
**UNIVERSIDADE FEDERAL DO CEARÁ**

**CENTRO DE TECNOLOGIA**

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA**

**DOUTORADO EM ENGENHARIA DE TELEINFORMÁTICA**

**JUNO VITORINO SARAIVA**

# OPTIMIZING POWER CONTROL IN CENTRALIZED AND DISTRIBUTED MIMO NETWORKS: STRATEGIES AND SOLUTIONS

**FORTALEZA**

**2024**

JUNO VITORINO SARAIVA

# OPTIMIZING POWER CONTROL IN CENTRALIZED AND DISTRIBUTED MIMO NETWORKS: STRATEGIES AND SOLUTIONS

Tese apresentada ao Doutorado em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Orientador: Prof. Dr. Walter da Cruz Freitas Júnior

Co-Orientador: Dr. Roberto Pinto Antonioli

FORTALEZA

2024

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)

---

S246o Saraiva, Juno Vitorino.
Optimizing power control in centralized and distributed mimo networks: strategies and solutions / Juno Vitorino Saraiva. – 2024.
149 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2024.
Orientação: Prof. Dr. Walter da Cruz Freitas Júnior.
Coorientação: Prof. Dr. Roberto Pinto Antonioli.

1. 5G. 2. MIMO Networks. 3. RRM Solutions. 4. UAV Communications. I. Título.
CDD 621.38

---

## ACKNOWLEDGMENTS

First, I want to express my deepest gratitude to my parents. Although words fall short of expressing how truly grateful I am, Mom and Dad, thank you for giving me the chance to fully commit to my studies.

I would also like to extend my thanks to my supervisors, Prof. Dr. Walter Freitas and Dr. Roberto Antonioli. A special thanks to Dr. Roberto Antonioli for his invaluable help and support throughout the years, and for the countless meetings and discussions that were crucial in developing this thesis.

I also want to extend my heartfelt thanks to Prof. Dr. Rodrigo Cavalcanti, Prof. Dr. Rafael Lima, Prof. Dr. Yuri Carvalho, and Prof. Dr. Walter Freitas, for giving me the chance to join the GTEL team and contribute to several research projects. Indeed, being a part of the GTEL research project team has been an incredibly enriching experience. Within the context of these projects, I extend my special thanks to Prof. Dr. Yuri Carvalho and Dr. Gábor Fodor for their significant contributions to the numerous studies that make up this thesis.

In this context, I would like to express my gratitude to the colleagues with whom I spent time and worked alongside at GTEL, particularly my friend Iran Mesquita, for the fruitful discussions and the collaborative work we have done together.

Finally, I acknowledge the technical and financial support from FUNCAP and Ericsson Research, Sweden, and Ericsson Innovation Center, Brazil, under UFC.47, UFC.48 and UFC.50 Technical Cooperation Contracts Ericsson/UFC. Also, this study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

## RESUMO

Sem dúvida, o uso de sistemas MIMO (do inglês, *multiple input, multiple output*) tem fornecido significantes benefícios para redes móveis desde gerações passadas. Nesse contexto, tanto arquiteturas MIMO centralizadas quanto distribuídas têm sido importantes para a evolução de tecnologias celulares, levando a significativos aperfeiçoamentos em métricas-chave tais como SE (do inglês, *spectral efficiency*), EE (do inglês, *energy efficiency*), justiça e QoS (do inglês, *quality of service*). Entretanto, dois desafios importantes para as atuais arquiteturas MIMO são gerenciar o grande número de dispositivos conectados e eficientemente integrar diferentes tipos de usuários, tais como GUEs (do inglês, *ground user equipments*) e UAVs (do inglês, *uncrewed aerial vehicles*), que possuem tanto padrões de mobilidade quanto qualidades de canal radicalmente distintos. Para abordar esses desafios, soluções de RRM (do inglês, *radio resource management*) são essenciais. Nesse contexto, nós propomos várias estratégias para lidar com diferentes objetivos importantes dentro de redes móveis. Mais especificamente, a primeira parte desta tese foca em redes MIMO centralizadas e introduz soluções de RRM baseadas na teoria de programação fracionária e teoria dos jogos. Além disso, nós consideramos cenários que têm sido menos explorados na literatura, incluindo cenários com modelos de tráfego *non-full buffer* e modelos de canais autorregressivos correlacionados no tempo. Na segunda parte desta tese, nós assumimos redes MIMO distribuídas. Nesse contexto, nós inicialmente adotamos uma abordagem sob a categoria de *potential games* para propor uma solução capaz de atender vários objetivos da rede, incluindo SE, EE e justiça. Subsequentemente, nós exploramos cenários envolvendo a coexistência de GUEs e UAVs dentro da mesma rede. Em seguida, empregando ferramentas de otimização convexa e *deep learning*, nós propomos soluções de RRM capazes de lidar com diferentes níveis de prioridade entre GUEs e UAVs baseados em seus requisitos individuais. Para todos os problemas de RRM abordados nesta tese, soluções descentralizadas são propostas. De fato, comparado a abordagens centralizadas, isso torna-se importante à medida que o número de usuários na rede cresce indefinidamente. Além disso, nós temos o objetivo de propor soluções flexíveis e adaptáveis que podem facilmente se adequar a mudanças de objetivos. Esta flexibilidade é crucial em ambientes de rede dinâmicos e complexos.



**Palavras-chave:** 5G; Redes MIMO; Soluções de RRM; Comunicações UAV.

## ABSTRACT

Definitely, the use of multiple input multiple output (MIMO) systems has provided significant benefits to mobile networks since previous generations. In this context, both centralized and distributed architectures of MIMO systems have been important to the evolution of cellular technologies, leading to significant improvements in key metrics such as spectral efficiency (SE) and energy efficiency (EE), fairness, and quality of service (QoS). However, two important challenges for current MIMO architectures are managing a massive number of connected users and efficiently integrating different types of users, such as ground user equipments (GUEs) and uncrewed aerial vehicles (UAVs), which have radically different mobility patterns and channel qualities. To address these challenges, radio resource management (RRM) solutions are essential. In this context, we propose various strategies to tackle different key objectives within mobile networks. More specifically, the first part of this thesis focuses on centralized MIMO networks and introduces RRM solutions leveraging fractional programming theory and game theory. Moreover, we consider scenarios that have been less explored in the literature, including scenarios with non-full buffer traffic models and time-correlated autoregressive channel models. In the second part of this thesis, we assume distributed MIMO networks. In this context, we initially adopt an approach under the category of potential games to propose a solution capable of attaining various network objectives, including SE, EE, and fairness. Subsequently, we explored scenarios involving the coexistence of GUEs and UAVs within the same network. Then, employing convex optimization and deep learning tools, we propose RRM solutions capable of accommodating diverse priority levels between GUEs and UAVs based on their individual requirements. For all RRM problems addressed in this thesis, decentralized solutions are proposed. Indeed, compared to centralized approaches, this becomes important as the number of users in the network grows indefinitely. Additionally, we aim to propose adaptable and flexible solutions that can easily adjust to changing objectives. This flexibility is crucial in dynamic and complex network environments.



**Keywords:** 5G; MIMO Networks; RRM Solutions; UAV Communications.

## LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| 3GPP | 3rd generation partnership project |
| 4G | 4th generation |
| 5G | 5th generation |
| 5G-StoRM | 5G stochastic radio channel for dual mobility |
| 6G | 6th generation |
| AI | artificial intelligence |
| AP | access point |
| AR | autoregressive |
| ARMA | autoregressive moving average |
| AWGN | additive white Gaussian noise |
| BS | base station |
| CDF | cumulative distribution function |
| CPU | central processing unit |
| CSI | channel state information |
| DQN | deep Q-network |
| DRL | deep reinforcement learning |
| EE | energy efficiency |
| FPC | fractional power control |
| GUE | ground user equipment |
| IBC | interference broadcast channel |
| ICT | information and communication technologies |
| IoE | internet of everything |
| IoT | internet of things |
| JSFRA | joint space-frequency resource allocation |
| KF | Kalman filter |
| KKT | Karush-Kuhn-Tucker |
| LOS | line of sight |
| LS | least squares |
| LTE | long term evolution |
| LTE-A | long term evolution advanced |
| MIMO | multiple input multiple output |

| | |
|---|---|
| MISO | multiple input single output |
| MMSE | minimum mean squared error |
| MRC | maximum ratio combining |
| MSE | mean squared error |
| MU-MIMO | multiuser MIMO |
| OFDM | orthogonal frequency division multiplexing |
| OFDMA | orthogonal frequency division multiple access |
| PDPC | pilot-and-data power control |
| PPO | proximal policy optimization |
| QoE | quality of experience |
| QoS | quality of service |
| RIS | reflective intelligent surface |
| RL | reinforcement learning |
| RRM | radio resource management |
| SCA | successive convex approximation |
| SCO | successive convex optimization |
| SE | spectral efficiency |
| SINR | signal to interference-plus-noise ratio |
| SIR | signal-to-interference ratio |
| SOCP | second-order cone programming |
| TDD | time division duplexing |
| TRPO | trust region policy optimization |
| TTI | transmission time interval |
| UAV | uncrewed aerial vehicle |
| UE | user equipment |
| UMi | urban micro |
| URLLC | ultra-reliable and low-latency communications |
| WMMSE | weighted minimum mean squared error |
| WPT | wireless power transfer |

## LIST OF SYMBOLS

### Chapter 2

| | |
|---|---|
| $(\cdot)^{\star}$ | Optimal value |
| $(\cdot)^{*}$ | Conjugate operator |
| $(\cdot)^{\mathrm{H}}$ | Hermitian operator |
| $\mathbb{E}\{\cdot\}$ | Expected value |
| $\mathbf{I}$ | Identity matrix |
| $\mathcal{B}$ | Set of all base stations (BSs) |
| $B$ | Number of BSs |
| $\mathcal{N}$ | Set of sub-channels |
| $N$ | Number of sub-channels |
| $N_{\mathrm{T}}$ | Number of antennas at the BS |
| $N_{\mathrm{R}}$ | Number of antennas at the user equipment (UE) |
| $\mathbf{U}_b$ | Set of UEs associated with BS $b$ |
| $U$ | Number of multi-antenna UEs |
| $\mathbf{H}_{b_i,u,n}$ | Channel matrix between UE $u$ and BS $b$ serving UE $i$ on sub-channel $n$ |
| $\hat{\mathbf{H}}_{b_i,s,n}$ | The estimated channel matrix between UE $u$ and BS $b$ serving UE $i$ on sub-channel $n$ |
| $\mathbf{y}_{u,s,n}$ | The downlink signal received by UE $u$ over spatial stream $s$ and sub-channel $n$ |
| $x_{u,s,n}$ | The mutually independent transmitted data symbol for UE $u$ and stream $s$ on sub-channel $n$ |
| $\mathbf{m}_{u,s,n}$ | Transmit beamforming vector for UE $u$ and stream $s$ on sub-channel $n$ |
| $\mathbf{w}_{u,s,n}$ | Receive beamformer vector for UE $u$ and stream $s$ on sub-channel $n$ |
| $\mathbf{E}$ | Error matrix with complex Gaussian independent and identically distributed (i.i.d.) entries with zero mean and unit variance |
| $\varrho$ | Denotes the reliability of the estimate |

| | |
|---|---|
| $\Gamma_{u,s,n}$ | The signal to interference-plus-noise ratio (SINR) for stream $s$ and sub-channel $n$ of UE $u$ |
| $\mathbf{n}_{u,n}$ | Noise at UE $u$ and sub-channel $n$ |
| $\sigma^2$ | Denotes the additive white Gaussian noise (AWGN) variance |
| $Q_u$ | Number of backlogged bits intended to UE $u$ |
| $S_{u,n}$ | Number of spatial streams allocated to UE $u$ on sub-channel $n$ |
| $\Delta_{\text{tti}}$ | Duration of one transmission time interval (TTI) |
| $A_u$ | Average number of packet arrivals for UE $u$ |
| $\lambda_u$ | Represent the instantaneous number of bits arriving for UE $u$ |
| $t_u$ | Number of transmitted bits per second for UE $u$ |
| $t_{u,s,n}$ | Denotes the number of transmitted bits per second over the $s$-th stream and $n$-th sub-channel of UE $u$ |
| $\epsilon_{u,s,n}$ | Denotes mean squared error (MSE) for UE $u$ and stream $s$ on sub-channel $n$ |
| $\tilde{\epsilon}_{u,s,n}$ | The MSE approximation point for UE $u$ and stream $s$ on sub-channel $n$ |
| $\alpha_u$ | Dual variable |
| $\beta_u$ | Dual variable |
| $\delta_b$ | Dual variable |
| $\lambda_{u,s,n}$ | Dual variable |
| $\mu_{u,s,n}$ | Dual variable |
| $\gamma_u$ | Denotes the priority weight of UE $u$ |
| $\xi_u$ | Models the rate requirement of UE $u$ |
| $P_b$ | Power budget of BS $b$ |
| $\zeta$ | Circuit power consumption |

