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

EDUARDO DE OLIVINDO CAVALCANTE

DECENTRALIZED STRATEGIES FOR BEAMFORMING OPTIMIZATION IN WIRELESS MOBILE NETWORKS

FORTALEZA
2023

EDUARDO DE OLIVINDO CAVALCANTE

DECENTRALIZED STRATEGIES FOR BEAMFORMING OPTIMIZATION IN 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. Yuri Carvalho Barbosa Silva

FORTALEZA
2023

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

C364d Cavalcante, Eduardo de Olivindo.
Decentralized Strategies for Beamforming Optimization in Wireless Mobile Networks / Eduardo de Olivindo Cavalcante. – 2023.
116 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2023.
Orientação: Prof. Dr. Yuri Carvalho Barbosa Silva.

1. Decentralização. 2. Formatação de feixes. 3. Escalonamento de usuários. 4. TDD dinâmico. 5. Aprendizagem de máquina. I. Título.

CDD 621.38

EDUARDO DE OLIVINDO CAVALCANTE

DECENTRALIZED STRATEGIES FOR BEAMFORMING OPTIMIZATION IN 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

Aprovada em: 30 de Janeiro de 2023.

BANCA EXAMINADORA

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

Prof. Dr. Walter da Cruz Freitas Júnior
Universidade Federal do Ceará

Dr. Roberto Pinto Antonioli
Instituto Atlântico

Prof. Dr. Vicente Angelo de Sousa Junior
Universidade Federal do Rio Grande do Norte

Prof. Dr. Aldebaro Barreto da Rocha Klautau Júnior
Universidade Federal do Pará

To you João and Joselita,
for all the sacrifices you have made to educate us.
Here is an outcome of your effort.

ACKNOWLEDGEMENTS

First of all, I give praise to God who guided my steps when I felt lost, gave me strength when I felt weak, provided me patience when I felt anxious, and granted me peace when I felt troubled.

I also express my appreciation to all of you who, in any form, supported me throughout all those years without hesitating or asking anything back. This would never be possible without you.

The apogee of my gratitude belongs to my family. I am grateful for the unconditional love, care and prayers from my parents, João and Joselita, and brothers, Eider and Erivan. I am utmost thankful to my beloved wife, Rebeca, who was by my side in the most troubled moments, your love, understanding and support made me stronger. I love you.

I am very grateful to my advisor Prof. Yuri Carvalho Barbosa Silva and to Prof. Dr. Walter da Cruz Freitas Júnior and Dr. Gábor Fodor, who have a great part, not only in this work but also in my development as a researcher and professor. Without your great effort none of this would be possible.

I am thankful to my colleagues from UFC and GTEL for the productive talks, advice and suggestions that helped me in the development of this work. Special thanks to my project team, your contribution to this work is immeasurable.

I am also thankful to my colleagues from IFCE who taught me to be patient and motivated me to persist.

Finnaly, I acknowledge the technical and financial support from the Ericsson Innovation Center, Brazil, under UFC.46 and UFC.48 Technical Cooperation Contracts Ericsson/UFC.

A educação, direito de todos e dever do Estado e da família, será promovida e incentivada com a colaboração da sociedade, visando ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania ...
(Constituição Federal de 1988)

RESUMO

O sempre-crescente aumento em demanda de redes móveis sem fio impulsiona a evolução e molda futuras tecnologias de comunicação. As próximas gerações de rede devem lidar com requisitos cada vez mais rigorosos em relação ao número de dispositivos conectados, taxa de transferência de dados, latência e confiabilidade. Dentre as estratégias para aumentar a eficiência da rede, destacamos a formatação de feixes, que visa focar um feixe de sinal em uma direção específica evitando outras, e o escalonamento de usuários, que otimiza a seleção de usuários a transmitir utilizando um determinado recurso. Algoritmos eficientes ao ponto de vista da rede devem considerar otimização cooperativa de forma a coordenar as entidades da rede em direção a um objetivo global comum. Devido à natureza distribuída dos sistemas em rede, uma coordenação centralizada em muitos casos não é adequada e soluções descentralizadas são desejadas. Com base nisso, o principal objetivo desta tese é projetar estratégias descentralizadas para otimizar a formatação de feixes e o escalonamento de usuários em cenários de sistemas móveis de próximas gerações. Dividimos o conteúdo desta tese em duas partes independentes: Na primeira, focamos em um cenário multi-célula, multi-usuário, multi-fluxo, MIMO (do inglês, *multiple-input multiple-output*) com duplexação por divisão de tempo (TDD) dinâmica, onde consideramos a otimização da formatação de feixes bidirecional com o objetivo de minimizar a soma do consumo de energia da rede com restrições de relação sinal-interferência mais ruído (SINR) por fluxo. Uma primeira abordagem para solução deste problem assume processamento centralizado e requer a disponibilidade de informação de estado de canal (CSI) global. Uma segunda abordagem é realizada de forma descentralizada, baseada em ADMM (do inglês, *alternating direction method of multipliers*) e requer CSI local e uma carga de sinalização reduzida. Ambas as abordagens convergem para um gasto mínimo de energia da rede, enquanto um desempenho próximo ao ideal pode ser obtido ao limitar o número de iterações. Na segunda parte, focamos no enlace direto (DL) de uma rede multi-célula, multi-usuário, MISO (do inglês, *multiple-input single-output*) na presença de erros de CSI, onde consideramos o desenvolvimento de dois esquemas de escalonamento de usuários com execução distribuída baseados em DQL (do inglês, *deep Q-learning*) multi-agente para resolver o problema de formatação de feixes com intuito de maximizar a soma das taxas com restrições de potência por base. Os resultados das simulações mostram que na presença de erros de CSI os esquemas propostos superam os algoritmos de última geração tanto em termos de eficiência espectral média, quanto de tempo de execução.

Palavras-chave: decentralização; formatação de feixes; escalonamento de usuários; TDD dinâmico; aprendizado de máquina.

ABSTRACT

The ever-growing demand increase of wireless mobile networks drives the evolution and shapes future communication technologies. Next network generations must deal with increasingly stringent requirements regarding the number of connected devices, data throughput, latency, and reliability. Among strategies to enhance network efficiency, we highlight beamforming, which aims to focus a signal beam in a specific direction while avoiding others, and user scheduling, which optimizes the selection of users to transmit using a given resource. Network-wide efficient algorithms must consider cooperative optimization in order to coordinate network entities into a common global objective. Due to the distributed nature of networked systems, centralized coordination is in many cases not suitable, and decentralized solutions are desired. Based on that, the main objective of this thesis is to design decentralized strategies for optimizing beamforming and user scheduling in scenarios of next-generation mobile systems. We divide this thesis contents into two independent parts: In the first, we focus on a multi-cell, multi-user, multi-stream, multiple-input multiple-output (MIMO), dynamic time division duplexing (TDD) scenario, where we consider bidirectional beamforming optimization with the objective of minimizing network sum-power expenditure with per-stream signal-to-interference-plus-noise ratio (SINR) constraints. A first approach to solve this problem assumes centralized processing and requires the availability of global channel state information (CSI). A second approach is performed in an iterative decentralized manner, based on the alternating direction method of multipliers (ADMM), which requires local CSI and has a reduced signaling load. Both approaches converge to a minimum network power expenditure, while close-to-optimum performance can be obtained when limiting the number of iterations. In the second part, we focus on the downlink (DL) of a multi-cell, multi-user multiple-input single-output (MISO) network in the presence of CSI errors, where we consider the development of two distributed-execution user scheduling schemes based on multi-agent deep Q-learning (DQL) to solve the beamforming problem of sum-rate maximization with per base station (BS) power constraints. Simulation results show that in the presence of CSI errors the proposed schemes outperform state-of-the-art algorithms both in terms of average spectral efficiency and execution time.

Keywords: decentralization; beamforming; user scheduling; dynamic TDD; machine learning.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

3GPP	3rd generation partnership project
4G	fourth generation
5G	fifth generation
6G	sixth generation
ADMM	alternating direction method of multipliers
AI	artificial inteligence
BF	beamforming
BS	base station
CDF	cumulative distribution function
CSI	channel state information
DL	downlink
DQL	deep Q-learning
DQN	deep Q-Network
eIMTA	enhanced interference mitigation and traffic adaptation
eMBB	enhanced mobile broadband
FDD	frequency division duplexing
IBC	interference broadcast channel
IoE	internet of everything
ITU	international telecommunication union
KKT	Karush-Kuhn-Tucker
KPI	key performance indicator
LTE	long term evolution
MBS	multicast and broadcast service
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
ML	machine learning
MMSE	minimum mean square error
mMTC	massive machine type communication

MRT	maximum ratio transmission
MSE	mean square error
MU-MIMO	multiuser multiple input multiple output
mULC	massive ultra-reliable low-latency communication
NP	nondeterministc polynomial-time
QoS	quality of service
ReLU	rectifier linear unit
SDMA	space-division multiple access
SDP	semidefinite programming
SINR	signal-to-interference-plus-noise ratio
SISO	single-input single-output
SLNR	signal-to-leakage-plus-noise ratio
SLNR-MAX	maximum signal-to-leakage-plus-noise ratio
SNR	signal-to-noise ratio
SOCP	second-order cone programming
SUS	signal-to-leakage-plus-noise ratio based user selection
TDD	time division duplexing
UE	user equipment
UL	uplink
uLBC	Ultra-reliable low-latency broadband communication
uMBB	ubiquitous mobile broadband
uRLLC	ultra-reliable and low latency communication
WMMSE	weighted minimum mean square error
ZF	zero-forcing

LIST OF SYMBOLS

$(\cdot)^*$	Optimal value
$\mathcal{A}$	Set of possible actions
a^t	Action at time t
a_b^t	Action of BS b at time t
B	Number of BSs
$\mathcal{B}_{ul}$	Set of BSs in UL mode
$\mathcal{B}_{dl}$	Set of BSs in DL mode
c	Penalty parameter
$\mathbf{C}_k^j$	Value defined in (3.62) and (3.63)
$C^{(t)}$	Sum-capacity at time t
$C_i^{(t)}$	Capacity achieved by i in time t
$\mathbf{d}_i$	Data symbols relative to UE i
$d_{i,s}$	Data symbols relative to UE i stream s
d_{k_b}	Data symbol of UE k in BS b
$\mathbf{D}_{i,s}$	Matrix defined in (3.10) for UE i stream s
$\mathcal{D}_b$	Convex constraints set regarding DL BS b, defined in (3.14)
$\mathbf{e}$	Error vector
$\mathbf{G}_{b_i,b_l}$	Channel matrix between the BS $b_i \in \mathcal{B}_{dl}$ and b_l
$\mathbf{g}_{k_b j}$	Small-scale fading vector between BS j and UE k served by BS b
$\mathbf{H}_{b_i,k}$	Channel matrix between the BS that serves UE i (b_i) and UE k
$\mathbf{h}_{k_b j}$	Channel vector from BS j to UE k served by BS b
$\hat{\mathbf{h}}_{k_b j}$	I/Q transformed channel vector from BS j to UE k served by BS b
$\mathbf{I}$	Identity matrix
K	Number of UEs
K_b	Number of UEs of BS b
$\mathcal{K}_{ul}$	Set of UEs in UL mode
$\mathcal{K}_{dl}$	Set of UEs in DL mode
$\mathcal{K}_b$	Set of UEs served by BS b
$\mathcal{L}(\cdot)$	Lagrange function
$L(\cdot)$	Loss function
$\mathbf{M}_i$	BF BS filter with respect to UE i

$\hat{\mathbf{M}}_i$	BF BS General-rank positive semidefinite matrix with regarding to UE i
$\mathbf{m}_{i,s}$	BF BS filter with respect to UE i stream s
$\hat{\mathbf{M}}_{i,s}$	BF BS General-rank positive semidefinite matrix regard UE i stream s
$\mathbf{m}_{k_b}$	Transmit beamforming vector of user equipment (UE) k in BS b
$\mathcal{M}^{(t)}$	Learning Mini-batch at time t
M_s	Mini-batch size
N_b	Number of antennas at BSs
N_u	Number of antennas at UEs
$\mathbf{n}_k$	Noise at the link of UE k
n_{k_b}	Noise in UE k of BS b
p_i	Auxiliary variable which accounts for the total power used by i.
$\mathbf{P}_b^{\chi}$	Linear mapping matrix that relates $\boldsymbol{\chi}_b$ and $\boldsymbol{\chi}$
$\mathbf{P}_{\chi}$	Combination of Linear mapping matrices that relate $\boldsymbol{\chi}_b \forall b \in \mathcal{B}_{dl}$
$\mathbf{P}_i^{\psi}$	Linear mapping matrix that relates $\boldsymbol{\psi}_i$ and $\boldsymbol{\psi}$
$\mathbf{P}_{\psi}$	Combination of mapping matrices regarding $\boldsymbol{\psi}_i \forall i \in \mathcal{B}_{dl} \cup \mathcal{K}_{ul}$
$\mathbf{P}_l^{\Omega}$	Linear mapping matrix that relates $\boldsymbol{\Omega}_l$ and $\boldsymbol{\Omega}$
$\mathbf{P}_{\Omega}$	Combination of mapping matrices regarding $\Omega_l \forall l \in \mathcal{K}_{ul}$
$\mathbf{P}_i^{\theta}$	Linear mapping matrix that relates $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}$
$\mathbf{P}_{\theta}$	Combination of mapping matrices regarding $\boldsymbol{\theta}_i \forall i \in \mathcal{B}_{dl} \cup \mathcal{K}_{ul}$
P_b	Maximum transmit power at BS b
p_{k_b}	Squared Power of UE k_b transmission
$\mathbb{P}_{ss'}^{a}$	Transition probability from state s to state s' with the action a
$\mathbf{Q}_{j,k}$	Channel matrix between UEs j and k
$Q(\cdot)$	Q-function
r^t	Reward at time t
r_b^t	Reward of BS b at time t
$R^{(t)}$	Expected cumulative discounted rewards
$\mathbb{R}(\cdot)$	Expected reward
$\mathcal{S}_k$	Set of streams of UE k
$\mathcal{S}$	Set composed by all states
$s^{(t)}$	State at time t
$s_b^{(t)}$	State of BS b at time t
T_u	Interval in which $\theta_{\text{target}}^{(t)}$ is updated to be equal to $\theta_{\text{train}}^{(t)}$

$\mathbf{U}_{i,s}$	Matrix defined in (3.11) for UE i stream s
$\mathcal{U}_l$	Convex constraints set regarding UL UE l, defined in (3.15)
$\mathbf{W}_i$	BF UE filter at UE i
$\hat{\mathbf{W}}_i$	BF UE General-rank positive semidefinite matrix with respect to at UE i
$\mathbf{w}_{i,s}$	BF UE filter at UE i stream s
$\hat{\mathbf{W}}_{i,s}$	BF UE General-rank positive semidefinite matrix at UE i stream s
W_l	Number of neurons in layer l
$\mathbf{y}_k^{dl}$	Signal of all streams received by UE $k \in \mathcal{K}_{dl}$
$y_{k,s}^{dl}$	Signal stream s received by UE $k \in \mathcal{K}_{dl}$
$\mathbf{y}_k^{ul}$	Signal received by BS $b_l \in \mathcal{B}_{ul}$, sent from its served UE $l \in \mathcal{K}_{ul}$
$y_{k,S}^{ul}$	Signal stream s received by BS $b_l \in \mathcal{B}_{ul}$, sent from its served UE $l \in \mathcal{K}_{ul}$
y_{k_b}	Signal received by single-antenna UE k of BS b
$y^{(t)}$	Learning target at time t
$\boldsymbol{\alpha}_b^{(t)}$	Activation patters of BS b
$\beta_{k_b j}$	Large-scale fading coefficient between BS j and UE k served by BS b
β	Learning rate
$\Gamma_{k,s}^{dl}$	SINR for the stream s measured at UE k in DL
$\Gamma_{l,s}^{ul}$	SINR for the stream s measured at the serving BS of UE l in UL
Γ_{k_b}	SINR in UE k in BS b
$\gamma_{i,s}$	Minimum SINR value for the link of stream s of UE i
γ	Discount factor
ϵ	Greedy learning process parameter
$\boldsymbol{\theta}$	uplink (UL) UE to DL UE interference set
$\theta_{l,k,s}$	Element of $\boldsymbol{\theta}$ - UL UE k to DL UE l, stream s interference variable
$\boldsymbol{\theta}_i$	DL BS and UL UE i specific interference set
$\theta_{l,k,s}^i$	Element of $\boldsymbol{\theta}_i$ - UL UE k to DL UE l, stream s interference variable
$\tilde{\boldsymbol{\theta}}_i$	Combination of all DL BS and UL UE i specific interference sets
θ	deep Q-Network (DQN) parameters
$\theta_{\text{train}}^{(t)}$	Training DQN parameters at time t
$\theta_{\text{target}}^{(t)}$	Target DQN parameters at time t
$\boldsymbol{\lambda}_b^{\chi}$	Dual variable regarding $\boldsymbol{\chi}_b$ equality constraint in (3.20)
$\tilde{\boldsymbol{\lambda}}_{\chi}$	Combination of all dual variables regarding $\boldsymbol{\chi}$
$\boldsymbol{\lambda}_l^{\Omega}$	Dual variable regarding $\boldsymbol{\Omega}_l$ equality constraint in (3.20)

$\tilde{\boldsymbol{\lambda}}_{\Omega}$	Combination of all dual variables regarding Ω
$\boldsymbol{\lambda}_i^{\theta}$	Dual variable regarding $\boldsymbol{\theta_i}$ equality constraint in (3.20)
$\tilde{\boldsymbol{\lambda}}_{\theta}$	Combination of all dual variables regarding $\boldsymbol{\theta}$
$\boldsymbol{\lambda}_i^{\psi}$	Dual variable regarding $\boldsymbol{\psi_i}$ equality constraint in (3.20)
$\tilde{\boldsymbol{\lambda}}_{\psi}$	Combination of all dual variables regarding $\boldsymbol{\psi}$
λ_i^{ρ}	Dual variable regarding ρ_i equality constraint in (3.20)
$\boldsymbol{\lambda}$	Lagrange multipliers vector
λ_i	Specific Lagrange multipliers for SINR constraint of node i.
μ_b	Lagrange multipliers with respect to the power constraint of BS b
ξ	Reliability of the channel estimation
π	Learning policy
ρ_i	Slack variable
σ^2	Noise variance
ϕ	Set of experiences
$\boldsymbol{\chi}$	DL BS to DL UE interference set
$\chi_{b,k,s}$	Element of $\boldsymbol{\chi}$ - DL BS b to DL UE k, stream s interference variable
$\boldsymbol{\chi}_b$	DL BS b specific interference set
$\chi_{b,k,s}^b$	Element of $\boldsymbol{\chi}_b$ - DL BS b to DL UE k, stream s interference variable
$\tilde{\boldsymbol{\chi}}_b$	Combination of all DL BS b specific interference sets
$\boldsymbol{\psi}$	DL BS to UL BS interference set
$\psi_{b,k,s}$	Element of $\boldsymbol{\psi}$ - DL BS b to UL BS user k, stream s interference variable
$\boldsymbol{\psi}_i$	DL BS and UL UE i specific interference set
$\psi_{b,k,s}^b$	Element of $\boldsymbol{\psi}_i$ - DL BS b to UL BS user k, stream s interference variable
$\tilde{\boldsymbol{\psi}}_i$	DL Combination of all BS and UL UE i specific interference sets
$\boldsymbol{\Omega}$	UL UE to UL BS interference set
$\Omega_{l,k,s}$	Element of $\boldsymbol{\Omega}$ - UL UE l to UL BS user k, stream s interference variable
$\boldsymbol{\Omega}_l$	UL UE l specific interference set
$\Omega_{l,k,s}^l$	Element of $\boldsymbol{\Omega}_l$ - UL UE l to UL BS user k, stream s interference variable
$\tilde{\boldsymbol{\Omega}}_l$	Combination of all UL UE l specific interference set

TABLE OF CONTENTS

1 INTRODUCTION

In this introductory chapter, we aim to familiarize the reader with the context in which this thesis is inserted and to briefly present the main contributions of the works here contained. Also in this chapter, we provide a description of how this document is organized and the list of main scientific productions developed within this thesis studies.

The remainder of this chapter is structured as follows. Section 1.1 discusses the technological context of this thesis. Section 1.2 describes this thesis organization and its main contributions. Section 1.3 lists the details of the related scientific production.

1.1 Thesis Context

In the present days, individuals in our society have the feeling that the ability to connect with anyone at any time and everywhere, via worldwide networks, is almost as a fully guaranteed resource. However, it was not without decades of research and development of communication systems that such perception was achieved. In fact, the ongoing progress of wireless cellular networks has a tremendous impact on our society's ability to communicate.

Over the last decades, the development of wireless communications progressed along society's demand, drastically changing from the first generation of wireless voice-only networks to current data-dominated wireless fifth generation (5G) systems (1). The works presented in this thesis belong in the context of the evolving mobile wireless communications of fifth generation (5G) and beyond.

Initial 5G network deployments are already a reality across the globe. 5G comes into existence as a technology revolution in wireless communication networks, with the promise of ultra-high data throughput, very low latency, and enhanced energy efficiency. 5G will deliver mobile broadband services anywhere to anyone at any time (2). First standardized in 2017 by the 3rd generation partnership project (3GPP) with Release 15, 5G's development continues as further releases are delivered (3). 5G's initial conception, however, dates back to almost a decade ago, in a continuous global coordinated effort from academia and industry to determine use cases and requirements, and to develop technologies and solutions to support its application (4).

Based on the discussions and agreements between industry and academia about the range of the application scenarios for 5G, the international telecommunication union (ITU) has categorized three primary use cases envisioned for the new radio as this generation's key performance indicators (KPIs) (5) (6). The use cases are: enhanced mobile broadband (eMBB), which will require high data rates and enhanced coverage; ultra-reliable and low latency communication (uRLLC), which will drive the requirement for very low latency; and massive machine type communication (mMTC), wich will need to deal with a large number of connected devices. These 5G major applications and requirements are further detailed as follows:

- **eMBB**: The everlasting growth in data rates required by the people's demand for

novel network-consuming applications, such as ultra-high resolution video and virtual reality, makes this use case of primary importance for 5G. The eMBB aims not only to provide improved data rates to highly demanding users but also to enhance user's experience by provisioning reliable quality of service (QoS) anywhere/anytime for users that require moderate rates (7).

- **uRLLC**: Emerging applications for wireless networks such as autonomous cars, remote driving, industrial automation, and remote surgery are the root of this use case. It is self-explanatory this application's necessity for low latency and for available and reliable networks, since a failure or a delay to take a crucial action can be costly. Granting both strict latency requirements and ultra-high reliability is crucial for these services, but it is a very challenging task (8).

- **mMTC**: The number of different devices that can connect and communicate in wireless networks without human intervention in upcoming services will climb. Devices such as sensors and actuators in health monitoring, automation, logistics, and security have become smaller and cheaper and are characterized by transmitting a low volume of data and by requiring high energy efficiency. Although low data rates are demanded, the large number of connected devices may lead to network congestion (9).

Each of the 5G primary services would present different requirements regarding data throughput, latency, coverage, power consumption, and others. Hence, the network must be able to adapt to the user's demand and manage to support such applications via the same 5G air interface. Such a broad set of capabilities of 5G systems are specified by the ITU (5) and 3GPP (10) as a list of some of the main KPIs for the downlink (DL) and uplink (UL) directions, detailed in Table 2. The values in Table 2 show how strict the requirements of 5G are in order to support current and future applications. Such values are orders of magnitude higher than the ones achieved by previous fourth generation (4G) (11).

Commercial deployments of the first 5G version have already started and were expected to reach its first billion subscriptions worldwide by the end of 2022, which is still only almost 12% of the total subscriptions. However, the number of users is expected to fiercely rise in the near future (12). Although already in use, 5G development is not over. 3GPP along with industry and academia are still in full commitment to the task of improving network capabilities in the current work on Release 18, also called 5G Advanced. The studies on the future 5G Advanced aim to further enhance previous releases' performance by working in the areas of network energy savings, coverage, mobility support, multiple-input multiple-output (MIMO) evolution, multicast and broadcast service (MBS), and positioning (13).

Along with the development of 5G Advanced, academia and industry start to envision what comes next in the future, regarding technologies beyond 5G and sixth generation (6G) (14, 15, 16, 17). The current 5G commercial version offers significant improvements over previous

Table 2 – Required KPIs for 5G networks (5).

KPI	Use case	DL requirement	UL requirement
Peak data rate	eMBB	20 Gbps	10 Gbps
5th Percentile user data rate	eMBB	100 Mbps	50 Mbps
Peak spectral efficiency	eMBB	30 bps/Hz	15 bps/Hz
5th Percentile user spectral efficiency (Dense Urban)	eMBB	0.225 bps/Hz	0.15 bps/Hz
Energy efficiency	eMBB	Efficient data transm. Low idle power consump.	Efficient data transm. Low idle power consump.
User plane latency	eMBB	4 ms	4 ms
	uRLLC	1 ms	1 ms
Reliability	uRLLC	$1 - 10^{-5}$ success prob.	$1 - 10^{-5}$ success prob.
Connection density	mMTC	1 M devices/km^2	1 M devices/km^2

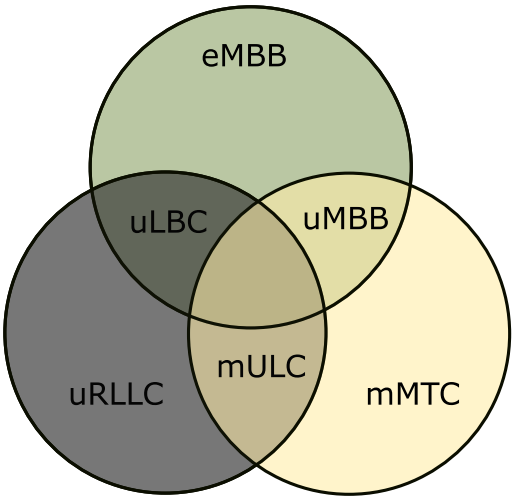
Source: Created by the author.

generations, and further 5G enhancements will improve them even more. However, it surely will not be able to fulfill the daunting demands of a future data-centric, data-dependent, and automated society (14). Tomorrow's needs will, in fact, relate to smart city internet of everything (IoE) applications, with an ongoing unprecedented proliferation of services, ranging from extended reality to telemedicine, flying vehicles, brain-computer interfaces, and connected autonomous systems. Such applications require even more from the long-discussed rate-reliability-latency tradeoff (15).

This way, when regarding use cases, 6G scenarios are seen as enhancements or extensions of 5G scenarios (17). As shown in Figure 1, the use cases are considered as an overlap of previous ones. The ubiquitous mobile broadband (uMBB) considers full availability of connection and a capacity boost to disruptive services, such as wearables and onboard communications. The Ultra-reliable low-latency broadband communication (uLBC) considers applications that require jointly uRLLC and extremely high throughput, such as immersive gaming and tactile internet. The massive ultra-reliable low-latency communication (mULC) combines characteristics of mMTC and uRLLC for the deployment of massive sensors in industries.

Such use cases enforce even more strict requirements on the main envisioned KPIs when compared to the 5G ones. The works in (15, 17) foresee the future 6G minimal requirements: the expected peak data rates are in the order of 1 Tpbs, with user experienced rate, i.e. the 5th percentile of the rate at levels of 1 Gbps or even higher. Spectral efficiency is predicted to be three times higher over the 5G requirements, with 10 to 100 times greater energy efficiency. Latency is envisioned to remain smaller than 1 ms (100 µs or even 10 µs). Reliability is expected to improve two orders of magnitude attaining around $1 - 10^{-7}$ of success probability, with connection density

Figure 1 – Evolution of 5G use cases into scenarios envisioned for 6G: uMBB, uLBC and mULC .



Source: Adapter from (17).

achieving 10^7 M devices per km^2.

The task of overcoming the rigorous requirements of 5G and beyond 5G is a great technological challenge. Thus, the scientific community has been massively active on such assignment, focused on developing novel strategies and techniques that are able to improve the network performance of mobile systems to the level of the required KPIs.

Among the strategies to improve system throughput and accommodate the increasing number of devices, we highlight the trend to densify the network by deploying more base stations (BSs) per area unit (18). Such solution seeks to increase the reuse of radio resources, with each BS serving a reduced set of user equipments (UEs). Also, communicating pairs would have smaller separation distances, making the effects of path loss less harmful. However, substantial challenges arise when network densification is applied, mainly regarding the implementation costs of further infrastructure and the performance limitation regarding inter-cell interference when the network entities are closer (19). In this sense, along with densification, other strategies must be implemented to deal with the arising challenges and enhance network efficiency and QoS. Two techniques are of high importance for this task: beamforming and user scheduling.

With multi-antenna transmitters, transmit beamforming is used to increase signal power at the intended receiver and reduce interference signals to non-intended users (20). High signal power can be achieved by transmitting the same data signal from all antennas, but with different amplitudes and phases such that signal components add coherently at the intended receiver. Low interference is accomplished by making the signal components add destructively at non-intended users (21). With multi-antenna receivers, receive beamforming can also be performed, in a similar way, by adjusting receive amplitude and phase of receive antenna gains so that it focus the reception to the desired direction (22) .

In multi-user environments, user scheduling algorithms have the responsibility to

decide which set of users should transmit at a given time slot and how much bandwidth resources should be allocated to each active user (23). Temporarily deactivating a user's transmission might be advantageous, for example, if it is strongly interfering with others or spending excessive power to overcome channel issues. Such a decision can be made based on different optimization choices, such as rate maximization or system power minimization, depending on service provider decision (24).

Efficient beamforming and user scheduling optimization techniques work in a complementary way and their design should consider the existence of each other. After all, a user scheduling decision might lead to different interference situations for the beamforming stage, depending on the optimization decision, and a specific beamforming strategy might affect scheduling performance. Also, the algorithms to optimize both techniques must consider the network capabilities regarding computational complexity and information sharing between optimizing entities. In this sense, in real-world applications, the less complex and distributable strategies are preferred to be implemented.

Another strategy to overcome beyond 5G requirements is the employment of artificial inteligence (AI). On the list of 6G technologies, AI is recognized as the one with the highest potential (17). The increasing complexity and heterogeneity of such future scenarios will probably prevent closed-form and manual optimizations since their solution becomes intractable. The machine learning (ML) algorithms use statistical techniques to allow machines to learn the execution of a particular task, with the goal of maintaining a specific performance metric, based on a particular experience and to improve its decision-making capabilities as they acquire more knowledge (25). ML is recognized as being a great strategy to deal with complex problems where existing solutions require considerable optimization effort, or for problems in which there is no solution with traditional approaches. By extracting information from previous data, ML can detect anomalies, predict future scenarios, adapt to fluctuating environments, get insights from complex problems and discover patterns (26).

For the future of wireless communications, ML has the role of creating more intelligent networks, by simplifying and improving the transport of real-time data, since AI increases efficiency and reduces the processing delay of the communication steps (27). This way, by using ML to perform system optimization, latency is expected to be significantly reduced and network performance is expected to become more resilient to environment imperfections and fluctuations.

Inserted in this context of wireless communication's evolution, the works presented in this thesis aim to develop strategies to improve performance for 5G and beyond systems. More specifically, we propose beamforming and user scheduling solutions for dense networks. We seek that the proposed solutions are tractable and applicable to real-world scenarios, by making them of distributed execution and by providing required signaling schemes. Also, one of the aimed solutions is derived using ML, in order to make its execution faster and robust to channel estimation imperfections.

1.2 Contributions and Thesis Structure

The presented technological context shows the main trends and possible tracks for wireless evolution. Based on such context, this thesis has as its main objective the design of decentralized strategies for the optimization of beamforming and user scheduling in 5G and beyond systems. More specifically, we divide this thesis' contents into two independent parts that aim to develop solutions to two different scenarios.

The first part aims at deriving a decentralized algorithm for the bidirectional beamforming optimization in multi-user multi-stream MIMO network functioning under a dynamic time division duplexing (TDD) regime, with the objective of minimizing network sum-power expenditure with per-stream signal-to-interference-plus-noise ratio (SINR) constraints. The main contributions of this investigation can be summarized as follows:

- The proposal of a centralized algorithm for the solution of the proposed problem based on alternate convex optimization of receivers and transmit vectors.
- The convergence and optimality analysis of the proposed centralized solution. The derived algorithm is proved to converge globally to the set of Karush-Kuhn-Tucker (KKT) conditions.
- The derivation of a decentralized algorithm for the solution of the proposed optimization problem, based on alternate convex optimization of receive and transmit vectors along with alternating direction method of multipliers (ADMM), that requires the exchange of reduced crucial information between nodes.
- The design, description and analysis of a lightweight signaling scheme to support the decentralized algorithm application.
- The convergence and optimality analysis of the proposed decentralized algorithm.
- The performance evaluation by means of simulations, where centralized and decentralized solutions are compared.

The second part aims at developing distributed-execution ML user scheduling schemes to solve the beamforming problem in DL multi-user multiple-input single-output (MISO) networks in the presence of channel state information (CSI) errors, with the objective of sum-rate maximization with per BS power constraints. The most relevant contributions of this research can be outlined as follows:

- The development of a distributed-execution algorithm based on deep Q-learning (DQL), that aims to obtain fast and robust computations in the presence of CSI errors. This solution requires local CSI knowledge and a reduced exchange of information between BSs;

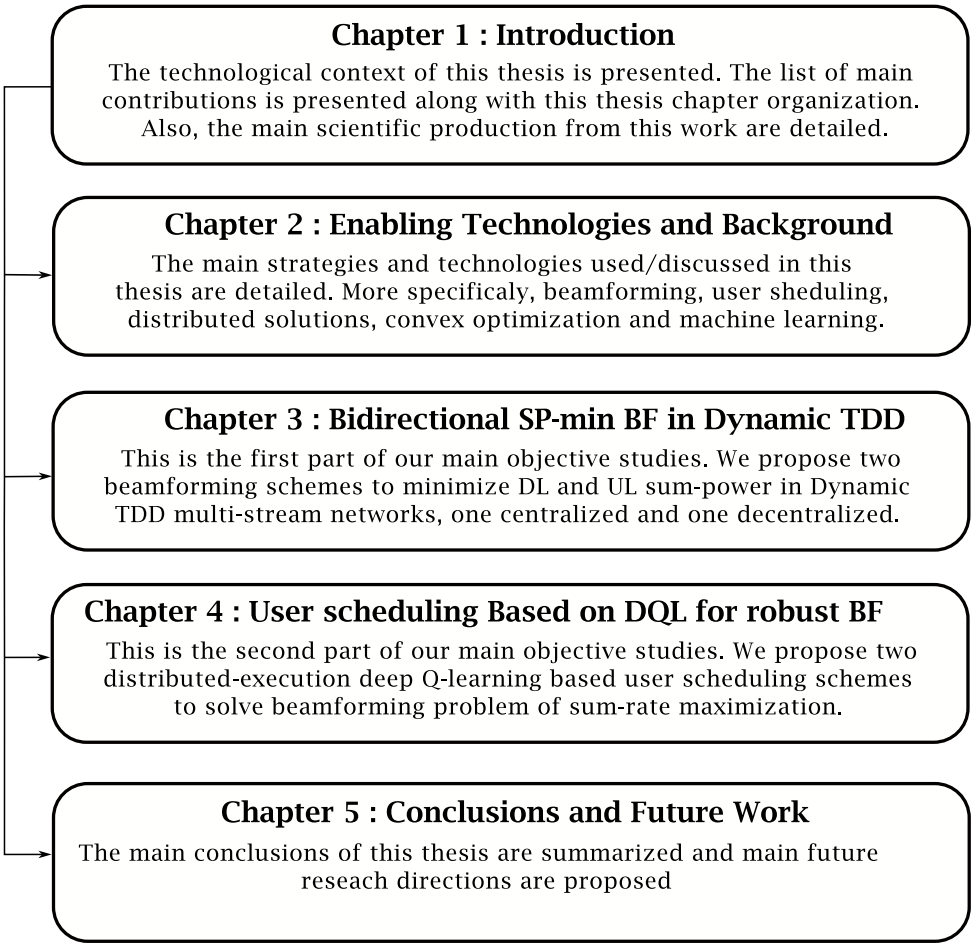
- The derivation of another distributed-execution algorithm based on DQL, that also aims to obtain fast and robust computations in the presence of CSI errors. This second algorithm assumes only local information during the execution phase, which reduces signaling overhead;
- The proposed algorithms learn a policy that extends the application of an optimization model which is based on perfect CSI to a more realistic scenario with CSI errors;
- The design of signaling schemes for the proposed solutions for both training and execution phases;
- The performance evaluation by means of simulations, where we compare the proposed solution with state-of-the-art algorithms.

The content of this thesis is organized into five chapters, including this introductory chapter. Figure 2 illustrates the summary of this thesis organization. Each chapter's content is made self-contained and chapters can be read in any sequence. The outline of the remaining chapters is presented as follows:

Chapter 2 addresses the central technologies that serve as a basis for the scenarios and for the solutions approached by the two major parts of this work. In its first sections, this chapter discusses the basis and main issues regarding beamforming and user scheduling schemes, which are the main focus of this thesis, summarizing the main state-of-the-art solutions and applications. The following section builds the reasoning on why decentralized/distributed algorithms, which are aimed in this work, are needed in multi-cell multi-user scenarios. The following two sections address the details of the theories used as basis for the solution of the optimization problems proposed in the first and second parts of this thesis, which are convex optimization and ML, respectively.

Chapter 3 considers the task of dealing with cross-link interference in dynamic TDD networks by optimizing the bidirectional beamforming in a multi-user multi-stream MIMO scenario, with the goal of minimizing the system sum-power while providing a minimum SINR for each stream. Two solution strategies are proposed, one assumes centralized computation with full network information, and serves as a benchmark for the other solution, which is a decentralized one, that assumes local information knowledge and some level of coordination among nodes. Both strategies are shown to iterate towards an optimal solution, the centralized one is proved to converge and the decentralized one approximates such performance. Since the decentralized algorithm requires coordination among nodes, the required signaling scheme is proposed and is shown to be lightweight when compared to centralized signaling requirements. The simulation results show that both strategies achieve their goals, reducing power while providing the required minimum SINR, the decentralized one, as expected, does not reach the same level of power minimization of the centralized, but efficiently approximates it while requiring a reduced amount of iterations.

Figure 2 – Thesis organization.



Source: Created by the author.

Chapter 4 proposes two distributed-execution multi-agent DQL user scheduling schemes to solve the beamforming problem of maximizing the rate in a MISO system with channel estimation errors. The main idea is that the beamforming is based on an optimal solution structure, derived by uplink-downlink duality, that assumes perfect CSI, while DQL has the task of adapting this structure to be robust by means of its actions and rewards from the real environment. Both algorithms require coordination among BSs and a master entity during the training phase. The required signaling scheme is provided. In the execution phase the first algorithm, DQL-F, requires a lightweight signaling between BSs while the second DQL-L does not rely upon any information share, only local knowledge. Results show that both algorithms are able to provide fast and robust solutions for the maximization problem, with DQL-L presenting a slight loss in performance when compared to DQL-F, as expected. When compared to state-of-the-art algorithms, both strategies were able to provide faster computations and to show gains when channel estimation errors were present.

Finally, Chapter 5 draws the main conclusions obtained along the discussions and results, and points towards the major research directions that arise as possible extensions and continuations from the studies presented in this document.

1.3 Scientific Production

The content and contributions of this thesis were published with the following information.

- Eduardo de O. Cavalcante; G. Fodor, Y. C. B. Silva and W. C. Freitas, "Bidirectional Sum-Power Minimization Beamforming in Dynamic TDD MIMO Networks," in IEEE Transactions on Vehicular Technology, vol. 68, no. 10, pp. 9988-10002, Oct. 2019, doi: 10.1109/TVT.2019.2937474.

- I. M. Braga, Eduardo de O. Cavalcante, G. Fodor, Y. C. B. Silva, C. F. M. e Silva and W. C. Freitas, "User Scheduling Based on Multi-Agent Deep Q-Learning for Robust Beamforming in Multicell MISO Systems," in IEEE Communications Letters, vol. 24, no. 12, pp. 2809-2813, Dec. 2020, doi: 10.1109/LCOMM.2020.3015462.

