNEDY, J.; EBERHART, R. C. **A discrete binary version of the particle swarm algorithm**. In: PROC. IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation. [S.l.: s.n.], Oct. 1997. v. 5, 4104–4108 vol.5.

73 OZCAN, E.; MOHAN, C. Analysis of a simple particle swarm optimization system. **Intelligent Engineering Systems Through Artificial Neural Networks**, v. 1998, p. 253–258, 1998.

74 AHMED, Q. Z. et al. Minimizing the Symbol-Error-Rate for Amplify-and-Forward Relaying Systems Using Evolutionary Algorithms. **IEEE Transactions on Communications**, v. 63, n. 2, p. 390–400, Feb. 2015. ISSN 0090-6778.

75 PTZOLD, M. **Mobile Radio Channels**. 2nd. [S.l.]: WILEY, 2012. ISBN 0470517476, 9780470517475.

76 FAN, J. et al. **MCS Selection for Throughput Improvement in Downlink LTE Systems**. In: INT. Conf. on Computer Commun. and Networks. [S.l.: s.n.], July 2011. p. 1–5.

77 YU, B.; ISHII, H.; YANG, L. **System Level Performance Evaluation of Dynamic TDD and Interference Coordination in Enhanced Local Area Architecture**. In: PROC. IEEE 77th Vehicular Technology Conference (VTC Spring). [S.l.: s.n.], June 2013. p. 1–6.

78 CHOI, Y. S.; SOHN, I.; LEE, K. B. **A novel decentralized time slot allocation algorithm in dynamic TDD system**. In: CCNC 2006. 2006 3rd IEEE Consumer Communications and Networking Conference, 2006. [S.l.: s.n.], Jan. 2006. v. 2, p. 1268–1272.

79 TAKAHASHI, H.; YOKOMAKURA, K.; IMAMURA, K. **A Transmit Power Control Based Interference Mitigation Scheme for Small Cell Networks Using Dynamic TDD in LTE-Advanced Systems**. In: PROC. IEEE Vehicular Technology Conference (VTC). [S.l.: s.n.], May 2014. p. 1–5.

80 NA, C.; HOU, X.; JIANG, H. **Interference alignment based dynamic TDD for small cells**. In: PROC. IEEE Globecom Workshops (GC Wkshps). [S.l.: s.n.], Dec. 2014. p. 700–705.

81 PATCHARAMANEEPAKORN, P.; ARMOUR, S.; DOUFEXI, A. Coordinated beamforming schemes based on modified signal-to-leakage-plus-noise ratio precoding designs. **IET Communications**, v. 9, n. 4, p. 558–567, 2015.

82 BENGTSSON, M.; OTTERSTEN, B. Optimal and suboptimal transmit beamforming. In: HANDBOOK of Antennas in Wireless Communications. [S.l.]: CRC Press, 2001.

83 ZHANG, Y. J. A.; SO, A. M. C. Optimal Spectrum Sharing in MIMO Cognitive Radio Networks via Semidefinite Programming. **IEEE Journal on Selected Areas in Communications**, v. 29, n. 2, p. 362–373, Feb. 2011.

84 PALOMAR, D. P.; CHIANG, M. A tutorial on decomposition methods for network utility maximization. **IEEE Journal on Selected Areas in Communications**, v. 24, n. 8, p. 1439–1451, Aug. 2006.

85 JOSHI, S. K. et al. Maximization of Worst-Case Weighted Sum-Rate for MISO Downlink Systems With Imperfect Channel Knowledge. **IEEE Transactions on Communications**, v. 63, n. 10, p. 3671–3685, Oct. 2015.

86 ZHANG, Y.; DALL'ANESE, E.; GIANNAKIS, G. B. **Distributed robust beamforming for MIMO cognitive networks**. In: PROC. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). [S.l.: s.n.], Mar. 2012. p. 2953–2956.

87 WONG, K.-K.; ZHENG, G.; NG, T.-S. **Convergence analysis of downlink MIMO antenna systems using second-order cone programming**. In: 62. PROC. IEEE Vehicular Technology Conference (VTC). [S.l.: s.n.], Sept. 2005. v. 1, p. 492–496.

88 STURM, J. F. Using SeDuMi 1.02, A MATLAB toolbox for optimization over symmetric cones. **Optimization Methods and Software**, v. 11, n. 1–4, p. 625–653, 1999.

89 GRANT, M.; BOYD, S.; YE, Y. **CVX: MATLAB software for disciplined convex programming**. [S.l.], 2008.

90 HE, J.; SALEHI, M. **Low-Complexity Coordinated Interference-Aware Beamforming for MIMO Broadcast Channels**. In: PROC. IEEE Vehicular Technology Conference (VTC). [S.l.: s.n.], Oct. 2007. p. 685–689.

91 AI, W.; HUANG, Y.; ZHANG, S. New Results on Hermitian Matrix Rank-one Decomposition. **Mathematical Programming**, v. 128, n. 1-2, p. 253–283, June 2011.

92 ALKHATEEB, A. et al. MIMO Precoding and Combining Solutions for Millimeter-Wave Systems. **IEEE Communications Magazine**, v. 52, n. 12, p. 122–131, Dec. 2014.