**Chapter 3**

| | |
|---|---|
| $(\cdot)^\star$ | Optimal value |
| $(\cdot)^*$ | Conjugate operator |
| $(\cdot)^{\mathrm{T}}$ | Transpose operator |
| $(\cdot)^{\mathrm{H}}$ | Hermitian operator |

| | |
|---|---|
| $\mathbb{E}\{\cdot\}$ | Expected value |
| $\mathbf{I}$ | Identity matrix |
| $N$ | Number of antennas at the BS |
| $\mathcal{K}$ | Set of all UEs |
| $K$ | Number of single-antenna UEs |
| $\tau_{\text{p}}$ | Number of subcarriers dedicated for pilot sequences |
| $\tau_{\text{d}}$ | Number of subcarriers dedicated for data symbols |
| $\rho_k^{\text{p}}$ | Pilot power of UE $k$ |
| $\rho_k$ | Data power of UE $k$ |
| $\boldsymbol{\rho}$ | Vector of data powers of all UEs |
| $\boldsymbol{\rho}_{-(k)}$ | Vector of data powers of all UEs except of UE $k$ |
| $P_{\text{tot}}$ | Maximum power value transmitted by a UE |
| $\mathbf{s}_k$ | Denotes the orthogonal pilot sequence of UE $k$ |
| $x_k$ | Transmitted data symbol by UE $k$ |
| $\alpha_k$ | Large fading between UE $k$ and the BS |
| $\mathbf{A}_k$ | Transition matrix of the $\mathrm{AR}(1)$ process of UE $k$ |
| $\mathbf{h}_k$ | Small scale fading between UE $k$ and the BS |
| $\hat{\mathbf{h}}_k$ | Estimated small scale fading between UE $k$ and the BS |
| $\mathbf{C}_k$ | Covariance matrix of $\mathbf{h}_k$ |
| $\mathbf{R}_k$ | Covariance matrix of $\hat{\mathbf{h}}_k$ |
| $\mathbf{Q}_k$ | Covariance matrix of $\boldsymbol{\vartheta}_k$ |
| $\boldsymbol{\vartheta}_k$ | Denotes a complex normal distributed noise vector for UE $k$ |
| $\mathbf{G}_k$ | Denotes a generic receiver vector for UE $k$ |
| $\mathcal{G}$ | Denotes a non-cooperative power control strategy game |
| $\boldsymbol{\Upsilon}$ | Denotes the strategy total space |
| $\mathcal{P}_k$ | Denotes a discrete and finite set of strategies or actions available for UE $k$ |
| $\mathbf{y}$ | Denotes the multiuser MIMO (MU-MIMO) received data signal at the BS |
| $\mathbf{N}$ | Denotes the AWGN matrix |

| | |
|---|---|
| $\sigma_{\mathrm{p}}^2/\sigma_{\mathrm{d}}^2$ | Denotes the AWGN variance |
| $\mathbf{n}_{\mathrm{d}}$ | Denotes the noise on the received data signal |
| $\mathbf{w}_k$ | Denotes the channel estimation error for UE $k$ |

**Chapter 4**

| | |
|---|---|
| $(\cdot)^\star$ | Optimal value |
| $(\cdot)^{\mathrm{H}}$ | Hermitian operator |
| $\mathbb{E}\{\cdot\}$ | Expected value |
| $\mathbf{I}$ | Identity matrix |
| $\mathcal{L}$ | Set of all access points (APs) |
| $L$ | Number of APs |
| $N$ | Number of antennas at the AP |
| $\mathcal{K}$ | Set of all UEs |
| $K$ | Number of UEs |
| $\tau_{\mathrm{p}}$ | Coherence interval dedicated to uplink channel estimation |
| $\tau_{\mathrm{u}}$ | Coherence interval designated for uplink data transmission |
| $\tau_{\mathrm{d}}$ | Coherence interval allocated for downlink data transmission |
| $\rho_k$ | Data power of UE $k$ |
| $\boldsymbol{\rho}$ | Vector of data powers of all UEs |
| $\boldsymbol{\rho}_{-(k)}$ | Vector of data powers of all UEs except of UE $k$ |
| $P_{\mathrm{max}}$ | Maximum power value transmitted by a UE |
| $\mathbf{h}_{kl}$ | Denotes the channel between AP $l$ and UE $k$ |
| $\hat{\mathbf{h}}_{kl}$ | Denotes the estimated channel between AP $l$ and UE $k$ |
| $\tilde{\mathbf{h}}_{kl}$ | Denotes the channel estimation error between AP $l$ and UE $k$ |
| $\mathbf{h}_k$ | Denotes the collective channel from all APs of UE $k$ |
| $\beta_{kl}$ | Denotes the the large-scale fading coefficient that describes pathloss and shadowing between UE $k$ and AP $l$ |
| $\mathbf{R}_{kl}$ | Describes the large-scale fading, including shadowing, pathloss, spatial channel correlation and antenna gains between UE $k$ and AP $l$ |

| | |
|---|---|
| $\mathbf{R}_k$ | Denotes the block-diagonal spatial correlation matrix for UE $k$ |
| $\mathbf{B}_{kl}$ | Covariance matrix of $\hat{\mathbf{h}}_{kl}$ |
| $\mathbf{C}_{kl}$ | Covariance matrix of $\tilde{\mathbf{h}}_{kl}$ |
| $\mathbf{D}_{kl}$ | If the $n$-th diagonal is 1, the $n$-th antenna of AP $l$ is allowed to transmit and to decode signals from UE $k$, and 0 otherwise |
| $\mathbf{D}_k$ | Denotes the block-diagonal matrix for UE $k$ |
| $\mathbf{A}$ | Defines the AP selection matrix |
| $\mathcal{D}_l$ | Represents the subset of UEs served by AP $l$ |
| $\mathcal{M}_k$ | Represents the subset of APs serving UE $k$ |
| $\mathbf{y}_{tl}^{\text{pilot}}$ | Denotes the received signal at AP $l$ when the subset of UEs is assigned to pilot $t$ |
| $\mathbf{\Psi}_{tl}$ | Denotes the correlation matrix for $\mathbf{y}_{tl}^{\text{pilot}}$ |
| $\mathbf{y}_l$ | Data signal received at AP $l$ |
| $s_k$ | Data symbol of UE $k$ |
| $\hat{s}_k$ | Estimated data symbol of UE $k$ |
| $\sigma^2$ | Denotes the AWGN variance |
| $\mathbf{v}_{kl}$ | Receive beamformer vector for UE $k$ and AP $l$ |
| $\mathbf{v}_k$ | Collective of combining vectors for UE $k$ |
| $\mathcal{G}$ | Denotes a non-cooperative power control strategy game |
| $\alpha$ | Denotes an input parameter of game $\mathcal{G}$ |
| $\mathbf{\Upsilon}$ | Denotes the strategy total space |
| $\mathcal{P}_k$ | Denotes a finite set of strategies or actions available for UE $k$ |
| $\mu_k(\cdot)$ | Denotes the payoff function for UE $k$ |

**Chapter 5**

| | |
|---|---|
| $(\cdot)^\star$ | Optimal value |
| $(\cdot)^{\text{H}}$ | Hermitian operator |
| $\mathbf{I}$ | Identity matrix |
| $\mathcal{A}$ | Set of all APs |

| | |
|---|---|
| $\mathcal{A}_k$ | Set containing all the APs serving UE $k$ |
| $L$ | Number of APs |
| $N$ | Number of antennas at the AP |
| $\mathcal{K}$ | Set of all UEs |
| $\mathcal{K}_a$ | Denotes the group of UEs that are serviced by AP $a$ |
| $K$ | Number of UEs |
| $\mathcal{G}$ | Set of all ground user equipments (GUEs) |
| $K_{\mathrm{g}}$ | Number of GUEs |
| $\mathcal{U}$ | Set of all uncrewed aerial vehicles (UAVs) |
| $K_{\mathrm{u}}$ | Number of UAVs |
| $\tau_{\mathrm{p}}$ | Coherence interval dedicated to uplink channel estimation |
| $\tau_{\mathrm{u}}$ | Coherence interval designated for uplink data transmission |
| $\tau_{\mathrm{d}}$ | Coherence interval allocated for downlink data transmission |
| $s_k$ | Data symbol of UE $k$ |
| $\hat{s}_k$ | Estimated data symbol of UE $k$ |
| $\boldsymbol{\phi}_k$ | Pilot sequence sent by UE $k$ |
| $\mathbf{y}_a$ | Data signal received at AP $a$ |
| $\hat{\mathbf{y}}_a$ | Estimated data signal received at AP $a$ |
| $\mathbf{Y}_a$ | Cumulative received pilot signal at AP $a$ |
| $\boldsymbol{\upsilon}_{k,a}$ | Uplink receive combining vector for UE $k$ and AP $a$ |
| $\mathbf{W}_a$ | Contains the thermal noise contribution at the AP $a$ |
| $\sigma^2$ | Denotes the AWGN variance |
| $\rho_k^{\mathrm{pilot}}$ | Pilot power of UE $k$ |
| $\rho_k$ | Data power of UE $k$ |
| $\boldsymbol{\rho}$ | Vector of data powers of all UEs |
| $\boldsymbol{\rho}_{-(k)}$ | Vector of data powers of all UEs except of UE $k$ |
| $P_{\mathrm{max}}^k$ | Maximum power value transmitted by a UE $k$ |
| $\mathbf{g}_{k,a}$ | The channel vector connecting UE $k$ to AP $a$ |