It is worth mentioning that this thesis was developed under the context of the following Ericsson/UFC technical cooperation projects:

- UFC.46 - Network Assisted Intelligent Vehicle-to-Everything communications (NAIVE), November/2018 - November/2020,

- UFC.48 - User Centric Networks and Reconfigurable Surfaces for Next Generation Wireless, December/2020 - December/2022,

in which a number of eight technical reports, four in the first project and four in the second project, have been delivered.

2 ENABLING TECHNOLOGIES AND BACKGROUND

In this chapter, we provide an overview of the major theories and technologies that enable the scenarios and solutions approached in this thesis. The main objective of this section is to familiarize the reader with the basic context of such technologies and their main state-of-the-art applications.

This chapter is structured as follows. Section 2.1 discusses the strategy of beamforming, which serves as a basis for the whole content of this thesis. Section 2.2 approaches the concepts of user scheduling, which is in focus in the second part of this thesis. Section 2.3 motivates the employment of distributed solutions in wireless networks, which are goals all along this work. Section 2.4 provides an overview of the convex optimization theories used as tools for the solution of the optimization problems proposed in the first part of this work and for the basic beamforming structure for the second part. Section 2.5 provides an overview of the machine learning theories used as tools for the solution of the optimization problems proposed in the second part of this thesis.

2.1 Multi-user beamforming

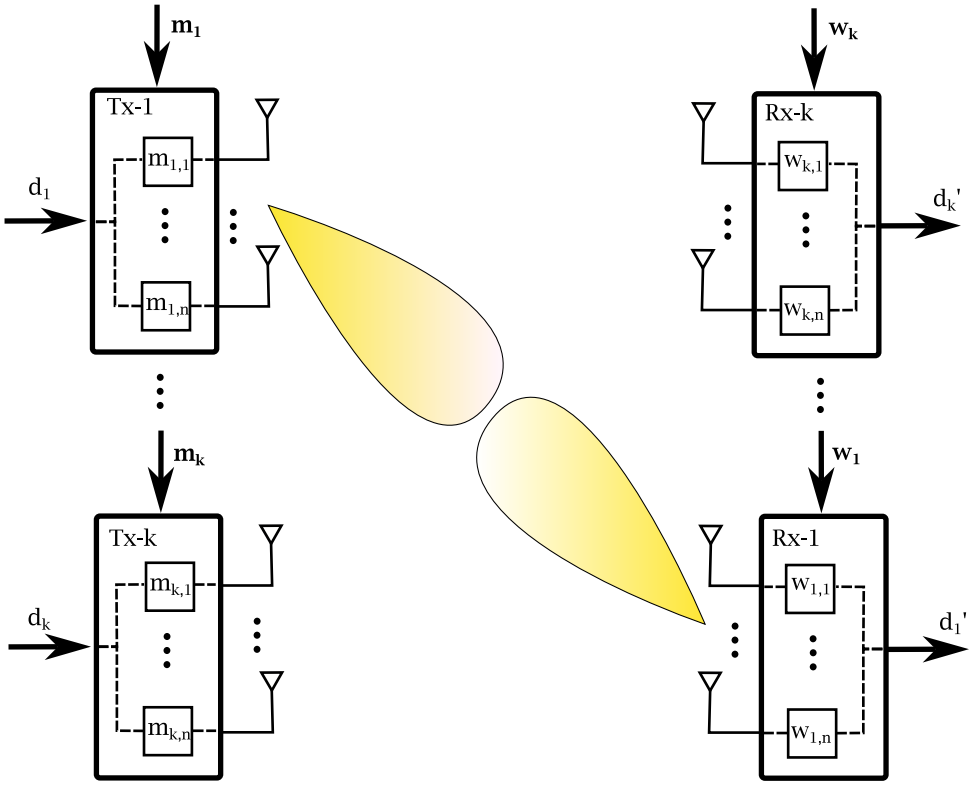
Spectral efficiency is a measure of the amount of information that can reliably be transmitted over a given link per time and frequency unities. In future communication systems, the improvement of such parameter is mandatory in order to achieve the required network performance. To increase system-wide spectral efficiency there must be an improvement in users' SINR, which can be achieved by a power increase of the received desired signal and an interference power reduction.

There are some ways to seek such SINR improvement in cellular networks. One could think that the more direct way to increase the desired signal received power would be simply by employing more power to the transmitters. However, in a multi-user scenario that could lead to stronger interference with other communication links, and system-wide capacity would suffer. Also, deliberately increasing transmit power would aggravate energy consumption.

Beamforming, on the other hand, is a versatile and powerful approach to receive and transmit signals of interest in a spatially selective way in the presence of interference and noise (22). Beamforming refers to a spatial filtering applied on multi-antenna transmitter and/or receiver in order to favor some directions over others (21). Such filtering is achieved by adjusting amplitudes and phases of the signal components transmitted by each antenna in a way to make them add coherently at the intended receiver and destructively at non-intended receivers (28). Receive beamforming would, by a similar approach, focus antenna gain in the direction of its intended transmitter, and avoid interfering transmitters. Both strategies can work simultaneously in MIMO scenarios.

Figure 3 illustrates transmit and receive beamforming, where it can be seen that transmitter 1 multiplies the data signal at each antenna circuit, by the corresponding weight in

Figure 3 – Transmit and receive beamforming process.



Source: Created by the author.

the precoder vector $\mathbf{m}_1$, with the goal of focusing signal in the direction of receiver 1. On the other hand, receiver 1 focuses its reception gains into the desired direction of transmitter 1 by multiplying the signal received at each antenna by the corresponding weight in the decoder vector $\mathbf{w}_1$, providing an estimate for the transmitted data signal. Thus, signal power and reception gains on the desired link directions are stronger, increasing received power while avoiding causing and receiving interference with regard to other links.

The ability to focus beams in specific directions while avoiding others allows for signal separation on the space domain with the so-called space-division multiple access (SDMA), in which users spatially separated can be served simultaneously using the same radio resource. Unfortunately, the finite number of transmit antennas only provides a limited amount of spatial directivity, which means that there are still energy leakages between the users that act as interference (28).

The task of designing beamforming vectors that maximize signal power at the intended user while minimizing interference leakage in a multi-user environment is generally a nondeterministc polynomial-time (NP) hard problem (29). However, for specific situations, classic closed-form transmit beamforming solutions are widely known:

- Maximum ratio transmission (MRT): has the objective to increase signal power at the intended user, i.e. maximizes the signal-to-noise ratio (SNR). It is also called matched filter, since its formulation aims that the beamforming vector matches the

channel vector, maximizing the inner product of the channel and beamforming vector. In noise-limited systems, it is regarded as an optimal beamforming strategy, whereas in the high SNR regime it might indiscriminately cause interference (30).

- Zero-forcing (ZF): has the goal of eliminating the interference leakage at unintended users in an altruistic approach. As the strategy name states, it tries to force interference to be zero by making the unintended user's channel orthogonal to the chosen precoder. By doing so, ZF might not transmit sufficient power to the desired user in the low SNR regime (31).

- Maximum signal-to-leakage-plus-noise ratio (SLNR-MAX): tries to balance signal power maximization and interference power minimization in a heuristic way. This is made by maximizing the signal-to-leakage-plus-noise ratio (SLNR), that is the ratio between the signal power at the intended user and the normalized noise plus the total interference power that leaks to non-intended users. Equivalent approaches are also known by other names such as transmit minimum mean square error (MMSE) (32) and transmit Wiener filter (33). Although being a generally suboptimal strategy, SLNR-MAX combines the benefits of MRT at low SNR with the benefit of ZF at high SNR (34).

In wireless multi-cell multi-user scenarios, system-wide beamforming optimization can be performed by two main approaches: uncoordinated and coordinated. As the name states, in the uncoordinated approach, each entity optimizes its beamformers independently without any form of cooperation, whereas in coordinated beamforming the entities cooperate to jointly optimize the beamformers. The uncoordinated approach offers low complexity, however, the coordination of the latter approach yields an improvement in system-wide overall performance at the cost of complication in coordination management and in the information share among network entities.

Beyond the classic beamforming solutions, different strategies to optimize the beamforming in multi-cell multi-user networks have been proposed for the more diverse scenarios and criteria. We highlight the beamforming schemes with the objectives of power minimization and sum-rate maximization.

Power minimization beamforming has as its main objective the design of precoders that minimize the power consumed by the network while guaranteeing a minimum QoS constraint, conventionally SINR. Such a design decision is well-motivated and desired since reducing power consumption is one of the main aspects of future system trends. It was introduced in (35) for the multi-cell MISO scenario, which proposed a solution that jointly finds a set of feasible beamforming weight vectors and downlink transmit power allocation, such that the SINR at each link is greater than a target value in both downlink and uplink directions. The work in (36) targets the solution of the power minimization under additional per-antenna power constraints by

employing uplink-downlink duality, and is extended by the work in (37) that achieves the joint globally optimal beamformers across all BSs in a decentralized manner.

Alternative power minimization beamforming solutions presented in (38) and (39) focus on recasting the optimization problem into a standard convex form relaxed semidefinite programming (SDP) and second-order cone programming (SOCP), respectively. Such reformulations allow the efficient solution through standard convex optimization numerical tools such as CVX (40). When considering the joint optimization of transmitters and receivers in MIMO scenarios, the power minimization problem can no longer be recast into convex forms. The authors in (41) proposed a solution for this network setting based on uplink-downlink duality, while authors in (42) and (43) derived solutions based on alternate transmit and receiver optimization with centralized and decentralized algorithms, respectively.

Sum-rate maximization beamforming aims at designing precoders that maximize the network rate while consuming a specific amount of power. Directly optimizing total system sum-rate may not be fair to all users, since users with good channel conditions may be favored. This way, to add fairness to the beamforming procedure, weights can be assigned to the sum-rate of different users. This optimization task is known to be NP-hard even for the single-antenna case, for which only local optima can be found. Nevertheless, that does not diminish interest in practical methods to achieve good rate maximization performance, since such optimization objective has the desirable properties of: 1) it can prioritize users in order to provide some fairness among them by adjusting the weights; 2) it has an implicit user and stream 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 (44). Therefore, it continues to be an important research topic that is still being extensively studied.

Solutions for the weighted sum-rate maximization problem using convex optimization have been proposed by some works, such as those in (45, 46, 47, 48). In (45) the sum-rate maximization problem was studied in multi-cell MISO systems and a solution based on the branch and bound method was presented. In (46) a local optimal solution was proposed exploiting an iterative weighted minimum mean square error (WMMSE) approach for a single-cell MIMO system. The authors in (47) provided proof of convergence for such an approach, as well as an extension for multi-cell scenarios. This same scenario was studied in (48), for which a solution based on fractional programming was proposed.

In both parts of this thesis, we approach and propose beamforming strategies. In the first part, we propose a bidirectional approach to optimize the MIMO beamformers with the goal of sum-power minimization with SINR constraints, in both centralized and decentralized manners. In the second part, we approach the sum-rate maximization beamforming, by working with a pre-defined model-based beamforming structure and a ML scheduling scheme.

2.2 User Scheduling

Resource allocation in general plays a crucial role in wireless communications, by having the task of managing the time/frequency and power resources. In single-input single-output (SISO) systems, those are the available assets and the managing strategies seek to improve performance by assigning each user to a specific resource. On the other hand, in MIMO systems, another level of possibilities for the allocation algorithms arises, as the spatial dimension appears as a new resource to be exploited (49).

User scheduling algorithms have the responsibility to decide which set of users should transmit at a given time slot and how much bandwidth resources should be allocated to each active user (23). Such a decision affects the whole network behavior and can lead to very different levels of user QoS. Conventionally, user scheduling strategies, just like beamforming, take the form of optimization tasks with diverse objectives.

In SISO systems, the major user scheduling solutions are the ones considering the link activation problem and the age of information minimization. In the first, the objective is to maximize the number of active links under SINR constraints. The works in (50) and (51) consider this problem. Variations and extensions of this problem include the works in (52) and (53), that take into consideration interference cancellation. In the latter, a metric called age of information, defined as the current (observation) time minus the time at which the observed state was generated is minimized. The works in (54) and (55) consider that objective with and without packet deadlines, respectively.

In MIMO systems, efficient beamforming and user scheduling optimization techniques work in a complementary way and their optimization should consider the existence of each other, since they directly affect each other. This way, in MIMO systems, in addition to scheduling decisions, the optimization must extend to transmit/receive beamforming vectors, transmit powers, etc. (56). Conventional beamforming transceiver design normally considers a fixed given scenario, with a pre-defined set of BSs and UEs. However, the achievable performance of such methods depends heavily on the channel characteristics of selected users. Thus user selection becomes a key approach to benefit from multi-user diversity and achieve full multiplexing gain (57). Due to this characteristic of dual optimization, in order to achieve a feasible outcome both strategies can tune their decisions: the beamforming scheme can relax QoS requirements, or the scheduling algorithm can reduce the number of users (57).

The major user scheduling optimization objectives in multi-antenna networks are maximizing the system sum-rate and minimizing the total transmit power. For the first objective, the work in (58) proposed centralized and semi-distributed algorithms for user scheduling focusing on rate-constrained sum-rate maximization for the MIMO interference broadcast channel (IBC), and the authors in (59) proposed a distributed signal-to-leakage-plus-noise ratio based user selection (SUS) which maximizes the weighted sum-rate (WSR) in multi-cell MISO downlink systems. A strategy that considers the sum-rate maximization problem with per-user

rate constraints in the MIMO IBC and proposes a solution that jointly optimizes transmit/receive beamforming vectors, transmit powers, and user scheduling is provided in (60). Considering the second objective, the works in (61) and (62) proposed user scheduling solutions for sum-power minimization under minimum SINR constraints. Beyond that, a user scheduling algorithm using DQL in a multi-antenna scenario was considered in (63).

The second part of this thesis considers the derivation of distributed-execution ML user scheduling to solve the beamforming problem in DL multi-user MISO networks in the presence of CSI errors, with the objective of sum-rate maximization with per BS power constraints.

2.3 Distributed Solutions

Networked systems consist of a large number of interconnected subsystems (agents), which are required to cooperate in order to achieve a desirable global objective (64). The cooperative control of wireless multi-cell multi-user scenarios itself is an application of such networked systems. The algorithms for optimizing beamforming and user scheduling have to be executed in a system-wide coordinated manner, so that the agents cooperate in order to achieve a global optimization objective and obtain an improved overall system performance.

For cooperative control strategies to be successful, numerous issues must be addressed, including the definition and management of shared information among a group of agents to facilitate the coordination of these agents. In cooperative optimization strategies, the agent (or agents) responsible for the optimization must be aware of the common objectives, constraints and variables (65). For instance, in wireless network optimization, a common assumption is that the optimizing agent must have CSI knowledge of the nodes affected by its optimization.

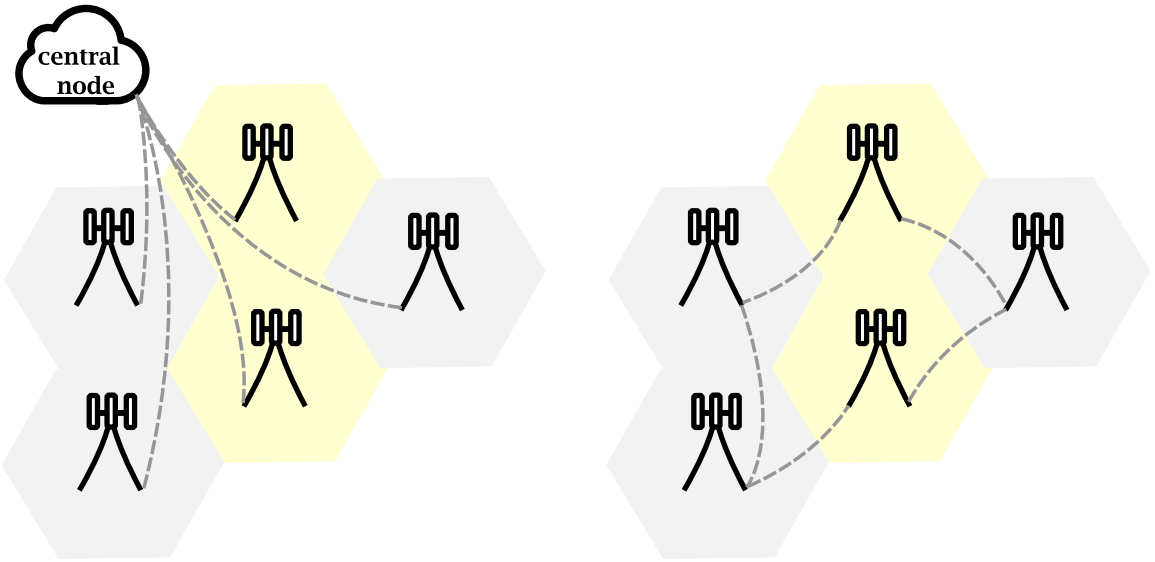
Coordination in networked systems can be approached by two distinctive means: centralized and decentralized/distributed. Figure 4 illustrates an example of the two corresponding setups for a given network with multiple communication nodes. In the first, agents are linked by a central node that coordinates them, while in the second, agents are connected to each other and coordinate themselves.

In centralized optimization, a central unit (a simple central node or a cloud radio access network (C-RAN)) is assumed to be present as a master entity and it is in charge of performing all the optimization, taking into account the relationship among the cells. In order to do this, the central unit must acquire knowledge of the channels between all BSs and all UEs in the system, i.e., global CSI, which can be achieved by each BS sending its local CSI to a central controlling unit via backhaul. Beyond that, the results from computations and corresponding commands must be sent back to the respective nodes. Alternatively, centralized approaches are also commonly defined with each node performing its computations. For this, every BS must have access to global CSI via backhaul sharing.

Due to the distributed nature of networked systems, such centralized strategies are,

Figure 4 – Coordination setups: centralized and decentralized/distributed.

(a) Centralized (b) Decentralized/distributed



Source: Created by the author.

in many cases, not suitable to solve optimization problems. Although centralized solutions are often able to find a network global optimum through joint optimization, they have an inherent high computational complexity and demand a prohibitive amount of signaling load in order to gather all required information and share results. Moreover, the centralized framework is subject to performance limitations, such as a single point of failure, high communication requirement, substantial computation burden, and limited flexibility and scalability. All of these have made imperative the use of distributed approaches to solve coordinated optimization problems (64).

In decentralized/distributed wireless optimization, each agent takes its own decisions relying on the availability of local CSI, i.e., knowledge of the channels between itself and other nodes in the system, in order to improve a specific goal. In coordinated designs, agents must cooperate to obtain a global optimum by minimizing or maximizing a global utility function with certain constraints, and, to do this, decentralized solutions allow some level of exchanging decision variables information with neighboring communication nodes (66). In wireless systems, this cooperation exchange can happen via backhaul and/or over-the-air signaling.

The outcome of a decentralized algorithm is usually sub-optimal due to the nature of conflict of interest among agents. Nevertheless, decentralized algorithms are often more viable for wireless optimization problems than centralized ones since the network itself is decentralized, and the whole process of gathering and broadcasting information required by the centralized process is not needed. Another gain for distributed solutions is that each subproblem is much less complex than the whole global optimization, and by distributing these smaller problems computation can enhance optimization and improve scalability.

Based on that discussion, the choice between centralized and decentralized optimization is dependent on the level of requirements and on the network information share capabilities.

This is true since decentralized algorithms are able to offer reduced signaling overhead and complexity, while there is a cost of degradation in performance. Also, different decentralized algorithms can provide varying levels of performance by changing levels of signaling cooperation, i.e. nodes can increase selfishness by cooperating less and reduce global performance, or act more altruistically by cooperating at high levels and increase performance.

In this thesis, we aim to propose decentralized/distributed algorithms to optimize beamforming and user scheduling with global objective functions. In the first part, we derive both a centralized and a decentralized version of the sum-power minimization beamforming algorithm in dynamic TDD scenario. The centralized algorithm is shown to be optimal, while the decentralized one can achieve close to optimal performance and is able to vary the levels of coordination at the cost of a slight performance loss. In the second part, we derive two centralized training/decentralized execution DQL user scheduling strategies for the sum-rate maximization beamforming. The first strategy is based on a reduced amount of cooperation signaling among agents in execution time, while the second does only rely on local information. Both strategies achieve good optimization performance, with the second one slightly below the first.

2.4 Convex Optimization

Convex optimization is a special class of mathematical problems that serve as a powerful tool for addressing the design and analysis of applications in a wide range of study fields. In signal processing and communications, convex optimization plays a crucial role, since many important problems can either be cast as or be converted into convex optimization problems, which can be solved numerically in a reliable and efficient way.

Convex optimization refers to the minimization of a convex objective function subject to convex constraints. In this class of problems, a very useful property is that a local optimum is also a global optimum. Also, in convex problems, a rigorous optimality condition and duality theory exist to verify the optimal solution (67).

Recognizing or reformulating a problem as convex offers the great advantage that it can be solved, very reliably and efficiently, using interior-point methods or other methods for convex optimization. Such solution methods are reliable enough to be embedded in computer-aided design and analysis tools, or even real-time control systems. Off-the-shelf software tools, such as the widely used CVX (40) and SEDUMI (68), offer a quite straightforward environment for solving convex optimization problems in an accurate and efficient way. Consequently, once a problem is expressed in a convex form, it may be regarded as "solved" from a numerical standpoint. Another great advantage is that the associated dual problem often has interesting interpretations in terms of the original problem, and sometimes leads to different efficient or distributed methods for solving it (69).

The theory of convex optimization permeates the entire content of this thesis. Therefore, in the sequence, we provide a very brief review of some of the key aspects of convex

optimization approached in this thesis, mainly based on the studies in (69) and (67).

The convex optimization problem

A generic convex optimization problem is one of the form

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0,\ i = 1, 2, \ldots, m, \\ & h_j(x) = 0,\ j = 1, 2, \ldots, r, \\ & x \in \mathcal{S}. \end{aligned} \tag{2.1}$$

where f_0 is the objective function to be minimized, $\{f_i\}_{i=1}^m$ and $\{h_j\}_{j=1}^r$ are the inequality and equality constraint functions, respectively, and $\mathcal{S}$ is a constraint set.

The optimization problem in (2.1) is convex if:

1. The objective function and inequality constraint functions $\{f_i\}_{i=0}^m$ are convex.

   A function is convex if the line segment between any two points lies above the graph between the two points, i.e. a convex function f_i: $\mathbb{R}^n \to \mathbb{R}$ must satisfy

   $$f_i(\theta x + (1-\theta)y) \leq \theta f_i(x) + (1-\theta) f_i(y), \quad \forall \theta \in [0, 1]. \tag{2.2}$$

2. The equality constraint functions $\{h_j\}_{j=1}^r$ are affine.

   A function h_j: $\mathbb{R}^n \to \mathbb{R}^m$ is affine if it is a sum of a linear function and a constant, i.e. it has the form $h_j(x) = Ax + b$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

3. The constraint set $\mathcal{S}$ is convex.

   A set $\mathcal{S}$ is convex if the line segment between any two points in $\mathcal{S}$ lies in $\mathcal{S}$, i.e. for any two points x and $y \in \mathcal{S}$, then

   $$\theta x + (1-\theta)y \in \mathcal{S}, \quad \forall \theta \in [0, 1]. \tag{2.3}$$

The problem in (2.1) is said to be feasible if there exists at least one feasible point that respects all constraints, and infeasible otherwise. The optimal value of (2.1) is denoted as p^* and x^* denotes the optimal point where $f_0(x^*) = p^*$.

Standard convex optimization forms

The convex optimization problems in engineering design applications commonly fall into the form of widely known pre-defined optimization models. In many wireless problems, such as the ones in this thesis and the problems in (39) and (38), authors focus on recasting the optimization problems into standard convex forms of second-order cone programming (SOCP) or semidefinite programming (SDP), which can be numerically solved by CVX and SEDUMI .

The standard form of an SOCP problem is (70)

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{f}^H\mathbf{x} \\ \text{subject to} \quad & \begin{bmatrix} \mathbf{c}_i^H\mathbf{x}+d_i \\ \mathbf{A}_i^H\mathbf{x}+\mathbf{b}_i \end{bmatrix} \succeq_K 0, \quad i=1,\ldots,N. \end{aligned} \tag{2.4}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable, and $\mathbf{f}, \mathbf{A}_i, \mathbf{b}_i, \mathbf{c}_i$ and d_i are the data parameters of appropriate sizes. In the second-order cone constraint, the notation $\succeq_K$ denotes the following generalized inequality

$$\begin{bmatrix} z \\ \mathbf{z} \end{bmatrix} \succeq_K 0 \leftrightarrow ||\mathbf{z}|| \leq z. \tag{2.5}$$

The standard form of an SDP problem is (71)

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{f}^H\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}(\mathbf{x}) \succeq 0. \end{aligned} \tag{2.6}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable, $\mathbf{A}(\mathbf{x}) = \mathbf{A}_0 + \sum_{i=1}^n x_i\mathbf{A}_i$, with problem data being $\mathbf{f}$ and the hermitian matrices $\mathbf{A}_i$ of appropriate sizes. The notation $\succeq$ denotes the positive semidefinite generalized inequality.

Duality

The principle of duality in optimization states that problems may be viewed from two different perspectives, the primal problem and the dual problem. Such principle is an important tool for the solutions in this thesis. Often primal problems such as the ones in (2.1), (2.4) and (2.6) come in a form of minimization, then the dual problems have the form of a maximization.

For a problem in the form of (2.1) (not necessarily convex), we can form the Lagrangian function

$$L(x,\lambda,\nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{r} \nu_j h_j(x). \tag{2.7}$$

where λ_i are the dual variables associated with its inequality constraints and ν_j are the dual variables associated with its equality constraints.

The dual function associated with (2.1) is defined as the minimum value of the Lagrangian over x as

$$g(\lambda,\nu) = \inf_{x\in\mathcal{S}} (L(x,\lambda,\nu)). \tag{2.8}$$

Since the dual function is the pointwise infimum of a family of affine functions, it is concave, even when the problem (2.1) is not convex. For each pair (λ,ν) with $\lambda \geq 0$ the dual function gives us a lower bound on the optimal value p^*. The search for the best lower bound that

can be obtained from the dual function defines the (always convex) dual optimization problem

$$\begin{aligned} \underset{\lambda,\nu}{\text{maximize}} \quad & g(\lambda,\nu) \\ \text{subject to} \quad & \lambda \geq 0. \end{aligned} \tag{2.9}$$

The optimal value of the dual problem, denoted as d^*, is by definition the best lower bound for p^*, i.e. $d^* \leq p^*$, even if the original problem is not convex. This property is called weak duality. Strong duality, on the other hand, implies that the duality gap is zero, i.e. $d^* = p^*$. Strong duality does not, in general, hold, but if the primal problem is convex in the forms of the problems above, we usually have strong duality.

Optimality conditions

In order to analyze optimality for a convex optimization problem in the form of (2.1) we must evaluate the KKT conditions. In such convex problems the KKT conditions are sufficient for the points to be primal and dual optimal. Then, x^* and (λ^*, ν^*) are primal and dual optimal, with zero duality gaps if

$$f_i(x^*) \leq 0, \forall i = 1, \ldots, m \tag{2.10}$$

$$h_j(x^*) = 0, \forall j = 1, \ldots, r \tag{2.11}$$

$$\lambda_i^* \geq 0, \forall i = 1, \ldots, m \tag{2.12}$$

$$\lambda_i^* f_i(x^*) = 0, \forall i = 1, \ldots, m \tag{2.13}$$

$$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{r} \nu_j^* \nabla h_j(x^*) = 0. \tag{2.14}$$

The first two conditions state that x^* is primal feasible. The third condition represents dual feasibility. The fourth means the complementary slackness for the primal and dual inequality constraint pair: $f_i(x) \leq 0$ and $\lambda_i \geq 0$. The last condition states that the gradient of the Lagrange function vanishes at $x = x^*$, i.e. x^* minimizes $L(x, \lambda^*, \nu^*)$. We employ such KKT analysis in the proof of convergence and optimality for the algorithms in the fist part of this thesis.

ADMM

Still in the context of convex optimization, now regarding the solving stage, we highlight a method that has recently drawn significant attention, namely the ADMM. ADMM is described by (72) as a simple but powerful algorithm that is well suited to distributed convex optimization and takes the form of a decomposition-coordination procedure, in which the solutions to small subproblems are coordinated to find a solution to a large global problem. The ADMM algorithm tries to mix the benefits of other two preceding optimization methods. It tries to implement the parallelism provided by the dual ascent method (72, Chapter 2.1) and the robustness of the method of multipliers (72, Chapter 2.2), by implementing a combination of both algorithms. A very brief review of ADMM is provided next.

Consider the following simple convex optimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{Ax} = \mathbf{b}, \end{aligned} \tag{2.15}$$

with $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ being a strictly convex function.

This problem can be rewritten by splitting the variable $\mathbf{x}$ in two, $\mathbf{x}$ and $\mathbf{z}$, related as shown in the constraint part of (2.16), with the objective function separable across this splitting.

$$\begin{aligned} \underset{\mathbf{x},\mathbf{z}}{\text{minimize}} \quad & f(\mathbf{x}) + g(\mathbf{z}) \\ \text{subject to} \quad & \mathbf{Ax} + \mathbf{Bz} = \mathbf{c}, \end{aligned} \tag{2.16}$$

with $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{z} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, $\mathbf{B} \in \mathbb{R}^{p \times m}$, and $\mathbf{c} \in \mathbb{R}^p$. Assuming that f and g are convex functions.

Writing the augmented Lagrangian for Problem (2.16) we obtain

$$L_\rho(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + g(\mathbf{z}) + \mathbf{y}^T(\mathbf{Ax} + \mathbf{Bz} - \mathbf{c}) + (\rho/2)||\mathbf{Ax} + \mathbf{Bz} - \mathbf{c}||_2^2, \tag{2.17}$$

where $\mathbf{y}$ is the dual variable and ρ is the penalty parameter.

The ADMM algorithm is similar to the dual decomposition and the method of multipliers and consists of the three steps shown bellow: the minimization of the augmented Lagrangian over each of the two variables, $\mathbf{x}$ and $\mathbf{z}$, and the update of the dual variable.

$$\begin{aligned} \mathbf{x}^{k+1} &:= \underset{\mathbf{x}}{\text{argmin}}\, L_\rho(\mathbf{x}, \mathbf{z}^k, \mathbf{y}^k), \\ \mathbf{z}^{k+1} &:= \underset{\mathbf{z}}{\text{argmin}}\, L_\rho(\mathbf{x}^{k+1}, \mathbf{z}, \mathbf{y}^k), \\ \mathbf{y}^{k+1} &:= \mathbf{y}^k + \rho(\mathbf{Ax}^{k+1} + \mathbf{Bz}^{k+1} - \mathbf{c}). \end{aligned} \tag{2.18}$$

The difference between the method of multipliers and ADMM is that the updates of $\mathbf{x}$ and $\mathbf{z}$ are done jointly in the former, but in an alternate fashion in the latter ($\mathbf{z}^{k+1}$ depends on $\mathbf{x}^{k+1}$), and this allows the decomposition when f or g are separable. In this sense, ADMM is a robust method for solving large-scale optimization problems in a decentralized and parallel way. In the first part of this thesis, ADMM is used as the strategy for the decentralization of the proposed algorithm.

2.5 Machine Learning

Machine learning ML, which is a subset of artificial intelligence, has a crucial role in the future of wireless communications. With ML as a driving technology, networks are expected to be more intelligent, by providing more flexible features such as self-configuration, self-optimization, and self-healing and by simplifying and improving the transport of real-time data (73) (27).

ML techniques allow machines to learn the execution of a particular task, with the goal of maintaining a specific performance metric, based on a particular experience and to improve its decision-making capabilities as they acquire more knowledge (25). ML is widely expected to become a key component of beyond 5G networks since it offers considerable potential benefits, from which we highlight (74):

- *Extracting information from data*: Channel and interference models are extremely complicated in reality due to the dynamic nature of wireless communication channels. ML techniques may automatically extract the unknown channel information by learning from the communication data and prior knowledge.
- *Large scale optimization*: As the density of wireless access points continues to increase, there is an urgent need for global optimization of communication resources. However, the size, number of variables, and coupling relation among entities make optimization a difficult task. On the other hand, ML algorithms (e.g., deep learning) may be able to model the highly nonlinear correlations and estimate system parameters.
- *Adapting to fluctuating environments*: The dynamic nature of wireless channels, traffic, and interference in multi-user scenarios makes the constant adaptation to the varying environment a burdensome task. ML will realize learning-based adaptive configuration of networks by finding out behavioral patterns and responding quickly and flexibly to various scenarios.

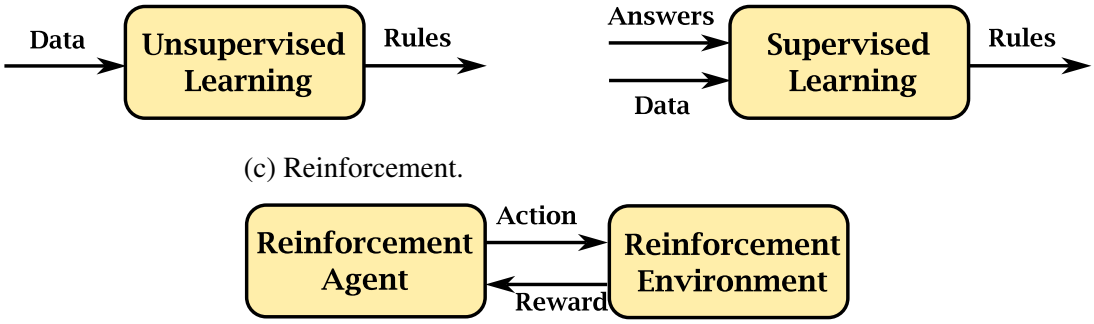
The way that an ML application takes form varies according to the nature of the agents, environment, and data, and the way they interact with each other. ML algorithms are typically classified into three main categories, namely supervised learning, unsupervised learning, and reinforcement learning, see Figure 5. In supervised learning, the goal is to predict an outcome $\mathbf{y}$ from an entry value $\mathbf{x}$ by looking at several examples of a random vector $\mathbf{x}$ and its label value of vector $\mathbf{y}$, by estimating $p(\mathbf{y}|\mathbf{x})$ or particular properties of that distribution. Unsupervised learning implicates observing different instances of a random vector $\mathbf{x}$ aiming to learn its probability distribution $p(\mathbf{x})$. In reinforcement learning an agent interacts with the surrounding environment to learn a policy that maximizes the rewards obtained by its actions on the environment (26).

Reinforcement learning is one of the most important research directions of machine learning, which has significant impacts on the development of AI over the last 20 years (75). In reinforcement learning, an agent interacts with the environment, taking actions and obtaining rewards from it. In general, the agents do not have full information about the environment, but they are considered to know a subset of parameters that delimits a state of the environment. The experience from rewards received from past actions defines which actions are good or bad for the agent. This way, the agent must keep exploring actions to discover the ones that maximize the rewards. Actions also change the state of the environment, affecting the actual state and

Figure 5 – Machine learning strategies: Unsupervised, supervised and reinforcement learning.

(a) Unsupervised.

(b) Supervised.



(c) Reinforcement.

Source: Adapted from (26).

future ones. Therefore the agent's goal is to learn over time a policy that defines the best possible actions to maximize future rewards over the long run, given the current state of the environment. (76).

In conventional reinforcement learning applications, a simple lookup table can be used to keep track of states, actions, and expected rewards. Therefore, the application of reinforcement learning in the optimization of modern networks that are large-scale and complicated, leads the computational complexity to rapidly become unmanageable. As a result, deep Q-learning has been developed to be an alternative solution to overcome this challenge (75). Deep Q-learning or deep Q-Network (DQN) corresponds to a merge between the Q-learning algorithm and a deep neural network. In summary, a lookup table is replaced by a deep neural network. Notably, recent advances in deep neural networks, in which several layers of nodes are used to buildup progressively more abstract representations of the data, have made it possible for artificial neural networks to learn concepts such as object categories directly from raw sensory data (77). A full overview of Deep Q-learning is provided in Section 4.5.

In the second part of this thesis, we employ reinforcement learning to efficiently solve a user scheduling problem for the optimization of beamforming. More specifically, we employ the deep Q-learning paradigm to perform multi-agent optimization in an intelligent, decentralized, robust, and fast way.

3 BIDIRECTIONAL SUM-POWER MINIMIZATION BEAMFORMING IN DYNAMIC TDD MULTI-USER MULTI-STREAM MIMO NETWORKS

Employing dynamic time division duplexing (TDD) can increase the system-wide spectral efficiency of applications with varying and unbalanced downlink (DL) and uplink (UL) data traffic requirements. However, in order to achieve this efficiency gain, it is necessary to manage the effects of cross-link interference, which are generated among cells transmitting in opposite link directions.

This chapter considers bidirectional sum-power minimization beamforming in a multi-user multi-stream multiple-input multiple-output (MIMO) network, as a means to deal with this cross-link interference, by forcing a minimum signal-to-interference-plus-noise ratio constraint for both UL and DL.

We propose two iterative approaches to solve this beamforming problem. The first approach assumes centralized processing and requires the availability of global channel state information. The second approach is performed in a decentralized manner, based on the alternating direction method of multipliers (ADMM) and requires only local channel state information and reduced signaling load. Both approaches are shown to converge to a minimum network power expenditure, whereas close-to-optimum performance can be obtained when limiting the number of iterations.

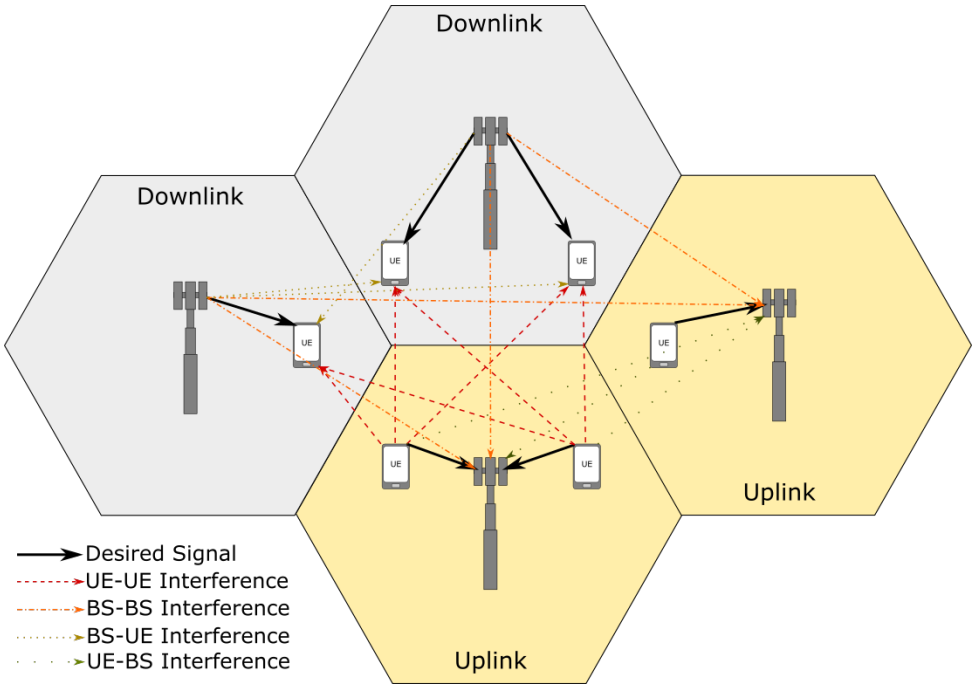
The remainder of the chapter is structured as follows. Section 3.1 provides a brief introduction to the topic of dynamic TDD and to the studies in this Chapter. Section 3.2 discusses related works and the contributions of this chapter. Section 3.3 introduces the system model. Section 3.4 formulates the optimization problem and describes the centralized beamforming (BF) solution. Section 3.5 reformulates the problem and describes the ADMM-based decentralized solution. Section 3.6 provides a complete proof of convergence for the proposed algorithms. Section 3.7 presents a signaling scheme to support the proposed decentralized algorithm. Section 3.8 provides numerical examples for the proposed algorithms before conclusions are drawn in Section 3.9.

3.1 Introduction

A promising way of accommodating the increasing number of mobile users and devices and their high-capacity requirements is the deployment of small cells (18). In such systems, using TDD is preferred over frequency division duplexing (FDD) since it can better match traffic asymmetry between the UL and DL. Also, TDD takes advantage of channel reciprocity between UL and DL, which reduces the complexity and the amount of required channel feedback signaling to acquire channel state information (CSI) compared with the FDD case (78).