| | |
|---|---|
| $\hat{\mathbf{g}}_{k,a}$ | The estimated channel vector connecting UE $k$ to AP $a$ |
| $\beta_{k,a}$ | Characterizes the scalar coefficient modeling the channel path-loss and shadowing effects between the UE $k$ and the AP $a$ |
| $\mathbf{h}_{k,a}$ | Small scale fading coefficients between UE $k$ and AP $a$ |
| $R_{k,a}$ | Ricean $R$-factor between UE $k$ and AP $a$ |
| $\mathbf{a}(\theta_{k,a})$ | Represents the steering vector for the angle $\theta_{k,a}$ between UE $k$ and AP $a$. |
| $\lambda$ | Represents an input parameter |
| $a_k^{(t)}$ | Individual action taken by UE $k$ for a time step $t$ |
| $\mathbf{a}^{(t)}$ | Action vector for a time step $t$ |
| $r^{(t)}$ | Reward function for a time step $t$ |
| $\pi\left(a\|s;\phi_\pi\right)$ | Denotes the conditional probability of taking each action $a$ when in state $s$ |
| $\mu\left(s;\phi_\mu\right)$ | Takes observation $s$ and returns the corresponding expectation of the discounted long-term reward. |
| $\phi_\pi$ | Denotes the parameters of the actor |
| $\phi_\mu$ | Denotes the parameters of the critic |
| $\mathbf{s}_k^{(t)}$ | Observation or state for a given UE $k$ for a time step $t$ |
| $D^{(t)}$ | Advantage function for a time step $t$ |
| $G^{(t)}$ | Return function for a time step $t$ |
| $\tilde{\gamma}$ | Discount factor |
| $\tilde{\lambda}$ | GAE factor |
| $\delta$ | Temporal difference error |
| $\mathcal{L}\left(\cdot\right)$ | Loss function |
| $q\left(\cdot\right)$ | Represents the ratio of probabilities between the current policy and the previous one |
| $c(\cdot)$ | Represents the clipped surrogate objective function |
| $\tilde{\epsilon}$ | Clip factor |
| $\mathcal{W}$ | Denotes an entropy term |
| $\mathcal{E}(\cdot)$ | Denotes the entropy loss |
| $w$ | Denotes the entropy loss weight |

## TABLE OF CONTENTS

# 1 GENERAL OVERVIEW

Multi-antenna systems are not entirely new, i.e., they have been pivotal in advancements from 4th generation (4G) onward, enhancing signal quality and network capacity by utilizing multiple antennas at both the transmitter and receiver ends. Undoubtedly, the implementation of multi-antenna systems, also referred to as multiple input multiple output (MIMO) systems, in cellular networks has sparked a true revolution in this field. Therefore, this technology is not only here to stay but is also poised to play an important role in future generations of cellular networks (Zheng *et al.*, 2015).

In general, this is largely attributed to the extensive evidence in the literature demonstrating that a significant increase in the number of antennas at the BS, a strategy known as massive MIMO, can lead to substantial improvements in key metrics for mobile networks, such as spectral efficiency (SE) (i.e., bits per second per hertz of bandwidth) and energy efficiency (EE) (i.e., data rate per unit of energy consumed). By leveraging these enhanced capabilities, massive MIMO not only supports more users simultaneously but also optimizes the power and bandwidth resources, which are vital for maintaining the sustainability of mobile network infrastructure. As a result, the adoption of massive MIMO is increasingly viewed as an essential advancement in the ongoing evolution of cellular technologies, promising even greater network performance and efficiency in the future (Larsson *et al.*, 2014).

Although the concept of massive MIMO is very promising for future mobile networks, the emergence of revolutionary applications, such as widespread connectivity in smart environments, augmented/virtual reality, and high-definition video streaming, necessitates fundamental changes in mobile networks from one generation to the next. These applications demand not only higher data rates but also lower latency, increased reliability, and more efficient energy usage. Furthermore, the internet of everything (IoE) concept, which results in a drastic increase in connected devices, presents an important challenge in the current context of mobile networks (Elhoushy *et al.*, 2022).

In the context of UAV communications, the use of UAVs, commonly known as drones, has steadily become crucial in various professional fields, marking a significant change in many practices and operations. This is because their wide range of applications covers several civilian areas, including aerial photography, precision agriculture, environmental monitoring, and search and rescue missions. Further, UAVs are increasingly used in infrastructure inspections, disaster management, delivery services, and even wildlife conservation efforts, showcasing their

versatility and impact across diverse industries (Hayat *et al.*, 2016; Tsouros *et al.*, 2019; Silvagni *et al.*, 2017; Wang *et al.*, 2020).

However, despite the promising potential of UAVs, efficiently integrating them with traditional terrestrial users also remains a challenge for current cellular networks. This is because UAVs, as aerial users, have vastly different characteristics compared to terrestrial users, such as mobility patterns, channel quality, quality of service (QoS) and quality of experience (QoE) requirements. As a result, planning a network that can efficiently serve both terrestrial and aerial users simultaneously is a complex task (Geraci *et al.*, 2018; Zhang *et al.*, 2019; Gong *et al.*, 2020).

To accommodate such requirements and challenges, networks must evolve to incorporate more advanced forms of MIMO, adaptive network architectures, and smarter resource allocation methods (Zhao *et al.*, 2020). Furthermore, these new network architectures must efficiently deal with traditional problems such as variations in QoS and cell-edge problems, which are inherent to conventional cellular network designs and represent significant obstacles for mobile network operators.

It is important to highlight that the principles of massive MIMO are not confined to concentrated or centralized antenna arrays, but could also be applied to distributed deployments. In this context, cell-free massive MIMO, a practical incarnation of distributed massive MIMO, has become an intensive research topic in industry and academia. This is because, differently from centralized cellular network architectures, cell-free massive MIMO can potentially mitigate significant pathloss variations and cell-edge problems (Chen *et al.*, 2021). However, despite the substantial benefits associated with cell-free massive MIMO networks, there is a growing demand for more ecological designs and green solutions (Chien *et al.*, ). Conversely, deploying a distributed antenna array in cell-free massive MIMO systems can significantly increase power consumption in wireless communication systems (Imoize *et al.*, 2022).

Therefore, power control is a critical issue in cell-free massive MIMO networks and acts as a performance enhancement metric for ensuring the achievement of network optimization while maintaining the QoS for all users (Chien *et al.*, ). Moreover, power control techniques are essential in mitigating the problems of interference and pilot contamination that are common in cell-free networks. In fact, regardless of network architecture and depending on available resources, power control enhances a wide variety of essential network functionalities, including the throughput, max-min fairness, SE and EE.

While sustainable energy usage is undoubtedly a critical aspect of modern and future mobile networks, other implementation issues are equally essential. For instance, an efficient power control solution that utilizes a centralized framework and suffers from scalability issues is less effective (Björnson; Sanguinetti, 2020). Similarly, solutions that are overly rigid and lack flexibility to adapt to changes also fall short of meeting the dynamic needs of advanced network environments. These examples highlight the necessity for both *scalability* and *adaptability* in developing effective solutions. In other words, the current concept of an *efficient solution* can encompass new dimensions that go beyond traditional measures, including not only performance in relation to optimal or near-optimal solutions and the computational costs involved (Calabrese *et al.*, 2018).

Nowadays, given the massive number of devices connected to the network, it is as crucial to propose near-optimal solutions for power control as it is to ensure that these solutions can scale effectively with the increasing number of network users (Chen *et al.*, 2021). Moreover, efficiency in performance and scalability can also be complemented by an adaptable and flexible framework. This adaptability is an important dimension in the current context of mobile networks because network objectives often evolve, and solutions that are too dependent on a specific structure/framework may struggle to meet new or shifting objectives with ease (Iliev *et al.*, ). In this sense, a versatile framework enables quick and effective adjustments to align with changing demands and technologies.

This thesis focuses on radio resource management (RRM) in multi-antenna mobile networks for 5th generation (5G) and beyond systems. More specifically, this work introduces both centralized and decentralized strategies and solutions aimed at effectively managing power control in the downlink and uplink phases of MIMO networks. These approaches are designed to optimize various objectives, including SE, EE, fairness, and QoS. Considering the context of increasing the number of antennas per site, we design power control solutions for MIMO interference broadcast channel (IBC), massive MIMO, and cell-free massive MIMO systems.

The next section provides more details about MIMO IBC, massive MIMO, cell-free massive MIMO systems, and RRM. Moreover, in Section 1.1.5, we provide an overview of the primary tools used in this thesis to address RRM problems. Then, Section 1.2 presents the objectives and the structure of this thesis. Finally, Section 1.3 details the scientific contributions of this thesis.

## 1.1 Background

### *1.1.1 MIMO IBC Networks*

In general, MIMO technology involves using multiple antennas at both the transmitter and receiver ends of a communication link, and it has been one of the most important advancements in wireless communications. This is because MIMO technology enables a transmitter to send multiple signals simultaneously through its array of antennas, facilitating the vectorization of the transmission process. In this context, the main advantages of MIMO systems usually are (Jungnickel *et al.*, 2009):

- **Increased Data Rate**: By using multiple antennas, MIMO systems can transmit several data streams simultaneously over the same radio frequency channel. This multiplexing significantly increases the throughput without requiring additional bandwidth.
- **Enhanced Signal Quality**: In general, MIMO systems can exploit multipath propagation, where transmitted signals bounce off objects and arrive at the receiver at different angles and times. This diversity can be used to improve signal robustness and reduce errors.
- **Increased Range**: It is also known that MIMO systems can focus the energy more efficiently through beamforming, thus extending the communication range.

On the other hand, IBC deals with scenarios where multiple transmitters are broadcasting to multiple receivers, but each transmitter is intended for a specific receiver or group of receivers. In this context, the main challenge in IBC is managing interference because signals from unintended transmitters can affect the receiver's ability to correctly interpret its intended signal.

Combining MIMO with IBC results in a complex system where multiple transmitters are sending multiple data streams to their respective receivers. In that scenario, each transmitter may have multiple antennas, and each receiver may also have multiple antennas. As a result, the system must manage and mitigate interference from other transmitters while maximizing the efficiency and throughput for each intended receiver. To efficiently deal with these challenges, the following key techniques are typically employed in MIMO IBC (Venkatraman *et al.*, 2016):

- **Precoding**: This technique involves adjusting the signals transmitted from the multiple antennas to ensure that the signal at the receiver is as clear as possible. Precoding can be used to mitigate interference at other receivers.
- **Beamforming**: This is another form of precoding that focuses the transmitted signal in

a specific direction to enhance signal strength at the receiver and reduce interference to others.

- **Spatial Multiplexing**: This involves transmitting separate and different information signals from each of the multiple antennas, effectively multiplying the capacity of the communication channel.

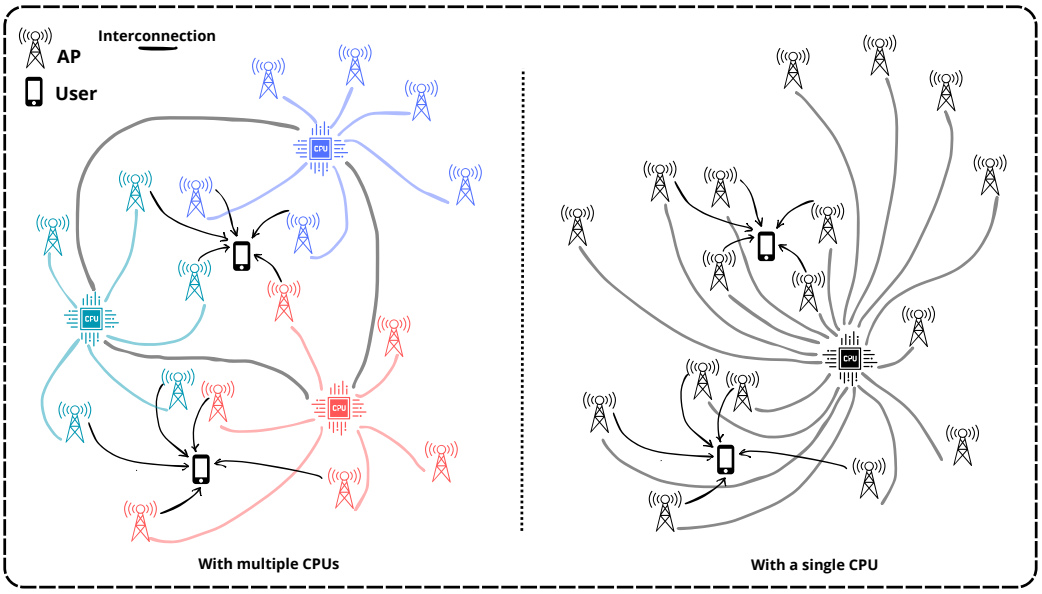
Therefore, although complex, MIMO IBC is a scenario particularly relevant in cellular networks where different BSs (transmitters) need to communicate with multiple mobile devices (receivers) in densely populated areas. Moreover, the principles of MIMO IBC are applicable any multi-user environment where high data rates and efficient spectrum use are critical. By optimizing how signals are transmitted and received, MIMO IBC significantly improves the performance of wireless networks, making it an important network architecture of current and future communication systems like 5G and beyond.

### *1.1.2 Massive MIMO Networks*

Massive MIMO is an advanced evolution of MIMO technology, which itself is a cornerstone of modern wireless communications. In this context, the term "massive" refers to the use of a significantly larger number of antennas at the BSs compared to traditional MIMO systems. This technology is one of the critical enablers of 5G networks and is also projected to be fundamental for future 6th generation (6G) networks. This is because the large scale use of antennas increases the capacity of a cell dramatically, as more data streams can be transmitted simultaneously without requiring additional bandwidth. Further, in massive MIMO, the scale of this multiplexing is increased exponentially, allowing for more users to be served simultaneously with high throughput. It is also true that with more antennas, beamforming becomes more precise. At last, it is known that massive MIMO can enhance both SE and EE (Lu *et al.*, 2014). These efficiencies are crucial for meeting the expanding data demands and environmental concerns.

However, massive MIMO architectures also face challenges. First, in terms of complexity in signal processing, handling the signal processing for hundreds of antennas increases the computational load and complexity. Also, although beamforming helps reduce interference, the large number of antennas and the density of users can still lead to complex interference patterns that need sophisticated algorithms to manage. In terms of hardware and cost, deploying and maintaining a large number of antennas increases the hardware requirements and the associated costs. Even so, massive MIMO is a great technology that can enhance network performance

Figure 1 – Illustration of a cell-free network with single and multiple CPUs.




Source: Created by the author.

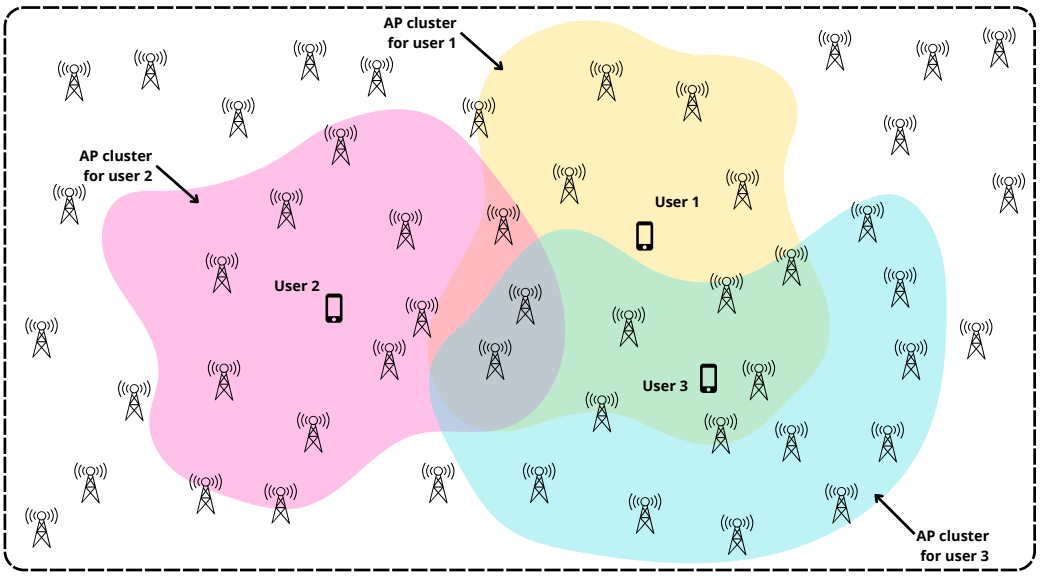
by orders of magnitude over traditional MIMO systems, efficiently addressing a wide range of challenges of modern mobile systems (Marzetta, 2015).

### *1.1.3 Cell-Free Massive MIMO Networks*

In contrast to traditional cellular networks, in cell-free systems there is no partitioning of geographical regions to define the concept of cells or cell boundaries and, therefore, users are not associated to a specific BS or cell. Instead, the network architecture known as cell-free massive MIMO combines two important and efficient concepts in mobile communications, namely massive MIMO and ultra-dense networks. As it is known, the conventional modality of massive MIMO systems is co-located massive MIMO in which BSs are equipped with a large number of antennas to communicate simultaneously with a smaller number of users on the same time-frequency resource through beamforming (Albreem *et al.*, 2019).

Nevertheless, differently from classical massive MIMO systems (i.e., a centralized framework), massive MIMO architectures can also be implemented in a distributed manner. This involves setting up numerous single or multiple-antenna APs that are spread out geographically. These APs are linked to a central processing unit (CPU) through high-speed fiber or wireless backhaul/fronthaul interconnections. In this novel architecture, differently from traditional massive MIMO systems, this distributed approach can offer better quality-of-service QoS to users at the cell-edge, thanks to the presence of more APs located closer to these users (Elhoushy

Figure 2 – Illustration of the user-centric concept in cell-free networks.



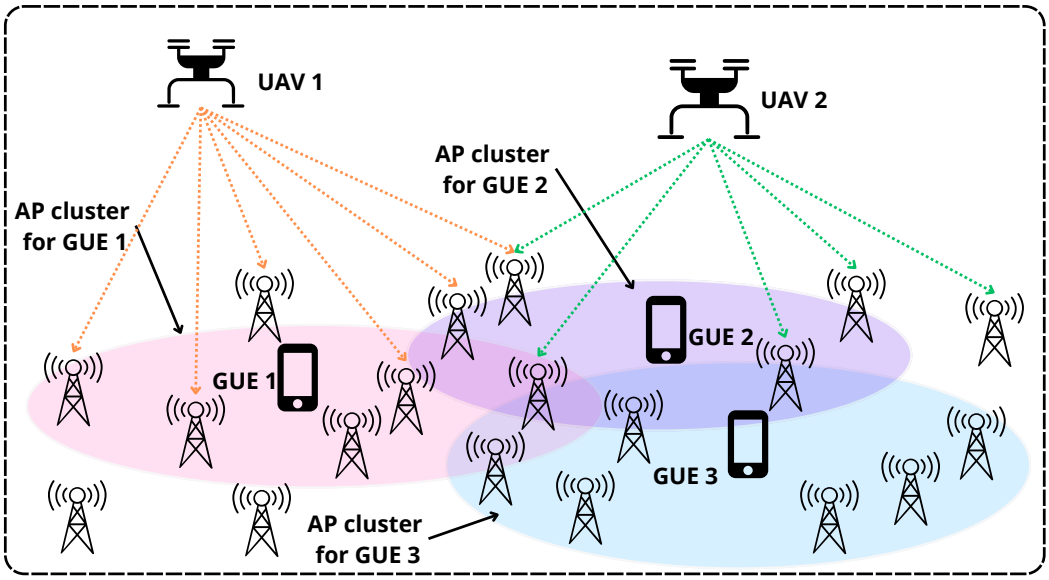

Source: Created by the author.

*et al.*, 2022). In this context, this modified practical embodiment of distributed massive MIMO systems has been emerged under the name of *cell-free massive MIMO systems* (Ngo *et al.*, 2017). Fig. 1 illustrates architectures featuring single or multiple CPUs for cell-free networks.

Therefore, from the perspective of the user, the quality of experience with centralized and decentralized MIMO architectures can vary significantly. A significant distinction between these architectures from the user's perspective is that the *user-centric* approach is a common strategy employed in cell-free networks (i.e., decentralized MIMO frameworks). More specifically, user-centric cell-free networks establish a cooperative serving cluster of spatially distributed transmitters tailored specifically for each user, hence the designation "user-centric" (Mohammadi *et al.*, 2023). This approach effectively eliminates traditional cell-edges, which is why it is termed "cell-free". According to this definition, each user in the network is served by a nearby group of transmitters (i.e., APs), placing each user at the effective center of their serving cluster.

Obviously, this introduces new challenges in the area of mobile networks, including the need to address questions like: "How many and which APs should each user choose to best serve their needs?". This issue, commonly referred to as *cluster formation*, has been extensively explored in literature, and it is usually closely linked to the objectives sought in the network, such as SE, EE, fairness, among others. Fig. 2 depicts the concept of a user-centric approach in cell-free networks, demonstrating the selection of APs by each network user and the quantity chosen. In this figure, note that the user-centric cell-free approach represents a significant shift from the

Figure 3 – Depiction of the cell-free scenario in which UAVs and GUEs coexist.




Source: Created by the author.

traditional cell concept, as now the users are always positioned in the center of their own "cell". Consequently, this reduces the *edge-effect*, a common problem in traditional architectures that diminished the performance of users located farther from the BS. Additionally, while not a fixed rule, note that an AP can often serve multiple users simultaneously, meaning approaches that involve overlapping in addressing the cluster formation problem are quite common (Björnson; Sanguinetti, 2020; Chen *et al.*, 2021).

In general, the user-centric philosophy of cell-free networks is highly beneficial, as it has facilitated, for example, efficient integration among various types or categories of users. More specifically, the integration of GUEs and UAVs has presented significant challenges and has been extensively studied in recent literature. This interest is driven by the growing competition for radio resources, such as APs, which is clearly illustrated in Fig. 3. Moreover, due to the fact that GUEs and UAVs exhibit significantly different behaviors and channel conditions, achieving overall network efficiency in environments where these two types of users coexist becomes even more complex (Fotouhi *et al.*, 2019). However, as mentioned earlier, cell-free networks have successfully achieved this integration through the use of a robust RRMs algorithms/strategies, especially in terms of power control. Indeed, in these scenarios, power control plays a crucial role in managing interference levels to ensure that the presence of UAVs does not significantly compromise the performance of GUEs.

### *1.1.4 Radio Resource Management*

The most common strategy in RRM methods involves implementing flexible scheduling of resources. This approach dynamically allocates available resources while considering multiple constraints, including system bandwidth, energy consumption, and QoS. In this context, one of the main tasks of RRM is commonly to ensure QoS for the users with an optimal use of resources. Particularly in cellular networks, RRM is a crucial aspect for several key reasons. This is because the overall performance of these systems directly depends on how efficiently the available radio resources are managed/optimized, e.g., subcarriers, time slots, transmit power, APs, among others (Calabrese *et al.*, 2018; Saraiva *et al.*, 2020).