In previous generations, systems that employ TDD have used fixed time intervals

Figure 6 – Different links in a dynamic TDD scenario: desired and interference links (UE-UE, BS-BS, BS-UE and UE-BS).



Source: Created by the author.

for each link direction (UL and DL), assuming that all cells transmit in the same direction at any given point in time. This design avoids creating interference between base stations (BSs) and between user equipments (UEs) transmitting in opposite directions (79). However, the reduced cell size allows the instantaneous traffic demands to vary significantly between cells (80). Therefore, a dynamic reconfiguration between UL and DL can increase the spectral efficiency of systems supporting varying or unbalanced data requirements in the UL and DL directions.

However, the coexistence of different link directions in neighbor cells gives rise to additional types of interference: between BSs and between UEs located in neighbor cells that transmit in opposite directions. Figure 6 illustrates these interferences in a four cell scenario. The BSs transmitting in DL generate BS-to-BS interference to the BS receiving in UL, while the UE transmitting in UL generates UE-to-UE interference to UEs operating in the opposite direction. Therefore, it is intuitively clear that in order to achieve the gains of dynamic reconfiguration between UL and DL, it is crucial to mitigate the negative effects of the additional interference.

Due to the presence of BS-to-BS and UE-to-UE interference, the performance of dynamic TDD systems depends critically on proper scheduling and time slot assignment along with an appropriate beam allocation. In practice, the cell and time slot configurations (i.e. whether the time slots in a cell operate in UL or DL direction) and the time slot allocation decisions (i.e. which time slot is assigned to which user for UL or DL transmissions) are made advantageously separately from the precoding decisions. That is, in practice, in order to manage the complexity of traffic and quality of service (QoS)-dependent scheduling – as part of the medium access control procedures – and configuring the UL and DL precoding weights, suboptimal schemes can

be used. Specifically, such suboptimal schemes assume that scheduling and UL/DL configuration and slot assignment have been made and develop BF solutions to manage UL/DL interference. Such strategy is found in (81), where the slot assignment is assumed prior to BF and in (82) that uses direct beamformer estimation in dynamic TDD.

In the long term evolution (LTE) release 12, the enhanced interference mitigation and traffic adaptation (eIMTA) concept was introduced, and 7 different UL-DL frame configurations have been standardized for dynamic TDD, as shown in table 3. In this table the letter "D" and "U" refer to downlink and uplink subframes, while the letter "S" refers to a special subframe used as guard interval. In this standard, each BS can select one of the configurations according to its decision, with a DL/UL ratio varying from 40/60 to 90/10 (83).

Table 3 – Dynamic TDD UL-DL configuration on LTE.

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

Source: Adapted from (83).

In this chapter, we consider a dynamic TDD system in which any BS can operate in UL or DL mode in a given time slot configured by a suitable scheduling or operation and management entity, as in, for example, (84) and (85). This means that we do not restrict our link direction decision to the possibilities in table 3. This scheduling decision may lead to an interference situation, similar to the one shown in Figure 6, illustrating the presence of diverse interference links.

In this work, we propose taking advantage of the typical multi-antenna deployments of cellular bases stations and employing BF as a way of mitigating the effects of interference in dynamic TDD networks. Such solutions seek to minimize the system-wide sum-power and to manage the effects of cross-link interferences by guaranteeing a minimum signal-to-interference-plus-noise ratio (SINR) threshold for each user in both the UL and DL directions. Sum-power minimization in densely deployed small-cell networks is a highly motivated objective, since reduced power consumption helps the large-scale sustainable deployment of such networks and reduces the operational expenditures.

In a multi-cell scenario, BF can be coordinated by a central node or in a decentralized manner. Accordingly, in this chapter, we propose two BF schemes to mitigate interference, and

compare their performance in terms of system-wide power consumption. The first solution, devised for benchmarking purposes, assumes the availability of a central entity that has access to global CSI. The second, a practically viable approach uses a decentralized algorithm based on ADMM, and requires local CSI available in each cell. It uses a lightweight backhaul and over-the-air signaling scheme in order to exchange the required run-time information between cells. We find that both approaches can achieve the optimum solution, which minimizes sum-power expenditure while maintaining the required SINR levels. In both schemes, the total signaling load can be controlled by limiting the number of iterations, at the cost of achieving close-to-optimal performance.

Throughout this chapter, the bold lowercase letter $\mathbf{x}$ represents a vector and the bold uppercase letter $\mathbf{X} \in \mathbb{C}^{M \times N}$ is used to denote a matrix drawn from the $M \times N$ matrix space defined on the complex field. $\mathbf{X}^T$, $\mathbf{X}^H$, $\mathbf{X}^{-1}$ and $\mathbf{X}^\dagger$ stand, respectively, for transpose, Hermitian, inverse and pseudo-inverse of a matrix $\mathbf{X}$. $\mathbf{X} \geq 0$ means that $\mathbf{X}$ is positive semidefinite. For a vector $\mathbf{x}$, $\|\mathbf{x}\|_2$ denotes the Euclidean norm. $\{x_i\}_{\forall i}$ represents a set of elements x_i in the values of i denoted by the subscript expression.

3.2 Related Works and Contributions

Some strategies to manage interference in dynamic TDD systems have been previously proposed, including clustering, time slot allocation and power control. Clustering schemes are proposed in (86), according to which neighboring cells are grouped into clusters that use the same link direction so as to eliminate cross-link interference. Time slot allocation algorithms that avoid allocating adjacent time slots to UEs located in the border area of neighboring cells have also been proposed to manage dynamic TDD interference, see for example (87) or (88) for a comparison between different time slot allocation strategies. Unfortunately, these solutions cannot accommodate unbalanced traffic demands in neighbor cells, because in these schemes neighbor cells with different traffic demands are forced to choose the same transmit direction or wait for a later time slot to transmit.

A power control solution is proposed in (89), which controls the BS-to-BS interference by reducing the transmission power of the aggressor BS. Another solution is proposed in (90), which increases the transmission power of the UE that has its communication affected by interference. However, neither of these approaches is optimal, since reducing the BS power directly affects the DL capacity, while increasing the UE transmit power leads to a high battery consumption, which is an important issue for mobile transceivers. These previous interference management techniques do not make use of the multi-antenna ability to provide more degrees of freedom in order to deal with interference.

Another very well known strategy to manage interference is beamforming (BF). Numerous BF solutions have been proposed for non-dynamic TDD multiple-input single-output (MISO) and MIMO systems and focusing on varied optimization objectives. For example, the

works presented by (47) and (91) consider a sum-rate or sum-utility maximization problem and impose a power budget at each transmitter. The work in (47) is shown to minimize the weighted sum of mean square error (MSE) and maximizes sum rate in the DL interference broadcast channel scenario, by using the block coordinate descent technique. Reference (91) minimizes the system-wide MSE and maximizes the sum rate for a K-user MIMO system using an iterative and reconfigurable distributed scheme. By contrast, the proposed methodology in this chapter aims to minimize the total transmission power, while guaranteeing a quality target for each user. Sum-power minimization BF is also considered as the optimization objective by (92) for the MISO scenario, by (93) for the MISO scenario with limited backhaul information exchange, by (42) and (94) for a multiuser multiple input multiple output (MU-MIMO) system and by (43) for the multi-cell MIMO system. A performance analysis of coordinated BF schemes considering multiple objectives is provided in (95) and (96) for the DL multi-cell MIMO scenarios considering CSI uncertainties and LTE specifications, respectively.

An important class of state-of-the-art techniques is machine learning-based solutions, which are recently being used to optimize several objectives in mobile networks, including BF with the objective of sum-rate maximization and sum-power minimization. The works presented in (97), (98), and (99), for example, propose learning-based solutions for transceiver design. The first two focus on downlink transmission in a single-cell scenario, while the last one extends the learning ideas of the others to a dynamic environment.

Although BF has been considered in a variety of scenarios, the results obtained by the previously mentioned works do not carry over to dynamic TDD systems, since they are not applicable in the presence of BS-to-BS and UE-to-UE interference. Recognizing this aspect, BF solutions that are applicable in dynamic TDD systems were previously proposed considering a MISO deployment. In (100), the authors propose a strategy to maximize the energy efficiency. However, that solution requires a central entity and is based on non-linear fractional programming, which has a high complexity, rendering it unpractical in dense small-cell networks. In contrast, the authors in (101) presented a solution using a pricing approach that penalizes the DL BSs which cause interference. Using that approach, the DL capacity suffers in order to reduce the interference in the UL. Finally, the authors of (102) proposed a sum-power minimization decentralized approach to optimize the DL BS transmitters with SINR constraints for the DL users and a maximum interference power constraint to control BS-to-BS interference.

It is important to realize that in a MIMO scenario, the interference management can be further improved by benefiting from the multi-antenna capability on the UEs side. In such scenario, a solution is presented in (81), where the authors propose a strategy to maximize the weighed sum rate, along with a signaling scheme that supports the optimization. Another related work (103) considers sum-power minimization BF, imposing SINR constrains for each DL transmission and maximum interference power constraints for BS-to-BS interference. However, in that work the UL transmitters are considered fixed, which unfortunately implies that only the BS power is minimized and that the UE power expenditure is not optimized. Also, in that

work the interference caused by the UEs to other nodes – including UE-to-UE and UE-to-BS interference – is not treated.

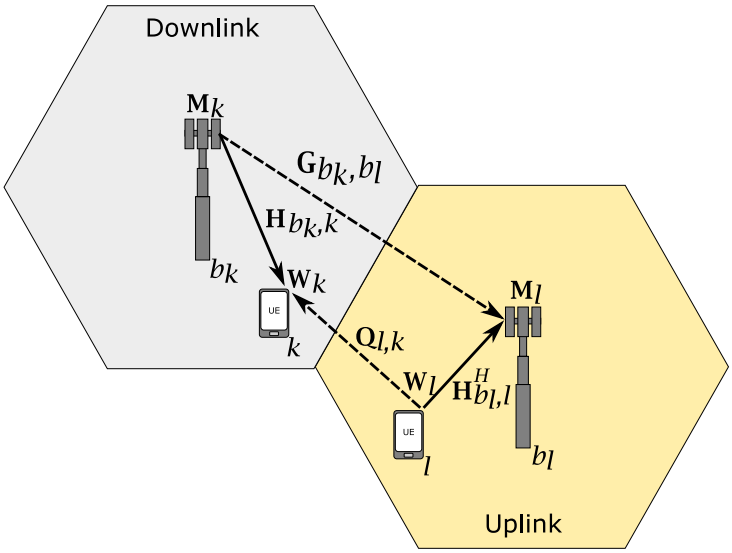
Along another line of research, a full-duplex deployment is considered, in which both BSs and UEs can transmit and receive simultaneously. In this scenario, the interference situation is similar to the one present in dynamic TDD networks, but it includes an additional self-interference term. BF solutions have also been proposed for such a scenario, see for example (104), (105) and (106). In the first two works the optimization objectives differ from the objective of the present chapter, since we consider sum-power minimization, while (104), (105) consider sum rate maximization and weighted MSE minimization. Also, from the architecture perspective, (105) focuses on a single-cell scenario, which lacks the effects of inter-cell interference. On the other hand, the work presented in (106) has the objective to minimize the sum-power in a K-link full-duplex MIMO scenario. Although (106) has a similar objective, it differs from our work with regards to the solution approach, which for that paper is based on a penalty method along with space-time scheduling, including the time dimension to the optimization objective. Another important difference is that these three works do not explicitly analyze a distributed implementation of the proposed solutions, which is one of the main contributions of our work. In this sense, although a full-duplex deployment may be similar to a dynamic TDD scenario, the full-duplex solutions cannot be easily extended to fulfill the requirements imposed by our objectives and the dynamic TDD scenario.

It is therefore intuitively clear that further gains in performance can be achieved in dynamic TDD MIMO scenarios employing sum-power minimization BF, by allowing both DL and UL transmitters to be optimized with respect to transmit power, and by imposing predetermined SINR constraints for DL and UL reception.

Accordingly, in the present chapter we consider a solution approach based on sum-power minimization BF in dynamic TDD multi-user multi-stream MIMO networks. Unlike previous works, we aim to manage all types of interference, including UE-to-UE interference, by enforcing SINR constraints to each user both in the UL and DL. This task is challenging, since the proposed schemes must also manage the inter-user inter-stream interference, which is inherently present in multi-user multi-stream MIMO systems. Also, the proposed schemes enable the UEs to minimize their transmit power while also minimizing the transmit power levels of the BSs. We believe that minimizing both the UE and the network transmit power levels in dynamic TDD multi-cell MIMO systems is a novel and important contribution.

In the present work we provide both centralized and decentralized solutions. The decentralized scheme is based on ADMM, and it is developed along with a light-weight signaling structure in order to facilitate the proposed decentralized coordination. With respect to the decentralization procedure, an important related work is (107). Such work also uses ADMM as a tool to allow a distributed implementation, and aims to minimize sum-power in the DL multi-cell MISO scenario. In contrast, our scheme considers multi-antenna UEs, and they can act as both receivers and transmitters. Therefore, the distributed transceiver design is significantly more

Figure 7 – Main notation symbols of the dynamic TDD system model.



Source: Created by the author.

complex and challenging, since the transceivers have a general structure and the interference pattern is markedly different from that in the DL of a regular cellular system.

Another main contribution of our work is the convergence analysis of the proposed algorithms. Related convergence analysis are presented by (94) and (108) for the iterative transmit and receiver optimization with the objective to minimize the sum-power in a DL MIMO multi-user scenario. In contrast, our work is inherently multi-cell, since we consider a multi-cell dynamic TDD system, in which not only multi-cell interference must be combated, but this multi-cell interference has a complex structure due to the UL-DL cross-interference.

In order to highlight, even more, the contributions of this work with respect to some of the main related papers, we summarize the differences in terms of the objective function, constraints, architecture, centralized or distributed solution approach, CSI availability and the merits of the proposed solution in Table 10 located in Appendix A in the end of this work.

3.3 System Model

We consider a multi-cell multi-stream dynamic TDD multi-user MIMO system with a group formed by B BSs, each equipped with N_b antennas, and K UEs, each equipped with N_u antennas. Figure 7 Illustrates the main notation used to describe this chapter's system model.

Let $\mathcal{B}_{ul}$, $\mathcal{B}_{dl}$, $\mathcal{K}_{ul}$ and $\mathcal{K}_{dl}$ represent the sets of BSs and UEs in the UL and DL mode, respectively, and $\mathcal{S}_k$ denote the set of streams of UE k. Then, the signal of all streams received

by UE $k \in \mathcal{K}_{dl}$, $\mathbf{y}_k^{dl} = [y_{k,1}^{dl}, \ldots, y_{k,S}^{dl}]^T \in \mathbb{C}^{S\times 1}$, can be written as

$$\mathbf{y}_k^{dl} = \underbrace{\mathbf{W}_k^H \mathbf{H}_{b_k,k} \mathbf{M}_k \mathbf{d}_k}_{\text{useful signal + ISI}} + \underbrace{\sum_{i\in\mathcal{K}_{dl}\setminus k} \mathbf{W}_k^H \mathbf{H}_{b_i,k} \mathbf{M}_i \mathbf{d}_i}_{\text{BS to UE interf.}} + \underbrace{\sum_{j\in\mathcal{K}_{ul}} \mathbf{W}_k^H \mathbf{Q}_{j,k} \mathbf{W}_j \mathbf{d}_j}_{\text{UE to UE interf.}} + \underbrace{\mathbf{W}_k^H \mathbf{n}_k}_{\text{noise}}, \tag{3.1}$$

where $\mathbf{H}_{b_i,k} \in \mathbb{C}^{N_u\times N_b}$ denotes the channel between the BS that serves UE i (denoted as b_i) and UE k; $\mathbf{M}_i = [\mathbf{m}_{i,1}, \ldots, \mathbf{m}_{i,S}] \in \mathbb{C}^{N_b\times S}$ denotes a BF BS filter with respect to UE i, while $\mathbf{W}_i = [\mathbf{w}_{i,1}, \ldots, \mathbf{w}_{i,S}] \in \mathbb{C}^{N_u\times S}$ denotes the BF filter at UE i, independently from the link direction; $\mathbf{d}_i = [d_{i,1}, \ldots, d_{i,S}]^T \in \mathbb{C}^{S\times 1}$ denotes the data symbols relative to UE i, which is assumed to have unit variance; $\mathbf{Q}_{j,k} \in \mathbb{C}^{N_u\times N_u}$ denotes the channel between UEs j and k, and $\mathbf{n}_k \in \mathbb{C}^{N_u\times 1} \sim N(0, N_0)$ represents the noise at the link of UE k. Note that the second term of (3.1) includes intra-cell and inter-cell interference.

On the other hand, the signal received by BS $b_l \in \mathcal{B}_{ul}$, sent from its served UE $l \in \mathcal{K}_{ul}$, $\mathbf{y}_k^{ul} = [y_{k,1}^{ul}, \ldots, y_{k,S}^{ul}]^T \in \mathbb{C}^{S\times 1}$, is written as

$$\mathbf{y}_l^{ul} = \underbrace{\mathbf{M}_l^H \mathbf{H}_{b_l,l}^H \mathbf{W}_l \mathbf{d}_l}_{\text{useful signal + ISI}} + \underbrace{\sum_{j\in\mathcal{K}_{ul}\setminus l} \mathbf{M}_l^H \mathbf{H}_{b_l,j}^H \mathbf{W}_j \mathbf{d}_j}_{\text{UE to BS interf.}} + \underbrace{\sum_{i\in\mathcal{K}_{dl}} \mathbf{M}_l^H \mathbf{G}_{b_i,b_l} \mathbf{M}_i \mathbf{d}_i}_{\text{BS to BS interf.}} + \underbrace{\mathbf{M}_l^H \mathbf{n}_l}_{\text{noise}}, \tag{3.2}$$

where $\mathbf{G}_{b_i,b_l} \in \mathbb{C}^{N_b\times N_b}$ denotes the channel between the BS $b_i \in \mathcal{B}_{dl}$ and b_l. Channel reciprocity is used to obtain the channels between UEs and BSs ($\mathbf{H}_{l,b_l} = \mathbf{H}_{b_l,l}^H$).

Considering the reception of each individual stream s, the received signal expressions can also be formulated as

$$\begin{aligned} y_{k,s}^{dl} = {} & \underbrace{\mathbf{w}_{k,s}^H \mathbf{H}_{b_k,k} \mathbf{m}_{k,s} d_{k,s}}_{\text{useful signal}} + \underbrace{\sum_{s'\neq s} \mathbf{w}_{k,s}^H \mathbf{H}_{b_k,k} \mathbf{m}_{k,s'} d_{k,s'}}_{\text{ISI}} \\ & + \underbrace{\sum_{i\in\mathcal{K}_{dl}\setminus k} \mathbf{w}_{k,s}^H \mathbf{H}_{b_i,k} \mathbf{M}_i \mathbf{d}_i}_{\text{BS to UE interf.}} + \underbrace{\sum_{j\in\mathcal{K}_{ul}} \mathbf{w}_{k,s}^H \mathbf{Q}_{j,k} \mathbf{W}_j \mathbf{d}_j}_{\text{UE to UE interf.}} + \underbrace{\mathbf{w}_{k,s}^H \mathbf{n}_k}_{\text{noise}}, \end{aligned} \tag{3.3}$$

$$\begin{aligned} y_{l,s}^{ul} = {} & \underbrace{\mathbf{m}_{l,s}^H \mathbf{H}_{b_l,l}^H \mathbf{w}_{l,s} d_{l,s}}_{\text{useful signal}} + \underbrace{\sum_{s'\neq s} \mathbf{m}_{l,s}^H \mathbf{H}_{b_l,l}^H \mathbf{w}_{l,s'} d_{l,s'}}_{\text{ISI}} \\ & + \underbrace{\sum_{j\in\mathcal{K}_{ul}\setminus l} \mathbf{m}_{l,s}^H \mathbf{H}_{b_l,j}^H \mathbf{W}_j \mathbf{d}_j}_{\text{UE to BS interf.}} + \underbrace{\sum_{i\in\mathcal{K}_{dl}} \mathbf{m}_{l,s}^H \mathbf{G}_{b_i,b_l} \mathbf{M}_i \mathbf{d}_i}_{\text{BS to BS interf.}} + \underbrace{\mathbf{m}_{l,s}^H \mathbf{n}_l}_{\text{noise}}. \end{aligned} \tag{3.4}$$

3.4 Centralized Approach

The BF optimization problem is to minimize the total transmit sum power, while satisfying a minimum SINR threshold for every stream, in both the UL and DL directions. The

power of each multi-antenna node is represented by the squared Frobenius norm of the transmit vectors of each transmission, which are represented as the matrix $\mathbf{M}_i$ for each user i in the DL direction, and as the matrix $\mathbf{W}_j$ for each user j in the UL direction. Since our objective is to minimize the sum-power, we sum the power of each transmitter in the objective function. This problem can be conveniently expressed as an optimization task as:

$$\begin{aligned} \min_{\{\mathbf{M}_i, \mathbf{W}_i\}_{\forall i}} \; & \sum_{i \in \mathcal{K}_{dl}} \|\mathbf{M}_i\|_F^2 + \sum_{j \in \mathcal{K}_{ul}} \|\mathbf{W}_j\|_F^2 \\ \text{s. t.} \quad & \Gamma_{k,s}^{dl} \geq \gamma_{k,s}, \quad \forall k \in \mathcal{K}_{dl}, \quad \forall s \in \mathcal{S}_k, \\ & \Gamma_{l,s}^{ul} \geq \gamma_{l,s}, \quad \forall l \in \mathcal{K}_{ul}, \quad \forall s \in \mathcal{S}_l, \end{aligned} \tag{3.5}$$

where $\Gamma_{k,s}^{dl}$ and $\Gamma_{l,s}^{ul}$ denote the SINR for the stream s that is measured at UE k in DL and at the serving BS of UE l in UL, respectively; $\gamma_{i,s}$ denotes the minimum SINR value for the communication link of stream s of user i, and can be written as:

$$\Gamma_{k,s}^{dl} = \frac{|\mathbf{w}_{k,s}^H \mathbf{H}_{b_k,k} \mathbf{m}_{k,s}|^2}{\sum\limits_{s' \neq s} |\mathbf{w}_{k,s}^H \mathbf{H}_{b_k,k} \mathbf{m}_{k,s'}|^2 + \sum\limits_{i \in \mathcal{K}_{dl} \setminus k} ||\mathbf{w}_{k,s}^H \mathbf{H}_{b_i,k} \mathbf{M}_i||^2 + \sum\limits_{j \in \mathcal{K}_{ul}} ||\mathbf{w}_{k,s}^H \mathbf{Q}_{j,k} \mathbf{W}_j||^2 + ||\mathbf{w}_{k,s}||^2 N_0}, \tag{3.6}$$

$$\Gamma_{l,s}^{ul} = \frac{|\mathbf{m}_{l,s}^H \mathbf{H}_{b_l,l}^H \mathbf{w}_{l,s}|^2}{\sum\limits_{s' \neq s} |\mathbf{m}_{l,s}^H \mathbf{H}_{b_l,l}^H \mathbf{w}_{l,s'}|^2 + \sum\limits_{j \in \mathcal{K}_{ul} \setminus l} ||\mathbf{m}_{l,s}^H \mathbf{H}_{b_l,j}^H \mathbf{W}_j||^2 + \sum\limits_{i \in \mathcal{K}_{dl}} ||\mathbf{m}_{l,s}^H \mathbf{G}_{b_i,b_l} \mathbf{M}_i||^2 + ||\mathbf{m}_{l,s}||^2 N_0}. \tag{3.7}$$

Problem (3.5) is not jointly convex in $\{\mathbf{M}_i\}_{\forall i}$ and $\{\mathbf{W}_i\}_{\forall i}$. However, as we will show in the sequel, it can be solved by an iterative approach, in which the transmit and receive beamformers are calculated in an alternate fashion until some stopping criterion.

To compute the transmit terms $\{\{\mathbf{M}_i\}_{i \in \mathcal{K}_{dl}}, \{\mathbf{W}_i\}_{i \in \mathcal{K}_{ul}}\}$, the receivers $\{\{\mathbf{W}_i\}_{i \in \mathcal{K}_{dl}}, \{\mathbf{M}_i\}_{i \in \mathcal{K}_{ul}}\}$ must be fixed. On the other hand, to compute the receivers $\{\{\mathbf{W}_i\}_{i \in \mathcal{K}_{dl}}, \{\mathbf{M}_i\}_{i \in \mathcal{K}_{ul}}\}$, the transmitters $\{\{\mathbf{M}_i\}_{i \in \mathcal{K}_{dl}}, \{\mathbf{W}_i\}_{i \in \mathcal{K}_{ul}}\}$ must be fixed. In the proposed centralized approach, these two steps are executed iteratively until some stopping criterion is reached, such as the relative power change between iterations or a predefined maximum number of iterations.

3.4.1 Transmit Filter Optimization

In the first step of this iterative process to solve Problem (3.5), the optimization variables are the transmit matrices $\{\mathbf{M}_i\}_{i \in \mathcal{K}_{dl}}$ and $\{\mathbf{W}_i\}_{i \in \mathcal{K}_{ul}}$. This sub-problem can be written as (3.8).

$$\begin{aligned} \min_{\substack{\{\mathbf{M}_i\}_{i \in \mathcal{K}_{dl}} \\ \{\mathbf{W}_i\}_{i \in \mathcal{K}_{ul}}}} \; & \sum_{i \in \mathcal{K}_{dl}} \|\mathbf{M}_i\|_F^2 + \sum_{j \in \mathcal{K}_{ul}} \|\mathbf{W}_j\|_F^2 \\ \text{s. t.} \quad & \Gamma_{k,s}^{dl} \geq \gamma_{k,s}, \quad \forall k \in \mathcal{K}_{dl}, \quad \forall s \in \mathcal{S}_k, \\ & \Gamma_{l,s}^{ul} \geq \gamma_{l,s}, \quad \forall l \in \mathcal{K}_{ul}, \quad \forall s \in \mathcal{S}_l. \end{aligned} \tag{3.8}$$

In the sequel we show how Problem (3.8) can be cast into a semidefinite programming (SDP) convex form.

Thus, Problem (3.8) can be written as the relaxed SDP formulation in (3.9) by replacing the rank-one matrices $\mathbf{m}_{i,s}\mathbf{m}_{i,s}^H$ and $\mathbf{w}_{i,s}\mathbf{w}_{i,s}^H$ by general-rank positive semidefinite matrices $\hat{\mathbf{M}}_{i,s} \succeq 0 \in \mathbb{C}^{N_b \times N_b}$ and $\hat{\mathbf{W}}_{i,s} \succeq 0 \in \mathbb{C}^{N_u \times N_u}$ respectively. Consequentially, we also write the multi-stream multiplications $\mathbf{M}_i\mathbf{M}_i^H$ and $\mathbf{W}_i\mathbf{W}_i^H$ as $\hat{\mathbf{M}}_i \in \mathbb{C}^{N_b \times N_b}$ and $\hat{\mathbf{W}}_i \in \mathbb{C}^{N_u \times N_u}$, respectively:

$$
\begin{aligned}
\min_{\substack{\{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_{dl}} \\ \{\hat{\mathbf{W}}_i\}_{i\in\mathcal{K}_{ul}}}} \quad & \sum_{b\in\mathcal{B}_{dl}} \mathrm{Tr}(\tilde{\mathbf{M}}_b) + \sum_{k\in\mathcal{K}_{ul}} \mathrm{Tr}(\hat{\mathbf{W}}_k) \qquad (3.9)\\
\text{s. t. } & \mathbf{w}_{k,s}^H \mathbf{H}_{b_k,k}\mathbf{D}_{k,s}\mathbf{H}_{b_k,k}^H \mathbf{w}_{k,s} \geq \sum_{b'\in\mathcal{B}_{dl}\setminus b_k} \mathbf{w}_{k,s}^H \mathbf{H}_{b',k}\tilde{\mathbf{M}}_{b'}\mathbf{H}_{b',k}^H \mathbf{w}_{k,s} \\
& + \sum_{b''\in\mathcal{B}_{ul}} \mathbf{w}_{k,s}^H \Bigg(\sum_{i\in\mathcal{K}_{b''}} \mathbf{Q}_{i,k}\hat{\mathbf{W}}_i\mathbf{Q}_{i,k}^H \Bigg)\mathbf{w}_{k,s} + N_0, \ \forall k \in \mathcal{K}_{dl}, \forall s \in \mathcal{S}_k, \\
& \mathbf{m}_{k,s}^H \mathbf{U}_{k,s}\mathbf{m}_{k,s} \geq \sum_{b'\in\mathcal{B}_{dl}} \mathbf{m}_{k,s}^H \mathbf{G}_{b',b_k}\tilde{\mathbf{M}}_{b'}\mathbf{G}_{b',b_k}^H \mathbf{m}_{k,s} \\
& + \sum_{b''\in\mathcal{B}_{ul}\setminus b_k} \mathbf{m}_{k,s}^H \Bigg(\sum_{i\in\mathcal{K}_{b''}} \mathbf{H}_{b_k,i}^H \hat{\mathbf{W}}_i \mathbf{H}_{b_k,i} \Bigg)\mathbf{m}_{k,s} + N_0, \ \forall k \in \mathcal{K}_{ul}, \forall s \in \mathcal{S}_k, \\
& \hat{\mathbf{M}}_{k,s} \succeq 0, \quad \forall k \in \mathcal{K}_{dl}, \forall s \in \mathcal{S}_k, \\
& \hat{\mathbf{W}}_{k,s} \succeq 0, \quad \forall k \in \mathcal{K}_{ul}, \forall s \in \mathcal{S}_k,
\end{aligned}
$$

where $\mathcal{K}_b$ denotes the set of users served by BS b and

$$
\mathbf{D}_{i,s} = \frac{1}{\gamma_{i,s}}\hat{\mathbf{M}}_{i,s} - \sum_{s'\neq s} \hat{\mathbf{M}}_{i,s'} - \sum_{j\in\mathcal{K}_{b_i}\setminus i} \hat{\mathbf{M}}_j, \qquad (3.10)
$$

$$
\tilde{\mathbf{M}}_b = \sum_{i\in\mathcal{K}_b} \hat{\mathbf{M}}_i, \qquad (3.11)
$$

$$
\mathbf{U}_{i,s} = \mathbf{H}_{b_i,i}^H \Bigg(\frac{1}{\gamma_{i,s}}\hat{\mathbf{W}}_{i,s} - \sum_{s'\neq s} \hat{\mathbf{W}}_{i,s'} \Bigg)\mathbf{H}_{b_i,i} - \sum_{j\in\mathcal{K}_{b_i}\setminus i} \mathbf{H}_{b_i,j}^H \hat{\mathbf{W}}_j \mathbf{H}_{b_i,j}. \qquad (3.12)
$$

By using the sets of variables presented in Table 4, the optimization problem becomes (3.13), with SINR constraints for DL and UL rewritten as the convex conic sets in (3.14) and (3.15).

$$
\begin{aligned}
\min_{\substack{\boldsymbol{\chi},\boldsymbol{\Omega},\boldsymbol{\theta},\boldsymbol{\psi}, \\ \{p_i\}_{i\in\mathcal{B}_{dl}\cup\mathcal{K}_{ul}} \\ \{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_{dl}},\{\hat{\mathbf{W}}_i\}_{i\in\mathcal{K}_{ul}}}} \quad & \sum_{i\in\mathcal{B}_{dl}\cup\mathcal{K}_{ul}} p_i \qquad (3.13)\\
\text{s. t.} \quad & \Big(\{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_b}, \boldsymbol{\chi}, \boldsymbol{\psi}, \boldsymbol{\theta}, p_b \Big) \in \mathcal{D}_b, \ \forall b \in \mathcal{B}_{dl}, \\
& \Big(\hat{\mathbf{W}}_l, \boldsymbol{\Omega}, \boldsymbol{\psi}, \boldsymbol{\theta}, p_l \Big) \in \mathcal{U}_l, \ \forall l \in \mathcal{K}_{ul},
\end{aligned}
$$

where p_i is an auxiliary variable which accounts for the total power used by i, and

$$\begin{aligned}
\mathcal{D}_b = \Bigg\{ \Bigg(& \{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_b}, \boldsymbol{\chi}, \boldsymbol{\psi}, \boldsymbol{\theta}, p_b \Bigg) \Bigg| \ \mathrm{Tr}\{\tilde{\mathbf{M}}_b\} = p_b, \\
& \mathbf{w}_{k,s}^H \mathbf{H}_{b,k} \mathbf{D}_{k,s} \mathbf{H}_{b,k}^H \mathbf{w}_{k,s} \geq \sum_{b'\in\mathcal{B}_{dl}\setminus b} \chi_{b',k,s} + \sum_{l\in\mathcal{K}_{ul}} \theta_{l,k,s} + N_0, \ \forall k \in \mathcal{K}_b, \forall s \in \mathcal{S}_k, \\
& \mathbf{w}_{j,s}^H \mathbf{H}_{b,j} \tilde{\mathbf{M}}_b \mathbf{H}_{b,j}^H \mathbf{w}_{j,s} \leq \chi_{b,j,s}, \ \forall j \in \mathcal{K}_{dl}\setminus\mathcal{K}_b, \forall s \in \mathcal{S}_j, \\
& \mathbf{m}_{l,s}^H \mathbf{G}_{b,b_l} \tilde{\mathbf{M}}_b \mathbf{G}_{b,b_l}^H \mathbf{m}_{l,s} \leq \psi_{b,l,s}, \ \forall l \in \mathcal{K}_{ul}, \forall s \in \mathcal{S}_l, \\
& \hat{\mathbf{M}}_{k,s} \succeq 0, \ \forall k \in \mathcal{K}_b, \forall s \in \mathcal{S}_k, \Bigg\}, \forall b \in \mathcal{B}_{dl},
\end{aligned} \tag{3.14}$$

$$\begin{aligned}
\mathcal{U}_l = \Bigg\{ \Bigg(& \hat{\mathbf{W}}_l, \boldsymbol{\Omega}, \boldsymbol{\psi}, \boldsymbol{\theta}, p_l \Bigg) \Bigg| \ \mathrm{Tr}\{\hat{\mathbf{W}}_l\} = p_l, \\
& \mathbf{m}_{l,s}^H \mathbf{H}_{b_l,l}^H \left(\frac{1}{\gamma_{i,s}} \hat{\mathbf{W}}_{i,s} - \sum_{s'\neq s} \hat{\mathbf{W}}_{i,s'} \right) \mathbf{H}_{b_l,l} \mathbf{m}_{l,s} \geq \sum_{b\in\mathcal{B}_{dl}} \psi_{b,l,s} + \sum_{l'\in\mathcal{K}_{ul}\setminus l} \Omega_{l',l,s} + N_0, \forall s \in \mathcal{S}_l, \\
& \mathbf{w}_{k,s}^H \mathbf{Q}_{l,k} \hat{\mathbf{W}}_l \mathbf{Q}_{l,k}^H \mathbf{w}_{k,s} \leq \theta_{l,k,s}, \ \forall k \in \mathcal{K}_{dl}, \forall s \in \mathcal{S}_k, \\
& \mathbf{m}_{j,s}^H \mathbf{H}_{b_j,l}^H \hat{\mathbf{W}}_l \mathbf{H}_{b_j,l} \mathbf{m}_{j,s} \leq \Omega_{l,j,s}, \ \forall j \in \mathcal{K}_{ul}\setminus l, \forall s \in \mathcal{S}_j, \\
& \hat{\mathbf{W}}_{l,s} \succeq 0, \forall s \in \mathcal{S}_l, \Bigg\}, \forall l \in \mathcal{K}_{ul}.
\end{aligned} \tag{3.15}$$

Table 4 – Inter-cell interference sets

Set	Inter-cell interference
$\boldsymbol{\chi} = \{\chi_{b,k,s}\}_{b\in\mathcal{B}_{dl},\ k\in\mathcal{K}_{dl}\setminus\mathcal{K}_b,\forall s\in\mathcal{S}_k}$	DL BS to DL UE
$\boldsymbol{\psi} = \{\psi_{b,k,s}\}_{b\in\mathcal{B}_{dl},\ k\in\mathcal{K}_{ul},\forall s\in\mathcal{S}_k}$	DL BS to UL BS
$\boldsymbol{\Omega} = \{\Omega_{l,k,s}\}_{l\in\mathcal{K}_{ul},\ k\in\mathcal{K}_{ul}\setminus\mathcal{K}_b,\forall s\in\mathcal{S}_k}$	UL UE to UL BS
$\boldsymbol{\theta} = \{\theta_{l,k,s}\}_{l\in\mathcal{K}_{ul},\ k\in\mathcal{K}_{dl},\forall s\in\mathcal{S}_k}$	UL UE to DL UE

Source: Created by the author.

If perfect CSI is available, the solution of an SDP problem similar to Problem (3.13) is shown by (107) to satisfy the rank-one requirement. The analysis presented in that work is also valid in our case and the optimal $\{\mathbf{M}_i\}_{i\in\mathcal{K}_{dl}}$ and $\{\mathbf{W}_i\}_{i\in\mathcal{K}_{ul}}$ can be extracted from $\{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_{dl}}$ and $\{\hat{\mathbf{W}}_i\}_{i\in\mathcal{K}_{ul}}$.

3.4.2 Receive Filter Optimization

In the second step, the receive vectors are calculated, while keeping the transmit filters fixed. The optimal unit norm receivers $\{\mathbf{w}_{i,s}\}_{i\in\mathcal{K}_{dl},s\in\mathcal{S}_i}$ and $\{\mathbf{m}_{i,s}\}_{i\in\mathcal{K}_{ul},s\in\mathcal{S}_i}$ are the ones that maximize each stream SINR, and can be calculated using the minimum mean square error (MMSE) approach as

$$\mathbf{w}_{k,s} = \frac{\bar{\mathbf{w}}_{k,s}}{\|\bar{\mathbf{w}}_{k,s}\|}, \forall k \in \mathcal{K}_{dl}, \forall s \in \mathcal{S}_k, \tag{3.16}$$

$$\bar{\mathbf{w}}_{k,s} = \left(\sum_{i \in \mathcal{K}_{dl}} \mathbf{H}_{b_i,k} \mathbf{M}_i \mathbf{M}_i^H \mathbf{H}_{b_i,k}^H + \sum_{j \in \mathcal{K}_{ul}} \mathbf{Q}_{j,k} \mathbf{W}_j \mathbf{W}_j^H \mathbf{Q}_{j,k}^H + N_0 \mathbf{I}_{N_u} \right)^{-1} \mathbf{H}_{b_k,k} \mathbf{m}_{k,s}, \tag{3.17}$$

$$\mathbf{m}_{k,s} = \frac{\bar{\mathbf{m}}_{k,s}}{\|\bar{\mathbf{m}}_{k,s}\|}, \forall k \in \mathcal{K}_{ul}, \forall s \in \mathcal{S}_k, \tag{3.18}$$

$$\bar{\mathbf{m}}_{k,s} = \left(\sum_{i \in \mathcal{K}_{dl}} \mathbf{G}_{b_i,b_k} \mathbf{M}_i \mathbf{M}_i^H \mathbf{G}_{b_i,b_k}^H + \sum_{j \in \mathcal{K}_{ul}} \mathbf{H}_{b_k,j}^H \mathbf{W}_j \mathbf{W}_j^H \mathbf{H}_{b_k,j} + N_0 \mathbf{I}_{N_b} \right)^{-1} \mathbf{H}_{b_k,k}^H \mathbf{w}_{k,s}. \tag{3.19}$$

3.4.3 Centralized Algorithm

In the centralized coordination approach a master entity (such as a central node or a cloud radio access network controller) is in charge of performing all the BF computation, taking into account the relationship among the cells. In order to do this, the central controller must acquire knowledge of the channels between all BSs and all users in the system, i.e., a global CSI, which can be achieved by each BS sending its local CSI to a central controlling unit via a backhaul network. In addition, the filters resulting from the BF computations must be sent back to the respective transmitters. Alternatively, centralized approaches are also commonly defined with each BS performing its computations. For this, every BS must have access to global CSI via backhaul sharing.