In this thesis, we design and evaluate RRM solutions aimed at enhancing the efficiency and performance of mobile networks. The developed algorithms are tailored to address a wide range of functionalities within these networks, as detailed below:

- **Spectral Efficiency**: Radio frequency spectrum is a limited radio resource that is expensive and regulated by government entities. Thus, RRM is fundamental to ensure that this scarce resource is used efficiently, maximizing the number of users and the amount of data that can be transmitted within the available bandwidth.
- **Energy Efficiency**: Undoubtedly, in the future, EE metrics will become increasingly crucial and informative. Thus, by managing the power used for transmission effectively, RRM not only can extend the battery life of mobile devices but also it can reduce the energy consumption of network equipment. This is increasingly important as the number of networked devices grows exponentially and energy efficiency becomes a critical consideration.
- **Quality of Service**: In the context of QoS, RRM plays a vital role in managing network resources to meet the varying quality of service requirements of different applications, e.g., voice, video, and data. It helps in prioritizing resources, ensuring that critical services like emergency calls or real-time video streaming maintain high quality despite fluctuations in network traffic.
- **Interference Management**: In a densely populated network environment, interference between devices can significantly degrade performance. In this context, RRM solutions can include techniques to minimize this interference, ensuring that users receive a reliable and stable connection.
- **Fairness**: Enhancing fairness through RRM solutions on mobile networks is also feasible.

Fairness, in this context, means ensuring equitable access to network resources for all users, regardless of their location, device capabilities, or the type of service they are using.

In general, optimizing resources to address the aforementioned points involves usually solving optimization problems within the practical constraints and limitations inherent to mobile network structures. In the literature, centralized RRM solutions for these problems are often derived using convex optimization tools when feasible, due to their optimal or near-optimal performance. Additionally, alternative solutions using various techniques with lower computational costs are proposed to facilitate practical implementation.

However, effective RRM solutions can extend beyond mere optimization of radio resources. More specifically, striving for both high SE and EE remains as a paramount goal in mobile networks, and similarly, ensuring fairness and QoS for all users also remains profoundly significant. Yet, solely optimizing radio resources to meet a given specific objective function at a low computational cost might not suffice in the contemporary context of mobile networks (Calabrese *et al.*, 2018).

As previously discussed, the massive number of users connected to the network is an established reality. Therefore, RRM solutions, even those with low computational cost, become impractical when their computational complexity directly depends on the total number of users, especially in scenarios with an exponentially large user base. Consequently, reducing the dependence of user numbers on the computational complexity of RRM solutions can be a crucial factor (Björnson; Sanguinetti, 2020). In this sense, distributing the solution across different nodes in the network appears to be a fundamental approach in the search for feasible solutions when the network has a massive number of users.

Furthermore, solutions that can easily adapt to changing objectives and do not rely on specific tools like convex optimization are especially valuable, given the high complexity and dynamic nature of current RRM problems.

### *1.1.5 Approaches for Solving RRM Problems*

In the context of mobile networks, RRM problems present complex and multifaceted challenges. Various methodologies have been explored and implemented to address these problems effectively. This section provides a general overview of the main approaches utilized in this thesis, including successive convex optimization (SCO)/successive convex approximation (SCA), fractional programming theory, game theory, and deep reinforcement learning (DRL).

Each of these methods offers unique advantages and insights, contributing to a comprehensive understanding and solution of RRM problems.

#### *1.1.5.1 Successive Convex Optimization*

SCO is an iterative optimization technique that solves a non-convex problem by approximating it as a sequence of convex problems. The idea is to tackle complex, non-convex optimization problems by breaking them down into simpler, convex subproblems that are easier to solve. This method is iterative, with each iteration refining the solution, and it typically converges to a local optimum. However, the quality of this local optimum depends on the initial approximation and the nature of the problem. If the non-convex problem has multiple local minima, the method may converge to different solutions based on the starting point (Razaviyayn *et al.*, ; Razaviyayn, 2014).

In general, for mobile networks, SCO-based methods are employed to manage the non-convex characteristics of RRM problems by iteratively solving a series of convex approximations. By leveraging the mathematical properties of convex functions, SCO enhances both the efficiency and accuracy of the resource allocation process, ultimately leading to optimal or near-optimal solutions (Luo; Yu, 2006).

Below, we outline the main characteristics of the SCO technique, highlighting its advantages as well as its disadvantages.

- By solving convex problems iteratively, SCO leverages efficient algorithms for convex optimization, making the approach computationally feasible. However, the choice of the initial point can significantly affect the convergence and quality of the solution.
- SCO can handle a wide range of non-convex problems by appropriately choosing the convex approximation techniques. However, the effectiveness of SCO depends on how well the convex approximation represents the original non-convex problem.
- SCO often converges to a local optimum, providing high-quality solutions even if the global optimum is hard to achieve. However, while each convex problem is easier to solve, the overall computational cost can be high if many iterations are needed for convergence.

Therefore, SCO is a robust and versatile technique for tackling non-convex optimization problems (Liu; Lu, 2014). By breaking down complex problems into manageable convex subproblems, SCO provides an effective strategy for finding high-quality solutions across various applications. In the context of mobile networks, due to their efficiency and high-quality solutions,

SCO-based methods are often used as upper bound benchmarks. Consequently, they serve as a crucial reference point for evaluating alternative solutions.

#### *1.1.5.2 Fractional Programming Theory*

Fractional programming theory is a branch of optimization theory focused on the properties and optimization of fractional functions (Zappone *et al.*, 2015). This approach provides a robust framework for balancing multiple objectives, facilitating the development of efficient and effective resource allocation strategies.

In this context, Dinkelbach's approach is a well-known iterative algorithm used to solve fractional programming problems. These problems involve optimizing a ratio of two functions, which is common in many real-world applications. Named after the German mathematician Werner Dinkelbach, who introduced it in 1967 (Dinkelbach, 1967), this method transforms a fractional programming problem into a series of simpler, parameterized optimization problems that can be solved more easily. In mobile networks, the Dinkelbach's approach is a widely recognized optimization technique employed to enhance EE (i.e., the ratio of the sum data rate to the total power consumption) (Zappone; Jorswieck, 2017).

This is because Dinkelbach's approach provides a systematic and effective approach for solving fractional programming problems. By iteratively adjusting a parameter and optimizing over the resulting convex subproblems, Dinkelbach's algorithm facilitates the optimization of ratios such as EE efficiently. Moreover, it is worth noting that the classical approach initially proposed in (Dinkelbach, 1967) has been significantly generalized. The underlying concept allows for solving not only traditional fractional problems with a single fraction but also more complex problems with multiple fractions. Consequently, techniques derived from Dinkelbach's classic approach can effectively address a wider range of problems (Shen; Yu, 2018).

In the context of green communications, to optimize the traditional definition of EE (i.e., bit/J) as a single fraction using Dinkelbach's classical approach, the following sequence of steps is commonly employed in the literature:

- **Problem Formulation**: The initial step involves formulating the EE optimization problem. This typically includes defining the objective function, which is the ratio of the total data rate to the total power consumption.
- **Transformation**: Using the Dinkelbach method, the fractional objective function is transformed into a sequence of parametric linear problems. This transformation hinges on

the introduction of a parameter that represents the current estimate of the optimal objective function value.

- **Iterative Optimization**: The method proceeds iteratively. In each iteration, the transformed linear problem is solved to find the optimal solution for the current parameter value. The parameter is then updated based on the obtained solution.
- **Convergence Analysis**: The iterative process continues until convergence is achieved. Convergence occurs when the parameter value stabilizes, indicating that the optimal solution for the original fractional problem has been found. In this step, due to the application of Newton's method, convergence typically happens rapidly.

Therefore, the Dinkelbach method serves as a powerful tool for maximizing EE in mobile networks. Its ability to simplify complex fractional programming problems and guarantee convergence makes it indispensable in the pursuit of greener and more cost-effective mobile communication systems (Isheden *et al.*, 2012a).

#### *1.1.5.3 Game Theory*

Typically, game theory is a mathematical framework for analyzing situations in which multiple parties, known as players, make decisions that are interdependent. This interdependence means that the outcome for each player depends on the actions of all involved. Below, we highlight some fundamental concepts in the context of game theory (Osborne; Rubinstein, 1994):

- **Players**: The decision-makers in the game.
- **Strategies**: The plans or actions available to each player.
- **Payoffs**: The outcomes or returns players receive from a combination of strategies.
- **Games**: The scenarios or models that define the interaction between players, which can be basically cooperative or non-cooperative.

Another important concept in game theory is the Nash equilibrium. It represents the solution of a game, defined as the combination of the best strategies for each rational player that maximizes their own utility, given the strategies chosen by other players. At the Nash equilibrium, no player has an incentive to deviate from their selected strategy, as doing so would decrease their payoff (Mkiramweni *et al.*, 2019).

In the context of mobile networks, game theory is applied to model the interactions between multiple users and network resources as strategic games (Han, 2012). By analyzing the behavior of rational players, this approach helps design mechanisms that promote fair and

optimal resource distribution, taking into account both competitive and cooperative dynamics within the network environment.

More specifically, one of the key advantages of game theory is its ability to capture the strategic interactions among users in a decentralized network. In the context of power control, for example, users aim to maximize their own utility, but their actions also affect the performance of other users due to interference. From that perspective, even if game theory allows us to model this interaction as a non-cooperative game, where each user is a player who selects their power allocation strategy to maximize their individual utility, it is possible to maintain an altruistic view of the system (Scutari *et al.*, 2010). Furthermore, game theory provides flexibility in capturing different system objectives, i.e., by formulating the utility functions appropriately, various performance metrics can be optimized, e.g., maximizing system throughput, minimizing total power consumption, or achieving fairness among users.

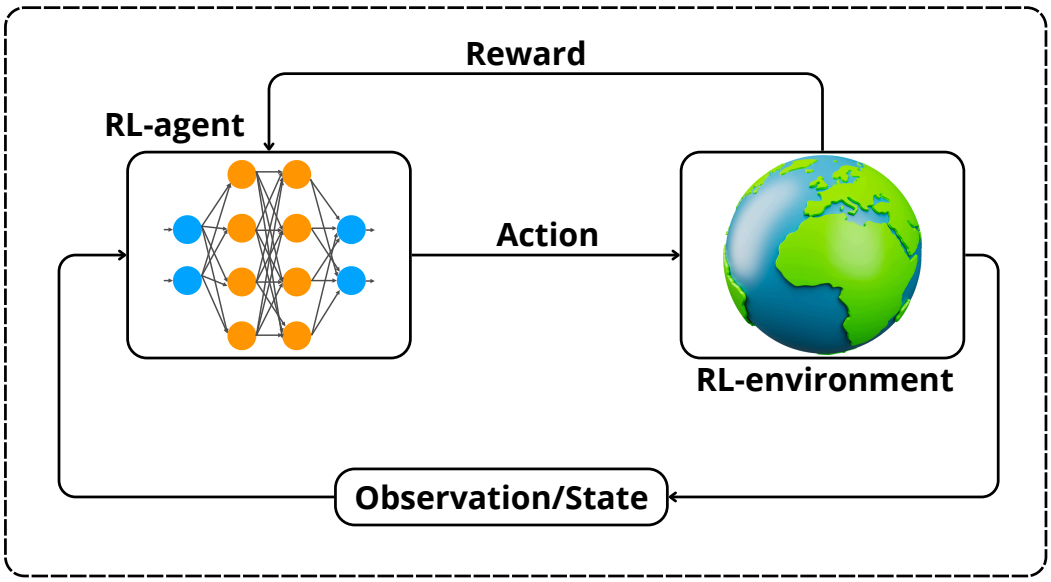
In summary, game theory is a valuable tool for power control in cellular systems since it enables the modeling of strategic interactions among users, facilitating decentralized optimization and achieving a balance between individual user performance and interference management. By applying different game models and solution concepts, game theory offers flexibility in capturing various objectives. Moreover, game theory allows for adaptability to dynamic network conditions, making it suitable for practical and feasible implementation in mobile networks.