To solve Problem (3.5), we propose the following iterative algorithm, denoted as Algorithm 1, which must be computed in a centralized way. This requirement is imposed by the interference terms in the sets $\boldsymbol{\chi}, \boldsymbol{\psi}, \boldsymbol{\Omega},$ and $\boldsymbol{\theta}$, which are variables of the transmit optimization Problem (3.13) and are each shared by two nodes (the interferer and the interfered), coupling their solutions. Thus, the optimal values of these variables are dependent on the decision of multiple nodes. Consequently, the Problem (3.13) cannot be solved independently by each node. Therefore, in a dynamic TDD network, coupled nodes are the BSs or UEs that cause interference or disturbance to one another, i.e., the selection of a BF filter in one node affects the selection of the (receive or transmit) filter at the other node.

Algorithm 1 can be conveniently initialized by setting the transmit vectors using some simple and suitable precoding scheme such as maximum ratio transmission (MRT). After the initialization, the algorithm performs the calculation of the receive and transmit filters iteratively until a predefined convergence criterion is met.

The alternate optimization of receive and transmit vectors is well known and has been used in a variety of scenarios. However, to the best of our knowledge this is the first attempt to use it in a dynamic TDD scenario to optimize both UL and DL transceivers. This is not a straightforward task, since the dynamic TDD scenario adds a large set of variables and a complex cross-interference structure to the problem. It is also important to state that the SDP optimization step (step 4 in algorithm 1) performed to compute the transmit BF vectors adds to the total

complexity of the solution, and it is more complex to solve such class of problems than to solve closed form equations, for example. However, such convex problems can be efficiently solved by standard convex optimization tools (for example CVX (40)), and a worst case complexity analysis which is also valid for our case is presented by (71).

Algorithm 1 Centralized dynamic TDD BF

1: Initialize the precoders $\{\mathbf{M}_i\}_{i\in\mathcal{K}_{dl}}$ and $\{\mathbf{W}_i\}_{i\in\mathcal{K}_{ul}}$.
2: **repeat**
3: Compute receivers $\{\mathbf{W}_i\}_{i\in\mathcal{K}_{dl}}$ and $\{\mathbf{M}_i\}_{i\in\mathcal{K}_{ul}}$ (3.16), (3.18).
4: Compute precoders $\{\mathbf{M}_i\}_{i\in\mathcal{K}_{dl}}$ and $\{\mathbf{W}_i\}_{i\in\mathcal{K}_{ul}}$ (3.13).
5: **until** Some convergence criterion

The convergence and optimality of Algorithm 1 as the solution for the Problem in (3.5) is discussed in Section 3.6.

3.5 Decentralized Approach via ADMM

In realistic scenarios, the requirement of full CSI knowledge at a central node, or in every node, is impractical, since the transmission of such channel information requires huge amounts of signaling between nodes. On the other hand, decentralized approaches are better suited for practical scenarios, since they require only local CSI knowledge.

In a coordinated decentralized approach to solve the proposed BF problem, Problem (3.13) must be divided into subproblems, each solved by one transmitter and each receiver must be calculated locally. In order to do this, one requirement is that each node must have access to channel state information about the links between itself and the other nodes, i.e. CSI at the transmitter and at the receiver. This can be acquired using a precoded pilot signaling scheme.

In addition to that, the nodes still need to share between themselves some additional parameters via backhaul in order to coordinate the optimization procedure. However, it is necessary to keep the amount of backhaul signaling at a minimum. The signaling strategy to achieve this goal is detailed in Section 3.7.

In order to reformulate (3.13) into a distributable form, let us first introduce a new set of variables in Table 5, which account for the node-specific inter-cell interference links. The variables in these sets are called local variables and will be used to decouple the global interference terms $(\boldsymbol{\chi}, \boldsymbol{\psi}, \boldsymbol{\Omega}, \boldsymbol{\theta})$.

For each of this node specific interference terms there are linear mapping matrices $\mathbf{P}_i^{\chi}$, $\mathbf{P}_i^{\Omega}$, $\mathbf{P}_i^{\theta}$ and $\mathbf{P}_i^{\psi} \in \{0,1\}$, which relate them to the global interference variables, such that $\boldsymbol{\chi}_i = \mathbf{P}_i^{\chi}\boldsymbol{\chi}$, $\boldsymbol{\Omega}_i = \mathbf{P}_i^{\Omega}\boldsymbol{\Omega}$, $\boldsymbol{\theta}_i = \mathbf{P}_i^{\theta}\boldsymbol{\theta}$ and $\boldsymbol{\psi}_i = \mathbf{P}_i^{\psi}\boldsymbol{\psi}$ $\forall i \in \mathcal{B}_{dl} \cup \mathcal{K}_{ul}$. This way, we can rewrite the problem of (3.13) in a form which each SINR constraint is expressed using independent convex

Table 5 – Node specific inter-cell interference sets.

Node	Inter-cell interference set
$b \in \mathcal{B}_{dl}$	$\boldsymbol{\chi}_b = \{\chi_{b,i,s}, \chi_{b',i',s}\}_{i\in\mathcal{K}_{dl}\setminus\mathcal{K}_b,\, b'\in\mathcal{B}_{dl}\setminus b, i'\in\mathcal{K}_b, \forall s}$
	$\boldsymbol{\psi}_b = \{\psi_{b,i,s}\}_{i\in\mathcal{K}_{ul}, \forall s}$
	$\boldsymbol{\theta}_b = \{\theta_{l,i,s}\}_{l\in\mathcal{K}_{ul},\, i\in\mathcal{K}_b, \forall s}$
$l \in \mathcal{K}_{ul}$	$\boldsymbol{\Omega}_l = \{\Omega_{l,i,s}, \Omega_{l',l,s}\}_{i\in\mathcal{K}_{ul}\setminus l, l'\in\mathcal{K}_{ul}\setminus l, \forall s}$
	$\boldsymbol{\psi}_l = \{\psi_{b,l,s}\}_{b\in\mathcal{B}_{dl}, \forall s}$
	$\boldsymbol{\theta}_l = \{\theta_{l,i,s}\}_{i\in\mathcal{K}_{dl}, \forall s}$

Source: Created by the author.

sets as (3.20).

$$
\begin{aligned}
\min_{\substack{\boldsymbol{\chi}_b,\boldsymbol{\Omega}_l,\boldsymbol{\theta}_i\boldsymbol{\psi}_i,\rho_i \\ \boldsymbol{\chi},\boldsymbol{\Omega},\boldsymbol{\theta},\boldsymbol{\psi},p_i \\ \{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_{dl}},\{\hat{\mathbf{W}}_i\}_{i\in\mathcal{K}_{ul}}}} & \sum_{i\in\mathcal{B}_{dl}\cup\mathcal{K}_{ul}} p_i \\
\text{s. t. } & \left(\{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_b}, \boldsymbol{\chi}_b, \boldsymbol{\psi}_b, \boldsymbol{\theta}_b, p_b\right) \in \mathcal{D}_b,\ \forall b \in \mathcal{B}_{dl}, \\
& \left(\hat{\mathbf{W}}_l, \boldsymbol{\Omega}_l, \boldsymbol{\psi}_l, \boldsymbol{\theta}_l, p_l\right) \in \mathcal{U}_l,\ \forall l \in \mathcal{K}_{ul}, \\
& \boldsymbol{\chi}_b = \mathbf{P}_b^{\chi}\boldsymbol{\chi},\ \forall b \in \mathcal{B}_{dl}, \\
& \boldsymbol{\Omega}_l = \mathbf{P}_l^{\Omega}\boldsymbol{\Omega},\ \forall l \in \mathcal{K}_{ul}, \\
& \boldsymbol{\theta}_i = \mathbf{P}_i^{\theta}\boldsymbol{\theta},\ i \in \mathcal{B}_{dl}\cup\mathcal{K}_{ul}, \\
& \boldsymbol{\psi}_i = \mathbf{P}_i^{\psi}\boldsymbol{\psi},\ i \in \mathcal{B}_{dl}\cup\mathcal{K}_{ul}, \\
& p_i = \rho_i,\ i \in \mathcal{B}_{dl}\cup\mathcal{K}_{ul},
\end{aligned}
\tag{3.20}
$$

where $\rho_i \geq 0, \forall\, i \in \mathcal{B}_{dl}\cup\mathcal{K}_{ul}$ are slack variables introduced in order to impose the power penalty parameter term.

The solution of optimization problem described in (3.20) requires centralized execution, since it still presents the global interference variables that couple different nodes in the network. However, it is written in a way that these global variables are organized to fit into the recently presented node-specific interference variables. Such formulation allow us to use a notorious iterative mathematical method for distributed optimization, the ADMM, as it is presented in the next section.

3.5.1 ADMM Decentralized Solution

A promising approach to solve, in a decentralized way, optimization problems like the one in (3.20) is ADMM (72). The ADMM solution consists of three iterative steps: 1) update of local primal variables $\boldsymbol{\chi}_b, \boldsymbol{\Omega}_l, \boldsymbol{\theta}_i, \boldsymbol{\psi}_i, p_i$ 2) update of global primal variables $\boldsymbol{\chi}, \boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{\rho}$ and 3) update of dual variables, $\boldsymbol{\lambda}_b^{\chi}, \boldsymbol{\lambda}_l^{\Omega}, \boldsymbol{\lambda}_i^{\theta}, \boldsymbol{\lambda}_i^{\psi}, \lambda_i^{\rho}$, related to each equality constraint in Problem (3.20).

In an ADMM decentralized solution, each node must compute its own local variables independently, and then share its values to the other nodes, which compute their versions of the global variables. After that, the dual variables are calculated to be used in the next iteration of the ADMM algorithm. The solutions iterates towards an optimum value of sum-power and interference terms.

3.5.1.1 Local Variables Update

In the first ADMM step for our problem, we seek the update of local primal variables $\boldsymbol{\chi}_b, \boldsymbol{\Omega}_l, \boldsymbol{\theta}_i, \boldsymbol{\psi}_i, p_i$ .The local variables update at iteration $(t+1)$ is given by the argument of the minimization of the augmented Lagrangian of (3.20) and is given for each DL BS b as

$$\begin{aligned}
&\{\boldsymbol{\chi}_b(t+1), \boldsymbol{\theta}_b(t+1), \boldsymbol{\psi}_b(t+1), p_b(t+1)\} = \\
&\underset{\substack{\boldsymbol{\chi}_b, \boldsymbol{\theta}_b \\ \boldsymbol{\psi}_b, p_b \\ \{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_b}}}{\arg\min} \left\{ \begin{array}{l} \left(p_b + \frac{c}{2}\|\mathbf{P}_b^{\chi}\boldsymbol{\chi}(t) - \boldsymbol{\chi}_b\|^2 + \frac{c}{2}\|\mathbf{P}_b^{\theta}\boldsymbol{\theta}(t) - \boldsymbol{\theta}_b\|^2 + \frac{c}{2}\|\mathbf{P}_b^{\psi}\boldsymbol{\psi}(t) - \boldsymbol{\psi}_b\|^2 + \right. \\ \left. \frac{c}{2}(\rho_b(t) - p_b)^2 - \boldsymbol{\lambda}_b^{\chi^T}(t)\boldsymbol{\chi}_b - \boldsymbol{\lambda}_b^{\theta^T}(t)\boldsymbol{\theta}_b - \boldsymbol{\lambda}_b^{\psi^T}(t)\boldsymbol{\psi}_b - \lambda_b^{\rho}(t)p_b \right) \end{array} \right\} \\
&\quad \text{s. t.} \quad \left(\{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_b}, \boldsymbol{\chi}_b, \boldsymbol{\psi}_b, \boldsymbol{\theta}_b, p_b \right) \in \mathcal{D}_b,
\end{aligned} \tag{3.21}$$

and for each UL UE l as

$$\begin{aligned}
&\{\boldsymbol{\Omega}_l(t+1), \boldsymbol{\theta}_l(t+1), \boldsymbol{\psi}_l(t+1), p_l(t+1)\} = \\
&\underset{\substack{\boldsymbol{\Omega}_l, \boldsymbol{\theta}_l \\ \boldsymbol{\psi}_l, p_l \\ \hat{\mathbf{W}}_l}}{\arg\min} \left\{ \begin{array}{l} \left(p_l + \frac{c}{2}\|\mathbf{P}_l^{\chi}\boldsymbol{\Omega}(t) - \boldsymbol{\Omega}_l\|^2 + \frac{c}{2}\|\mathbf{P}_l^{\theta}\boldsymbol{\theta}(t) - \boldsymbol{\theta}_l\|^2 + \frac{c}{2}\|\mathbf{P}_l^{\psi}\boldsymbol{\psi}(t) - \boldsymbol{\psi}_l\|^2 + \right. \\ \left. \frac{c}{2}(\rho_l(t) - p_l)^2 - \boldsymbol{\lambda}_l^{\Omega^T}(t)\boldsymbol{\Omega}_l - \boldsymbol{\lambda}_l^{\theta^T}(t)\boldsymbol{\theta}_l - \boldsymbol{\lambda}_l^{\psi^T}(t)\boldsymbol{\psi}_l - \lambda_l^{\rho}(t)p_l \right) \end{array} \right\} \\
&\quad \text{s. t.} \quad \left(\hat{\mathbf{W}}_l, \boldsymbol{\Omega}_l, \boldsymbol{\psi}_l, \boldsymbol{\theta}_l, p_l \right) \in \mathcal{U}_l,
\end{aligned} \tag{3.22}$$

where $c \geq 0$ is the penalty parameter.

3.5.1.2 Global Variables Update

In the second ADMM step for our problem, we perform the update of the global primal variables $\boldsymbol{\chi}, \boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\psi}$, and $\boldsymbol{\rho}$. The global variables update is also given by the argument of the minimization of the augmented Lagrangian of (3.20), which has closed form solution given

as

$$\boldsymbol{\chi}(t+1) = \mathbf{P}_{\chi}{}^{\dagger}\left(\tilde{\boldsymbol{\chi}}(t+1) - \frac{1}{c}\tilde{\boldsymbol{\lambda}}_{\chi}(t)\right), \tag{3.23}$$

$$\boldsymbol{\Omega}(t+1) = \mathbf{P}_{\Omega}{}^{\dagger}\left(\tilde{\Omega}(t+1) - \frac{1}{c}\tilde{\boldsymbol{\lambda}}_{\Omega}(t)\right), \tag{3.24}$$

$$\boldsymbol{\theta}(t+1) = \mathbf{P}_{\theta}{}^{\dagger}\left(\tilde{\boldsymbol{\theta}}(t+1) - \frac{1}{c}\tilde{\boldsymbol{\lambda}}_{\theta}(t)\right), \tag{3.25}$$

$$\boldsymbol{\psi}(t+1) = \mathbf{P}_{\psi}{}^{\dagger}\left(\tilde{\boldsymbol{\psi}}(t+1) - \frac{1}{c}\tilde{\boldsymbol{\lambda}}_{\psi}(t)\right), \tag{3.26}$$

$$\rho_i(t+1) = p_i(t+1) - \frac{1}{c}\lambda_i^{\rho}(t), \quad \forall i \in \mathcal{B}_{dl} \cup \mathcal{K}_{ul}, \tag{3.27}$$

where $\mathbf{P}_{\chi} = [\mathbf{P}_1^{\chi T}, \ldots, \mathbf{P}_B^{\chi T}]^T$, $\tilde{\boldsymbol{\chi}} = [\boldsymbol{\chi}_1{}^T, \ldots, \boldsymbol{\chi}_B{}^T]^T$, $\tilde{\boldsymbol{\lambda}}_{\chi} = [\boldsymbol{\lambda}_1^{\chi T}, \ldots, \boldsymbol{\lambda}_B^{\chi T}]^T$ and the other variables are defined similarly.

This step can still be further simplified. Note that a specific node does not need to compute every element in the global sets. Each node only requires elements of the global sets which have information about interference links related to itself. The reason for that is because the other two ADMM steps (described in Subsections 3.5.1.1 and 3.5.1.3) multiply the global variables by the node specific linear mapping matrices, what yields a node-specific version of the global variable, since the terms not related to the node are multiplied by 0.

This way, each of the elements in the global sets can be computed as

$$\chi_{b,i,s}(t+1) = \frac{1}{2}\left(\chi^{b}_{b,i,s}(t+1) - \frac{1}{c}\lambda_b^{\chi_{b,i,s}}(t)\right) + \frac{1}{2}\left(\chi^{b_i}_{b,i,s}(t+1) - \frac{1}{c}\lambda_{b_i}^{\chi_{b,i,s}}(t)\right), \tag{3.28}$$

$$\Omega_{l,i,s}(t+1) = \frac{1}{2}\left(\Omega^{l}_{l,i,s}(t+1) - \frac{1}{c}\lambda_l^{\Omega_{l,i,s}}(t)\right) + \frac{1}{2}\left(\Omega^{i}_{l,i,s}(t+1) - \frac{1}{c}\lambda_i^{\Omega_{l,i,s}}(t)\right), \tag{3.29}$$

$$\theta_{l,i,s}(t+1) = \frac{1}{2}\left(\theta^{l}_{l,i,s}(t+1) - \frac{1}{c}\lambda_l^{\theta_{l,i,s}}(t)\right) + \frac{1}{2}\left(\theta^{i}_{l,i,s}(t+1) - \frac{1}{c}\lambda_i^{\theta_{l,i,s}}(t)\right), \tag{3.30}$$

$$\psi_{b,i,s}(t+1) = \frac{1}{2}\left(\psi^{b}_{b,i,s}(t+1) - \frac{1}{c}\lambda_b^{\psi_{b,i,s}}(t)\right) + \frac{1}{2}\left(\psi^{i}_{b,i,s}(t+1) - \frac{1}{c}\lambda_i^{\psi_{b,i,s}}(t)\right), \tag{3.31}$$

where $\chi^{b}_{b,i,s}$ denotes the element $\chi_{b,i,s}$ in the local variables of BS b, and $\lambda_b^{\chi_{b,i,s}}$ stands for dual element of BS b corresponding to $\chi_{b,i,s}$. The other variables are defined similarly.

By using this approach, in order to calculate the global variables that are necessary for it, each node only needs to have access to the result of the local expression $\frac{1}{2}(\text{primal} - \frac{1}{c}\text{dual})$ from the coupled nodes. For example, BS b needs to have access to its own local expression and to the result of the expression $(\chi^{b_i}_{b,i,s}(t+1) - \frac{1}{c}\lambda_{b_i}^{\chi_{b,i,s}}(t))$ computed locally at BS b_i, in order to compute $\chi_{b,i,s}$.

3.5.1.3 Dual Variables Update

In the third, and last, ADMM step for our problem, we perform the update of the dual variables $\boldsymbol{\lambda}_b^{\chi}, \boldsymbol{\lambda}_l^{\Omega}, \boldsymbol{\lambda}_i^{\theta}, \boldsymbol{\lambda}_i^{\psi}, \lambda_i^{\rho}$, which can be done by the subgradient method (109) expressed

as

$$\boldsymbol{\lambda}_b^{\chi}(t+1) = \boldsymbol{\lambda}_b^{\chi}(t) + c(\boldsymbol{P}_b^{\chi}\boldsymbol{\chi}(t+1) - \boldsymbol{\chi}_b(t+1)), \tag{3.32}$$

$$\boldsymbol{\lambda}_l^{\Omega}(t+1) = \boldsymbol{\lambda}_l^{\Omega}(t) + c(\boldsymbol{P}_l^{\Omega}\boldsymbol{\Omega}(t+1) - \boldsymbol{\Omega}_l(t+1)), \tag{3.33}$$

$$\boldsymbol{\lambda}_i^{\theta}(t+1) = \boldsymbol{\lambda}_i^{\theta}(t) + c(\boldsymbol{P}_i^{\theta}\boldsymbol{\theta}(t+1) - \boldsymbol{\theta}_i(t+1)), \tag{3.34}$$

$$\boldsymbol{\lambda}_i^{\psi}(t+1) = \boldsymbol{\lambda}_i^{\psi}(t) + c(\boldsymbol{P}_i^{\psi}\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}_i(t+1)), \tag{3.35}$$

$$\lambda_i^{\rho}(t+1) = \lambda_i^{\rho}(t) + c(\rho_i(t+1) - p_i(t+1)). \tag{3.36}$$

3.5.2 Decentralized Algorithm

A decentralized version of Algorithm 1 can be written based on the proposed ADMM solution as Algorithm 2.

Algorithm 2 Decentralized dynamic TDD BF

1: Initialize $\boldsymbol{\chi}$, $\boldsymbol{\Omega}$, $\boldsymbol{\theta}$, $\boldsymbol{\psi}$ and $\boldsymbol{\rho}$.
2: BS $b, \forall b \in \mathcal{B}_{dl}$: Initialize DL precoders $\{\mathbf{M}_i\}_{i\in\mathcal{K}_b}$;
   UE $l, \forall l \in \mathcal{K}_{ul}$: Initialize UL precoders $\mathbf{W}_l$.
3: **repeat**
4: BS $b, \forall b \in \mathcal{B}_{dl}$: Use $\{\mathbf{M}_i\}_{i\in\mathcal{K}_b}$ to transmit pilots;
   UE $l, \forall l \in \mathcal{K}_{ul}$: Use $\mathbf{W}_l$ to transmit pilots.
5: BS $b, \forall b \in \mathcal{B}_{ul}$: Compute receivers $\{\mathbf{M}_i\}_{i\in\mathcal{K}_b}$ (3.18);
   UE $k, \forall k \in \mathcal{K}_{dl}$: Compute receivers $\mathbf{W}_k$ (3.16).
6: BS $b, \forall b \in \mathcal{B}_{ul}$: send precoded pilots;
   UE $k, \forall k \in \mathcal{K}_{dl}$: send precoded pilots.
7: **repeat**
8: BS $b, \forall b \in \mathcal{B}_{dl}$: Compute $\boldsymbol{\chi}_b, \boldsymbol{\theta}_b, \boldsymbol{\psi}_b$ and p_b (3.21);
   UE $l, \forall l \in \mathcal{K}_{ul}$: Compute $\boldsymbol{\Omega}_l, \boldsymbol{\theta}_l, \boldsymbol{\psi}_l$ and p_l (3.22).
9: BS $b, \forall b \in \mathcal{B}_{dl}$: Share the results of $\frac{1}{2}(\text{primal} - \frac{1}{c}\text{dual})$ to coupled nodes;
   UE $l, \forall l \in \mathcal{K}_{ul}$: Share the results of $\frac{1}{2}(\text{primal} - \frac{1}{c}\text{dual})$ to coupled nodes.
10: BS $b, \forall b \in \mathcal{B}_{dl}$: Update its global variables terms in $\boldsymbol{\chi}$, $\boldsymbol{\theta}$, $\boldsymbol{\psi}$, and $\boldsymbol{\rho}$ (3.27) to (3.31);
   UE $l, \forall l \in \mathcal{K}_{ul}$: Update its global variables terms in $\boldsymbol{\Omega}$, $\boldsymbol{\theta}$, $\boldsymbol{\psi}$, and $\boldsymbol{\rho}$ (3.27) to (3.31).
11: BS $b, \forall b \in \mathcal{B}_{dl}$: Update the dual variables $\boldsymbol{\lambda}_b^{\chi}, \boldsymbol{\lambda}_b^{\theta}, \boldsymbol{\lambda}_b^{\psi}$ and λ_b^{ρ} (3.32) to (3.36);
   UE $l, \forall l \in \mathcal{K}_{ul}$: Update the dual variables $\boldsymbol{\lambda}_l^{\Omega}, \boldsymbol{\lambda}_l^{\theta}, \boldsymbol{\lambda}_l^{\psi}$ and λ_l^{ρ} (3.32) to (3.36).
12: **until** Some stop criterion
13: BS $b, \forall b \in \mathcal{B}_{dl}$: Solve (3.21) for $\{\hat{\mathbf{M}}_i\}_{i\in\mathcal{K}_b}$;
   UE $l, \forall l \in \mathcal{K}_{ul}$: Solve (3.22) for $\{\hat{\mathbf{W}}_l\}$.
14: **until** Some stop criterion.

Note that Algorithm 2 performs an alternate optimization of transmitters and re-

ceivers as in the centralized approach. This can be noted in steps 5 and 13, where receivers and transmitters are computed, respectively. It can also be noted that the ADMM approach is used in order to compute the inter-cell interference terms ($\boldsymbol{\chi}, \boldsymbol{\Omega}, \boldsymbol{\theta}$ and $\boldsymbol{\psi}$), which are used in the decentralized transmitter optimization. Steps 8, 10 and 11 relate to the local, global and dual variables update in the ADMM approach. However, in order for Algorithm 2 to be fully functional, it needs to guarantee that the information required by each node to perform the necessary optimization is available. For that, steps 6 and 9 describe the required signaling procedure, with over-the-air precoded pilot transmission and backhaul communication. These steps are further detailed and discussed in Section 3.7.

In our later discussions, we refer to the loop for the alternate optimization, defined by steps 3 and 14, as outer loop, and to the loop for the ADMM solution, defined by steps 7 and 12, as inner loop.

3.6 Convergence Analysis

In this section we present the convergence analysis for both algorithms 1 and 2. In Subsection 3.6.1 we present a proof of convergence for the alternate optimization of transmitters and receivers, which defines Algorithm 1, but also serves as basis for Algorithm 2. In Subsection 3.6.2 the convergence of the ADMM decentralized approach in Algorithm 2 is discussed.

3.6.1 Convergence Analysis for the Alternate Optimization

In this subsection we show that the proposed alternate optimization of transmitters and receivers, as presented in Algorithm 1, converges globally to the set of Karush-Kuhn-Tucker (KKT) conditions for the proposed optimization Problem (3.5). This proof is based on the convergence analysis for a MIMO multi-user DL scenario presented by (108).

For the sake of simplicity, the convergence analysis is performed for the case in which each user transmits one stream, however, it also applies to the multi-stream case. This way the transmit and receive matrices $\mathbf{M}_i$ and $\mathbf{W}_i$ are written as the vectors $\mathbf{m}_i$ and $\mathbf{m}_i$, respectively, and our main optimization Problem (3.5) becomes

$$\begin{aligned} \min_{\{\mathbf{m}_i, \mathbf{m}_i\}_{\forall i}} \; & \sum_{i \in \mathcal{K}_{dl}} \|\mathbf{m}_i\|^2 + \sum_{j \in \mathcal{K}_{ul}} \|\mathbf{m}_j\|^2 \\ \text{s. t.} \quad & \Gamma_k^{dl} \geq \gamma_k, \quad \forall k \in \mathcal{K}_{dl}, \\ & \Gamma_l^{ul} \geq \gamma_l, \quad \forall l \in \mathcal{K}_{ul}. \end{aligned} \tag{3.37}$$

Let us first define the Lagrange function associated with Problem (3.37) as

$$\begin{aligned}\mathcal{L}(\mathbf{m},\mathbf{w},\boldsymbol{\lambda}) \triangleq & \sum_{i\in\mathcal{K}_{dl}} \|\mathbf{m}_i\|_2^2 + \sum_{j\in\mathcal{K}_{ul}} \|\mathbf{w}_j\|_2^2 \\ & + \sum_{k\in\mathcal{K}_{dl}} \lambda_k \left(\sum_{i\in\mathcal{K}_{dl}\setminus k} |\mathbf{w}_k^H \mathbf{H}_{b_i,k}\mathbf{m}_i|^2 + \sum_{j\in\mathcal{K}_{ul}} |\mathbf{w}_k^H \mathbf{Q}_{j,k}\mathbf{w}_j|^2 + ||\mathbf{w}_k||^2 N_0 - \frac{1}{\gamma_k}|\mathbf{w}_k^H \mathbf{H}_{b_k,k}\mathbf{m}_k|^2 \right) \\ & + \sum_{l\in\mathcal{K}_{ul}} \lambda_l \left(\sum_{j\in\mathcal{K}_{ul}\setminus l} |\mathbf{m}_l^H \mathbf{H}_{b_l,j}^H \mathbf{w}_j|^2 + \sum_{i\in\mathcal{K}_{dl}} |\mathbf{m}_l^H \mathbf{G}_{b_i,b_l}\mathbf{m}_i|^2 + ||\mathbf{m}_l||^2 N_0 - \frac{1}{\gamma_l}|\mathbf{m}_l^H \mathbf{H}_{b_l,l}^H \mathbf{w}_l|^2 \right), \end{aligned} \tag{3.38}$$

where $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_K]$ is the Lagrange multiplier vector. Thus, the KKT conditions for Problem (3.5) are given by

$$\lambda_k \left(\sum_{i\in\mathcal{K}_{dl}\setminus k} \mathbf{H}_{b_i,k}\mathbf{m}_i\mathbf{m}_i^H \mathbf{H}_{b_i,k}^H + \sum_{j\in\mathcal{K}_{ul}} \mathbf{Q}_{j,k}\mathbf{w}_j\mathbf{w}_j^H \mathbf{Q}_{j,k}^H + \mathbf{I}N_0 - \frac{1}{\gamma_k}\mathbf{H}_{b_k,k}\mathbf{m}_k\mathbf{m}_k^H \mathbf{H}_{b_k,k}^H \right)\mathbf{w}_k = 0,\ \forall k \in \mathcal{K}_{dl}, \tag{3.39}$$

$$\lambda_l \left(\sum_{j\in\mathcal{K}_{ul}\setminus l} \mathbf{H}_{b_l,j}^H \mathbf{w}_j\mathbf{w}_j^H \mathbf{H}_{b_l,j} + \sum_{i\in\mathcal{K}_{dl}} \mathbf{G}_{b_i,b_l}\mathbf{m}_i\mathbf{m}_i^H \mathbf{G}_{b_i,b_l}^H + \mathbf{I}N_0 - \frac{1}{\gamma_l}\mathbf{H}_{b_l,l}^H \mathbf{w}_l\mathbf{w}_l^H \mathbf{H}_{b_l,l} \right)\mathbf{m}_l = 0,\ \forall l \in \mathcal{K}_{ul}, \tag{3.40}$$

$$\left(\mathbf{I} - \frac{\lambda_k}{\gamma_k}\mathbf{H}_{b_k,k}^H \mathbf{w}_k\mathbf{w}_k^H \mathbf{H}_{b_k,k} + \sum_{i\in\mathcal{K}_{dl}\setminus k} \lambda_i \mathbf{H}_{b_k,i}^H \mathbf{w}_i\mathbf{w}_i^H \mathbf{H}_{b_k,i} + \sum_{l\in\mathcal{K}_{ul}} \lambda_l \mathbf{G}_{b_k,b_l}^H \mathbf{m}_l\mathbf{m}_l^H \mathbf{G}_{b_k,b_l} \right)\mathbf{m}_k = 0,\ \forall k \in \mathcal{K}_{dl}, \tag{3.41}$$

$$\left(\mathbf{I} - \frac{\lambda_l}{\gamma_l}\mathbf{H}_{b_l,l}\mathbf{m}_l\mathbf{m}_l^H \mathbf{H}_{b_l,l}^H + \sum_{j\in\mathcal{K}_{ul}\setminus l} \lambda_j \mathbf{H}_{b_j,l}\mathbf{m}_j\mathbf{m}_j^H \mathbf{H}_{b_j,l}^H + \sum_{k\in\mathcal{K}_{dl}} \mathbf{Q}_{l,k}^H \mathbf{w}_k\mathbf{w}_k^H \mathbf{Q}_{l,k} \right)\mathbf{w}_l = 0,\ \forall l \in \mathcal{K}_{ul}, \tag{3.42}$$

$$\lambda_k \left(\sum_{i\in\mathcal{K}_{dl}\setminus k} |\mathbf{w}_k^H \mathbf{H}_{b_i,k}\mathbf{m}_i|^2 + \sum_{j\in\mathcal{K}_{ul}} |\mathbf{w}_k^H \mathbf{Q}_{j,k}\mathbf{w}_j|^2 + ||\mathbf{w}_k||^2 N_0 - \frac{1}{\gamma_k}|\mathbf{w}_k^H \mathbf{H}_{b_k,k}\mathbf{m}_k|^2 \right) = 0,\ \forall k \in \mathcal{K}_{dl}, \tag{3.43}$$

$$\lambda_l \left(\sum_{j\in\mathcal{K}_{ul}\setminus l} |\mathbf{m}_l^H \mathbf{H}_{b_l,j}^H \mathbf{w}_j|^2 + \sum_{i\in\mathcal{K}_{dl}} |\mathbf{m}_l^H \mathbf{G}_{b_i,b_l}\mathbf{m}_i|^2 + ||\mathbf{m}_l||^2 N_0 - \frac{1}{\gamma_l}|\mathbf{m}_l^H \mathbf{H}_{b_l,l}^H \mathbf{w}_l|^2 \right) = 0,\ \forall l \in \mathcal{K}_{ul}, \tag{3.44}$$

$$\gamma_k \left(\sum_{i\in\mathcal{K}_{dl}\setminus k} |\mathbf{w}_k^H \mathbf{H}_{b_i,k}\mathbf{m}_i|^2 + \sum_{j\in\mathcal{K}_{ul}} |\mathbf{w}_k^H \mathbf{Q}_{j,k}\mathbf{w}_j|^2 + ||\mathbf{w}_k||^2 N_0 \right) \leq |\mathbf{w}_k^H \mathbf{H}_{b_k,k}\mathbf{m}_k|^2,\ \forall k \in \mathcal{K}_{dl}, \tag{3.45}$$

$$\gamma_l\left(\sum_{j\in\mathcal{K}_{ul}\setminus l}|\mathbf{m}_l^H\mathbf{H}_{b_l,j}^H\mathbf{w}_j|^2+\sum_{i\in\mathcal{K}_{dl}}|\mathbf{m}_l^H\mathbf{G}_{b_i,b_l}\mathbf{m}_i|^2+||\mathbf{m}_l||^2N_0\right)\leq|\mathbf{m}_l^H\mathbf{H}_{b_l,l}^H\mathbf{w}_l|^2,\ \forall l\in\mathcal{K}_{ul},\tag{3.46}$$

$$\lambda_k\geq 0,\ \forall k,\tag{3.47}$$

where equations (3.39) to (3.42) represent the first-order optimality conditions with respect to DL and UL receive and DL and UL transmitters, respectively. Equations (3.43) and (3.44) are the complementary conditions for DL and UL. Equations (3.45) and (3.46) are the primal feasibility conditions for DL and UL and (3.47) is the dual feasibility condition.

Proposition 1. *Let* $\{(\mathbf{m}^r,\mathbf{w}^r,\boldsymbol{\lambda}^r)\}_{r=1}^{\infty}$ *denote a sequence generated by Algorithm 1. Suppose* $(\mathbf{m}^0,\mathbf{w}^0,\boldsymbol{\lambda}^0)$ *is a feasible solution for Problem* (3.37)*, then every limit point of* $\{(\mathbf{m}^r,\mathbf{w}^r,\boldsymbol{\lambda}^r)\}_{r=1}^{\infty}$ *is a KKT point of Problem* (3.37)*.*

Proof. The state of the transmitters and receivers at the steps of Algorithm 1 is illustrated as $\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r-1},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{r-1}\}\rightarrow\{\{\mathbf{w}_i\}_{i\in\mathcal{K}_{dl}}^{r},\{\mathbf{m}_i\}_{i\in\mathcal{K}_{ul}}^{r}\}\rightarrow\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{r}\}$, where the two arrows correspond to the update rules of Algorithm 1. It can be noted that the objective function is bounded from below, thus the transmitters $\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{r}\}$ are bounded and consequently the receivers $\{\{\mathbf{w}_i\}_{i\in\mathcal{K}_{dl}}^{r},\{\mathbf{m}_i\}_{i\in\mathcal{K}_{ul}}^{r}\}$ are bounded as well. Thus, the sequence $\{(\mathbf{m}^r,\mathbf{w}^r)\}$ has at least one limit point. Consider a subsequence $\{(\mathbf{m}_j^r,\mathbf{w}_j^r)\}_{j=1}^{\infty}$ converging to the limit point $\{\mathbf{m}^*,\mathbf{w}^*\}$. Additionally, consider another subsequence, we assume that $\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r-1},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{r-1}\}$ converges to a limit point $\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{**},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{**}\}$.

First, we prove the equality of the two possible limit points of the transmitters $\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{*},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{*}\}=\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{**},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{**}\}$ (up to a phase rotation). Since the objective is decreasing and it is bounded from below, we can state that

$$\sum_{i\in\mathcal{K}_{dl}}\|\mathbf{m}_i\|_2^2+\sum_{j\in\mathcal{K}_{ul}}\|\mathbf{w}_j^*\|_2^2=\sum_{i\in\mathcal{K}_{dl}}\|\mathbf{m}_i^{**}\|_2^2+\sum_{j\in\mathcal{K}_{ul}}\|\mathbf{w}_j^{**}\|_2^2.\tag{3.48}$$

Now consider fixed transmit beamformers $\{\{\dot{\mathbf{m}}_i\}_{i\in\mathcal{K}_{dl}},\{\dot{\mathbf{w}}_i\}_{i\in\mathcal{K}_{ul}}\}$ so that

$$\Gamma_k^{dl}(\{\dot{\mathbf{m}}_i\}_{i\in\mathcal{K}_{dl}},\{\dot{\mathbf{w}}_i\}_{i\in\mathcal{K}_{ul}},\mathbf{w}_k^*)\geq\gamma_k,\quad\forall k\in\mathcal{K}_{dl},\tag{3.49}$$

$$\Gamma_l^{ul}(\{\dot{\mathbf{m}}_i\}_{i\in\mathcal{K}_{dl}},\{\dot{\mathbf{w}}_i\}_{i\in\mathcal{K}_{ul}},\mathbf{m}_l^*)\geq\gamma_l,\quad\forall l\in\mathcal{K}_{ul}.\tag{3.50}$$

Due to the continuity of the SINR function, there exists an index q so that for all $j>q$ we can state that $\Gamma_k^{dl}(\{\dot{\mathbf{m}}_i\}_{i\in\mathcal{K}_{dl}},\{\dot{\mathbf{w}}_i\}_{i\in\mathcal{K}_{ul}},\mathbf{w}_k^{r_j})\geq\gamma_k,\forall k\in\mathcal{K}_{dl}$ and $\Gamma_l^{ul}(\{\dot{\mathbf{m}}_i\}_{i\in\mathcal{K}_{dl}},\{\dot{\mathbf{w}}_i\}_{i\in\mathcal{K}_{ul}},\mathbf{m}_l^{r_j})\geq\gamma_l,\forall l\in\mathcal{K}_{ul}$. Since transmit beamformers are updated after receive beamformers, for all $j>q$ we have

$$\sum_{i\in\mathcal{K}_{dl}}\|\mathbf{m}_i^{r_j}\|_2^2+\sum_{j\in\mathcal{K}_{ul}}\|\mathbf{w}_j^{r_j}\|_2^2\leq\sum_{i\in\mathcal{K}_{dl}}\|\dot{\mathbf{m}}_i\|_2^2+\sum_{j\in\mathcal{K}_{ul}}\|\dot{\mathbf{w}}_j\|_2^2.\tag{3.51}$$

Let $j\rightarrow\infty$, it implies

$$\sum_{i\in\mathcal{K}_{dl}}\|\mathbf{m}_i^{*}\|_2^2+\sum_{j\in\mathcal{K}_{ul}}\|\mathbf{w}_j^{*}\|_2^2\leq\sum_{i\in\mathcal{K}_{dl}}\|\dot{\mathbf{m}}_i\|_2^2+\sum_{j\in\mathcal{K}_{ul}}\|\dot{\mathbf{w}}_j\|_2^2.\tag{3.52}$$

Furthermore, according to the update rule we have

$$\Gamma_k^{dl}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r_j},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{r_j},\mathbf{w}_k^{r_j})\geq\gamma_k,\quad\forall k\in\mathcal{K}_{dl},\tag{3.53}$$

$$\Gamma_l^{ul}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r_j},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{r_j},\mathbf{m}_l^{r_j})\geq\gamma_l,\quad\forall l\in\mathcal{K}_{ul}.\tag{3.54}$$

and thus

$$\Gamma_k^{dl}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{*},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{*},\mathbf{w}_k^{*})\geq\gamma_k,\quad\forall k\in\mathcal{K}_{dl},\tag{3.55}$$

$$\Gamma_l^{ul}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{*},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{*},\mathbf{m}_l^{*})\geq\gamma_l,\quad\forall l\in\mathcal{K}_{ul}.\tag{3.56}$$

Combining (3.55) and (3.56) with the fact that (3.52) holds for any fixed transmitters satisfying (3.49) and (3.50), we obtain

$$\begin{aligned}\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{*},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{*}\}=\underset{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}}{\arg\min}&\sum_{i\in\mathcal{K}_{dl}}\|\mathbf{m}_i\|_2^2+\sum_{j\in\mathcal{K}_{ul}}\|\mathbf{w}_j\|_2^2\\ \text{s. t.}\quad&\Gamma_k^{dl}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}},\mathbf{w}_k^{*})\geq\gamma_k,\quad\forall k\in\mathcal{K}_{dl},\\ &\Gamma_l^{ul}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}},\mathbf{m}_l^{*})\geq\gamma_l,\quad\forall l\in\mathcal{K}_{ul}.\end{aligned}\tag{3.57}$$

On the other hand, since the update of receive beamformers using MMSE receiver maintains the SINR feasibility, we have $\Gamma_k^{dl}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r_j-1},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{r_j-1},\mathbf{w}_k^{r_j})\geq\gamma_k,\forall k\in\mathcal{K}_{dl}$ and $\Gamma_l^{ul}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{r_j-1},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{r_j-1},\mathbf{m}_l^{r_j})\geq\gamma_l,\forall l\in\mathcal{K}_{ul}$. Letting $j\to\infty$ we obtain

$$\Gamma_k^{dl}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{**},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{**},\mathbf{w}_k^{*})\geq\gamma_k,\quad\forall k\in\mathcal{K}_{dl},\tag{3.58}$$

$$\Gamma_l^{ul}(\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{**},\{\mathbf{w_i}\}_{i\in\mathcal{K}_{ul}}^{**},\mathbf{m}_l^{*})\geq\gamma_l,\quad\forall l\in\mathcal{K}_{ul}.\tag{3.59}$$

Combining (3.57), (3.58), (3.59) and (3.48), we infer that $\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{**},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{**}\}$ is also a solution to Problem (3.57). Thus, since Problem (3.57) has a strictly convex objective function, it has a unique solution and we have $\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{*},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{*}\}=\{\{\mathbf{m}_i\}_{i\in\mathcal{K}_{dl}}^{**},\{\mathbf{w}_i\}_{i\in\mathcal{K}_{ul}}^{**}\}$ up to a phase rotation.

Next, we prove that the limit point $\{\mathbf{m}^*,\mathbf{w}^*\}$ is a KKT point for Problem (3.37). Based on the receivers update rule the unnormalized MMSE receivers are written as

$$\mathbf{w}_k^{r_j}=(\mathbf{C}_k^{r_j-1})^{-1}\mathbf{H}_{b_k,k}\mathbf{m}_k^{r-1},\forall k\in\mathcal{K}_{dl},\tag{3.60}$$

$$\mathbf{m}_k^{r_j}=(\mathbf{C}_k^{r_j-1})^{-1}\mathbf{H}_{b_k,k}^H\mathbf{w}_k^{r_j-1},\forall k\in\mathcal{K}_{ul},\tag{3.61}$$

with

$$\mathbf{C}_k^{r_j-1}=\sum_{i\in\mathcal{K}_{dl}}\mathbf{H}_{b_i,k}\mathbf{m}_i^{r_j-1}(\mathbf{m}_i^{r_j-1})^H\mathbf{H}_{b_i,k}^H+N_0\mathbf{I}_{N_u}+\sum_{j\in\mathcal{K}_{ul}}\mathbf{Q}_{j,k}\mathbf{w}_j^{r_j-1}(\mathbf{w}_j^{r_j-1})^H\mathbf{Q}_{j,k}^H,\forall k\in\mathcal{K}_{dl},\tag{3.62}$$

$$\mathbf{C}_k^{r_j-1}=\sum_{i\in\mathcal{K}_{dl}}\mathbf{G}_{b_i,b_k}\mathbf{m}_i^{r_j-1}(\mathbf{m}_i^{r_j-1})^H\mathbf{G}_{b_i,b_k}^H+N_0\mathbf{I}_{N_b}+\sum_{j\in\mathcal{K}_{ul}}\mathbf{H}_{b_k,j}^H\mathbf{w}_j^{r_j-1}(\mathbf{w}_j^{r_j-1})^H\mathbf{H}_{b_k,j},\forall k\in\mathcal{K}_{ul}.\tag{3.63}$$

Letting $j \to \infty$ implies

$$\mathbf{w}_k^* = (\mathbf{C}_k^*)^{-1}\mathbf{H}_{b_k,k}\mathbf{m}_k^*, \forall k \in \mathcal{K}_{dl}, \tag{3.64}$$

$$\mathbf{m}_k^* = (\mathbf{C}_k^*)^{-1}\mathbf{H}_{b_k,k}^H\mathbf{w}_k^*, \forall k \in \mathcal{K}_{ul}, \tag{3.65}$$

with

$$\mathbf{C}_k^* = \sum_{i\in\mathcal{K}_{dl}} \mathbf{H}_{b_i,k}\mathbf{m}_i^*(\mathbf{m}_i^*)^H\mathbf{H}_{b_i,k}^H + N_0\mathbf{I}_{N_u} + \sum_{j\in\mathcal{K}_{ul}} \mathbf{Q}_{j,k}\mathbf{w}_j^*(\mathbf{w}_j^*)^H\mathbf{Q}_{j,k}^H, \forall k \in \mathcal{K}_{dl}, \tag{3.66}$$

$$\mathbf{C}_k^* = \sum_{i\in\mathcal{K}_{dl}} \mathbf{G}_{b_i,b_k}\mathbf{m}_i^*(\mathbf{m}_i^*)^H\mathbf{G}_{b_i,b_k}^H + N_0\mathbf{I}_{N_b} + \sum_{j\in\mathcal{K}_{ul}} \mathbf{H}_{b_k,j}^H\mathbf{w}_j^*(\mathbf{w}_j^*)^H\mathbf{H}_{b_k,j}, \forall k \in \mathcal{K}_{ul}. \tag{3.67}$$

On the other hand, (3.57) implies that there exists a set of multipliers $\lambda_k^* \geq 0$ so that

$$\left(\mathbf{I} - \frac{\lambda_k^*}{\gamma_k}\mathbf{H}_{b_k,k}^H\mathbf{w}_k^*(\mathbf{w}_k^*)^H\mathbf{H}_{b_k,k} + \sum_{i\in\mathcal{K}_{dl}\setminus k} \lambda_i\mathbf{H}_{b_k,i}^H\mathbf{w}_i^*(\mathbf{w}_i^*)^H\mathbf{H}_{b_k,i} + \sum_{l\in\mathcal{K}_{ul}} \lambda_l\mathbf{G}_{b_k,b_l}^H\mathbf{m}_l^*(\mathbf{m}_l^*)^H\mathbf{G}_{b_k,b_l}\right)\mathbf{m}_k^* = 0, \forall k \in \mathcal{K}_{dl}, \tag{3.68}$$

$$\left(\mathbf{I} - \frac{\lambda_l^*}{\gamma_l}\mathbf{H}_{b_l,l}\mathbf{m}_l^*(\mathbf{m}_l^*)^H\mathbf{H}_{b_l,l}^H + \sum_{j\in\mathcal{K}_{ul}\setminus l} \lambda_j\mathbf{H}_{b_j,l}\mathbf{m}_j^*(\mathbf{m}_j^*)^H\mathbf{H}_{b_l,j}^H + \sum_{k\in\mathcal{K}_{dl}} \mathbf{Q}_{l,k}^H\mathbf{w}_k^*(\mathbf{w}_k^*)^H\mathbf{Q}_{l,k}\right)\mathbf{w}_l^* = 0, \forall l \in \mathcal{K}_{ul}, \tag{3.69}$$

$$\gamma_k\left(\sum_{i\in\mathcal{K}_{dl}\setminus k} |\mathbf{w}^{*H}{}_k\mathbf{H}_{b_i,k}\mathbf{m}_i^*|^2 + \sum_{j\in\mathcal{K}_{ul}} |(\mathbf{w}_k^*)^H\mathbf{Q}_{j,k}\mathbf{w}_j^*|^2 + ||\mathbf{w}_k^*||^2N_0\right) - |(\mathbf{w}_k^*)^H\mathbf{H}_{b_k,k}\mathbf{m}_k^*|^2 = 0, \ \forall k \in \mathcal{K}_{dl}, \tag{3.70}$$

$$\gamma_l\left(\sum_{j\in\mathcal{K}_{ul}\setminus l} |(\mathbf{m}_l^*)^H\mathbf{H}_{b_l,j}^H\mathbf{w}_j^*|^2 + \sum_{i\in\mathcal{K}_{dl}} |(\mathbf{m}_l^*)^H\mathbf{G}_{b_i,b_l}\mathbf{m}_i^*|^2 + ||\mathbf{m}_l^*||^2N_0\right) - |(\mathbf{m}_l^*)^H\mathbf{H}_{b_l,l}^H\mathbf{w}_l^*|^2 = 0, \ \forall l \in \mathcal{K}_{ul}, \tag{3.71}$$

Using (3.66), (3.67), we can rearrange (3.70) and (3.71) as

$$(\mathbf{w}_k^*)^H\left(\mathbf{C}_k^* - (1+\frac{1}{\gamma_k})\mathbf{H}_{b_k,k}\mathbf{m}_k^*(\mathbf{m}_k^*)^H\mathbf{H}_{b_k,k}^H\right)\mathbf{w}_k^* = 0, \forall k \in \mathcal{K}_{dl}, \tag{3.72}$$

$$(\mathbf{m}_l^*)^H\left(\mathbf{C}_l^* - (1+\frac{1}{\gamma_k})\mathbf{H}_{b_l,l}^H\mathbf{w}_l^*(\mathbf{w}_l^*)^H\mathbf{H}_{b_l,l}\right)\mathbf{m}_l^* = 0, \forall l \in \mathcal{K}_{ul}. \tag{3.73}$$

By substituting the MMSE receivers (3.64) and (3.65) in (3.72) and (3.73) we obtain

$$(\mathbf{m}_k^*)^H\mathbf{H}_{b_k,k}^H(\mathbf{C}_k^*)^{-1}\mathbf{H}_{b_k,k}\mathbf{m}_k^*\left(1 - (1+\frac{1}{\gamma_k})(\mathbf{m}_k^*)^H\mathbf{H}_{b_k,k}^H(\mathbf{C}_k^*)^{-1}\mathbf{H}_{b_k,k}\mathbf{m}_k^*\right) = 0, \ \forall k \in \mathcal{K}_{dl}, \tag{3.74}$$

$$(\mathbf{w}_l^*)^H \mathbf{H}_{b_l,l} \mathbf{C}_l^* \mathbf{H}_{b_l,l}^H \mathbf{w}_l^* \left(1 - (1+\frac{1}{\gamma_k})(\mathbf{w}_l^*)^H \mathbf{H}_{b_l,l} (\mathbf{C}_l^*)^{-1} \mathbf{H}_{b_l,l}^H \mathbf{w}_l^*\right) = 0,\ \forall l \in \mathcal{K}_{ul}. \tag{3.75}$$

Since $(\mathbf{m}_k^*)^H \mathbf{H}_{b_k,k}^H (\mathbf{C}_k^*)^{-1} \mathbf{H}_{b_k,k} \mathbf{m}_k^* \neq 0, \forall l \in \mathcal{K}_{dl}$ and $(\mathbf{w}_l^*)^H \mathbf{H}_{b_l,l} \mathbf{C}_l^* \mathbf{H}_{b_l,l}^H \mathbf{w}_l^* \neq 0, \forall l \in \mathcal{K}_{ul}$, it follows that

$$(1+\frac{1}{\gamma_k})(\mathbf{m}_k^*)^H \mathbf{H}_{b_k,k}^H (\mathbf{C}_k^*)^{-1} \mathbf{H}_{b_k,k} \mathbf{m}_k^* = 1,\ \forall k \in \mathcal{K}_{dl}, \tag{3.76}$$

$$(1+\frac{1}{\gamma_k})(\mathbf{w}_l^*)^H \mathbf{H}_{b_l,l} (\mathbf{C}_l^*)^{-1} \mathbf{H}_{b_l,l}^H \mathbf{w}_l^* = 1,\ \forall l \in \mathcal{K}_{ul}. \tag{3.77}$$

This implies that for each $k \in \mathcal{K}_{dl}$, the matrix $\mathbf{C}_k^* - (1+\frac{1}{\gamma_k})\mathbf{H}_{b_k,k}\mathbf{m}_k^*(\mathbf{m}_k^*)^H \mathbf{H}_{b_k,k}^H$ is positive semidefinite, and for each $l \in \mathcal{K}_{dl}$ the matrix $\mathbf{C}_l^* - (1+\frac{1}{\gamma_k})\mathbf{H}_{b_l,l}^H \mathbf{w}_l^*(\mathbf{w}_l^*)^H \mathbf{H}_{b_l,l}$ is positive semidefinite. Hence from (3.72) and (3.73) we get

$$\left(\mathbf{C}_k^* - (1+\frac{1}{\gamma_k})\mathbf{H}_{b_k,k}\mathbf{m}_k^*(\mathbf{m}_k^*)^H \mathbf{H}_{b_k,k}^H\right)\mathbf{w}_k^* = 0, \forall k \in \mathcal{K}_{dl}, \tag{3.78}$$

$$\left(\mathbf{C}_l^* - (1+\frac{1}{\gamma_k})\mathbf{H}_{b_l,l}^H \mathbf{w}_l^*(\mathbf{w}_l^*)^H \mathbf{H}_{b_l,l}\right)\mathbf{m}_l^* = 0,\ \forall l \in \mathcal{K}_{ul}. \tag{3.79}$$

Thus, from equations (3.68), (3.69), (3.70), (3.71), (3.78) and (3.79) we infer that the limit point $\{\mathbf{m}^*, \mathbf{w}^*\}$ is a KKT point of Problem (3.37). □

This way, we could prove that the alternate optimization of transmitters and receivers as presented in Algorithm 1 converges globally to the set of KKT conditions for the proposed optimization Problem (3.5).

3.6.2 Convergence Analysis for the Decentralized Algorithm

Since Algorithm 2 also implements the alternate optimization of transmitters and receivers, the proof in Subsection 3.6.1 also applies to it. In order for Algorithm 2 to fully converge to the optimal point it is required that the solution for the optimization of transmitters, performed by ADMM, also fully converges. ADMM is shown, by (72), to converge to an optimal point when applied to convex problems. That is, in order to achieve the same solution as in the centralized case, Algorithm 2 must iterate until convergence in both inner and outer loops.

Algorithm 2 requires signaling exchange at each iteration of the ADMM steps and also at each alternate transmit and receive optimization. It is, thus, impractical, to let the algorithms iterate until full convergence. However, ADMM converges to modest accuracy sufficient for many applications within few iterations (72). In Sections 3.7 and 3.8 we further discuss this characteristic, and show that for our application, good performance can be achieved with only few iterations.

An interesting point is that since the ADMM solution does not require that the local variables be equal to the global ones before convergence, an intermediate solution may

not strictly fulfill all the SINR constraints. Alternatively, the ADMM solution approximates the desired SINR constraints. However, a solution that strictly respects the minimum SINR constraints can be obtained by solving (3.21) and (3.22) including an additional constraint that forces the local variables to be equal to the global ones. In this case feasibility is not guaranteed in each iteration, and the algorithm is required to iterate more in order to reach a possible feasible solution. Additionally, the network can employ a mechanism that stops the algorithm after a fixed number of infeasible iterations and restarts with a different initialization or, if a feasible iteration is not found, it can declare the problem infeasible, and the systems requirements must be relaxed by lowering SINR targets or by decreasing the number of active users.

Another important convergence issue is that the centralized solution requires feasibility at all iterations, and consequently, if the first iteration is not feasible, the centralized algorithm fails to provide a solution. Due to this, a correct precoder initialization is crucial to enhance feasibility probability. In our simulations, we use the simple and suitable MRT scheme as a starting point for the algorithms, which leads to very few unfeasible cases ($< 0.1\%$). On the other hand, an interesting feature of the ADMM decentralized solution comes from the fact that it does not require feasibility at all iterations. This way, it is able to converge, or at least provide approximate solutions, even when the centralized algorithm fails to find a feasible point.

3.7 Signaling Scheme

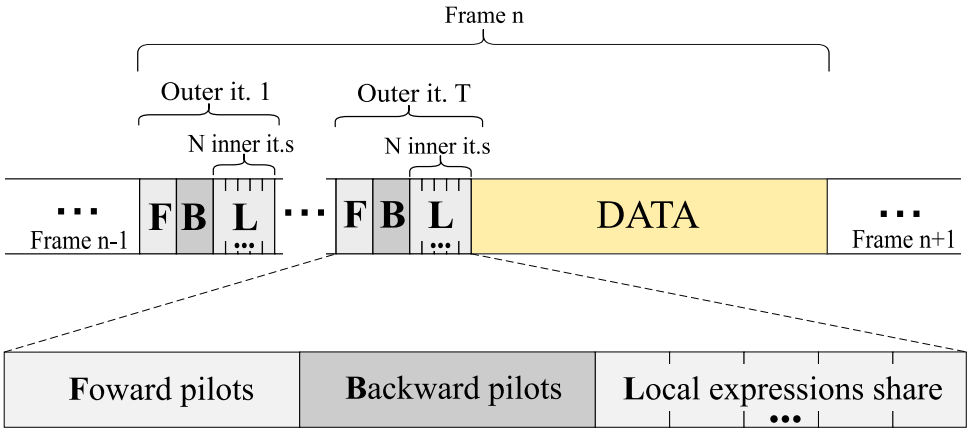
In a centralized approach, the proposed Algorithm 1 could be executed by a master entity (a central node or a cloud radio access network controller) that needs to have full CSI knowledge and is capable of sending the computed transmit and receive filters for the UEs and BSs. Another possible way of executing Algorithm 1 is by each BS performing its own computation. For that, each BS must have access to global CSI. Both ways are highly unpractical, given the required amount of information share.

For the decentralized approach we consider that each node only has access to local CSI. However, information still needs to be shared between the nodes, in order to coordinate them. Algorithm 2 relies on two types of information sharing: over-the-air precoded pilot signaling and local expressions share. Figure 8 illustrates the proposed frame structure, which is based on the bidirectional signaling structure proposed by (81).

The frame is divided in two main parts, the BF setup and the data transmission in both uplink and downlink directions. All cells must be synchronized in both parts, assuming that the system complies with the requirements of 3GPP (110). Indeed, (110) specifies the requirements for TDD systems and serve as the basis for dynamic TDD deployments, as described in (80), for example.

During the BF setup, the over-the-air signaling is denoted by the letter F, for the forward (transmitters to receivers) pilot signaling, corresponding to step 4 in Algorithm 2, and B for the backward (receivers to transmitters) pilot signaling, corresponding to step 5. This pilot

Figure 8 – Frame structure for the decentralized execution of Algorithm 2.



Source: Created by the author.

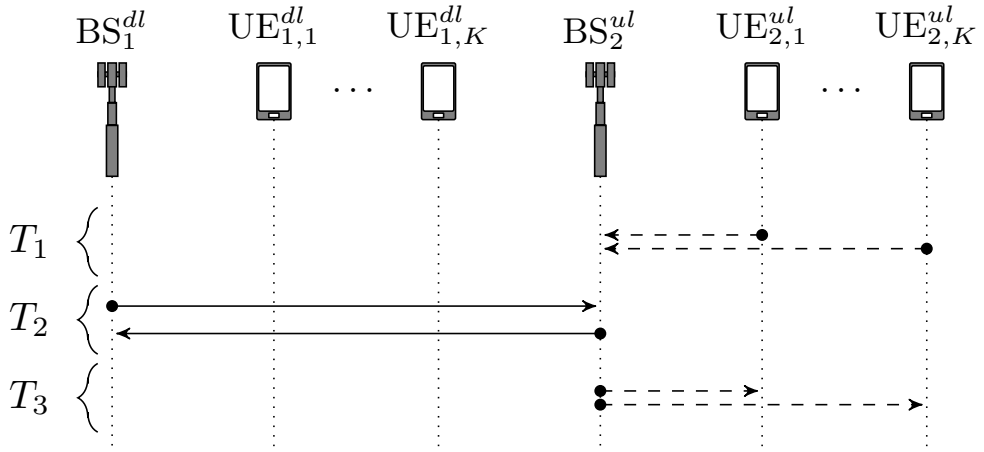
signaling exchange is required for the update of receivers (3.18) and (3.16) and for the update of local variables (3.21) and (3.22), since both require the knowledge of effective channels (that incorporate the transmit and receive filters) from opposite nodes. We consider that all nodes use known orthogonal precoded pilot symbols, allowing perfect signal separation and estimation of effective channels, and consequently perfect estimation of the precoders and receivers by the opposite nodes.

The sharing of the local expressions for the global variables update is denoted by the letter L, and occurs in step 9 of Algorithm 2. This procedure is required since, in order to compute the global variables in the ADMM solution, each node needs to have access to the result of the local expression $\frac{1}{2}(\text{primal} - \frac{1}{c}\text{dual})$ from the coupled nodes. This signaling stage is proposed to take place via both backhaul and control over-the-air channel, since information needs to be shared among all transmitting nodes (DL BSs and UL UEs). This way, all BSs must share information between themselves by using the backhaul link and each UL UE must send its data to its respective BS, and receive the updates from the other nodes also hearing from its BS. The proposed signaling scheme is shown in Figure 9 for a 2-cell scenario.

In this procedure, in time T_1 each UL UE share information about its local variables expressions (regarding $\boldsymbol{\Omega}_l, \boldsymbol{\psi}_l$ and $\boldsymbol{\theta}_l$) to its respective BS. In Time T_2 each DL BS sends the information about its local variables expressions (regarding $\boldsymbol{\chi}_b, \boldsymbol{\psi}_b$ and $\boldsymbol{\theta}_b$) to the coupled BSs, and each UL BS sends the terms from its users to each coupled BS. In time T_3 each UL BS sends the information received from other nodes to each of its UEs. By doing this, we guarantee that each node has the information required for the global variables update in the ADMM approach.

Note, in Figure 8, that in the proposed signaling scheme the steps F, B and L are repeated at each outer iteration of algorithm 2. Also note that in step L, the proposed local expression sharing, illustrated in Figure 9, is repeated at each inner iteration of algorithm 2. Due to this, the number of elements shared at each step must be analyzed and the number of iterations must be kept in a controlled amount, so that the decentralized algorithm is practically viable. These practical considerations are addressed next.

Figure 9 – Local expressions signaling strategy for the decentralized execution of Algorithm 2.



Source: Created by the author.

Figure 9 illustrates a network of 2 BSs, whereas a real deployment would support multiple BSs in a geographical area, with scheduling and UL/DL setup configured prior to starting the execution of Algorithm 2. In such situation, while the time slot structure shown in Figure 9 is still the same, practical issues arise for high capacity requirements. Therefore, in practice, BSs of a dynamic TDD network form a high speed network supported by a multi-Gbps backhaul infrastructure (see for example (111) or (112)). This type of backhaul network readily facilitates all necessary information exchange that is necessary for the smooth operation of dynamic TDD networks, including synchronization and other types of information, and allows for the information exchange required by the proposed algorithm.

3.8 Results and Practical Considerations

The main simulation scenario is formed by 4 cells (2 DL, 2 UL) with 100 m radius each, and 2 UEs per cell. Each BS has 6 antennas and each UE has 3 antennas. For the sake of simplicity, and without loss of generality, the simulations consider that each UE transmits or receives one single stream. This scenario is used since it is simple and presents all types of intra-cell and inter-cell interference links. However, performance gains can be obtained for a variety of interference-limited scenarios. Flat Rayleigh fading is considered, in which each element of the channels is an i.i.d. complex Gaussian random variable with zero mean and unit variance. The path-loss model is based on (113) and is provided in Table 6, where R is the distance between two nodes, measured in kilometers. The simulation results shown are obtained from 100 Monte-Carlo simulations. The algorithms are allowed to iterate 50 times before the results are harvested.

In the following subsections, we present and analyze the main simulation results. In Subsection 3.8.1, we analyze the basic results regarding the solution of the proposed optimization problem. In Subsection 3.8.2 we present an analysis of the results regarding the

convergence behavior of the proposed algorithms. In Subsection 3.8.3 we approach the main practical considerations of the application of the proposed algorithms, mainly regarding signaling overhead.

Table 6 – Propagation characteristics for the different link directions (R in km).

Link	**Path-loss**
BS to BS	If $R < 2/3$: $\mathrm{PL}(R) = 98.4 + 20\log_{10}(R)$. If $R \geq 2/3$: $\mathrm{PL}(R) = 101.9 + 40\log_{10}(R)$.
BS to UE UE to BS	LOS: $\mathrm{PL}(R) = 103.8 + 20.9\log_{10}(R)$. NLOS: $\mathrm{PL}(R) = 145.4 + 37.5\log_{0}(R)$. $\mathrm{Prob}_{\mathrm{LOS}}(R) = 0.5 - \min(0.5, 5\exp(-0.156/R))$ $+\min(0.5, 5\exp(-R/0.03))$.
UE to UE	If $R \leq 0.05$: $\mathrm{PL}(R) = 98.45 + 20\log_{10}(R)$. If $R > 0.05$: $\mathrm{PL}(R) = 175.78 + 40\log_{10}(R)$.

Source: Created by the author.

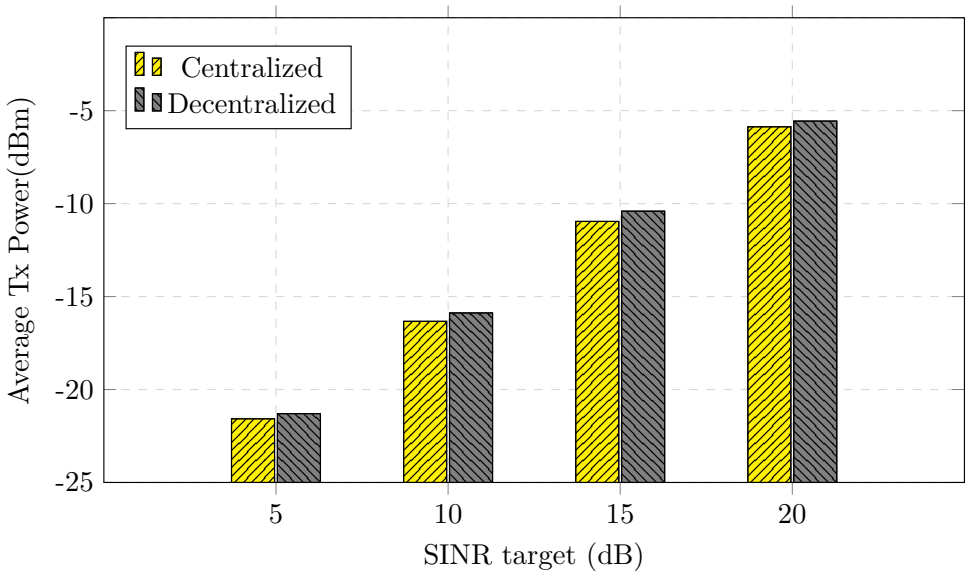
3.8.1 Optimization results

Our optimization problem seeks to minimize power while guaranteeing minimum SINR constraints. Therefore, in Figures 10 and 11 we show the main outcomes from such optimization when employing the centralized algorithm 1 and the decentralized algorithm 2 with 1 inner iteration.

In Figure 10, we show the average power achieved by the centralized and decentralized algorithms for multiple values of SINR target. It can be seen that, for all targets, the decentralized solution was able to well approximate the performance of the optimal centralized algorithm, with only a slightly higher power demand. It can also be noted, for all cases, that as the SINR targets are increased, power consumption rises, as expected.

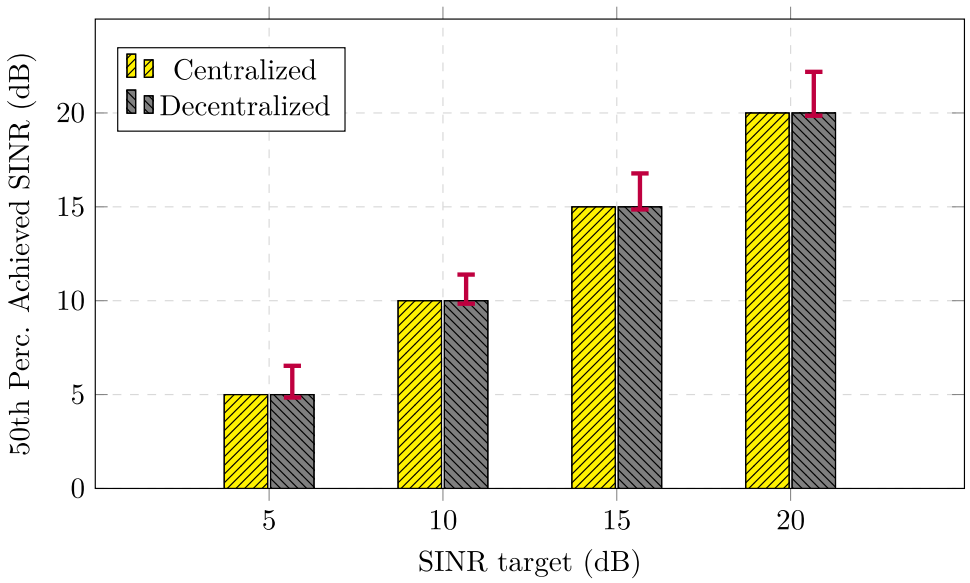
In Figure 11, we illustrate the 50th percentile of the SINR achieved by centralized and decentralized solutions for multiple values of SINR target. It can be seen that, for all targets, the centralized and decentralized solutions achieved very similar values, according to the required minimum target. The centralized algorithm forces the exact minimum SINR levels for all feasible instances. On the other hand, the decentralized algorithm approximates this value allowing some variation from the target. This is due to the fact that ADMM does not require feasibility at each iteration. Therefore, in this plot, we also show error bars for the decentralized results, in red color. These error bars illustrate the mean range of achieved SINR values deviating from the required target, defined as $\frac{1}{N}\sum_{i=0}^{N}|\text{Achieved Value - Target Value}|$. It can be seen that, for all cases, the

Figure 10 – Average power achieved by centralized and decentralized solutions for multiple values of SINR target.



Source: Created by the author.

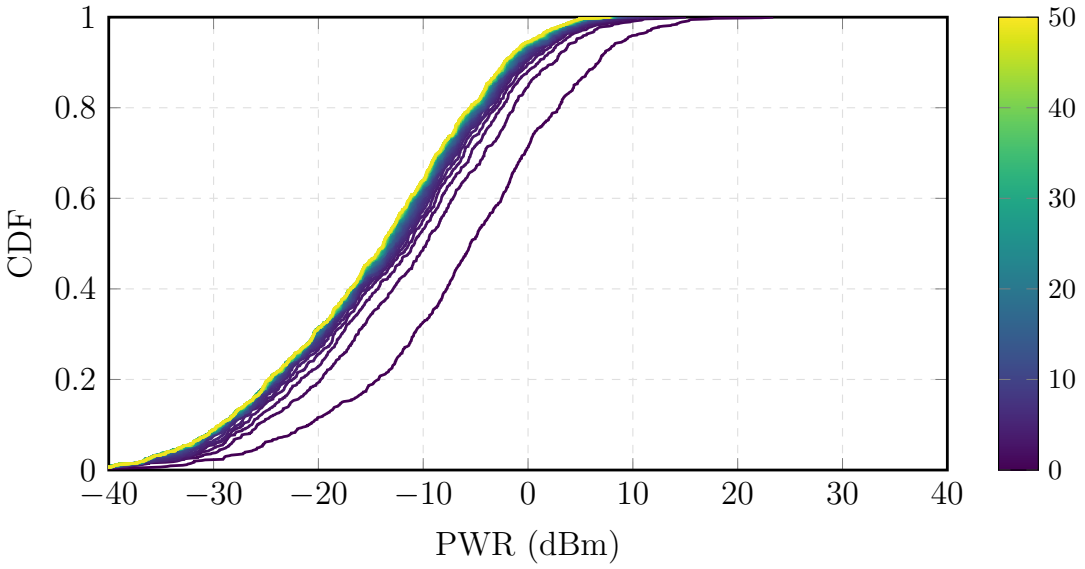
Figure 11 – 50th percentile of SINR achieved by centralized and decentralized solutions for multiple values of SINR target.



Source: Created by the author.

achieved SINRs lie in close proximity to the target, more specifically with values greater than the target, although few values that are slightly smaller than the target can occur. In the discussions of Figure 13, in the next subsection, we further discuss such behavior.

Figure 12 – CDFs of achieved power at each iteration of the centralized solution.



Source: Created by the author.

3.8.2 Convergence results

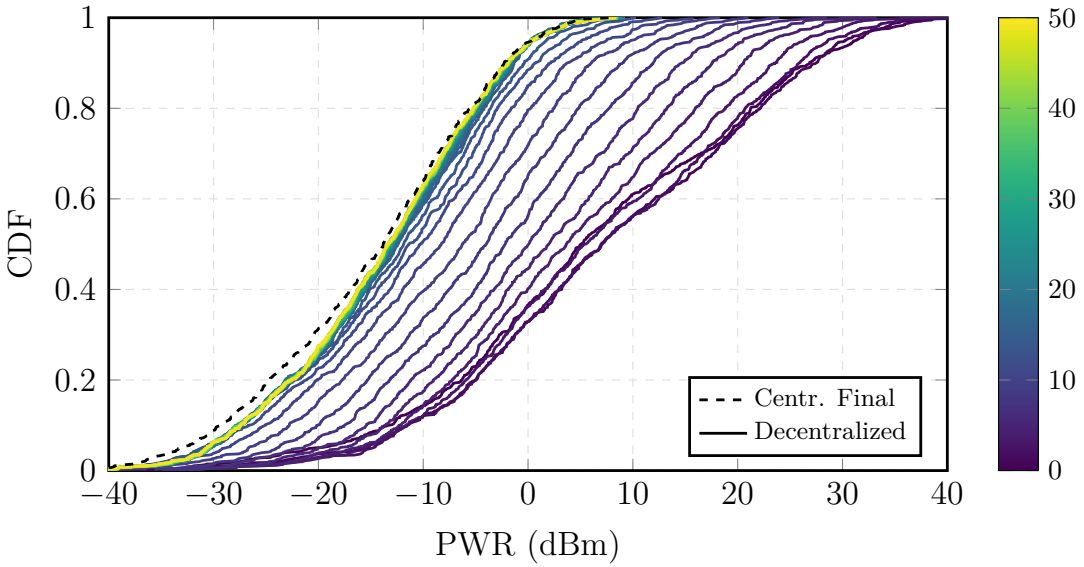
Not only showing that the solutions reach desirable optimization outcomes is important, but it also is crucial to show how the proposed algorithms achieve their results. In Figures from 12 to 15 we illustrate how the proposed algorithms converge. In the graphs shown here, we consider an SINR target of 20 dB.

In Figure 12 we show the cumulative distribution functions (CDFs) of the achieved power at each iteration of the centralized solution. The curves are shown in a color-scale, as illustrated in the provided color-bar, which indicates the iteration number. In this figure, we see that as the algorithm iterates, less power is demanded. Also, we note that in the initial iterations, the power reduction between steps is larger and the gains diminish as the number of iterations increase. Although the algorithms have run 50 iterations, from the figure we note that if much fewer iterations were taken, still very close performance would be perceived.

In Figure 13 we illustrate the CDFs of achieved power at each outer iteration of the decentralized solution. Also, we plot the final centralized outcome, to serve as a benchmark. In this figure, we see that as the algorithm iterates, less power is demanded, similar to the centralized solution. However, in this decentralized scheme, the initial iterations have a larger power consumption when compared to the initial steps of the centralized approach. On the other hand, as the decentralized algorithm iterates, it becomes very close to the optimal centralized final solution, as expected. Also, in this plot, if the total number of iterations were reduced to, for example, 15 or more, still close power-performance would be perceived.

For the decentralized algorithms it is also important to show how the SINR behaves at each iteration, and how it converges to the required target, since the ADMM solution does not guarantee that the SINR constraints are respected at all iterations. For this reason, we show in

Figure 13 – CDFs of achieved power at each outer iteration of the decentralized solution, comparing with final centralized outcome.



Source: Created by the author.

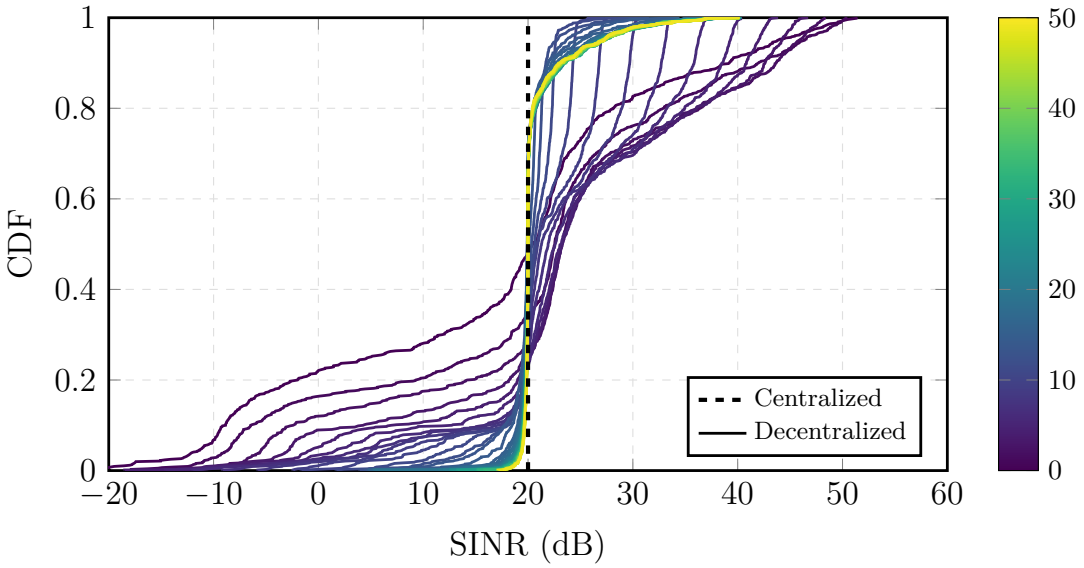
Figure 14 curves that illustrate the CDF of the SINRs at each outer iteration. Also, we plot the final centralized outcome, that forces the exact minimum SINR levels for all feasible instances, to serve as a benchmark. We see in this graph that as the algorithm iterates, the SINR variance is reduced and the results go in the direction of the desired value (20 dB). We also note that, in the final SINR curve, the great majority of results are very close to the target value (20dB) and only very few cases have smaller SINR value, the minimum case is 16dB. Values larger than the target can also be seen in the very top part of the graph. This fact relates to the error bars in Figure 11. Additionally, just like in the power convergence shown in Figure 13, in this plot, we also note that if the total number of iterations were reduced to 15 or more, still close SINR performance would be perceived.

We also point out that, even in cases in which the centralized algorithm fails to converge due to the requirement of feasibility in every iteration, the decentralized solution achieves a similar behavior as in Figures 13 and 14. Beyond that, a minimum SINR target can still be attained for the decentralized approach by forcing the local variables to be equal to the global ones.

The results shown up to this point consider the decentralized algorithm with only 1 inner iteration. However, full convergence for this algorithm is proven if the inner ADMM steps are performed until convergence. Thus it is important to analyze the behavior of the solution with more inner iterations. In Figure 15 we compare the power and SINR performance of the decentralized solutions with 1 and 10 inner iterations.

In Figure 15a we show the mean power at each outer iteration. We see that both approaches iterate in the direction of a very close minimum power point, and that from the 15th iteration onwards, both solutions obtain good approximations of the optimum value. The

Figure 14 – CDFs of achieved SINR at each outer step of the decentralized solution, comparing with centralized outcome.



Source: Created by the author.

decentralized approach with 1 inner iteration seems to be slower at the very initial iterations, but it provided a slightly smaller final power. The approach with 10 inner iterations obtains the smallest powers at the initial steps, but afterwards its convergence is slower.

In Figure 15b we plot the mean relative SINR error with respect to the target at each outer iteration. It can be seen that both solutions iterate towards close points, and the version with 1 inner iteration presents a slightly reduced final error. It can also be seen that the algorithm with more inner iterations obtains the best target approximation at the initial iterations, but its gains diminish with further iterations.