#### *1.1.5.4 Deep Reinforcement Learning*

When deep neural networks are used to approximate components of reinforcement learning (RL), such as value functions and/or policy functions, the approach is referred to as DRL. By leveraging the power of deep neural networks, DRL methods can handle high-dimensional input spaces, and learn complex policies directly from these inputs (Lee *et al.*, 2023). This ability to approximate intricate functions allows DRL to achieve remarkable performance in complex tasks, where traditional reinforcement learning methods would struggle.

In general, DRL is in contrast to “shallow” learning. This is because DRL leverages advanced machine learning techniques to dynamically adapt resource allocation policies based on real-time network conditions. Utilizing neural networks and RL algorithms, DRL learns and optimizes complex decision-making processes, significantly improving adaptability and performance in resource management (Li, 2017). Here are some important elements of any

Figure 4 – Schematic structure of deep reinforcement learning.




Source: Created by the author.

method based on RL (Sutton; Barto, 1998):

- **Agent**: The learner or decision-maker that interacts with the environment.
- **Environment**: The external system with which the agent interacts. The environment provides feedback in the form of rewards based on the agent's actions.
- **State**: A representation of the current situation of the environment.
- **Action**: The set of all possible moves the agent can make.
- **Reward**: Feedback from the environment following an action, indicating the immediate benefit of the action.

More specifically, the general idea behind RL-based techniques is to learn from experiences, as shown in Fig. 4. In the case of DRL, the agent is built using neural networks. Thus, through observations of a given environment, the agent decides which action to take in response to a specific state. By evaluating the reward obtained from that action, the agent determines whether the action was beneficial or effective in solving the problem at hand. Over time, this process enables the agent to identify the best actions for each state. As the agent gains more experience, it continuously refines its decision-making, ultimately arriving at the optimal policy and achieving excellent solutions.

In the context of mobile networks, DRL has proven to be a powerful tool for tackling complex RRM problems and has received significant attention in recent literature, see, e.g., (Huang *et al.*, 2023; Zhang *et al.*, 2023; Ye *et al.*, 2024; Nagib *et al.*, 2024). This is due to

DRL's remarkable adaptability and flexibility in addressing a wide range of challenges, including non-convex, combinatorial, and mixed optimization problems that involve both integer and continuous variables. Moreover, it is possible to use multiple independent agents, making decentralized solutions with this approach feasible.

Additionally, it is important to note that the reward in DRL does not necessarily need to be calculated using closed-form mathematical expressions; it can be empirically collected. This demonstrates that DRL can optimize functions that lack closed mathematical formulations (Lillicrap *et al.*, 2015). This capability highlights the abstraction power of DRL, which could be crucial for enabling significant advances in wireless networks in the future.

Therefore, DRL represents a powerful paradigm for training intelligent agents capable of making complex decisions. By combining the strengths of RL and deep learning, DRL has achieved remarkable successes across various domains.

## 1.2 Objectives and Thesis Structure

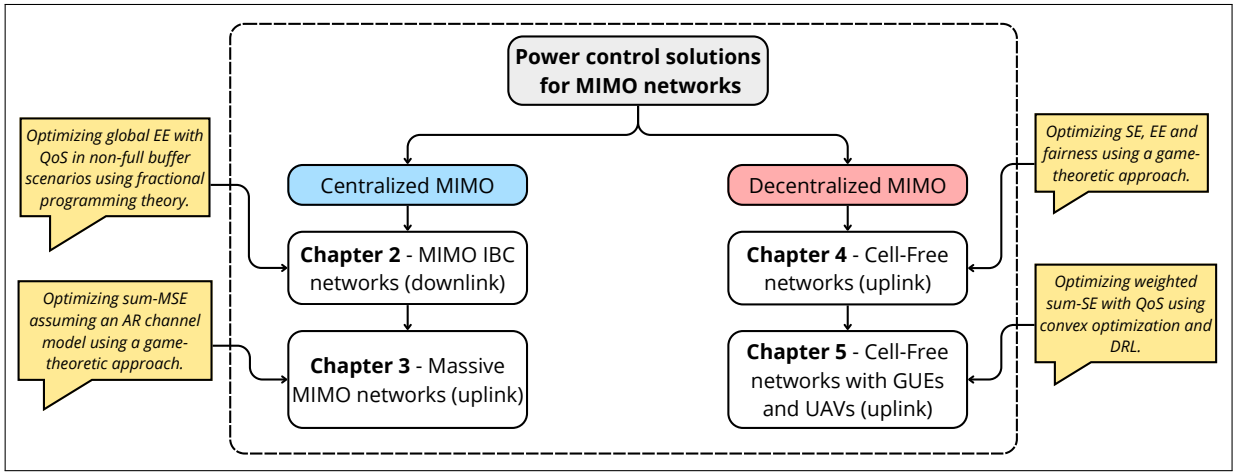
Considering the overview about MIMO networks and RRM presented in the previous sections of this chapter, the main objective of this thesis is to design power control algorithms for centralized and decentralized MIMO systems.

A block diagram illustrating the thesis structure is presented in Figure 5. Observe that the thesis is divided into two independent parts. In the first part, we address power control aspects related to uplink and downlink centralized MIMO systems, including MIMO IBC and uplink massive MIMO systems, which are discussed in Chapter 2 and Chapter 3, respectively. In the second part, we address power control aspects related to uplink decentralized MIMO systems, i.e., uplink cell-free networks, which are discussed in Chapter 4 and Chapter 5.

More specifically, in Chapter 2 the design of transceivers in MIMO IBC networks to optimize the global EE, while maintaining minimum rate constraints for each user, is investigated. However, differently from the existing literature, we assume a non-full buffer traffic model. Additionally, we consider a more realistic channel model that accounts for space and time correlations. In this context, we propose both centralized and decentralized solutions, leveraging fractional programming theory, particularly based on the Dinkelbach method.

Chapter 3 tackles the intricate problem of managing pilot and data power usage in uplink massive MIMO scenarios. In this study, we propose a novel game theory-based decentralized solution that effectively leverages first-order autoregressive (AR) channel models.

Figure 5 – Block diagram for thesis organization.




Source: Created by the author.

This approach offers significant performance benefits, which are thoroughly discussed in that chapter.

In Chapter 4, continuing with a game theory framework, we propose a potential game for data power control in uplink cell-free networks. This decentralized solution is versatile and can address various network objectives, including SE, EE, and fairness.

Chapter 5 aims to efficiently integrate aerial and terrestrial users into uplink cell-free networks. Specifically, we investigate the problem of maximizing the weighted sum of SEs while ensuring minimum SE requirements for each user. To tackle this problem, we propose both centralized and decentralized solutions, utilizing convex optimization and DRL techniques.

Finally, Chapter 6 draws the main conclusions taken from the solutions and results presented in this thesis along with some directions for possible future works.

## 1.3 Scientific Contributions

Currently, the content of this thesis is published or is under review with the following bibliographic information:

The contents of Chapters 1, 2, 3 and 4 are published or under review with the following bibliographic information:

[J1] Saraiva, J. V.; Antonioli, R. P.; Fodor, G.; Braga, I. M.; Freitas, W. C.; Silva, Y. C. B.; Silva, C. F. M. e. *Energy Efficiency Maximization Under Minimum Rate Constraints in Multi-Cell MIMO Systems With Finite Buffers*. **IEEE Transactions on Green Communications and Networking**, v. 5, n. 1, p. 174–189, 2021. DOI: 10.1109/ TGCN.2020.3043049.

[J2] Saraiva, J. V.; Antonioli, R.; Fodor, G.; Braga, I.; Freitas, W.; Silva, Y. *A Network-Assisted Game-Theoretic Design to Power Control in Autoregressive Fading Channels*. **IEEE Communications Letters**, v. 26, n. 7, p. 1663–1667, 2022. DOI: 10.1109/LCOMM.2022.3167367.

[C1] Saraiva, J. V.; Antonioli, R. P.; Fodor, G.; Freitas,W. C.; Silva, Y. C. B. *A Distributed Game-Theoretic Solution for Power Management in the Uplink of Cell-Free Systems*. **Proceedings of the IEEE Globecom Workshops (GC Wkshps)**, Rio de Janeiro, Brazil, 04-08 Dec., 2022. DOI: 10.1109/ GCWkshps56602.2022.10008611.

[C2] Saraiva, J. V.; Antonioli, R. P.; Fodor, G.; Freitas, W. C.; Silva, Y. C. B. *Adaptive Power Management for Maximizing Spectral Efficiency and QoS in Cell-Free Systems*. **Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS)**, Rio de Janeiro, Brazil, 14-17 Jul., 2024. DOI: 10.1109/ISWCS61526.2024.10639066.

[J3] Saraiva, J. V.; Antonioli, R. P.; Fodor, G.; Freitas, W. C.; Silva, Y. C. B. *Integrating Aerial and Ground Users: Deep Reinforcement Learning and Convex Optimization for Data Power Control in Cell-Free Networks*, 2024 (*To be submitted*).

[J4] Saraiva, J. V.; Antonioli, R. P.; Fodor, G.; Freitas, W. C.; Silva, Y. C. B. *Optimizing Data Power and Clustering for Terrestrial and Aerial Users in Cell-Free Systems*, 2024 (*To be submitted*).

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

- **UFC.47** - *Network Assisted Intelligent Vehicle-to-Everything communications (NAIVE)*, November/2018 - April/2020,
- **UFC.48** - *User Centric Networks and Reconfigurable Surfaces for Next Generation Wireless*, December/2020 - December/2022,
- **UFC.50** - *User-Centric Cell-free Massive MIMO for Airspace Coverage*, December/2022 - December/2024.

Within the framework of these projects, the author also collaborated in the following scientific works that are not directly related to the content of this thesis:

***Journal Papers***

[J5] Saraiva, J. V.; Braga Junior, I. M.; Monteiro, V. F.; Lima, F. R. M.; Maciel, T. F.; Freitas Junior, W. d. C.; Cavalcanti, F. R. P. *Deep Reinforcement Learning for QoS-Constrained*

*Resource Allocation in Multiservice Networks*. **Journal of Communication and Information Systems**, v. 35, n. 1, p. 66–76, Apr. 2020. DOI: 10.14209/jcis.2020.7 (*Invited Paper*).

[J6] Saraiva, J. V.; Lima, F. R. M.; Pessoa, A. M.; Maciel, T. F.; Freitas Junior, W. d. C.; Cavalcanti, F. R. P. *Energy-Efficient Radio Resource Allocation For Dual-Hop Relay-Assisted Orthogonal Frequency Division Multiple Access Systems With Quality of Service Provisioning*. **Transactions on Emerging Telecommunications Technologies**, v. 32, n. 10, e4293, 2021. DOI: https://doi.org/10.1002/ett.4293.