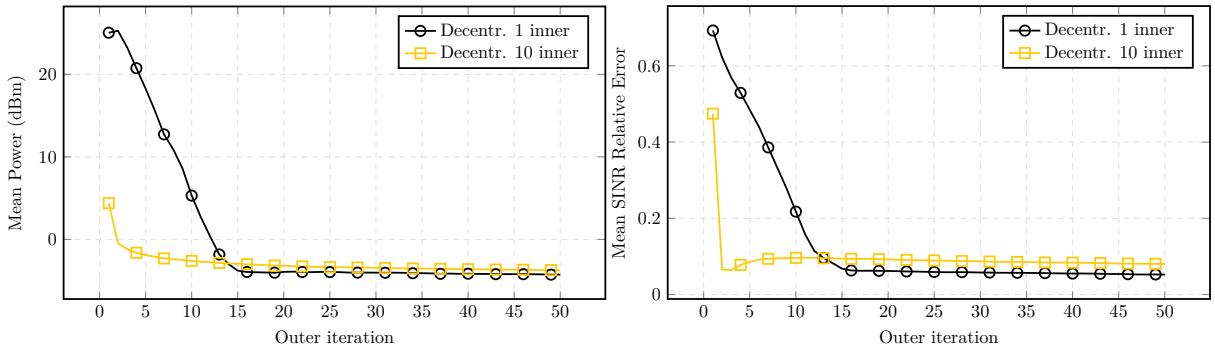
Such behaviors corroborate with the statement that says that ADMM can be very slow to converge to high accuracy, but it converges to modest accuracy within a few iterations (72). Therefore, the decentralized solution with 1 inner step is shown to be efficient and sufficient to obtain good-enough performance.

3.8.3 Practical considerations

From the convergence results we can see that the proposed decentralized algorithms can achieve good approximations of the centralized results with a reduced amount of iterations. However, it is still necessary to consider whether this amount of iterations represents a reduction in the quantity of information shared between nodes, since this is one of our main reasons to decentralize the computation. In order to perform such analysis we recollect the centralized and decentralized signaling schemes discussed in Section 3.7 and compute the number of scalars shared in each stage of both.

In a centralized approach, if we assume that each BS exchanges its local CSI with

Figure 15 – Comparison of Power and SINR convergence of decentralized solutions with 1 and 10 inner iterations.



(a) Mean power convergence

(b) SINR error convergence

Source: Created by the author.

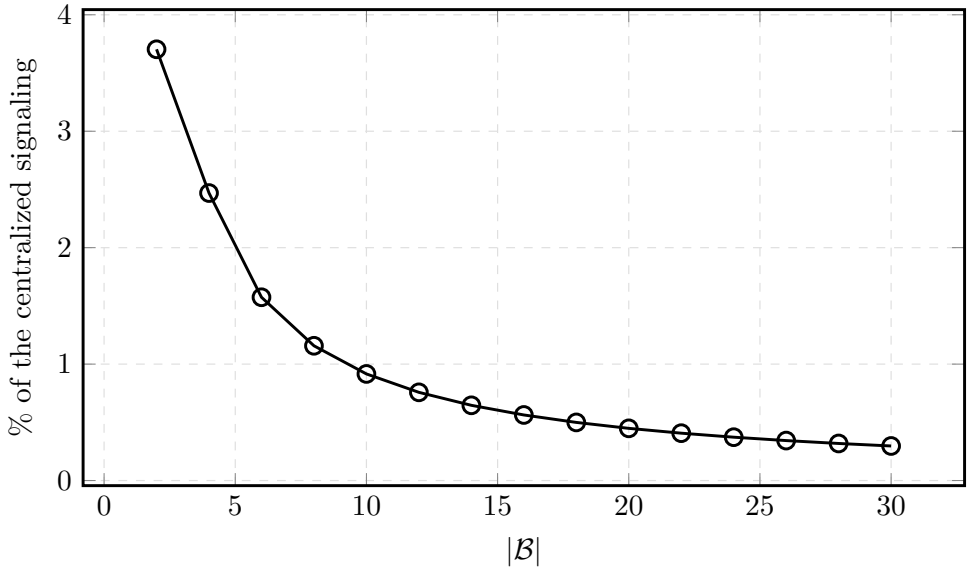
all other BSs via backhaul at setup time, an estimate of the total number of real valued scalars required to be sent is $2KN_bN_u(B-1)B+2|\mathcal{B}_{dl}|N_b^2(B-1)|\mathcal{B}_{ul}|+2K_b^2|\mathcal{B}_{ul}|N_u^2(B-1)|\mathcal{B}_{dl}|$, with each term being related to the exchange of the channels $\mathbf{H}$, $\mathbf{G}$ and $\mathbf{Q}$, respectively. Beyond that, the calculated precoders and receivers must still be sent to its respective node.

In the decentralized approach, by looking into the scheme presented in Figure 9 we can note that in T_1, a total of $2(|\mathcal{K}_{ul}|-1)+|\mathcal{B}_{dl}|+|\mathcal{K}_{dl}|$ scalars need to be sent per stream by each UE. This number is due to the sharing of the results of the local expressions regarding the update of the global variables related to the interference terms in the links UE to BS ($\boldsymbol{\Omega}_l$), BS to UE ($\boldsymbol{\psi}_l$) and UE to UE ($\boldsymbol{\theta}_l$), respectively. In T_2, the number of scalars shared per user stream by each DL BS is $2(|\mathcal{K}_{dl}|-|\mathcal{K}_b|)+|\mathcal{K}_{ul}|+|\mathcal{K}_{ul}||\mathcal{K}_b|$. Each of these terms are related to the interference terms in the links BS to UE ($\boldsymbol{\chi}_b$), BS to BS ($\boldsymbol{\psi}_b$) and UE to UE ($\boldsymbol{\theta}_b$), respectively.

By these expressions we can note that the signaling load depends on the number of BSs, UEs and antennas deployed in the scenario. Figure 16 shows a graph of the percentage of signaling that an iteration of the decentralized algorithm represents when compared to the amount of signaling required by the centralized one (scalars per iteration of decentralized/ total scalars of centralized $*100\%$) versus the number of BSs. In this graph, half of the BSs are functioning in each link direction. It can be seen in this graph that as the number of BSs increases the decentralized algorithm becomes lighter when compared to the centralized one in terms of signaling. This fact is crucial, since the larger the networks is, the more important it is to keep signaling on controlled levels. For instance, in a network with a total of 10 BSs, the centralized algorithm requires a total of 97200 scalars shared and the decentralized one requires 890 at each iteration, what represents less than 1%. This means that for this configuration, as long as the number of iterations is lower than 110, there is a control overhead reduction when using the decentralized approach.

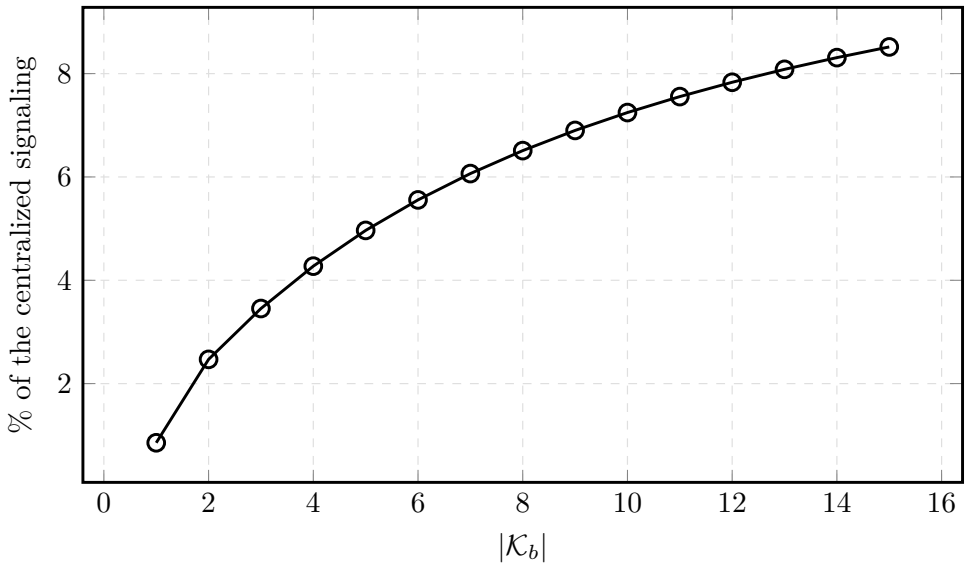
Figure 17 shows a graph of the percentage of signaling that an iteration of the decentralized algorithm represents when compared to the amount of signaling required by the

Figure 16 – Comparison of the amount of signaling required by each iteration of the decentralized algorithm, varying the number of BSs, $\{B_{ul}, B_{dl}, K_b, N_b, N_u\} = \{\frac{B}{2}, \frac{B}{2}, 2, 6, 3\}$.



Source: Created by the author.

Figure 17 – Comparison of the amount of signaling required by each iteration of the decentralized algorithm, varying number of UEs at each BS, $\{B_{ul}, B_{dl}, K_b, N_b, N_u\} = \{2, 2, K_b, 6, 3\}$.

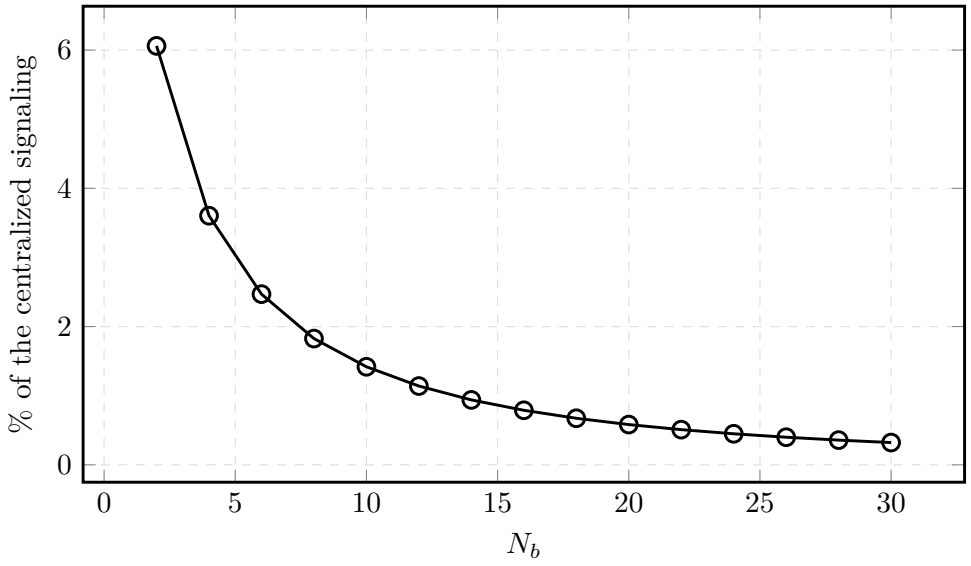


Source: Created by the author.

centralized one versus the number of UEs at each BS. It can be seen that reduced signaling levels are kept, but as the number of UEs at each BS increases, the decentralized algorithm becomes slightly less compelling.

Figure 18 shows a graph of the percentage of signaling that an iteration of the decentralized algorithm represents when compared to the amount of signaling required by the centralized one versus the number of BS antennas. As the number of antennas increases,

Figure 18 – Comparison of the amount of signaling required by each iteration of the decentralized algorithm, varying number of antennas at each BS, $\{B_{ul}, B_{dl}, K_b, N_b, N_u\} = \{2, 2, 2, N_b, 3\}$.



Source: Created by the author.

the centralized algorithm needs more signaling, since the channel matrices increase, whereas the decentralized signaling does not rely on that and remains unchanged. This means that as the number of antennas increases, the decentralized algorithm becomes more valuable when compared to the centralized one (this conclusion also applies to the UE antennas).

The total number of Forward-Backward and local variables sharing stages for the decentralized approach is determined by the number of outer and inner iterations in Algorithm 2. At each outer iteration a Forward-Backward step is performed and at each inner iteration the local variables are shared. In order to achieve the same solution as in the centralized case, Algorithm 2 must iterate until convergence in both (i.e. inner and outer) loops. However, as seen in the presented results, approximate solutions can be extracted at intermediate iterations, at the cost of suboptimal sum-power and SINR. Thus, the amount of iterations can be limited by the network operator in order to keep the control overhead within supported levels by fixing the number of inner and outer iterations, with the cost of close-to-optimum sum-power performance. Additionally, simulation results have shown that very close to optimal performance can be obtained with a small number of iterations and that the total amount of signaling is reduced compared with the centralized case.

3.9 Conclusions

In this chapter, we have proposed a BF solution for dynamic TDD Networks, which guarantees a minimum SINR for each link while minimizing the total transmitter sum-power. Two solutions were presented, a centralized version that requires full CSI, and a decentralized version based on ADMM that requires local CSI and a lightweight signaling procedure. We

have shown that both approaches converge to the same solution, and that the decentralized one can approximate the optimal solution with a reduced amount of iterations. Thus, the higher interference in dynamic TDD can be overcome, by using any of the proposed approaches, since users in DL and UL can have their SINR requirements guaranteed.

4 USER SCHEDULING BASED ON MULTI-AGENT DEEP Q-LEARNING FOR ROBUST BEAMFORMING IN MULTI-CELL MISO SYSTEMS

Maximizing the rate in multiple-input single-output (MISO) systems using distributed algorithms is an important task that typically incurs high computational cost. In this chapter, we propose two deep Q-learning-based user scheduling schemes to solve the beamforming problem of sum-rate maximization with per base station power constraints in multicell MISO scenarios.

The two key features of the proposed algorithms are that they are executed in a distributed fashion and are robust with respect to channel state information (CSI) errors. Simulation results show that in the presence of CSI errors the proposed schemes outperform state-of-the-art algorithms both in terms of average spectral efficiency and execution time.

The remainder of the chapter is structured as follows. Section 4.1 provides an introduction and discusses related works and the contributions. Section 4.2 introduces the system model and defines the main optimization problem. Section 4.3 describes the channel estimation strategy and defines the CSI error model. Section 4.4 presents the optimal solution structure for the main optimization problem, based on the uplink-downlink duality. Section 4.5 presents an overview about deep reinforcement learning and the proposed framework using deep Q-learning (DQL). Section 4.6 presents the two proposed multi-agent DQL strategies, along with the required signaling schemes. Section 4.7 provides and discusses the numerical results for the proposed algorithms before conclusions are drawn in Section 4.8.

4.1 Introduction

The sum-rate maximization problem with maximum power constraint is a classical problem in telecommunications systems, and it is known to be NP-hard (114). Solutions using convex optimization have been proposed by some works, such as those in (45, 47, 48). A solution based on the branch and bound method in multi-cell MISO systems was presented in (45). In (47), the authors proposed a local optimal solution by exploiting an iterative weighted minimum mean square error (WMMSE) approach for a multiple-input multiple-output (MIMO) system, while in (48) a solution based on fractional programming was proposed.

Solutions based on convex optimization – although interesting in terms of performance – usually come in the form of iterative mathematical methods, which often incur high computational cost. To address this issue, other studies assume closed-form solutions, which compute the beamforming without iterating resulting in low computation time, such as maximum ratio transmission, zero forcing and maximum signal-to-leakage-plus-noise ratio (SLNR-MAX) (34). However, this time reduction is obtained at the cost of performance loss. Consequently, the trade-off between execution time and performance becomes the main drawback of these solutions.

Some previous works focused on machine learning techniques in order to obtain a better trade-off between execution time and performance (97, 98). In these works, the authors proposed a supervised learning scheme based on the WMMSE algorithm outputs and an unsupervised learning strategy in order to improve the performance of the solution, while maintaining reasonably low computational complexity. Nevertheless, their approaches are centralized, which might compromise scalability and make practical deployment problematic. Furthermore, they also depend on a massive dataset of the WMMSE algorithm outputs. In (115), a distributed-execution scheme for power control using DQL was proposed in a single-antenna scenario.

User scheduling techniques play an important role when multiple communication links are competing for resources in a shared wireless medium. In this sense, some works focus on user scheduling to improve system performance. In (59) the authors proposed a distributed signal-to-leakage-plus-noise ratio based user selection (SUS) for multicell MISO systems. A user scheduling using DQL in a multi-antenna scenario was considered in (63). However, the authors assume only a centralized approach for single-cell scenarios.

The aforementioned works assume perfect CSI in their modeling. However, in practical applications there are channel estimation errors, which reduces the performance and applicability of these solutions. Imperfect CSI is considered by other works (116, 117, 118). In (116), the authors proposed a robust WMMSE algorithm based on the statistical knowledge of the links, while (117) proposed an efficient stochastic sum-rate maximization algorithm based on iterative optimization. Finally, in (118), the authors presented a low complexity beamforming scheme for MISO systems under imperfect CSI. However, these works assume the knowledge of parameters that characterize channel estimation, both at the receiver and transmitter, which depends on the channel dynamics and channel estimation schemes, and are difficult to obtain in practical implementations. In addition, to the best of our knowledge, no distributed multi-antenna scheme using DQL in multicell scenarios with imperfect CSI has been proposed in the literature.

In this chapter, we propose user scheduling solutions based on distributed-execution multi-agent DQL for solving the beamforming problem of maximizing the total rate in a multicell MISO system with per-base station (BS) power constraints assuming channel estimation errors. Differently from previous works in the literature, we focus on showing that the proposed learning scheme is able to efficiently adapt and improve the performance of a model based on perfect CSI in scenarios with the presence of channel estimation errors.

In summary, the main contributions of this work are: (1) Two new distributed-execution algorithms based on DQL (77), that aim to obtain fast and robust computations in the presence of CSI errors. The first solution requires local CSI knowledge and the exchange of information between BSs, while the second assumes only local information during the execution phase, which reduces signaling overhead at the cost of performance loss; (2) The proposed algorithms learn a policy that extends the application of an optimization model that is based on perfect CSI to a more realistic scenario with channel estimation errors; (3) The description of signaling aspects of the proposed solutions for both training and execution phases; (4) The

performance evaluation by means of simulations, where we compare the proposed solution with state-of-the-art algorithms.

In order to highlight, even more, the contributions of this work with respect to some of the main related papers, we summarize the differences in terms of the objective function, constraints, architecture, centralized or distributed solution approach, CSI availability and the merits of the proposed solution in Table 11 located in Appendix B in the end of this work.

4.2 System Model and Problem Formulation

We consider the downlink of a multicell MISO[1] system with B BSs, each with N_b antennas, serving K_b single-antenna user equipments (UEs). In addition, $\mathcal{B}$ and $\mathcal{K}_b$ represent the set of BSs and set of UEs of BS b, respectively. In this system, K is the total number of UEs within the network and $\mathcal{K}$ is its respective set. Let $\mathbf{h}_{k_b j} \in \mathbb{C}^{N_b \times 1}$ be the channel vector from BS j to UE k served by BS b, then the signal received by UE k in BS b can be written as:

$$y_{k_b} = \underbrace{\mathbf{h}_{k_b b}^{\mathrm{H}} \mathbf{m}_{k_b} d_{k_b}}_{\text{desired signal}} + \underbrace{\sum_{\substack{i=1 \\ i \neq k_b}}^{K_b} \mathbf{h}_{k_b b}^{H} \mathbf{m}_{i_b} d_{i_b}}_{\text{intracell interference}} + \underbrace{\sum_{\substack{j=1 \\ j \neq b}}^{B} \sum_{i=1}^{K_j} \mathbf{h}_{k_b j}^{H} \mathbf{m}_{i_j} d_{i_j}}_{\text{intercell interference}} + \underbrace{n_{k_b}}_{\text{noise}}, \tag{4.1}$$

where $\mathbf{m}_{k_b} \in \mathbb{C}^{N_b \times 1}$, $d_{k_b} \in \mathbb{C}$ and $n_{k_b} \in \mathbb{C}$ denote the transmit beamforming vector, data symbol and noise relative to UE k in the b-th cell, in which d_{k_b} and n_{k_b} are zero-mean complex Gaussian random variables with variances equal to one and σ^2, respectively.

In a scenario like this one, we have that the signal-to-interference-plus-noise ratio (SINR) at UE k in the cell b can be written as

$$\Gamma_{k_b} = \frac{|\mathbf{h}_{k_b b}^{H} \mathbf{m}_{k_b}|^2}{\sum\limits_{\substack{i=1 \\ i \neq k_b}}^{K_b} |\mathbf{h}_{k_b b}^{H} \mathbf{m}_{i_b}|^2 + \sum\limits_{\substack{j=1 \\ j \neq b}}^{B} \sum\limits_{i=1}^{K_j} |\mathbf{h}_{k_b j}^{H} \mathbf{m}_{i_j}|^2 + \sigma^2}. \tag{4.2}$$

Such system model can be used to represent a very broad set of scenarios. Our final goal is to find the optimum set of transmit beamformers, $\{\mathbf{m}_{k_b}\}_{\forall k \in \mathcal{K}_b, \forall b \in \mathcal{B}}$, that maximize the sum-rate while respecting a maximum power budget. Thus, we can write the sum-rate maximization subject to maximum power budget per base station as:

$$\underset{\{\mathbf{m}_{k_b}\}_{\forall k \in \mathcal{K}_b, \forall b \in \mathcal{B}}}{\text{maximize}} \sum_{b=1}^{B} \sum_{k=1}^{K_b} \log_2(1 + \Gamma_{k_b}), \tag{4.3a}$$

$$\text{s.t.} \sum_{k=1}^{K_b} ||\mathbf{m}_{k_b}||^2 \leq P_b, \quad \forall b \in \mathcal{B}, \tag{4.3b}$$

where $\|\cdot\|$ is the Euclidean norm and P_b is the maximum available power at BS b.

[1] The proposed learning-based scheme can be extended to MIMO scenarios by considering the receive filter (combiner) to be computed by known receive beamforming strategies.

4.3 Channel Model

In addition, we also consider a scenario where nodes only have access to imperfect CSI. Let us assume that $\mathbf{h}_{k_b j} = \sqrt{\beta_{k_b j}}\mathbf{g}_{k_b j}$, where $\mathbf{g}_{k_b j} \in \mathbb{C}^{N_b \times 1}$ is a vector with the small-scale fading channel coefficients, which are assumed to be quasi-static Gaussian independent and identically distributed (i.i.d.) with zero mean and unit variance. Also, let $\beta_{k_b j}$ represent the large-scale fading channel coefficient between the k-th UE at the b-th cell and BS j. Considering that the BS estimates the channel using a minimum mean square error (MMSE) approach, we have that the estimated channel vector satisfies (119):

$$\hat{\mathbf{h}}_{k_b j} = \sqrt{\beta_{k_b j}}\left(\xi \mathbf{h}_{k_b j} + \sqrt{1-\xi^2}\mathbf{e}\right), \tag{4.4}$$

where $0 \leq \xi \leq 1$ denotes the reliability of the channel estimation and $\mathbf{e} \in \mathbb{C}^{N_b \times 1}$ is an error vector with Gaussian i.i.d. entries with zero mean and unit variance. Note that if $\xi = 1$ the channel is perfectly estimated, while $\xi = 0$ means that the BS channel estimates are fully wrong. Moreover, the channel error modeling is done for both communication and interfering links.

4.4 Optimal Solution Structure Based on Uplink-Downlink Duality

In this chapter, we focus on distributed-execution multi-agent DQL in which our goal consists of showing that the proposed learning scheme is able to efficiently adapt and improve the performance of a model based on perfect CSI in scenarios with channel estimation errors. Thus, to obtain a solution robust to CSI errors, in the following we start by obtaining an optimal beamforming structure via uplink-downlink duality to problem (4.3) assuming perfect CSI.

An optimal solution for the sum rate maximization problem (4.3) is still unknown. However, as discussed in (28), there is a direct relation between this problem and the optimization task of minimizing the transmit sum-power with minimum SINR constraints, which has its optimal solution established by means of uplink-downlink duality. This relation is such that the solution approach for the sum-power minimization problem can be extended in order to derive a structure for the solution of problem (4.3).

In fact, the relation between these problems is that the vectors which solve the sum-power minimization problem, and thus achieve the required SINRs per UE using the minimum possible power, must also satisfy the power constraints of the sum-rate maximization problem. Therefore, by assuming that we know the optimal SINR $\Gamma^*_{k_b}$ for each UE k_b of each BS b, which maximizes the total data rate in (4.3), we can use the optimal solution for the sum-power minimization with SINR constraints set to $\Gamma^*_{k_b}$ for each UE k_b of each BS b. Since the beamforming vectors derived by the sum-power minimization problem are also feasible for problem (4.3) and achieve the required SINR values, they are an optimal solution.

However, in practice, we do not know the optimal SINR values, unless we solve problem (4.3). Indeed, the main difference between these problems is that the SINR constraints

are predefined in the sum-power minimization, while in sum-rate maximization the optimal SINR values need to be found along with the beamforming vectors. Nevertheless, the connection between these two problems implies that the optimal beamforming for one may hold for the other. Therefore, based on this relation, Proposition 2 can be used in order to establish a structure for the solution of problem (4.3).

Proposition 2. *Assuming perfect CSI, the beamforming vector for UE k in BS b,* $\mathbf{m}_{k_b}$*, is given by* $\mathbf{m}_{k_b} = \sqrt{p_{k_b}}\tilde{\mathbf{m}}_{k_b}$ *for all* k_b*, where* $\sqrt{p_{k_b}}$ *and* $\tilde{\mathbf{m}}_{k_b}$ *are the beamforming power and direction, respectively, with*

$$\tilde{\mathbf{m}}_{k_b} = \frac{\left(\frac{\mu_b}{P_b}\mathbf{I} + \sum_{j=1}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_jb}\mathbf{h}_{i_jb}^H\right)^{-1}\mathbf{h}_{k_bb}}{\left\|\left(\frac{\mu_b}{P_b}\mathbf{I} + \sum_{j=1}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_jb}\mathbf{h}_{i_jb}^H\right)^{-1}\mathbf{h}_{k_bb}\right\|} \tag{4.5}$$

for some nonnegative, λ_{i_j}*,* μ_b*,* $\forall b \in \mathcal{B}$ *and* $\forall i_j \in \mathcal{K}$*.*

Proof. The proof is presented in the Appendix C at the end of this work. □

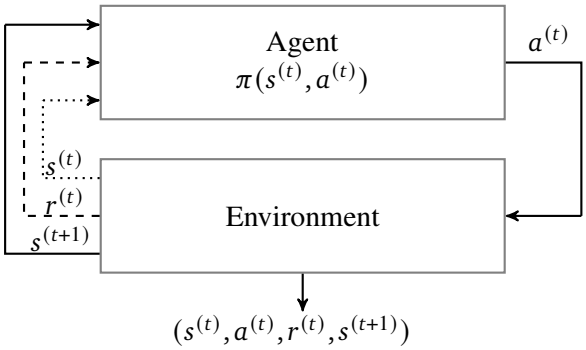
This proposition provides a structure of the beamforming as function of the Lagrange multipliers λ_{i_j} and μ_b and the power values p_{k_b} with $k, i \in \mathcal{K}$ and $b, j \in \mathcal{B}$. In addition, according to (34, Corollary 3.6), we have that $\sum_{j=1}^{B}\sum_{i=1}^{K_j}\lambda_{i_j} = 1$ and $\sum_{b=1}^{B}\mu_b = 1$. However, it is worth mentioning that this provides only a necessary condition for optimal solution, but not sufficient, except for special cases, such as single-cell transmission with a total power constraint, or any multi-cell scenario with only one power constraint (34). Also note that in order to compute (4.5) a BS needs to have access to CSI of the channels from itself to all served UEs in the own cell and to interfered UEs in neighbor cells, i.e., $\hat{\mathbf{h}}_{i_jb}, \forall i_j \in \mathcal{K}_j$ and $\forall j \in \mathcal{B}$

Next, we will show that the proposed DQL is able to adapt this structure by means of its actions, resulting in a robust beamforming scheme.

4.5 Deep Q-Learning Overview

Reinforcement learning is a machine learning technique characterized by the interaction of an agent with the surrounding environment to learn a policy that maximizes the rewards obtained by its actions on the environment. Note that the learning process is done by trial and error, where the agent gets a reward for each taken action (76). Thus, we define $\mathcal{S}$ as the set composed by all states, which is responsible for characterizing the environment of the agent, $\mathcal{A}$ is defined as the set of actions and the policy $\pi(s, a)$ is the probability of taking action a conditioned on the current state being s, with $\sum_{a\in\mathcal{A}}\pi(s, a) = 1$. The learning process occurs iteratively, where, assuming discrete time steps, the agent observes, at the time t, the current state $s^{(t)} \in \mathcal{S}$ from the environment and selects an action $a^{(t)} \in \mathcal{A}$ based on the policy $\pi(s^{(t)}, a^{(t)})$. Next, the agent gets

Figure 19 – Flowchart of the relation among agent and environment in reinforcement learning.



Source: Created by the author.

a reward $r^{(t)}$ by the taken action $a^{(t)}$ at state $s^{(t)}$ and the environment moves to the next state, $s^{(t+1)}$. Note that the reward must be able to show how beneficial the action is for the objective to be achieved. Finally, an experience is obtained in the tuple form as $(s^{(t)}, a^{(t)}, r^{(t)}, s^{(t+1)})$ and the process will be repeated until an optimal policy, π^*, is obtained. This learning process can be seen in Figure 19.

In general, reinforcement learning techniques aim to find a policy that maximizes the expected cumulative discounted rewards,

$$R^{(t)} = \mathbb{E}\left[\sum_{n=0}^{\infty} \gamma^{(n)} r^{(t+n)}\right], \tag{4.6}$$

where $\gamma \in [0, 1]$ is the discount factor (76). Moreover, in the stationary setting, the policy π can be determined by a function $Q(s, a)$, commonly called Q-function, where $Q(s, a)$ is defined as the expected accumulated discounted rewards once action a is taken under state s, assuming that the policy π is used, i.e,

$$Q(s, a) = \mathbb{E}\left[R^{(t)} | s^{(t)} = s, a^{(t)} = a\right]. \tag{4.7}$$

Thus, according to (115), the Q-function satisfies the Bellman equation:

$$Q(s, a) = \mathbb{R}(s, a) + \gamma \sum_{s' \in \mathcal{S}} \mathbb{P}^{a}_{ss'} \left(\sum_{a' \in \mathcal{A}} \pi(s', a') Q(s', a') \right), \tag{4.8}$$

where $\mathbb{R}(s, a) = \mathbb{E}[r^{(t)} | s^{(t)} = s, a^{(t)} = a]$ is the expected reward of taking action a at state s, and $\mathbb{P}^{a}_{ss'} = Pr(s^{(t+1)} = s' | s^{(t)} = s, a^{(t)} = a)$ is the transition probability from a given state s to state s' with the action a. The convergence to the action value function (4.7) has been proved for some iterative approaches (120). Moreover, note that the optimal policy consists of letting $\pi^*(s, a)$ equal to 1 for the most suitable action. Therefore, we can express the optimal Q-function associated with the optimal policy as

$$Q^*(s, a) = \mathbb{R}(s, a) + \gamma \sum_{s' \in \mathcal{S}} \mathbb{P}^{a}_{ss'} \left(\max Q(s', a') \right). \tag{4.9}$$

The well-known Q-learning algorithm can be used to obtain an optimal policy that maximizes the long-term expected accumulated discounted rewards (76). In its classical form, Q-learning constructs a lookup table, $Q(s,a)$, as a surrogate of the optimal Q-function. At the beginning, the lookup table is initialized with arbitrary values. Next, it is iteratively updated in order to find the optimal policy. Then, for each time step, an action is taken by the agent according to the ϵ-greedy policy. The ϵ-greedy policy consists in selecting a greedy action with probability $1-\epsilon$ or a random action with probability ϵ, where a greedy action on state s is given by $a = \arg\max_{\forall a \in \mathcal{A}} Q(s,a)$. The advantage of this policy is that it allows agents to explore through random action selection in order to make better action selections in the future, while avoiding to get stuck at non-optimal policies by greedy action selections (77). Moreover, note that ϵ decreases with time. After that, at the current time step t, a new experience $(s^{(t)}, a^{(t)}, r^{(t)}, s^{(t+1)})$ is obtained as result of the taken action and $Q(s^{(t)}, a^{(t)})$ is updated as follows

$$Q(s^{(t)}, a^{(t)}) \leftarrow\ Q(s^{(t)}, a^{(t)}) + \beta[r^{(t)} + \gamma \max_{\forall a \in \mathcal{A}} Q(s^{(t+1)}, a) - Q(s^{(t)}, a^{(t)})], \tag{4.10}$$

where β is the learning rate.

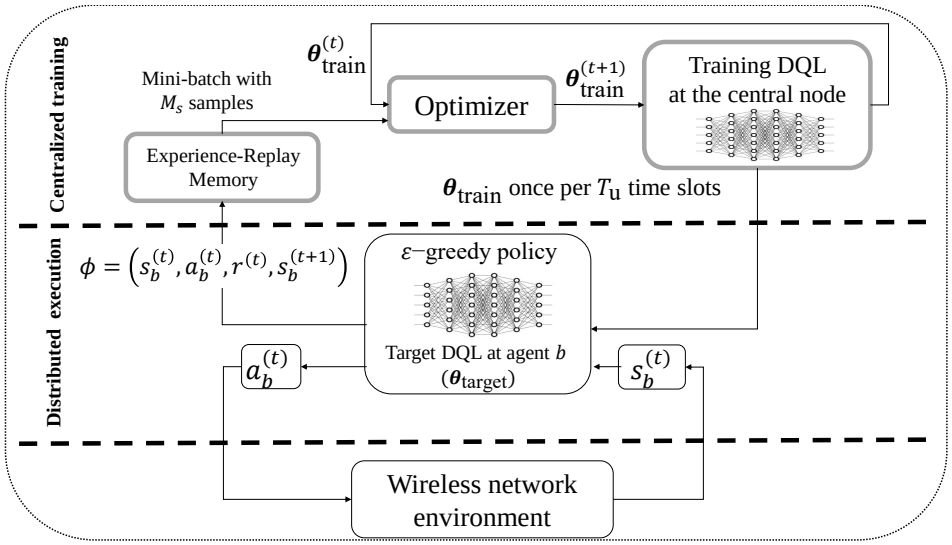
According to (76), the Q-learning algorithm converges to an optimal policy π^* if each action in the action space is executed under each state for an infinite number of times on an infinite run and the learning rate β decays appropriately. However, the Q-learning algorithm is only suitable when the state-action space is small, which is not our case. The reason is that the storage of the lookup table related to (4.10) becomes impractical as the state-action space increases and several states will be rarely visited, consequently creating holes in the lookup table.

Therefore, to deal with this issue, the use of deep Q-learning (77) should be properly adapted. Roughly speaking, deep Q-leaning or deep Q-Network (DQN) corresponds to a merge between the Q-learning algorithm and a deep neural network. In summary, the lookup table is replaced by a deep neural network. Therefore, the $Q(\cdot,\cdot)$ function can be expressed as $Q(\cdot,\cdot,\theta)$, where θ represents the DQN parameters. Obviously, this considerably reduces the search space, since in DQN the space is reduced to searching the best θ of finite dimensions, while in Q-learning a search is done within a lookup table of huge dimensions.

Like the classical Q-learning algorithm, for each time step t, the agent obtains experiences from taken actions on the environment in the form $\phi = (s^{(t)}, a^{(t)}, r^{(t)}, s^{(t+1)})$. A set of experiences are stored by the agent in a memory with size M. Note that the memory size is limited, where the oldest experiences are overwritten by the most recent ones when the number of experiences exceeds the capacity. Furthermore, according to (77), in order to improve the DQN stability, two DQNs are defined: the target DQN with parameters $\theta^{(t)}_{\text{target}}$ and the training DQN with parameters $\theta^{(t)}_{\text{train}}$. At every T_u steps, $\theta^{(t)}_{\text{target}}$ is updated to be equal to $\theta^{(t)}_{\text{train}}$. Also, the memory replay method can avoid oscillations and divergence in the parameters (77). Therefore, at time t, the least squares loss of the training DQN for a random mini-batch $\mathcal{M}^{(t)}$ with M_s samples is

$$L(\theta^{(t)}_{\text{train}}) = \sum_{\phi \in \mathcal{M}^{(t)}} \left(y^{(t)} - Q(s^{(t)}, a^{(t)}, \theta^{(t)}_{\text{train}}) \right)^2, \tag{4.11}$$

Figure 20 – Flowchart of the proposed multi-agent deep reinforcement learning algorithm with centralized training and distributed execution.



Source: Adapted from (115).

where the target is

$$y^{(t)} = r^{(t)} + \gamma \max_{\forall a \in \mathcal{A}} Q(s^{(t+1)}, a, \theta_{\text{target}}^{(t)}). \tag{4.12}$$

Observe that $y^{(t)}$ is calculated using the target DQN.

Finally, the loss function (4.11) is minimized by a stochastic gradient algorithm in order to train the mini-batch $\mathcal{M}^{(t)}$. Then, the training DQN updates its parameters with the new parameters provided by training. According to (121), the convergence to a set of good parameters occurs quickly.

4.6 Proposed multi-agent Deep Q-learning

We propose a multi-agent DQL with B agents, i.e., each BS is an agent. We propose an approach in which the training is performed by a centralized node, while the actions are taken distributively by agents, such as in (115, 122) in order to improve stability of our multi-agent optimization. In Figure 20 we illustrate the structure of the learning scheme. Specifically, we define $s_b^{(t)} \in \mathcal{S}$ as the state of agent b at time slot t, which is responsible for characterizing the environment of agent b and points out relevant information to agent b's action $a_b^{(t)} \in \mathcal{A}$. The actions are taken based only on local information employing the ϵ-greedy policy at each agent, where all actions are taken simultaneously and no agent has previous knowledge of actions taken by other agents. In addition, each agent receives a reward, $r_b^{(t)}$, based on the taken action from the environment and move on to the next state, $s_b^{(t+1)}$. Finally, an experience $\phi = (s_b^{(t)},\ a_b^{(t)},\ r_b^{(t)},\ s_b^{(t+1)})$ is obtained.

After that, the agents report their experiences to the central node, where they are stored within one same memory. This technique is denoted as replay memory (77, 115). Moreover,

note that the memory size, M, is limited, where the oldest experiences are overwritten by the most recent ones when the number of experiences exceeds the capacity. The DQLs can be expressed as a Q-function, $Q(s,a,\theta)$, which is defined as the expected accumulated discounted rewards once action a is taken under state s, and θ are the DQL parameters. Besides, two DQLs are defined with parameters $\theta^{(t)}_{\text{target}}$ and $\theta^{(t)}_{\text{train}}$. Each agent has the same copy of $\theta^{(t)}_{\text{target}}$, while $\theta^{(t)}_{\text{train}}$ are located at the central node. In addition, at time t, the central node randomly selects a mini-batch $\mathcal{M}^{(t)}$ with M_s samples from memory and, using a stochastic gradient algorithm, it minimizes the least squares loss of the training DQL, given by

$$L(\theta^{(t)}_{\text{train}}) = \sum_{\phi \in \mathcal{M}^{(t)}} \left(r^{(t)} + \gamma \max_{a} Q(s^{(t+1)}, a, \theta^{(t)}_{\text{target}}) - Q(s^{(t)}, a^{(t)}, \theta^{(t)}_{\text{train}}) \right)^2, \tag{4.13}$$

where $\gamma \in [0,1]$ is the discount factor (115). Finally, at every T_u steps, $\theta^{(t)}_{\text{target}}$ is set to be equal to $\theta^{(t)}_{\text{train}}$.

The behavior of the central node during the training phase is detailed in Algorithm 3. The behavior of the agents are detailed in sections 4.6.1 and 4.6.2.

Algorithm 3 Multi-Agent DQL: Central node training

Initialize θ_{train} and θ_{target}.
Send copies of θ_{target} to each agent.
repeat
    if Memory size M is full **then**.
        Delete B oldest experiences.
    end if
    Receive from each agent b its experience $\phi^{(t)}$.
    Store the agents experiences in the memory.
    Select a mini-batch $\mathcal{M}^{(t)}$ with M_s samples from memory.
    Update $\theta^{(t)}_{\text{train}}$ using the stochastic gradient algorithm regarding (4.13).
    if $t \pmod{T_u} = 0$ **then**
        Update $\theta^{(t)}_{\text{target}}$ to be equal to $\theta^{(t)}_{\text{train}}$ and send copies to each agent.
    end if
until Some convergence criterion.