***Conference Papers***

[C3] Saraiva, J. V.; Monteiro, V. F.; Lima, F. R. M.; Maciel, T. F.; Cavalcanti, F. R. P. *A Q-learning Based Approach to Spectral Efficiency Maximization in Multiservice Wireless Systems*. **Proceedings of the Simpósio Brasileiro de Telecomunicações e Processamento de Sinais**, Petrópolis, Brazil, 29 Sept. to 2 Out, 2019.

[C4] Saraiva, J. V.; Antonioli, R. P.; Fodor, G.; Lima, F. R. M.; Freitas Jr,W. C.; Silva, Y. C. *Fair Adaptive Power Allocation for Cell-Free Massive MIMO Systems*. **Proceedings of the Simpósio Brasileiro de Telecomunicações e Processamento de Sinais**, Fortaleza, Brazil, 26-29 Sept., 2021.

***Patents***

[P1] Antonioli, R. P.; Lobão, M. C. R.; Saraiva, J. V.; Fodor, G.; Silva, Y. C. B.; Jr,W. C. F. Transmission and Reception Point Clustering of Aerial Vehicles. Apr. 2024. Patent Application.

# 2 MAXIMIZING THE GLOBAL ENERGY EFFICIENCY IN MIMO IBC NETWORKS

This chapter addresses the multiple input multiple output (MIMO) transceiver design problem for energy efficiency (EE) maximization in the downlink of finite-buffer multicell systems. Unlike previous works, our problem formulation takes into account per-user minimum rate requirements. We arrive at a nonconvex fractional optimization problem, which is hard to tackle. By exploiting the properties of fractional programming, and using Dinkelbach's method, the resulting fractional form optimization problem is transformed to an equivalent optimization problem in subtractive form.

Next, the nonconvexity of this problem is handled using successive convex approximation, leading to iterative centralized and decentralized resource allocation solutions. Finally, considering a realistic channel model with space, frequency and time correlations, numerical results confirm the effectiveness of the proposed algorithms and indicate significant performance gains in terms of achieved EE over existing solutions for full and finite-buffer models.

## 2.1 Introduction

The rapid development of wireless communications has seen a dramatic and inevitable increase in the number of mobile devices, subscriptions and deployed infrastructure nodes. By $2025$, 5th generation (5G) mobile networks are expected to serve $2.6$ billion subscriptions covering up to $65\%$ of the world's population and generating $45\%$ of the world's total mobile data traffic (Ericsson, 2019). This new mobile world could lead to excessive energy demand and impose a heavy burden on the environment. Indeed, the energy consumption and gas emissions of information and communication technologies (ICT) infrastructures are reaching worrying proportions. The study in (Huq *et al.*, 2015) suggests that $3\%$ of the worldwide energy is consumed by communication networks. Meanwhile, ICT infrastructures are responsible for approximately $2\%$ of the worldwide $CO_2$ emissions (Busari *et al.*, ). Although these percentages may seem small, they are expected to increase due to the advent of 5G, which may contribute to a sharp growth of energy usage and greenhouse gas emissions unless proper design approaches and deployment practices are implemented (I *et al.*, 2014).

In this context, an important part of the solution lies in optimizing the EE of ICT systems, which can be defined as the ratio between the system throughput and the corresponding consumed energy (Isheden; Fettweis, ; Isheden *et al.*, 2012b). The advantages of focusing on EE

include making ICT systems sustainable and reducing considerably operational expenditures. Recognizing this objective, designing energy-aware architectures and green ICT systems has attracted considerable interest in the existing literature (Kolawole *et al.*, 2020; Chang *et al.*, 2017).

As an essential technology in long term evolution advanced (LTE-A) or 4th generation (4G) of cellular networks, MIMO systems provide high peak data rates, low latency and improved spectral efficiency (SE) and EE in modern mobile systems. Due to these capabilities, MIMO technology continues to play an important role in the next generations of wireless networks, such as 5G and beyond (Gogoi *et al.*, ; Misilmani; El-Hajj, ; Ngo *et al.*, 2018). Furthermore, the use of MIMO combined with other technologies, such as orthogonal frequency division multiplexing (OFDM), proves to be an even more attractive and robust system for currently deployed and future wireless broadband systems. One of the main advantages of MIMO-OFDM systems is their structure, which allows radio resources, such as subcarrier, power and bit, to be dynamically adapted (Zhang; Letaief, ; Pengfei Xia *et al.*, 2004; Venkatraman *et al.*, 2016; Basturk; Chen, 2020).

Generally, in multi-cell MIMO systems, transceiver designs that maximize a certain system utility are essential to provide good system performance. Therefore, MIMO transceiver designs with various objectives have been widely investigated in the literature. Conventional resource allocation schemes, such as the solution of sum-rate utility maximization problem subject to constraints including total transmit power, packet delay and/or quality of service (QoS) (for example, a minimum per-user rate) were considered in many recent works (see Section 2.1.1).

However, taking EE metrics into account, and striking a good balance between SE and EE targets, while meeting engineering constraints arising in decentralized network architectures, requires further research (Mahapatra *et al.*, 2016). Furthermore, the finite-buffer case is an important aspect for mobile network applications. Nevertheless, many works do not take the buffer size into account and make the unrealistic assumption that the packet queue length is infinitely large. In practice, however, network buffers are finite, and neglecting this aspect is clearly an engineering problem.

Finite-buffer traffic can be modeled using a simulation parameter called packet arrival rate, which is commonly used to control the offered traffic intensity in the system. When the traffic intensity in the system increases, the number of packets to be transmitted from the

base stations to the users also increases. At high traffic intensity, it becomes more challenging to satisfy QoS requirements in terms of minimum rate or maximum latency, since more packets should be transmitted, and there is only a limited number of sub-channels that can be used for downlink transmission. Furthermore, since there are more packets to be transmitted until buffer saturation, the base stations need to use higher transmission power to increase the data rate of the users, which may also have an impact on the energy efficiency of the system. The opposite occurs when the traffic intensity decreases.

Therefore, our objective in this work is to address the MIMO transceiver design problem for energy efficiency maximization in the downlink of finite-buffer multicell systems. Unlike previous works, our problem formulation takes into account per-user minimum rate requirements and utilizes finite-buffer models. Next, we discuss the closely related works and our contributions to the state of the art.

### *2.1.1 Related Works*

Although EE is becoming the mainstream for future wireless networks design, improving EE without impairing SE has been a challenge. On one hand, due to the scarce bandwidth resources, many relevant works in the literature have considered scenarios with multiple antennas, focusing on the SE, see, for example, (Shi *et al.*, 2011; Razaviyayn *et al.*, 2014; Joshi *et al.*, 2012; Oguejiofor; Zhang, ; Kaleva *et al.*, 2016; Antonioli *et al.*, 2019). In particular, in (Shi *et al.*, 2011) the MIMO interference broadcast channel (IBC) was considered, in which the authors proposed a linear transceiver design algorithm for weighted sum-rate maximization based on the iterative minimization of the weighted mean squared error (MSE). The scheme from (Shi *et al.*, 2011) was then combined with a user grouping algorithm in (Razaviyayn *et al.*, 2014).

Alternative approaches to solve the sum-rate maximization problem, such as via branch and bound in multiple input single output (MISO) systems, were proposed in (Joshi *et al.*, 2012) and (Oguejiofor; Zhang, ). However, (Shi *et al.*, 2011; Razaviyayn *et al.*, 2014; Joshi *et al.*, 2012; Oguejiofor; Zhang, ) focused on the sum-rate maximization problem without requiring per-link minimum rate demands, which is a critical issue for modern communication networks. In this context, (Kaleva *et al.*, 2016) and (Antonioli *et al.*, 2019) proposed frameworks considering QoS constraints in terms of per-link minimum rate in MIMO IBC systems. Therein, the authors proposed centralized, distributed and/or semi-distributed solutions that employed methods such

as successive convex approximation (SCA), difference of convex functions program, Lagrangian relaxation and branch and bound.

On the other hand, as previously noted, EE is attracting more and more attention in both industry and academia due to the growing energy consumption in telecommunications. Nevertheless, all of the above-mentioned works generally proposed algorithms in which the energy consumption is extremely high, which causes low EE, in contrast to the concept of green communications. Therefore, recent efforts have shifted towards dealing with EE-oriented optimization problems.

In general, these problems consist of either minimizing the total transmit power or maximizing the ratio between the transmission rate and the corresponding consumed energy. In MIMO systems, problems related to sum-power minimization were investigated in (Cavalcante *et al.*, 2018) and (Cavalcante *et al.*, 2019). Basically, these papers considered sum-power minimization using time division duplexing (TDD) while imposing signal to interference-plus-noise ratio (SINR) constraints for each downlink transmission, as well as maximum power constraints for base station (BS)-to-BS interference.

Minimizing the sum-power has been adopted as a conventional metric to improve the greenness of cellular systems. However, this metric focuses on the amount of power that is used to transmit data, which means that the power consumption due to electronic circuits is generally not taken into account. On the other hand, EE, as a ratio of throughput to energy consumption, becomes a critical performance metric for analyzing how efficiently energy is used for each transmitted bit. In other words, improving EE can be viewed as finding the optimal trade-off between the sum-rate and total energy usage. Obviously, this leads to more complex optimization problems, which in this case are typically nonlinear fractional problems in nature. General frameworks to solve this type of problem were presented in (Isheden; Fettweis, ; Isheden *et al.*, 2012b). In more general approaches, parametric solutions based on Dinkelbach's method (Dinkelbach, 1967) have been a common mathematical tool in the research of mobile networks. The efficiency of Dinkelbach's approach mainly depends on the efficiency to solve a subproblem for a given parameter, i.e., it depends on whether the parameterized problem can be effectively handled or not.

For example, along the direction of orthogonal multiple access without multiple antennas, that strategy was successfully adopted in (Saraiva *et al.*, ) and (Venturino *et al.*, 2015). In (Saraiva *et al.*, ) in particular, the authors focused only on QoS-constrained global

EE maximization, and the proposed solution was based on the Dinkelbach and branch and bound methods. In (Venturino *et al.*, 2015), three definitions of EE were considered for system design, including global EE, but without QoS requirements. In the proposed solutions, the authors used iterative algorithms based on Dinkelbach's approach and Karush-Kuhn-Tucker (KKT) conditions. Considering multi-antenna transmissions, in (Basturk; Chen, 2020), the EE maximization problem with minimum per-user rate requirement was defined for downlink MISO-orthogonal frequency division multiple access (OFDMA)-based cellular networks. The proposed solution was divided into two parts, which combined the Dinkelbach's method and a heuristic-based resource allocation scheme.

Regarding multi-user MIMO systems, the EE maximization problem can be more challenging, since the obtained subproblem can still be nonconvex even for a fixed Dinkelbach's parameter, which means that optimal solutions can be difficult to find. In (He *et al.*, 2013) and (Li *et al.*, ), only suboptimal solutions were achieved by an approach based on an equivalence between spectral efficiency maximization and MSE minimization. However, due to the ability to deal with nonconvexity, SCA techniques have attracted wide attention in various contexts, including EE-driven mobile communications. SCA approximates the original non-convex problem by convex subproblems, which are generally much easier to tackle than the original problem. Thus, by successively solving the approximation problems, a smooth point can be finally achieved (Dong *et al.*, 2020). For the case of EE optimization, SCA techniques were recently employed in (Yang *et al.*, 2019), where the authors proposed novel iterative algorithms that were also based on Dinkelbach's method. However, few studies have applied SCA-based techniques to solve EE-oriented optimization problems, especially considering multi-cell MIMO scenarios.