4.6.1 Multi-agent DQL with Full Information

To proceed we assume that each BS has access to CSI of the channels from itself to all served UEs in the own cell and to interfered UEs in neighbor cells, i.e., $\hat{\mathbf{h}}_{i_j b}, \forall i_j \in \mathcal{K}_j$ and $\forall j \in \mathcal{B}$. This requirement is due to the need to solve (4.5) in order to compute the transmit vectors. This knowledge can be acquired by means of pilot signaling, and may be imperfect due to estimation errors, as discussed in Section 4.2. In this section, we also assume that the nodes

can communicate via control link during training and execution phases in order to coordinate themselves. The structure for that signaling strategy is discussed in section 4.6.1.4. Next, we discuss and define the proposed structure for the specific DQL, by characterizing state, action and reward function.

4.6.1.1 State

As aforementioned, the state is responsible to characterize the environment for each agent and it is crucial for determining the best choice of an action. Therefore, we consider that $s_b^{(t+1)}$ is composed by three aspects: the estimated channel for all users served by agent b, $\hat{\mathbf{h}}_{k_b b}^{(t+1)}$, the previous total capacity, $C^{(t)}$, and the previous activation pattern, $\boldsymbol{\alpha}_b^{(t)}$, which will be explained in the next subsection.

However, note that $\hat{\mathbf{h}}_{k_b b}$ is composed by complex numbers, which are not supported by the current neural network software. To deal with this issue, two approaches can be used. The first consists of separating the complex vector $\hat{\mathbf{h}}_{k_b b}$ into the in-phase (real part) and quadrature (imaginary part) components, denoted as I/Q transformation, $\Re(\hat{\mathbf{h}}_{k_b b})$ and $\Im(\hat{\mathbf{h}}_{k_b b})$, respectively. The second approach, on the other hand, consists of mapping the complex vector $\hat{\mathbf{h}}_{k_b b}$ using the phase and magnitude information, $\mathfrak{P}(\hat{\mathbf{h}}_{k_b b})$ and $\mathfrak{M}(\hat{\mathbf{h}}_{k_b b})$, respectively, which is denoted as P/M transformation. Without loss of generality, we adopt the I/Q transformation. In addition, all $\Re(\hat{\mathbf{h}}_{k_b b}^{(t+1)})$ and $\Im(\hat{\mathbf{h}}_{k_b b}^{(t+1)})$, $\forall k_b \in \mathcal{K}_b$ are organized into only one vector $\hat{\mathbf{h}}_b^{(t+1)} \in \mathbb{C}^{2K_b N_b \times 1}$. Thus, we have that the state of agent b at time $t+1$ is defined as

$$s_b^{(t+1)} = \{\hat{\mathbf{h}}_b^{(t+1)}, C^{(t)}, \boldsymbol{\alpha}_b^{(t)}\}. \tag{4.14}$$

4.6.1.2 Action

To fulfill real-world demands, such as low latency and complexity constraints, we limit the set of actions to activating a subset of UEs. Therefore, we define an activation pattern $\boldsymbol{\alpha}_b = [\alpha_1, \ldots, \alpha_{K_b}]$ as a binary vector, in which α_k is equal to 1 if the UE k is selected to be active and zero otherwise. Thus, the action of the agent b is given by $\boldsymbol{\alpha}_b$ and the set of actions $\mathcal{A}_b$ is given by all the possible combinations for a given value of K_b. For example, assuming $K_b = 3$ we have

$$\mathcal{A}_b = \{[0,0,0],\ [0,0,1],\ [0,1,0],\ [0,1,1],\ [1,0,0],\ [1,0,1],\ [1,1,0],\ [1,1,1]\}. \tag{4.15}$$

Note that the number of possibilities is equal to 2^{K_b}, thus, the action space increases exponentially with the number of users per BS. However, according to (123), standardized systems can multiplex up to 4 UEs on the same time-frequency resource, resulting in a confined action space.

Once the action is taken and the active users are defined, the agents calculate the beamforming direction for all active users according to (4.5), using as input the estimated

channels, $\hat{\mathbf{h}}_{k_b j},\ \forall k \in \mathcal{K}_b$ and $\forall b, j \in \mathcal{B}$. Moreover, the parameters μ_b and λ_{k_b} are heuristically calculated such as in the SLNR-MAX beamforming proposed in (34):

$$\mu_b = \frac{K_b}{\sum_{j=1}^{B} K_j} = \frac{K_b}{K}, \tag{4.16}$$

and

$$\lambda_{k_b} = \frac{\alpha_{k_b}}{\sum\limits_{b=1}^{B} \|\boldsymbol{\alpha}_b\|_1}, \tag{4.17}$$

where $\|\cdot\|_1$ is the l_1 norm operator.

These parameters are related to the relative importance of enforcing the power constraint in a BS and the UE priority, respectively. Recall that the structure in (4.5) is based on perfect CSI, thus, the UEs' priority that maximizes the sum-rate can be suboptimal for scenarios with CSI errors. However, we can set the UEs' priority based on the taken action. Therefore, as the agents learn the best policy, they can select a better priority for each user in the presence of CSI errors. The power allocation, on the other hand, is performed by the well-known water-filling algorithm (34). Also note that, in order to compute λ_{k_b} in (4.17), the knowledge about other agents' activation patterns $\boldsymbol{\alpha}_b \forall b \in \mathcal{B}$ is needed.

4.6.1.3 Reward Function

Finally, the reward function must be designed to optimize the objective. Thus, the more obvious choice is the total capacity. Also, the agents' actions aim to increasingly obtain rewards, thus, the knowledge of the real quantity of transmitted data allows the solution to be more robust to the imperfect CSI, because the actions can be adjusted to minimize the channel estimation errors and obtain higher capacity. Therefore,

$$r^{(t)} = C^{(t)} = \sum_{b=1}^{B} \sum_{k=1}^{K_b} \log_2(1 + \Gamma_{k_b}). \tag{4.18}$$

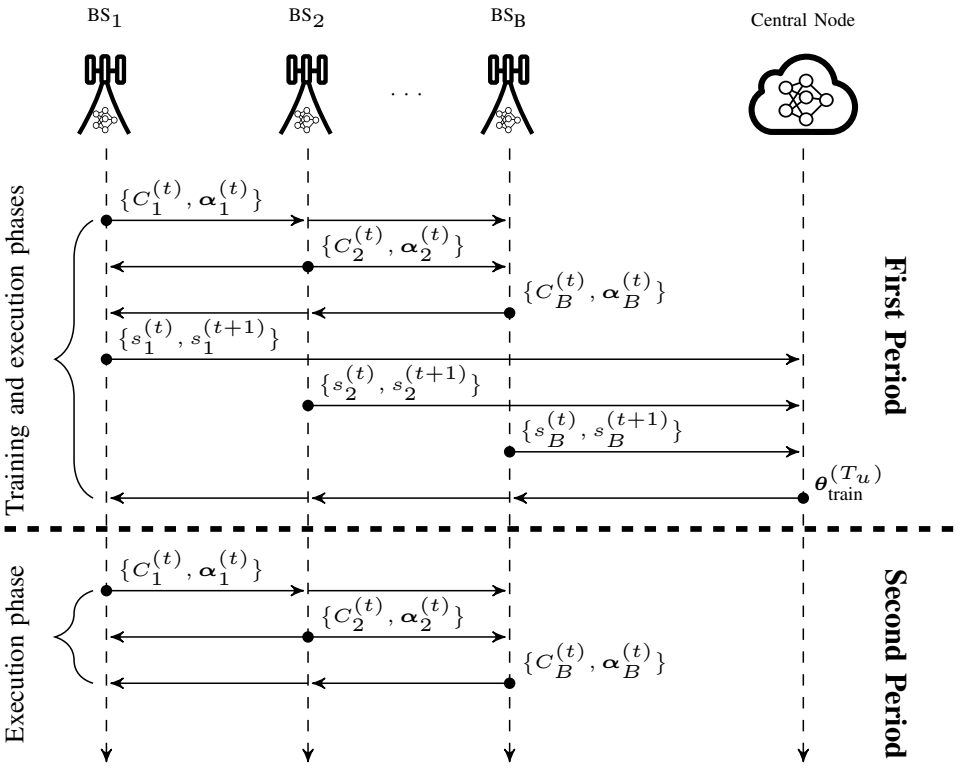
Note that, to effectively compute the reward in a real world deployment, we assume that each UE estimates the actual amount of transmitted data and, next, it is reported to the BS.

4.6.1.4 Signaling Scheme

Additionally, the proposed multi-agent DQL relies on a specific signaling scheme in order to supply the required information during execution. As it can be seen in Figure 21, we have two periods along the time. In the first period, both training (update of parameters of training DQL at the central node) and execution (agents' action and beamformers computation at the BSs) phases are performed in parallel.

Note that the CSI is constant during a time slot, the periods represent several time slots and each phase corresponds to an iteration of the proposed scheme, which fits within a time

Figure 21 – Signaling strategy for the training and execution phases for the DQL approaches.



Source: Created by the author.

slot. Consequently, the CSI changes during the training. Moreover, once the training is finalized, the DQL can continue taking greedy actions and new training is required only when there are changes in the neural network architecture, e.g., changes in the number of antennas

It is worth highlighting that it is not necessary to end the training to obtain the actions and, consequently, the beamformers. In addition, since the training is performed at the central node, the beamformers computation at the BSs reduces to linear and nonlinear operations, avoiding iterative processes. Thus, the computational complexity is reduced.

That said, at time slot $t+1$, each agent requires knowledge about the overall capacity achieved in the previous interval, $C^{(t)}$, since it is part of its input state, as in (4.14). The capacity $C^{(t)}$ consists of the sum of the rates of each BS, suggesting that each agent must share via backhaul its local capacity with other agents and acquire the local capacity from other ones. Also, during execution phase each agent needs the action pattern chosen by each BS, $\boldsymbol{\alpha}_b$, in order to compute the parameters λ_{k_b}, as in (4.17). In addition, the central node needs to receive via backhaul at each interval the experience from the agents, ϕ. Since $a_b^{(t)}$ and $r_b^{(t)}$ are already represented in the next state set $s_b^{(t+1)}$, each BS shares with the central node the tuple $(s_b^{(t)}, s_b^{(t+1)})$. During training it is also needed that at each T_u intervals the neural network parameters $\theta_{\text{train}}^{(t)}$ be broadcast from the central node to other agents via backhaul. In summary, during training, each agent must broadcast a scalar number representing the local capacity and K_b binary numbers for

the action pattern, besides reporting $2K_bN_b$ scalars regarding channel parameters to the central node. In addition, the central node broadcasts at each T_u time a scalar for each parameter of the neural network. Since the proposed neural network is a fully-connected DQL consisting of five layers, with three hidden layers, the number of parameters is given by

$$|\theta| = \sum_{l=0}^{3} W_{l+1}(W_l + 1), \tag{4.19}$$

where W_l is the number of neurons in layer l (115).

Once training is ended, only the execution phase is performed in the second period. Therefore, the central node leaves the network and the backhaul link will be used only for the information share between agents, i.e., the number of parameters is reduced to a scalar plus K_b binary digits, which refer to the local capacity and the action pattern, respectively.

In order to illustrate the amount of required training and execution signaling control overhead for the proposed solutions, we consider a scenario with parameters as described in Table 9 assuming a 1 ms time slot with $T_u = 100$ and that scalars are encoded as single precision floats (32 bits). In such network the signaling overhead generated during training by each agent is estimated as around 1.06 Mbps, while the central node generated training overhead is estimated as around 10.37 Mbps. On the other hand, the execution time overhead is only 36 bps. This way, we see that although we have a considerable amount of training overhead, the execution overhead is very reduced. We recall that the cooperation among agents in order to coordinate the network in the direction of a common capacity goal motivates our training scheme. However, we highlight that completely decentralized schemes with no need for control signaling are highly desirable and indeed motivators for future research.

The behavior of the agents, including signaling, during the training phase and execution phases are detailed in algorithm 4.

4.6.2 Multi-agent DQL with Local Information

Although the previous solution is an execution-distributed framework, the agents need to exchange information with each other, even when the training is finalized. Moreover, each BS needs to estimate the channel vectors from itself to the interfered UEs in order to compute the transmit vectors as in (4.5). In order to reduce such requirements, we also propose a solution based only on local information, where the BS has knowledge exclusively of the channel vectors to its served UEs. Thus, we assume the beamforming structure for single-cell case (98), where

$$\tilde{\mathbf{m}}_{k_b} = \frac{\left(\mathbf{I} + \sum\limits_{i=1}^{K_b} \frac{\lambda_{i_b}}{\sigma^2} \mathbf{h}_{i_b b} \mathbf{h}_{i_b b}^H\right)^{-1} \mathbf{h}_{k_b b}}{\left\| \left(\mathbf{I} + \sum\limits_{i=1}^{K_b} \frac{\lambda_{i_b}}{\sigma^2} \mathbf{h}_{i_b b} \mathbf{h}_{i_b b}^H\right)^{-1} \mathbf{h}_{k_b b} \right\|}. \tag{4.20}$$

Algorithm 4 Multi-agent DQL with Full Information: Agent b

Initialize the local copy of θ_{target} to be equal to the one shared by the central node.
Estimate channels, $\hat{\mathbf{h}}_{i_j b}^{(1)}, \forall i_j \in \mathcal{K}_j$ and $\forall j \in \mathcal{B}$, by means of pilot signaling and MMSE.
Initialize the values of the tuple $\{C_b^{(0)}, \boldsymbol{\alpha}_b^{(0)}\}$.
Build the initial state vector $s_b^{(1)} = \{\hat{\mathbf{h}}_b^{(1)}, C^{(0)}, \boldsymbol{\alpha}_b^{(0)}\}$.
repeat
    Run the local DQN and update the activation pattern, $\boldsymbol{\alpha}_b^{(t)}$, following the ϵ-greedy policy.
    Compute parameters μ_b and $\lambda_{k_b} \forall k \in \mathcal{K}_b$ as in (4.16) and (4.17), respectively.
    Update the beamforming vectors, $\mathbf{m}_{k_b} \forall k \in \mathcal{K}_b$, using (4.5).
    Use computed $\mathbf{m}_{k_b} \forall k \in \mathcal{K}_b$ to transmit.
    Compute the reward, $C_b^{(t)}$, as the achieved rate.
    Share with other agents the achieved capacity and activation pattern $\{C_b^{(t)}, \boldsymbol{\alpha}_b^{(t)}\}$.
    Estimate channels, $\hat{\mathbf{h}}_{i_j b}^{(t+1)}, \forall i_j \in \mathcal{K}_j$ and $\forall j \in \mathcal{B}$, by means of pilot signaling and MMSE.
    Build the state vector $s_b^{(t+1)} = \{\hat{\mathbf{h}}_b^{(t+1)}, C^{(t)}, \boldsymbol{\alpha}_b^{(t)}\}$ from gathered information.
    if Training phase = true **then**
        Send to central node its experience set ϕ as the tuple $(s_b^{(t)}, s_b^{(t+1)})$.
        if $t \pmod{T_u} = 0$ **then**
            Receive $\theta_{\text{train}}^{(t)}$ from central node and update local $\theta_{\text{target}}^{(t)}$ to be equal to $\theta_{\text{train}}^{(t)}$.
        end if
    end if
until Some convergence criterion.

Note that it is a special case of (4.5) and is also based on perfect CSI. Moreover, the μ_b parameters are not calculated and the action patterns do not need to be transmitted to other BSs. Here, we also assume that the input channels are the estimated ones, instead of the actual values. Moreover, the actions and the reward function remain the same as those in the previous solution and only the state is slightly changed, where instead of the previous total capacity, $C^{(t)}$, we assume the previous capacity achieved by BS b, $C_b^{(t)}$. Therefore, differently from Figure 21, the backhaul link is used only for the exchange of information with the central node for the training phase (first period), while in the second period no information is exchanged among the agents.

The behavior of the agents, including signaling, during the training phase and execution phases are detailed in algorithm 5.

In Table 7 we summarize the main differences of the two versions of the proposed Multi-agent DQL, one with full information and the other with only local information, with respect to the state, action, reward function and signaling requirement.

Algorithm 5 Multi-agent DQL with Local Information: Agent b

Initialize the local copy of θ_{target} to be equal to the one shared by the central node.
Estimate channels, $\hat{\mathbf{h}}_{i_b b}^{(1)}, \forall i_b \in \mathcal{K}_b$, by means of pilot signaling and MMSE.
Initialize the values of the tuple $\{C_b^{(0)}, \boldsymbol{\alpha}_b^{(0)}\}$.
Build the initial state vector $s_b^{(1)} = \{\hat{\mathbf{h}}_b^{(1)}, C_b^{(0)}, \boldsymbol{\alpha}_b^{(0)}\}$.
repeat
    Run the local DQN and update the activation pattern, $\boldsymbol{\alpha}_b^{(t)}$, following the ϵ-greedy policy.
    Compute parameters $\lambda_{k_b} \forall k \in \mathcal{K}_b$ as in (4.17).
    Update the beamforming vectors, $\mathbf{m}_{k_b} \forall k \in \mathcal{K}_b$, using (4.20).
    Use computed $\mathbf{m}_{k_b} \forall k \in \mathcal{K}_b$ to transmit.
    Compute the reward, $C_b^{(t)}$, as the achieved rate.
    Estimate channels, $\hat{\mathbf{h}}_{i_b b}^{(t+1)}, \forall i_b \in \mathcal{K}_b$, by means of pilot signaling and MMSE.
    Build the state vector $s_b^{(t+1)} = \{\hat{\mathbf{h}}_b^{(t+1)}, C_b^{(t)}, \boldsymbol{\alpha}_b^{(t)}\}$.
    if Training phase = true **then**
        Send to central node its experience set ϕ as the tuple $(s_b^{(t)}, s_b^{(t+1)})$.
        if $t \pmod{T_u} = 0$ **then**
            Receive $\theta_{\text{train}}^{(t)}$ from central node and update local $\theta_{\text{target}}^{(t)}$ to be equal to $\theta_{\text{train}}^{(t)}$.
        end if
    end if
until Some convergence criterion.

Table 7 – Strategy comparison of Full information (DQL-F) and Local information (DQL-L).

Strategy	State	Action	Reward	Signaling
DQL-F	$\{\hat{\mathbf{h}}_b^{(t+1)}, C^{(t)}, \boldsymbol{\alpha}_b^{(t)}\}$	$\boldsymbol{\alpha}_b$	$C^{(t)}$	Both periods in Figure 21
DQL-L	$\{\hat{\mathbf{h}}_b^{(t+1)}, C_b^{(t)}, \boldsymbol{\alpha}_b^{(t)}\}$	$\boldsymbol{\alpha}_b$	$C^{(t)}$	Only exchanges with central node in Figure 21

Source: Created by the author.

4.7 Results and Performance Evaluation

4.7.1 Simulation Parameters

We evaluate the performance of the proposed solution by means of numerical simulations, where we compare the proposed algorithm with benchmark solutions from the literature. The simulations are conducted in the downlink of a multicell MISO scenario with parameters $\{B, K_b, N_b\} = \{4, 4, 4\}$. The BSs are located at the center of each hexagonal cell and the UEs are uniformly distributed within each cell. The cell radius is equal to 500 m and we define an inner region with a 100 m radius from the cell center in which no UE can be placed.

In this work, we consider the main propagation mechanisms in the modeling. Therefore, the propagation model consists of a distance-dependent path loss model, a log-normal shadowing component, and a Rayleigh-distributed fast fading component. The path loss is given by $128.1 + 37.6\log_{10}(d)$ (in dB), where d is transmitter-to-receiver distance in kilometers and the log-normal shadowing standard deviation is 10 dB (124). It is also assumed that each time slot has a duration shorter than the channel coherence time and at each 10 time slots the UE positions are changed randomly. The power budget (P_b) for all BSs is equal to 46 dBm and we assume 10 MHz bandwidth. Finally, we set the noise power spectral density equal to -174 dBm/Hz. The main network parameters are presented in Table 8.

Table 8 – Main simulations parameters regarding network configuration.

Parameter	**Value**	**Unit**
$\{B, K_b, N_b\}$	$\{4, 4, 4\}$	–
Cell radius	500	m
Bandwidth	10	MHz
Power budget per BS	46	dBm
Shadowing standard deviation	10	dB
Path loss	$128.1 + 37.6\log_{10}(d)$	dB
Noise power spectral density	-174	dBm/Hz

Source: Created by the author.

Both DQL schemes were implemented using Tensorflow (125) and have the form of a five-layer fully connected neural network with one input layer, three hidden layers and one output layer. The numbers of neurons in the three hidden layers are 200, 100 and 40, respectively. The input size is equal to $K_b(2N_b + 1) + 1$ for each agent and the output size is equal to 2^{K_b}. Since K_b is set equal to 4, we have that the number of inputs and outputs are $(8N_b + 5)$ and 16, respectively.

We use the rectifier linear unit as DQL's activation function. The memory parameters M and M_s are equal to 10,000 and 256, respectively, and we use the Adam algorithm (126) with a learning rate equal to 10^{-4}. Moreover, the ϵ-greedy policy is given by $\epsilon^{(t+1)} = \max\{\epsilon_{\min}, (1 - \eta_\epsilon)\epsilon^{(t)}\}$ with $\epsilon^0 = 0.2$, $\epsilon_{\min} = 10^{-2}$ and $\eta_\epsilon = 10^{-4}$. Also, we set γ equal to 0, since the correlation between the agent's actions and its future rewards tends to be smaller for our application due to fading. The main learning parameters are summarized in Table 9.

In our framework, the DQL is trained once. Therefore, the number of time slots for training must be large enough to overcome the generalization problem. In this work, we dedicate 40,000 time slots for training. Moreover, during training, we set $T_u = 100$, thus, once every 100 time slots the central node will broadcast the trained parameters to all agents, which are available in the next time slot for usage. Once the DQL is trained, the central node leaves the network and the ϵ-greedy algorithm is finalized, i.e., the agents take only greedy actions. In order to

Table 9 – Main simulations parameters regarding learning scheme configuration.

Parameter	**Value**
Neural network type	Fully connected
Neural network size	$\{(8N_b+5), 200, 100, 40, 16\}$
Activation function	rectifier linear unit (ReLU)
Memory parameters $\{M, M_s\}$	$\{10000, 256\}$
Learning rate (126)	10^{-4}
ϵ-greedy policy	$\epsilon^{(t+1)} = \max\{10^{-2}, (1-10^{-4})\epsilon^{(t)}\}$, with $\epsilon^0 = 0.2$
γ	0

Source: Created by the author.

validate the training, another 5,000 time slots are used for the testing phase, in addition to the 40,000 dedicated to the training. Thus, we assume a total of 45,000 time slots. Note that the number of time slots for the testing phase was selected only for statistical reliability reasons, thus, in practical scenarios the DQL could continue taking greedy actions without requiring a new training.

In the simulation results, the DQL with full information is denoted as DQL-F and the one with only local information is called DQL-L. Moreover, we have three benchmarks to evaluate the performance of our algorithms, assuming that all solutions are fed with the estimated channel. The first is the WMMSE algorithm (47), the second is the SLNR-MAX beamforming (34) using the water-filling algorithm for power allocation and the third is the SUS algorithm presented in (59).

Finally, in order to evaluate the performance of solutions with respect to execution time, we compute the gain in relation to the execution time of WMMSE, which is used as reference to maximum execution time, given by

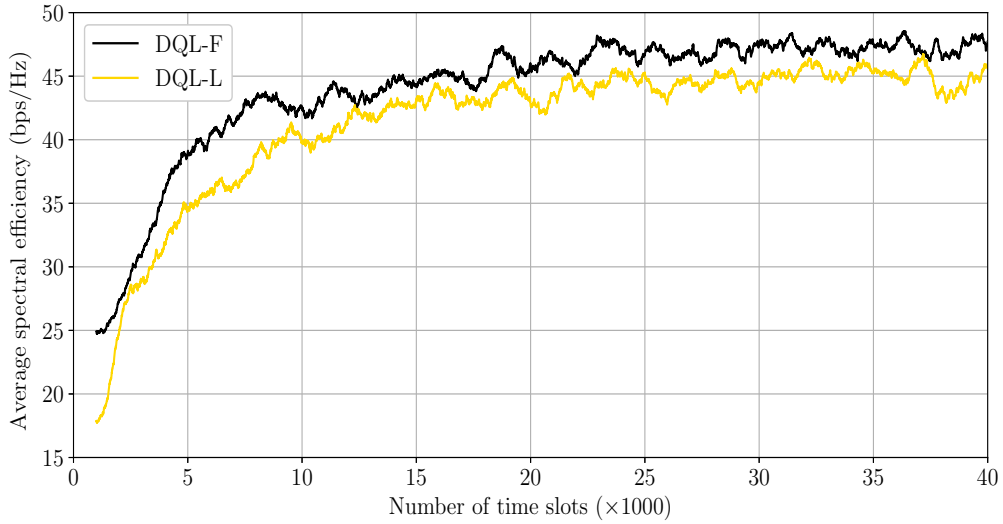
$$\text{Gain } (\%) = 100 \cdot \left(1 - \frac{\text{execution time of the solution}}{\text{execution time of WMMSE}}\right). \tag{4.21}$$

4.7.2 Optimization results

In Figure 22 we present the average spectral efficiency as a function of time in the first period, with channel estimation reliability $\xi = 0.95$ for both DQL-F and DQL-L solutions. As we can see, the achieved spectral efficiency increases with time, which shows that the agents learn to take better actions over time, selecting the action patterns that obtain higher data rates. Moreover, we observe that the DQL-F solution achieves higher spectral efficiency than DQL-L, as expected. We also see in Figure 22 that convergence speed is much greater in the beginning of the training procedure, so that the number of time slots dedicated for the training can be further reduced without great impact on the spectral efficiency.

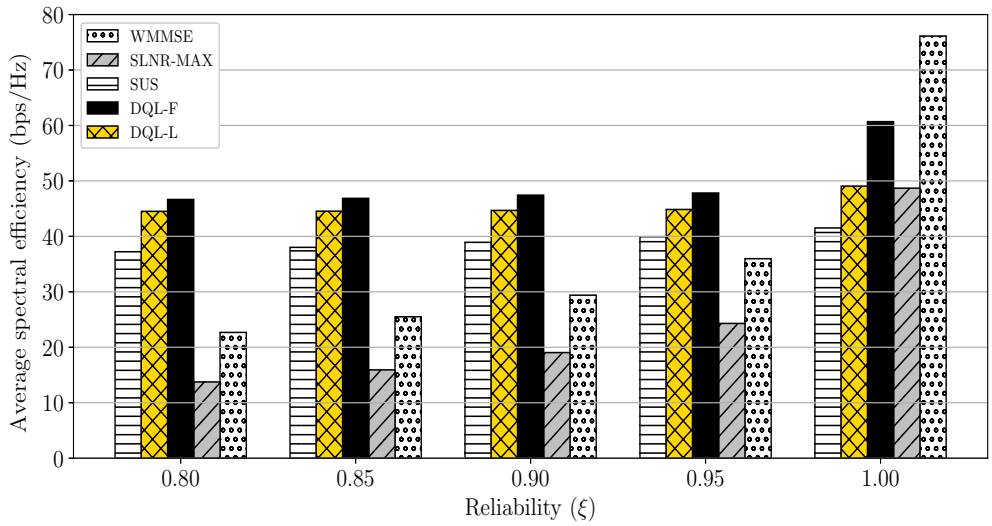
In Figure 23, we show the spectral efficiency versus the reliability of channel estima-

Figure 22 – Comparison of average spectral efficiency during the first period (training).



Source: Created by the author.

Figure 23 – Comparison of multiple strategies regarding spectral efficiency vs. reliability in the execution phase.

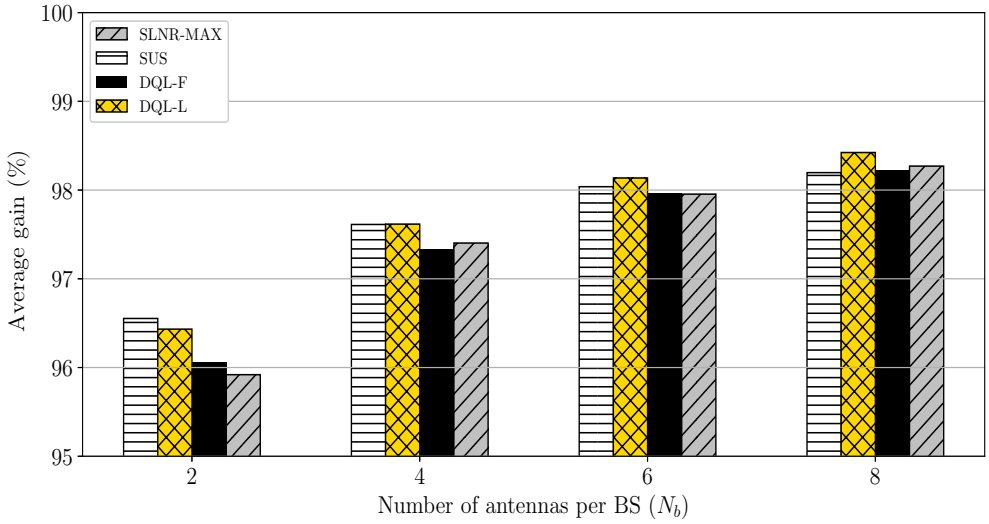


Source: Created by the author.

tion for all solutions in the second period. Note that for each value of ξ the DQLs were trained from scratch. As we can observe, the DQLs generalize the training for new channel realizations in the second period, where the training is finalized and only greedy actions are allowed. Indeed, the average spectral efficiency in the execution phase is similar to the one achieved in the training phase. In addition, spectral efficiency increases as reliability increases.

Another important observation from Figure 23 is related to the gain of the proposed solutions in comparison to the reference algorithms when there are estimation errors. Assuming ξ equal to 0.9, we have spectral efficiency gains for the DQL-F of 22%, 61% and 150%, when compared to SUS, WMMSE, SLNR-MAX, respectively. The reason for this is that the proposed

Figure 24 – Comparison of multiple strategies regarding the average execution time gain compared to the WMMSE algorithm.



Source: Created by the author.

solutions, in spite of being fed with the estimated channel, receive the real data rate achieved by the system as a reward. Thus, they learn to adjust their actions in order to maximize the obtained reward and, consequently, achieve higher data rates compared to the reference solutions. We highlight that in this result both training and execution were performed considering the same channel reliability ξ and do not consider scenarios in which there is a mismatch in reliability between training and execution.

We also observe that DQL-L presents a very close performance to that of DQL-F, which shows that the signaling can be reduced at the cost of a slight performance loss. We can also observe that WMMSE outperforms both DQL-F and DQL-L when the channel is perfectly estimated. The reason is that the dual variables λ_{k_b} and μ_b for all UE k and BS b, as well as the transmit power allocation, are calculated heuristically based on the taken actions. Consequently, the proposed solution has lower degrees of freedom than the WMMSE algorithm. Even so, our results clearly indicate that the proposed solutions yield gains compared to SLNR-MAX and SUS solutions.

Finally, in Figure 24 we present the average execution time gain compared to the WMMSE algorithm versus the number of antennas, assuming ξ equal to 0.9. The simulations were written in python, using equal configurations and all tests were performed using one same hardware setup. As we can see, the DQL is, approximately, 96% faster than the WMMSE algorithm when the number of antennas is 2 and increases as the number of antennas increases. This occurs because WMMSE calculates beamforming iteratively. On the other hand, we have that DQL presents an execution time similar to SLNR-MAX and SUS, considering that these solutions use closed-form expressions to compute the beamformers without iterations. In addition, DQL-L has a slight gain compared to the DQL-F solution, since the beamforming structure requires less calculations. As an overall analysis, the proposed solutions reach a good trade-off

between performance and execution time when compared to other solutions selected from recent literature.

4.8 Conclusions

In this work, we focused on the sum-rate maximization problem in a multi-cell MISO scenario with per-BS power constraints and imperfect CSI using multi-agent DQL. The proposed solutions are distributedly executed, with the transmit beamformers calculated based on agents' actions and expertise on the problem structure. From the results, we concluded that the proposed DQL schemes are able to learn a policy that makes a beamforming structure based on perfect CSI robust to scenarios with channel estimation errors. We also observed that signaling overhead can be reduced at the cost of a slight performance loss. In addition, the proposed schemes outperformed state-of-the-art solutions in scenarios with imperfect CSI and presented a low execution time, which is an important feature in real-world applications. As perspectives for future works we can mention the joint power and user scheduling or to consider a grid of beams approach, in which the BS selects a precoder out of a discrete set of beamformers (codebook).

5 CONCLUSIONS AND FUTURE WORK

In this thesis, we addressed the design of novel strategies to improve the performance of future mobile wireless systems. The main objective of this work was to propose decentralized strategies for the optimization of beamforming and user scheduling in scenarios from 5G and beyond systems. Along with that, we aimed for solutions that were tractable and applicable to real-world scenarios, by pursuing robust designs of distributed execution and by providing the required signaling schemes. This objective, along with the technological context that motivates it, and a list contributions of this thesis were presented in Chapter 1.

In Chapter 2, we discuss the main aspects regarding the major technologies and theories that generate the questions, scenarios and solution approaches studied in this thesis. First, we discuss the aspects of beamforming, which is the technology that permeates the whole context of this thesis, and we motivate the design of improved beamforming algorithms to enhance network performance. Then, we also discuss the aspects and motivate the design of refined user scheduling strategies as a means of improving efficiency. Next, we argue that the employment of distributed algorithms is desired and necessary for cooperative optimization in wireless mobile networks, by comparing aspects of centralized and decentralized optimization. After that, we provide an overview of the solution methods employed in this thesis: convex optimization and machine learning.

In Chapter 3, we considered the bidirectional sum-power minimization beamforming in a multi-user, multi-stream, MIMO network, as a means to deal with interference generated by the dynamic TDD scenario, by forcing a minimum signal-to-interference-plus-noise ratio constraint for both UL and DL. In order to solve this beamforming problem, we proposed two iterative approaches. The first presented approach is based on alternate convex optimization of receive and transmit vectors, assumes centralized processing, and requires the availability of global CSI. For that centralized solution, we derived and presented the proof that the proposed algorithm converges globally to the set of KKT conditions. The second approach is performed in a decentralized manner, based on ADMM and requires only local channel state information and a reduced signaling load for coordination. For that decentralized solution, we designed a lightweight signaling scheme to support the decentralized algorithm application and provided a convergence and optimality analysis. We also were able to provide numerical results which showed that both centralized and decentralized approaches were capable of converging to a minimum network power expenditure while guaranteeing SINR performance, and that the decentralized solution can well approximate the performance of the centralized one with a reduced amount of iterations and control signaling.

In Chapter 4 we addressed the development of distributed-execution ML user scheduling schemes to solve the beamforming problem in DL multi-cell MISO networks in the presence of CSI errors, with the objective of sum-rate maximization with per BS power constraints. The main essence of the proposed learning schemes is that they are able to efficiently adapt and

improve the performance of an optimal beamforming structure based on perfect CSI in scenarios with the presence of channel estimation errors. Both learning strategies are based on a multi-agent DQL algorithm, in which the training is performed by a centralized node, while the actions are taken distributively by agents. In that learning strategy, each agent acts based on a specific state, that characterizes its environment and points out relevant information, and receives back a reward as a response for that action. To fulfill real-world demands, such as low latency and complexity constraints, we limited the set of actions to activating a subset of UEs, i. e. user scheduling. The first proposed decentralized learning algorithm DQL-F assumes that each BS has access to local CSI and that the nodes can communicate via control link during training and execution phases in order to coordinate themselves. For that solution, we designed a lightweight signaling scheme to support the decentralized algorithm application. The second proposed decentralized learning algorithm DQL-L only assumes local CSI knowledge. We also provided numerical results which showed that the proposed DQL schemes were able to learn a policy that makes a beamforming structure based on perfect CSI robust to scenarios with channel estimation errors. We also observed that signaling overhead can be reduced at the cost of a slight performance loss. In addition, the proposed schemes outperformed state-of-the-art solutions in scenarios with imperfect CSI and presented a low execution time, which is an important feature in real-world applications.

Overall, throughout this thesis, we were able to propose the desired novel decentralized beamforming and user scheduling strategies to improve the performance of wireless mobile networks, along with the required schemes to support their application. Also, we were able to assess the performance of the proposed algorithms through simulation and analysis of results, which reinforced the quality of such solutions.

Future Research

As direct extensions of the works presented in this thesis we can enlist the following:

- *Imperfect CSI for the dynamic TDD solution*: The assumption of perfect CSI knowledge does not hold in reality. Although state-of-the-art channel estimation methods can provide good estimation, some levels of incorrectness still hold. This way, methods to optimize the beamforming which are robust to CSI errors are relevant in the scenario of dynamic TDD.
- *Other optimization objectives for the dynamic TDD solution*: we believe that the strategies for the decentralization of the solution proposed here can be extended to other network beamforming objectives such as sum-rate maximization and energy efficiency and different constraints such as rate.
- *ML based optimization in the context of Dynamic time division duplexing (TDD):* As a mix of both chapters, we consider to be of interest the application of similar

machine learning (ML) strategies to the scenario in which multiple interference links are generated by the MIMO Dynamic TDD scenario.

- *Fully distributed ML design:* In the solution proposed here we need coordination among users as means of cooperation in the direction of a common goal, which leads to a signaling overhead in the training stage. A completely distributed ML strategy, with good global performance, is a very desirable technology.

- *Multi-armed bandits ML design:* In the solution proposed here we employ a multi-agent DQL strategy. Another promising ML approach to deal with decisions is known as multi-armed bandits (127). The study of the application of such strategy to our problem is of high interest.

- *ML based joint optimization of power and user scheduling:* In the solution proposed here the power allocation is performed by the well-known water-filling algorithm, we believe that gains can be achieved if both user scheduling and power allocation are jointly optimized.

- *ML based optimization of the beamforming vectors*: We envision an extension in which the learning scheme is able to optimize the beamforming vector itself, not only the scheduling. However, there are concerns about this, regarding complexity and latency constraints,

As previously discussed, this thesis is inserted in the context of the always-evolving wireless networks. That means that the requirements of higher data rates, lower latency, and more connected users will continue to be the drivers of evolution. Based on that, new scenarios arise and we envision cell-free networks as possible grounds for the evolution of decentralized optimization solutions related to/evolved from the ones presented here. Cell-free Massive MIMO is considered as a promising technology for satisfying the increasing number of users and high rate expectations in beyond-5G networks (128). The main principle is that many distributed access points cooperate to simultaneously serve all the users within the network without creating cell boundaries (129). This way, due to the decentralized nature of such scenarios, there is a need to find practical architectures and ways to decentralize the processing.

REFERENCES

1 AGIWAL, M.; ROY, A.; SAXENA, N. Next generation 5G wireless Networks: a comprehensive survey. **IEEE Communications Surveys and Tutorials**, [*s. l.*], v. 18, n. 3, p. 1617–1655, Feb. 2016.

2 DOGRA, A.; JHA, R. K.; JAIN, S. A survey on beyond 5G network with the advent of 6G: architecture and emerging technologies. **IEEE Access**, [*s. l.*], v. 9, p. 67512–67547, 2020.

3 3GPP. **Release 15**: TR 21.915. [*S. l.*], 2018. Available from: `https://www.3gpp.org/specifications-technologies/releases/release-15`. Visited on: 12 Dec. 2022.

4 WANG, C.-X. *et al.* Cellular architecture and key technologies for 5G wireless communication networks. **IEEE communications magazine**, [*s. l.*], v. 52, n. 2, p. 122–130, 2014.

5 ITU. **ITU-R M.2410-0**: minimum requirements related to technical performance for IMT-2020 radio interface(s). [*S. l.*], 2017. Available from: `https://www.itu.int/pub/R-REP-M.2410-2017`. Visited on: 12 Dec. 2022.

6 TULLBERG, H. *et al.* The METIS 5G system concept: meeting the 5G requirements. **IEEE Communications magazine**, [*s. l.*], v. 54, n. 12, p. 132–139, 2016.

7 NAVARRO-ORTIZ, J. *et al.* A survey on 5G usage scenarios and traffic models. **IEEE Communications Surveys & Tutorials**, [*s. l.*], v. 22, n. 2, p. 905–929, 2020.

8 CHEN, H. *et al.* Ultra-reliable low latency cellular networks: use cases, challenges and approaches. **IEEE Communications Magazine**, [*s. l.*], v. 56, n. 12, p. 119–125, 2018.

9 SHARMA, S. K.; WANG, X. Toward massive machine type communications in ultra-dense cellular IoT networks: current issues and machine learning-assisted solutions. **IEEE Communications Surveys & Tutorials**, [*s. l.*], v. 22, n. 1, p. 426–471, 2019.

10 3GPP. **Study on scenarios and requirements for next generation access technologies**. [*S. l.*], 2017. Available from: `https://www.etsi.org/deliver/etsi_tr/138900_138999/138913/14.03.00_60/tr_138913v140300p.pdf`. Visited on: 12 Dec. 2022.

11 BUSARI, S. A. *et al.* Millimeter-wave massive MIMO communication for future wireless systems: a survey. **IEEE Communications Surveys & Tutorials**, [*s. l.*], v. 20, n. 2, p. 836–869, 2017.

12 ERICSSON. **Ericsson mobility report**. [S.l.]: Ericsson, 2022. Available from: `https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2022`. Visited on: 12 Dec. 2022.

13 LIN, X. An overview of 5G advanced evolution in 3GPP release 18. **IEEE Communications Standards Magazine**, [*s. l.*], v. 6, n. 3, p. 77–83, 2022.

14 GIORDANI, M. *et al.* Toward 6G networks: use cases and technologies. **IEEE Communications Magazine**, [*s. l.*], v. 58, n. 3, p. 55–61, 2020.

15 SAAD, W.; BENNIS, M.; CHEN, M. A vision of 6G wireless systems: applications, trends, technologies, and open research problems. **IEEE network**, [*s. l.*], v. 34, n. 3, p. 134–142, 2019.

16 YANG, P. *et al.* 6G wireless communications: vision and potential techniques. **IEEE network**, [*s. l.*], v. 33, n. 4, p. 70–75, 2019.

17 JIANG, W. *et al.* The road towards 6G: a comprehensive survey. **IEEE Open Journal of the Communications Society**, [*s. l.*], v. 2, p. 334–366, 2021.

18 BHUSHAN, N. *et al.* Network densification: the dominant theme for wireless evolution into 5G. **IEEE Communications Magazine**, IEEE, [*s. l.*], v. 52, n. 2, p. 82–89, Feb. 2014.

19 INTERDONATO, G. *et al.* Ubiquitous cell-free massive MIMO communications. **EURASIP Journal on Wireless Communications and Networking**, Springer, [*s. l.*], v. 2019, n. 1, p. 1–13, 2019.

20 TELATAR, E. Capacity of multi-antenna Gaussian channels. **European transactions on telecommunications**, Wiley Online Library, [*s. l.*], v. 10, n. 6, p. 585–595, 1999.

21 VAN VEEN, B. D.; BUCKLEY, K. M. Beamforming: A versatile approach to spatial filtering. **IEEE assp magazine**, [*s. l.*], v. 5, n. 2, p. 4–24, 1988.

22 GERSHMAN, A. B. *et al.* Convex optimization-based beamforming. **IEEE Signal Processing Magazine**, [*s. l.*], v. 27, n. 3, p. 62–75, 2010.

23 DAHLMAN, E.; PARKVALL, S.; SKOLD, J. **5G NR: the next generation wireless access technology**. [*S. l.*]: Academic Press, 2020.

24 SHEIKH, T. A.; BORA, J.; HUSSAIN, A. A survey of antenna and user scheduling techniques for massive MIMO-5G wireless system. *In:* INTERNATIONAL CONFERENCE ON CURRENT TRENDS IN COMPUTER, ELECTRICAL, ELECTRONICS AND COMMUNICATION. **Proceedings [...]**. [*S. l.*]: IEEE, 2017. p. 578–583.

25 JIANG, C. *et al.* Machine learning paradigms for next-generation wireless networks. **IEEE Wireless Communications**, [*s. l.*], v. 24, n. 2, p. 98–105, 2016.

26 MOROCHO-CAYAMCELA, M. E.; LEE, H.; LIM, W. Machine learning for 5G/B5G mobile and wireless communications: potential, limitations, and future directions. **IEEE access**, [*s. l.*], v. 7, p. 137184–137206, 2019.

27 CHOWDHURY, M. Z. *et al.* 6G wireless communication systems: applications, requirements, technologies, challenges, and research directions. **IEEE Open Journal of the Communications Society**, [*s. l.*], v. 1, p. 957–975, 2020.

28 BJÖRNSON, E.; BENGTSSON, M.; OTTERSTEN, B. Optimal multiuser transmit beamforming: a difficult problem with a simple solution structure. **IEEE Signal Processing Magazine**, [*s. l.*], v. 31, n. 4, p. 142–148, 2014.

29 LIU, Y.-F.; DAI, Y.-H.; LUO, Z.-Q. Coordinated beamforming for MISO interference channel: complexity analysis and efficient algorithms. **IEEE Transactions on Signal Processing**, [*s. l.*], v. 59, n. 3, p. 1142–1157, 2010.

30 LO, T. K. Maximum ratio transmission. *In:* 1999 IEEE international conference on communications (Cat. No. 99CH36311). [*S. l.*]: IEEE, 1999. v. 2, p. 1310–1314.

31 WIESEL, A.; ELDAR, Y. C.; SHAMAI, S. Zero-forcing precoding and generalized inverses. **IEEE Trans. Signal Process.**, [*s. l.*], v. 56, n. 9, p. 4409–4418, Sept. 2008.

32 VOJCIC, B. R.; JANG, W. M. Transmitter precoding in synchronous multiuser communications. **IEEE transactions on communications**, [*s. l.*], v. 46, n. 10, p. 1346–1355, 1998.

33 JOHAM, M.; UTSCHICK, W.; NOSSEK, J. A. Linear transmit processing in MIMO communications systems. **IEEE Transactions on signal Processing**, [*s. l.*], v. 53, n. 8, p. 2700–2712, 2005.

34 BJÖRNSON, E.; JORSWIECK, E. Optimal resource allocation in coordinated multi-Cell systems. **Foundations and Trends® in Communications and Information Theory**, [*s. l.*], v. 9, 2–3, p. 113–381, 2013.

35 RASHID-FARROKHI, F.; LIU, K. J. R.; TASSIULAS, L. Transmit beamforming and power control for cellular wireless systems. **IEEE Journal on Selected Areas in Communications**, [*s. l.*], v. 16, n. 8, p. 1437–1450, Oct. 1998.

36 YU, W.; LAN, T. Transmitter optimization for the multi-antenna downlink with per-antenna power constraints. **IEEE Transactions on signal processing**, [*s. l.*], v. 55, n. 6, p. 2646–2660, 2007.

37 DAHROUJ, H.; YU, W. Coordinated beamforming for the multicell multi-antenna wireless system. **IEEE transactions on wireless communications**, [*s. l.*], v. 9, n. 5, p. 1748–1759, 2010.

38 BENGTSSON, M.; OTTERSTEN, B. Optimal downlink beamforming using semidefinite optimization. *In:* ANNU. ALLERTON CONF. COMMUNICATION, CONTROL, AND COMPUTING, 37. **Proceedings [...]**. [*S. l.: s. n.*], 1999. p. 987–996.

39 WIESEL, A.; ELDAR, Y. C.; SHAMAI, S. Linear precoding via conic optimization for fixed MIMO receivers. **IEEE Trans. Signal Process.**, [*s. l.*], v. 54, n. 1, p. 161–176, Jan. 2006.

40 GRANT, M.; BOYD, S.; YE, Y. **CVX**: MATLAB software for disciplined convex programming. [S.l.], 2009. Available from: `http://web.cvxr.com/cvx/doc/`. Visited on: 26 Aug. 2023.

41 CHANG, J.-H.; TASSIULAS, L.; RASHID-FARROKHI, F. Joint transmitter receiver diversity for efficient space division multiaccess. **IEEE transactions on wireless communications**, [*s. l.*], v. 1, n. 1, p. 16–27, 2002.

42 BENGTSSON, M. A pragmatic approach to multi-user spatial multiplexing. *In:* SENSOR ARRAY AND MULTICHANNEL SIGNAL PROCESSING WORKSHOP. **Proceedings [...]**. [S.l.]: IEEE, 2002. p. 130–134.

43 PENNANEN, H.; TÖLLI, A.; LATVA-AHO, M. Pragmatic multi-cell MIMO beamforming with decentralized coordination. *In:* ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS. **Proceedings [...]**. [*S. l.*]: IEEE, 2012. p. 1996–2000.

44 ARDAH, K.; SILVA, Y.; CAVALCANTI, F. Decentralized linear transceiver design in multicell MIMO broadcast channels. **Journal of Communication and Information Systems**, [*s. l.*], v. 32, n. 1, 2017. Available from: `https://jcis.sbrt.org.br/jcis/article/download/461/342`. Visited on: 26 Aug. 2023.

45 JOSHI, S. K. *et al.* Weighted sum-rate maximization for MISO downlink cellular networks via branch and bound. **IEEE Transactions on Signal Processing**, [*s. l.*], v. 60, n. 4, p. 2090–2095, Dec. 2011.

46 CHRISTENSEN, S. S. *et al.* Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design. **IEEE Transactions on Wireless Communications**, [*s. l.*], v. 7, n. 12, p. 4792–4799, Dec. 2008.

47 SHI, Q. *et al.* An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel. **IEEE Transactions on Signal Processing**, [*s. l.*], v. 59, n. 9, p. 4331–4340, Sept. 2011.

48 SHEN, K.; YU, W. Fractional programming for communication systems—part I: power control and beamforming. **IEEE Transactions on Signal Processing**, [*s. l.*], v. 66, n. 10, p. 2616–2630, May 2018.

49 DOTTLING, M. *et al.* Integration of spatial processing in the WINNER B3G air interface design. *In:* VEHICULAR TECHNOLOGY CONFERENCE, 63. **Proceedings [...]**. [*S. l.*]: IEEE, 2006. v. 1, p. 246–250.

50 ANDREWS, M.; DINITZ, M. Maximizing capacity in arbitrary wireless networks in the SINR model: complexity and game theory. *In:* IEEE INFOCOM. **Proceedings [...]**. [*S. l.*]: IEEE, 2009. p. 1332–1340.

51 CAPONE, A. *et al.* A new computational approach for maximum link activation in wireless networks under the SINR model. **IEEE Transactions on Wireless Communications**, [*s. l.*], v. 10, n. 5, p. 1368–1372, May 2011.

52 YUAN, D. *et al.* On optimal link activation with interference cancelation in wireless networking. **IEEE Transactions on Vehicular Technology**, [*s. l.*], v. 62, n. 2, p. 939–945, Feb. 2013.

53 HE, Q.; YUAN, D.; EPHREMIDES, A. Maximum Link Activation with Cooperative Transmission and Interference Cancellation in Wireless Networks. **IEEE Transactions on Mobile Computing**, [*s. l.*], v. 16, n. 2, p. 408–421, Feb. 2017.

54 KAM, C. *et al.* On the age of information with packet deadlines. **IEEE Transactions on Information Theory**, [*s. l.*], v. 64, n. 9, p. 6419–6428, Sept. 2018.

55 HE, Q.; YUAN, D.; EPHREMIDES, A. Optimal link scheduling for age minimization in wireless systems. **IEEE Transactions on Information Theory**, IEEE, [*s. l.*], v. 64, n. 7, p. 5381–5394, July 2018.

56 WEERADDANA, P. C. *et al.* Multicell MISO downlink weighted sum-rate maximization: a distributed approach. **IEEE Transactions on Signal Processing**, [*s. l.*], v. 61, n. 3, p. 556–570, Feb. 2013.

57 CASTAÑEDA, E. *et al.* An overview on resource allocation techniques for multi-user MIMO systems. **IEEE Communications Surveys and Tutorials**, [*s. l.*], v. 19, n. 1, p. 239–284, Firstquarter 2017.

58 ANTONIOLI, R. P. *et al.* User scheduling for sum-rate maximization under minimum rate constraints for the MIMO IBC. **IEEE Wireless Communications Letters**, [*s. l.*], v. 8, n. 6, p. 1591–1595, 2019.

59 AHN, M. *et al.* **SLNR-based user scheduling for MISO downlink cellular systems**. *In:* VEHICULAR TECHNOLOGY CONFERENCE. **Proceedings [...]**. [*S. l.*]: IEEE, 2013. p. 1–5.

60 ANTONIOLI, R. P. *et al.* Decentralized user scheduling for rate-constrained sum-utility maximization in the MIMO IBC. **IEEE Transactions on Communications**, [*s. l.*], v. 68, n. 10, p. 6215–6229, 2020.

61 STRIDH, R.; BENGTSSON, M.; OTTERSTEN, B. System evaluation of optimal downlink beamforming with congestion control in wireless communication. **IEEE Transactions on Wireless Communications**, [*s. l.*], v. 5, n. 4, p. 743–751, Apr. 2006.

62 YU, L.; KARIPIDIS, E.; LARSSON, E. G. Coordinated scheduling and beamforming for multicell spectrum sharing networks using branch and bound. *In:* EUROPEAN SIGNAL PROCESSING CONFERENCE, 20. **Proceedings [...]**. [*S. l.*]: IEEE, 2012. p. 819–823.

63 HE, Y.; ZHANG, Z., *et al.* Deep-reinforcement-learning-based optimization for cache-enabled opportunistic interference alignment wireless networks. **IEEE Transactions on Vehicular Technology**, [*s. l.*], v. 66, n. 11, p. 10433–10445, Nov. 2017.

64 YANG, T. *et al.* A survey of distributed optimization. **Annual Reviews in Control**, Elsevier, [*s. l.*], v. 47, p. 278–305, 2019.

65 REN, W.; BEARD, R. W.; ATKINS, E. M. A survey of consensus problems in multi-agent coordination. *In:* AMERICAN CONTROL CONFERENCE. **Proceedings [...]**. [*S. l.*]: IEEE, 2005. p. 1859–1864.

66 LI, T.-F.; LI, H. A distributed optimization algorithm for multi-agent systems with limited communication. *In:* CHINESE CONTROL AND DECISION CONFERENCE. **Proceedings [...]**. [*S. l.*]: IEEE, 2020. p. 622–625.

67 LUO, Z.-Q.; YU, W. An introduction to convex optimization for communications and signal processing. **IEEE Journal on selected areas in communications**, [*s. l.*], v. 24, n. 8, p. 1426–1438, 2006.

68 STURM, J. F. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. **Optimization methods and software**, Taylor & Francis, [*s. l.*], v. 11, n. 1-4, p. 625–653, 1999.

69 BOYD, S.; VANDENBERGHE, L. **Convex optimization**. [S.l.]: Cambridge university press, 2004.

70 LOBO, M. S. *et al.* Applications of second-order cone programming. **Linear algebra and its applications**, Elsevier, [*s. l.*], v. 284, n. 1-3, p. 193–228, 1998.

71 VANDENBERGHE, L.; BOYD, S. Semidefinite programming. **SIAM review**, SIAM, [*s. l.*], v. 38, n. 1, p. 49–95, 1996.

72 BOYD, S. *et al.* Distributed optimization and statistical learning via the alternating direction method of multipliers. **Found. Trends Mach. Learn.**, Now Publishers Inc., [*s. l.*], v. 3, n. 1, p. 1–122, Jan. 2011.

73 LI, R. *et al.* Intelligent 5G: when cellular networks meet artificial intelligence. **IEEE Wireless communications**, [*s. l.*], v. 24, n. 5, p. 175–183, 2017.

74 WANG, C.-X. *et al.* Artificial intelligence enabled wireless networking for 5G and beyond: recent advances and future challenges. **IEEE Wireless Communications**, [*s. l.*], v. 27, n. 1, p. 16–23, 2020.

75 LUONG, N. C. *et al.* Applications of deep reinforcement learning in communications and networking: a survey. **IEEE Communications Surveys & Tutorials**, [*s. l.*], v. 21, n. 4, p. 3133–3174, 2019.

76 SUTTON, R. S.; BARTO, A. G. **Reinforcement learning: an introduction**. 2. ed. [*S. l.*]: MIT press, 2018.

77 MNIH, V.; KAVUKCUOGLU, K., *et al.* Human-level control through deep reinforcement learning. **Nature**, [*s. l.*], v. 518, n. 7540, p. 529, Feb. 2015.

78 ACCELLERAN. **The essential importance of LTE TDD for small cell deployments**. [*S. l.*], 2013.

79 PAULI, V.; LI, Y.; SEIDEL, E. Dynamic TDD for LTE-A and 5G. **Nomor Research GmbH, Tech. Rep**, [*s. l.*], 2015. Available from: `https://www.slideshare.net/eikoseidel/white-paper-dynamic-tdd-for-ltea-eimta-and-5g`. Visited on: 26 Aug. 2023.

80 SHEN, Z. *et al.* Dynamic uplink-downlink configuration and interference management in TD-LTE. **IEEE Commun. Mag.**, [*s. l.*], v. 50, n. 11, p. 51–59, Nov. 2012.

81 JAYASINGHE, P. *et al.* Bi-directional signaling for dynamic TDD with decentralized beamforming. *In:* INT. CONF. COMMUN. WORKSHOP. **Proceedings [...]**. [*S. l.*]: IEEE, 2015. p. 185–190.

82 JAYASINGHE, P. *et al.* Direct beamformer estimation for dynamic TDD networks with forward-backward training. *In:* INT. WORKSHOP SIGNAL PROCESSING ADVANCES IN WIRELESS COMMUNICATIONS. **Proceedings [...]**. [*S. l.*]: IEEE, 2017. p. 1–6.

83 3GPP. **TS 36.300**: evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN) overall description. [*S. l.: s. n.*], 2015.

84 ELBAMBY, M. S.; BENNIS, M.; LATVA-AHO, M. UL/DL decoupled user association in dynamic TDD small cell networks. *In:* INT. SYMP. WIRELESS COMMUNICATION SYSTEMS. **Proceedings [...]**. [*S. l.*]: IEEE, 2015. p. 456–460.

85 ARDAH, K. *et al.* A novel cell reconfiguration technique for dynamic TDD wireless networks. **IEEE Wireless Commun. Lett.**, [*s. l.*], v. 7, n. 3, p. 320–323, June 2018.

86 INTEL. **Performance analysis of DL-UL interference management and traffic adaptation in multi-cell Pico-Pico deployment scenario**. Jeju, Korea, 2012.

87 CHOI, Y. S.; SOHN, I.; LEE, K. B. A novel decentralized time slot allocation algorithm in dynamic TDD system. *In:* CONSUMER COMMUN, AND NETW. CONF. **Proceedings [...]**. [*S. l.*]: IEEE, 2006. v. 2, p. 1268–1272.

88 SOHN, I.; LEE, K. B.; CHOI, Y. S. Comparison of decentralized time slot allocation strategies for asymmetric traffic in TDD systems. **IEEE Trans. Wireless Commun.**, [*s. l.*], v. 8, n. 6, p. 2990–3003, June 2009.

89 3GPP. **R1-132350**: DL power control based interference mitigation for eIMTA. Fukuoka, Japan, 2013.

90 3GPP. **R1-132351**: UL power control based interference mitigation for eIMTA. Fukuoka, Japan, May 2013.

91 ALEXANDROPOULOS, G. C.; PAPADIAS, C. B. A reconfigurable distributed algorithm for K-user MIMO interference networks. *In:* INT. CONF. COMMUNICATIONS. **Proceedings [...]**. [*S. l.*]: IEEE, 2013. p. 3063–3067.

92 PENNANEN, H.; TOLLI, A.; LATVA-AHO, M. Decentralized coordinated downlink beamforming via primal decomposition. **IEEE Signal Process. Lett.**, [*s. l.*], v. 18, n. 11, p. 647–650, Nov. 2011.

93 TOLLI, A.; PENNANEN, H.; KOMULAINEN, P. Decentralized minimum power multi-cell beamforming with limited backhaul signaling. **IEEE Trans. Wireless Commun.**, [*s. l.*], v. 10, n. 2, p. 570–580, Feb. 2011.

94 WONG, K.-K.; ZHENG, G.; NG, T.-S. Convergence analysis of downlink MIMO antenna systems using second-order cone programming. *In:* VEHICULAR TECHNOLOGY CONFERENCE. **Proceedings [...]**. [*S. l.*]: IEEE, 2005. v. 1, p. 492–496.

95 ALEXANDROPOULOS, G. C.; FERRAND, P.; PAPADIAS, C. B. On the robustness of coordinated beamforming to uncoordinated interference and CSI uncertainty. *In:* WIRELESS COMMUNICATIONS AND NETWORKING CONF. **Proceedings [...]**. [*S. l.*]: IEEE, 2017. p. 1–6.

96 ALEXANDROPOULOS, G. C. *et al.* Advanced coordinated beamforming for the downlink of future LTE cellular networks. **IEEE Commun. Mag.**, [*s. l.*], v. 54, n. 7, p. 54–60, July 2016.

97 HUANG, H. *et al.* Unsupervised Learning-Based Fast Beamforming Design for Downlink MIMO. **IEEE Access**, [*s. l.*], v. 7, p. 7599–7605, Dec. 2019.

98 XIA, W. *et al.* A deep learning framework for optimization of MISO downlink beamforming. **IEEE Transactions on Communications**, [*s. l.*], v. 68, n. 3, p. 1866–1880, 2020.

99 ZHOU, H. *et al.* Continual learning-based fast beamforming adaptation in downlink MISO systems. **IEEE Wireless Communications Letters**, [*s. l.*], v. 12, n. 1, p. 36–39, 2023.

100 YOON, C.; CHO, D. H. Energy Efficient Beamforming and Power Allocation in Dynamic TDD Based C-RAN System. **IEEE Commun. Lett.**, [*s. l.*], v. 19, n. 10, p. 1806–1809, Oct. 2015.

101 GUIMARÃES, F. R. V. *et al.* Pricing-based distributed beamforming for dynamic time division duplexing systems. **IEEE Trans. on Veh. Technol.**, [*s. l.*], v. 67, n. 4, p. 3145–3157, Apr. 2018.

102 ARDAH, K. *et al.* An ADMM approach to distributed coordinated beamforming in dynamic TDD networks. *In:* INT. WORKSHOP COMPUT. ADVANCES IN MULTI-SENSOR ADAPTIVE PROCESS. **Proceedings [...]**. [*S. l.*]: IEEE, 2017. p. 705–709.

103 O. CAVALCANTE, E. de *et al.* Distributed beamforming in dynamic TDD MIMO networks with BS to BS interference constraints. **IEEE Wireless Commun. Lett.**, [*s. l.*], v. 7, n. 5, p. 788–791, Oct. 2018.

104 CIRIK, A. C. *et al.* Linear transceiver design for full-duplex multi-cell MIMO systems. **IEEE Access**, [*s. l.*], v. 4, p. 4678–4689, 2016.

105 TAGHIZADEH, O. *et al.* Hardware impairments aware transceiver design for bidirectional full-duplex MIMO OFDM systems. **IEEE Trans. Veh. Technol.**, [*s. l.*], v. 67, n. 8, p. 7450–7464, Aug. 2018.

106 CIRIK, A. C. *et al.* Sum-power minimization under rate constraints in full-duplex MIMO interference-channels. *In:* INT. ITG WORKSHOP SMART ANTENNAS. **Proceedings [...]**. [*S. l.*]: IEEE, 2017. p. 1–5.

107 SHEN, C. *et al.* Distributed robust multicell coordinated beamforming with imperfect CSI: An ADMM approach. **IEEE Trans. Signal Process.**, [*s. l.*], v. 60, n. 6, p. 2988–3003, June 2012.

108 SHI, Q. *et al.* SINR Constrained Beamforming for a MIMO Multi-User Downlink System: Algorithms and Convergence Analysis. **IEEE Trans. Signal Process.**, [*s. l.*], v. 64, n. 11, p. 2920–2933, June 2016.

109 BOYD, S. *et al.* Notes on decomposition methods. **EE364B, Stanford University**, [*s. l.*], p. 1–36, 2007.

110 3GPP. **TS 36.133 v10.3.3,0**: evolved universal terrestrial radio access (EUTRA): requirements for support of radio resource management. [*S. l.*], 2019.

111 HUANG, X. *et al.* A multi-gigabit microwave backhaul. **IEEE Commun. Mag.**, [*s. l.*], v. 50, n. 3, p. 122–129, Mar. 2012.

112 DAT, P. T. *et al.* High-capacity wireless backhaul network using seamless convergence of radio-over-fiber and 90-GHz millimeter-wave. **J. Lightw. Technol.**, [*s. l.*], v. 32, n. 20, p. 3910–3923, Oct. 2014.

113 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. [*S. l.*], 2012. (Release 11, TR 36.828).

114 LUO, Z.; ZHANG, S. Dynamic spectrum management: complexity and duality. **IEEE Journal of Selected Topics in Signal Processing**, [*s. l.*], v. 2, n. 1, p. 57–73, Feb. 2008.

115 NASIR, Y. S.; GUO, D. Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks. **IEEE Journal on Selected Areas in Communications**, [*s. l.*], v. 37, n. 10, p. 2239–2250, Oct. 2019.

116 LEE, H.; KO, Y.; YANG, H. On robust weighted sum rate maximization for mimo interfering broadcast channels with imperfect channel knowledge. **IEEE Communications Letters**, [*s. l.*], v. 17, n. 6, p. 1156–1159, June 2013.

117 RAZAVIYAYN, M.; BOROUJENI, M. S.; LUO, Z. A stochastic weighted MMSE approach to sum rate maximization for a MIMO interference channel. *In:* INT. WORKSHOP ON SIGNAL PROCESS. ADVANCES IN WIRELESS COMMUN. **Proceedings [...]**. [*S. l.*]: IEEE, 2013. p. 325–329.

118 MEDRA, M. *et al.* Low-complexity robust MISO downlink precoder design under imperfect CSI. **IEEE Transactions on Signal Processing**, [*s. l.*], v. 64, n. 12, p. 3237–3249, June 2016.

119 RUSEK, F.; PERSSON, D., *et al.* Scaling up MIMO: opportunities and challenges with very large arrays. **IEEE Signal Processing Magazine**, [*s. l.*], v. 30, n. 1, p. 40–60, Jan. 2013.

120 SINGH, S. *et al.* Convergence results for single-step on-policy reinforcement-learning algorithms. **Machine learning**, [*s. l.*], v. 38, n. 3, p. 287–308, Mar. 2000.

121 LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, [*s. l.*], v. 521, n. 7553, p. 436, 2015.

122 SARAIVA, J. V.; BRAGA JR., I. M., *et al.* Deep reinforcement learning for QoS-constrained resource allocation in multiservice networks. **Journal of Commun. and Information Syst.**, [*s. l.*], v. 35, n. 1, p. 66–76, Apr. 2020.

123 3GPP. **TS 36.211 v15.7.0**: evolved universal terrestrial radio access (E-UTRA); physical channels and modulation. [*S. l.: s. n.*], 2019.

124 3GPP. **TS 36.931 v15.0.0**: evolved universal terrestrial radio access (E-UTRA) RF requirements for LTE pico Node B. [*S. l.: s. n.*], 2018.

125 ABADI, M.; BARHAM, P., *et al.* TensorFlow: a system for large-scale machine learning. *In:* SYMPOSIUM ON OPERATING SYSTEMS DESIGN AND IMPLEMENTATION. **Proceedings [...]**. [*S. l.: s. n.*], Nov. 2016. p. 265–283.

126 KINGMA, D. P.; BA, J. Adam: a method for stochastic optimization. *In:* INT. CONF. ON LEARNING REPRESENTATIONS. **Proceedings [...]**. [*S. l.: s. n.*], 2015.

127 SLIVKINS, A. *et al.* Introduction to multi-armed bandits. **Foundations and Trends® in Machine Learning**, Now Publishers, Inc., [*s. l.*], v. 12, n. 1-2, p. 1–286, 2019.

128 BJÖRNSON, E.; SANGUINETTI, L. Making cell-free massive MIMO competitive with MMSE processing and centralized implementation. **IEEE Transactions on Wireless Communications**, [*s. l.*], v. 19, n. 1, p. 77–90, 2019.

129 SHAIK, Z. H.; BJÖRNSON, E.; LARSSON, E. G. Cell-free massive MIMO with radio stripes and sequential uplink processing. *In:* INTERNATIONAL CONFERENCE ON COMMUNICATIONS WORKSHOPS. **Proceedings [...]**. [*S. l.*]: IEEE, 2020. p. 1–6.

APPENDIX A – MAIN RELATED WORKS OF CHAPTER 3

Table 10 – Comparison of the most relevant references of Chapter 3.

Reference	Objective	Constraint	Architecture	Decentralization	CSI availability	Solution scheme	Merits of solution
Chapter 3	Sum-power min.	SINR for UL and DL	Multi-cell multi-stream MIMO dynamic TDD	Centr. and decentr.	Global for centr. local for decentr.	MMSE-SDP for centr. ADMM for decentr.	Converges to mininimum network power
(103)	DL sum-power min.	DL SINR BS-BS interf.	Multi-cell MIMO dynamic TDD	Centr. and decentr.	Global for centr. local for decentr.	MMSE-SOCP for centr. ADMM for decentr.	UL not optimized
(102)	DL sum-power min.	DL SINR BS-BS interf.	Multi-cell MISO dynamic TDD	Centr. and decentr.	Global for centr. local for decentr.	MMSE-SDP for centr. ADMM for decentr.	UL not optimized
(101)	DL rate max. penalty on interf.	DL sum-power	Multi-cell MISO dynamic TDD	Decentr.	local	Interference pricing	DL avoids interf. to UL
(107)	Weighted sum-power min.	SINR	DL Multi-cell MISO	Decentr.	Imperfect local	ADMM	Different levels of coordination
(43)	Sum-power min.	SINR	DL Multi-cell MISO	Decentr.	Local	Primal decomposition	Performance similar to centr.
(94)	Sum-power min.	SINR	DL Multi-user MIMO	Decentr.	Global	MMSE-SOCP	Convergence is proved
(106)	Sum-power min.	Rate	Full-duplex K-link MIMO	Centr.	Global	Penalty method	Space-time scheduling
(91)	Sum-rate max.	Power	K-link MIMO	Decentr.	Local	Iterative precoding	Combines MMSE and waterfilling

Source: Created by the author.

APPENDIX B – MAIN RELATED WORKS OF CHAPTER 4

Table 11 – Comparison of the most relevant references of Chapter 4.

Reference	Objective	Constraint	Architecture	Decentralization	CSI availability	Solution scheme	Merits of solution
Chapter 4	Sum-rate max.	Power	DL multi-cell MISO	Decentr.	Imperfect local	Multi-agent DQL user scheduling	Beamforming based on a opt. structure
(115)	Weighted sum-rate max.	Power	K-link single-antenna	Decentr.	Local	Multi-agent DQL power allocation	Considers CSI delays
(47)	Weighted sum-rate max.	Power	downlink (DL) Multi-cell MIMO	Decentr.	Local	Iterative min. of weighted MSE	Considers CSI delays
(34)	SLNR max.	Power	DL Multi-cell MISO	Decentr.	Local	Structured on uplink/downlink duality	Heuristically max. SNR min. interf. power
(59)	Weighted sum-rate max.	Power	DL Multi-cell MISO	Decentr.	Local	SLNR based user scheduling	Low complexity algorithm
(60)	Sum-rate max. penalized dissatisfaction	Rate	DL Multi-cell MIMO	Centr. and Decentr.	Global for centr. local for decentr.	Joint user scheduling and transceivers	Practical design low overhead
(63)	Interference Alignment	–	K-link MIMO	Centr.	Global	DQL user scheduling	Cache-enabled IA
(116)	Weighted sum-rate max.	Power	DL Multi-cell MIMO	Centr.	Imperfect Global	Robust weighted MMSE	Resilient to imperfect CSI
(98)	Three BF objectives	–	DL single-cell MISO	Centr.	Global	Supervised Unsupervised	Neural networks expert knowledge
(97)	Weighted sum-rate max	Power	DL single-cell MIMO	Centr.	Global	Unsupervised learning	Perform close to weighted MMSE

Source: Created by the author.

APPENDIX C – PROOF OF PROPOSITION 2

The related sum-power optimization problem can be conveniently expressed as an optimization task as

$$\min_{\substack{\{\mathbf{m}_{k_b}\}\forall k_b\in\mathcal{K}\\ \{P_b\}\forall b\in\mathcal{B}}} \quad \sum_{b=1}^{B} P_b, \tag{C.1a}$$

$$\text{s. t.} \sum_{k=1}^{K_b} ||\mathbf{m}_{k_b}||^2 \leq P_b, \quad \forall b \in \mathcal{B}, \tag{C.1b}$$

$$\Gamma_{k_b} \geq \Gamma_{k_b}^{\min}, \quad \forall b \in \mathcal{B}, \forall k \in \mathcal{K}_b \tag{C.1c}$$

where P_b is a variable that reflects the total power of BS b and $\Gamma_{k_b}^{\min}$ is the minimum SINR threshold of UE k of BS b.

In sum-rate maximization problems, the maximum power of each BS, P_b, are given. Thus, there is a direct relation between problem (4.3) and the feasibility problem (C.2):

$$\text{find} \quad \{\mathbf{m}_{k_b}\}_{\forall k_b\in\mathcal{K}} \tag{C.2a}$$

$$\text{s. t.} \sum_{k=1}^{K_b} ||\mathbf{m}_{k_b}||^2 \leq P_b, \quad \forall b \in \mathcal{B}, \tag{C.2b}$$

$$\Gamma_{k_b} \geq \Gamma_{k_b}^{\min}, \quad \forall b \in \mathcal{B}, \forall k \in \mathcal{K}_b \tag{C.2c}$$

The Lagrangian for problem (C.2) is

$$\begin{aligned}\mathcal{L}(\{\mathbf{m}_{k_b},\lambda_{k_b}\}_{\forall k_b\in\mathcal{K}},\{\mu_b\}_{\forall b\in\mathcal{B}}) &= \sum_{b=1}^{B}\mu_b\left(\frac{1}{P_b}\sum_{k=1}^{K_b}||\mathbf{m}_{k_b}||^2-1\right)\\ +\sum_{b=1}^{B}\sum_{k=1}^{K_b}\frac{\lambda_{k_b}}{\sigma^2}&\left(\sum_{\substack{i=1\\ i\neq k}}^{K_b}|\mathbf{h}_{k_b b}^H\mathbf{m}_{i_b}|^2+\sum_{\substack{j=1\\ j\neq b}}^{B}\sum_{i=1}^{K_j}|\mathbf{h}_{k_b j}^H\mathbf{m}_{i_j}|^2+\sigma^2-\frac{1}{\Gamma_{k_b}^{\min}}|\mathbf{h}_{k_b b}^H\mathbf{m}_{k_b}|^2\right),\end{aligned} \tag{C.3}$$

where, μ_b and λ_{k_b} are Lagrange multipliers with respect to the power constraint of BS b and to the SINR constraint of UE k_b, respectively.

This way, the dual problem of (C.2) can be written as

$$\max_{\substack{\{\mathbf{m}_{k_b},\lambda_{k_b}\}_{\forall k_b\in\mathcal{K}}\\ \{\mu_b\}_{\forall b\in\mathcal{B}}}} \sum_{b=1}^{B}\sum_{k=1}^{K_b}\lambda_{k_b}-\sum_{b=1}^{B}\mu_b, \tag{C.4a}$$

$$\text{s. t.} \quad \frac{\mu_b}{P_b}\mathbf{I}+\sum_{\substack{i=1\\ i\neq k}}^{K_b}\frac{\lambda_{i_b}}{\sigma^2}\mathbf{h}_{i_b b}\mathbf{h}_{i_b b}^H+\sum_{\substack{j=1\\ j\neq b}}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_j b}\mathbf{h}_{i_j b}^H-$$

$$\frac{\lambda_{k_b}}{\sigma^2\Gamma_{k_b}^{\min}}\mathbf{h}_{k_b b}\mathbf{h}_{k_b b}^H\geq 0, \quad \forall b\in\mathcal{B},\forall k\in\mathcal{K}_b, \tag{C.4b}$$

$$\mu_b\geq 0,\ \lambda_{k_b}\geq 0, \quad \forall b\in\mathcal{B},\forall k\in\mathcal{K}_b. \tag{C.4c}$$

Strong duality of problem (C.2) implies that in the optimum point of the primal and dual problems have the same objective value, i.e. $\sum_{b=1}^{B}\sum_{k_b=1}^{K_b}\lambda_{k_b}-\sum_{b=1}^{B}\mu_b=0$. Additionally, the stationarity Karush-Kuhn-Tucker (KKT) conditions imply that $\partial\mathcal{L}/\partial\mathbf{m}_{k_b}=\mathbf{0},\forall b\in\mathcal{B},\forall k\in\mathcal{K}_b$:

$$\frac{\mu_b}{P_b}\mathbf{m}_{k_b}+\sum_{\substack{i=1\\ i\neq k}}^{K_b}\frac{\lambda_{i_b}}{\sigma^2}\mathbf{h}_{i_bb}\mathbf{h}_{i_bb}^H\mathbf{m}_{k_b}+\sum_{\substack{j=1\\ j\neq b}}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_jb}\mathbf{h}_{i_jb}^H\mathbf{m}_{k_b}-\frac{\lambda_{k_b}}{\sigma^2\Gamma_{k_b}^{\min}}\mathbf{h}_{k_bb}\mathbf{h}_{k_bb}^H\mathbf{m}_{k_b}=0, \tag{C.5}$$

$$\Leftrightarrow\quad\left(\frac{\mu_b}{P_b}\mathbf{I}+\sum_{j=1}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_jb}\mathbf{h}_{i_jb}^H\right)\mathbf{m}_{k_b}=\frac{\lambda_{k_b}}{\sigma^2}\left(1+\frac{1}{\Gamma_{k_b}^{\min}}\right)\mathbf{h}_{k_bb}\mathbf{h}_{k_bb}^H\mathbf{m}_{k_b}, \tag{C.6}$$

$$\Leftrightarrow\quad\mathbf{m}_{k_b}=\underbrace{\left(\frac{\mu_b}{P_b}\mathbf{I}+\sum_{j=1}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_jb}\mathbf{h}_{i_jb}^H\right)^{-1}\mathbf{h}_{k_bb}}_{\text{vector}}\underbrace{\frac{\lambda_{k_b}}{\sigma^2}\left(1+\frac{1}{\Gamma_{k_b}^{\min}}\right)\mathbf{h}_{k_bb}^H\mathbf{m}_{k_b}}_{\text{scalar}} \tag{C.7}$$

Thus the beamforming vectors are

$$\mathbf{m}_{k_b}=\sqrt{p_{k_b}}\bar{\mathbf{m}}_{k_b},\quad\forall b\in\mathcal{B},\forall k\in\mathcal{K}_b, \tag{C.8}$$

where $\sqrt{p_{k_b}}$ denotes the beamforming power of UE k_b and

$$\bar{\mathbf{m}}_{k_b}=\frac{\left(\frac{\mu_b}{P_b}\mathbf{I}+\sum_{j=1}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_jb}\mathbf{h}_{i_jb}^H\right)^{-1}\mathbf{h}_{k_bb}}{\left\|\left(\frac{\mu_b}{P_b}\mathbf{I}+\sum_{j=1}^{B}\sum_{i=1}^{K_j}\frac{\lambda_{i_j}}{\sigma^2}\mathbf{h}_{i_jb}\mathbf{h}_{i_jb}^H\right)^{-1}\mathbf{h}_{k_bb}\right\|} \tag{C.9}$$

denotes the beamforming direction.

The connection between the sum power minimization and sum rate maximization problems implies that the beamforming solution for (4.3) is also given by (C.8), with $\sum_{k=1}^{K_b}p_{k_b}\leq P_b$, $\sum_{b=1}^{B}\sum_{k_b=1}^{K_b}\lambda_{k_b}=1$ and $\sum_{b=1}^{B}\sum_{k_b=1}^{K_b}\lambda_{k_b}-\sum_{b=1}^{B}\mu_b=0$.

This way, (C.8) provides the structure of the optimal beamforming as function of the Lagrange multipliers λ_{k_b} and μ_b and the power values p_{k_b}, with $k\in\mathcal{K}_b$ and $b\in\mathcal{B}$.