Moreover, although SE and EE are essentially contradictory objectives, a common point in all of the aforementioned works is the full-buffer model used in their analyses. While the full-buffer traffic model has been widely adopted in the literature for theoretical investigations, it may lead to an inaccurate assessment of the actual performance benefits in realistic scenarios (Ameigeiras *et al.*, 2012). Thus, adopting a finite-buffer model is more appropriate for practical scenarios. In the context of traffic-aware resource allocation, some studies have been concerned with the queue minimization problem. In (Seong *et al.*, 2006) a power allocation problem was addressed for minimizing the number of backlogged packets using geometric programming.

Considering wireless networks, this same problem was addressed and extended in (Weeraddana *et al.*, 2011) by formulating the corresponding user queues as the weights in the weighted sum-rate maximization problem. In MIMO IBC systems, the problem of efficient precoder design was investigated in (Venkatraman *et al.*, 2016), with the purpose of minimizing the number of queued packets in the coordinating BSs. Since the problem is nonconvex, the authors used the combination of SCA and alternating optimization to handle nonconvex constraints in the formulation of the problem. Other approaches, such as the equivalence between the SINR and the MSE, were also exploited in the solutions presented therein. However, from the point of view of energy consumption, approaches that minimize buffer size and maximize data rate are similar, and, consequently, lead to low EE.

In single-antenna scenarios, traffic-aware energy optimization was addressed in (Saxena *et al.*, 2014) and (Pan *et al.*, 2018). More specifically, in (Saxena *et al.*, 2014) a cooperative framework was proposed for reducing energy usage in long term evolution (LTE) networks, while optimizing traffic awareness. In (Pan *et al.*, 2018) two energy optimization strategies were proposed to improve the greenness of future cellular systems. One of the schemes was used to eliminate the power consumption impact of the BS, while keeping satisfactory QoS in terms of instantaneous traffic variance by a dynamic recovery mechanism, whereas the other aimed to enhance the channel utilization and EE for the edge BSs under neighbor cell cooperation. In this context, multi-cell MIMO and single-cell MISO systems were investigated in (Lakshminarayana *et al.*, ) and (Akra; Assaad, ), respectively. However, neither of these works considered EE appropriately modeled as the ratio of rate to power.

### *2.1.2 Main Contributions*

Motivated by the aforementioned observations and by the fact that recent works on MIMO systems have suggested that it is important to use finite-buffer models, because they may lead to better transceiver designs and more accurate performance analyses than full-buffer traffic models (3GPP, 2017b; Asplund, 2020), the main contributions of our work are:

1. Formulation of an optimization problem to improve the ratio between the sum-rate and the power consumption, which is referred to as global EE, subject to minimum data rate constraints for multi-cell MIMO-OFDM systems with a finite-buffer model.
2. Proposal of a centralized solution based on SCA in combination with Dinkelbach's approach, which solves the analyzed optimization problem up to a locally optimal solution by

performing a non-trivial SCA over the nonconvex constraints.

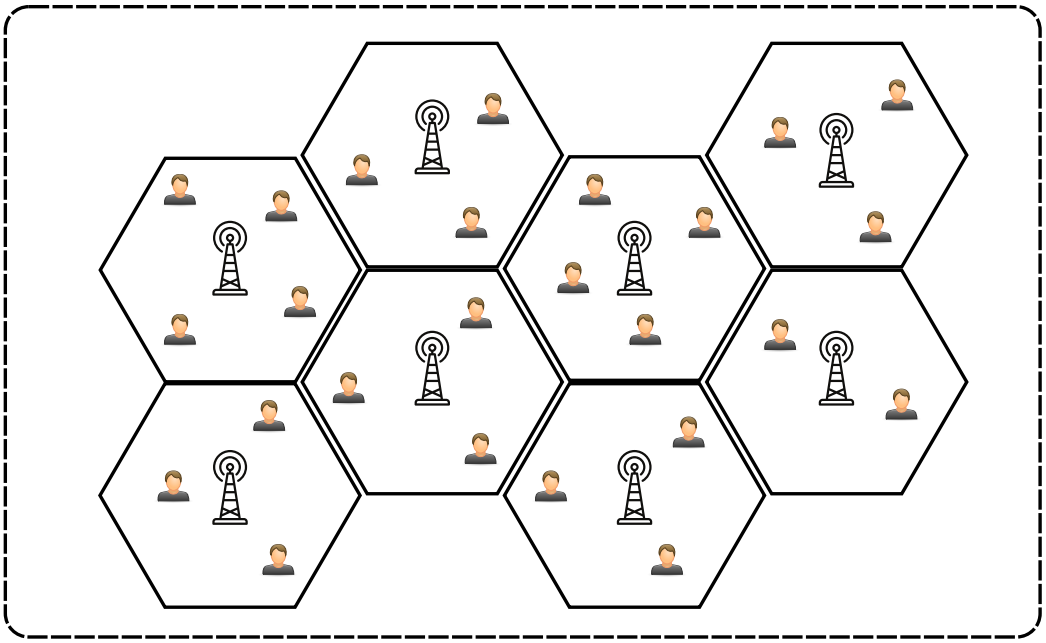
3. In addition, a practical decentralized algorithm based on a non-trivial dual decomposition is also provided, in which feasible initialization requirements are relaxed by applying the Lagrangian relaxation of the rate constraints. We also describe a practical over-the-air and backhaul signaling scheme to support the decentralized solution in practical scenarios.
4. Converge analysis showing that the combination of the SCA and Dinkelbach's approaches used in both centralized and decentralized algorithms converge to a KKT point of the formulated problem, which required us to organize the *inner* and *outer loops* of both algorithms such that convergence could be proven.
5. Performance evaluation by means of simulations, in which we compare the proposed solution with state-of-the-art heuristic and optimization-based algorithms. Unlike previous works, the proposed framework is analyzed using a realistic channel modeling based on the 3rd generation partnership project (3GPP) stochastic channel model with spatial, frequency and time correlations. We compare the performance of the proposed centralized and decentralized solutions and show the performance gap between them. Moreover, we evaluate our proposed solution considering perfect channel state information (CSI) estimation in terms of EE, number of backlogged bits in the buffer and consumed power for different system parameters, such as the number of sub-channels, average number of packet arrivals, maximum allowable transmit power, dynamic circuit power and per-user QoS demands. Besides that, we investigate the performance of the proposed and benchmarking algorithms under imperfect CSI estimation to better understand its impact on the global EE.

## 2.2 Network Model

The scenario examined in this chapter involves a MIMO IBC system with multiple users according to Figure 6.

Specifically, we consider the downlink of a multi-cell MIMO-OFDM system with $N$ sub-channels, in which a total of $B$ BSs equipped with $N_{\mathrm{T}}$ antennas serve in total $U$ multi-antenna user equipments (UEs), each one equipped with $N_{\mathrm{R}}$ antennas. We denote by $\mathcal{N} = \{1, 2, \ldots, N\}$ the set of sub-channel indices and by $\mathcal{B} = \{1, 2, \ldots, B\}$ the set of BS indices available in the system. The set of UEs associated with BS $b$ is denoted by $\mathcal{U}_b$, with $U_b = |\mathcal{U}_b|$, where each UE $u$ is served by a single BS $b_u$, i.e., $\mathcal{U}_b \cap \mathcal{U}_{\tilde{b}} = \emptyset,\ \forall b, \tilde{b} \in \mathcal{B}$ and $b \neq \tilde{b}$. Let $S_{u,n}$ denote a fixed

Figure 6 – A MIMO IBC system with multiple users.



Source: Created by the author.

number of spatial streams allocated to UE $u$ on sub-channel $n$.

The downlink signal received by UE $u$ over spatial stream $s$ and sub-channel $n$ can be expressed as

$$\mathbf{y}_{u,s,n} = \mathbf{H}_{b_u,u,n}\mathbf{m}_{u,s,n}x_{u,s,n} + \sum_{i=1}^{U}\sum_{\substack{j=1,\\(i,j)\neq(u,s)}}^{S_{i,n}} \mathbf{H}_{b_i,u,n}\mathbf{m}_{i,j,n}x_{i,j,n} + \mathbf{n}_{u,n}, \tag{2.1}$$

where $\mathbf{H}_{b_i,u,n} \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{T}}$ is the channel matrix between UE $u$ and BS $b$ serving UE $i$ on sub-channel $n$, $\mathbf{m}_{u,s,n} \in \mathbb{C}^{N_\mathrm{T}}$ is the transmit beamforming vector of the corresponding data stream, $x_{u,s,n}$ is the mutually independent transmitted data symbol with $\mathbb{E}\left[|x_{u,s,n}|^2\right] = 1$ and $\mathbf{n}_{u,n} \in \mathbb{C}^{N_\mathrm{R}} \sim \mathcal{CN}(0,\sigma^2)$ is the noise at UE $u$ and sub-channel $n$. UE $u$ decodes the signal $\mathbf{y}_{u,s,n}$ via a receive beamformer $\mathbf{w}_{u,s,n} \in \mathbb{C}^{N_\mathrm{R}}$.

The SINR for stream $s$ and sub-channel $n$ of UE $u$ is

$$\Gamma_{u,s,n} = \frac{|\mathbf{w}_{u,s,n}^\mathrm{H}\mathbf{H}_{b_u,u,n}\mathbf{m}_{u,s,n}|^2}{\sum\limits_{i=1}^{U}\sum\limits_{\substack{j=1,\\(i,j)\neq(u,s)}}^{S_{j,n}} |\mathbf{w}_{u,s,n}^\mathrm{H}\mathbf{H}_{b_i,u,n}\mathbf{m}_{i,j,n}|^2 + \sigma^2\|\mathbf{w}_{u,s,n}\|^2}. \tag{2.2}$$

Our derivations assume perfect CSI estimation at the transmitters and receivers. This has also been common in many relevant works in the literature, see, e.g., (Shi *et al.*, 2011; Kaleva *et al.*, 2016; Pennanen *et al.*, 2016; Venkatraman *et al.*, 2016).

Regarding our finite-traffic model, we denote as $Q_u$ the number of backlogged bits intended to UE $u$ at a given scheduling instant. Furthermore, in order to mathematically model the queue dynamics of UE $u$, we employ a packet arrival process based on the Poisson distribution[1]. Let $\lambda_u(i)$ represent the instantaneous number of bits arriving for UE $u$ at the $i$-th time instant, then the total number of queued bits at the $(i+1)$-th instant for UE $u$, denoted as $Q_u(i+1)$, is given by

$$Q_u(i+1) = [Q_u(i) - \Delta_{\text{tti}} t_u(i)]^+ + \lambda_u(i), \tag{2.3}$$

where $\Delta_{\text{tti}}$ is the duration of one transmission time interval (TTI) and $t_u$ denotes the number of transmitted bits per second for UE $u$. At the $i$-th instant, the transmission rate normalized over a unit bandwidth of UE $u$ is given by

$$t_u(i) = \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n}(i), \tag{2.4}$$

where $t_{u,s,n} = \log_2(1 + \Gamma_{u,s,n})$ denotes the number of transmitted bits per second over the $s$-th stream and $n$-th sub-channel of UE $u$ for the SINR $\Gamma_{u,s,n}$. This traffic model is similar to the one used in (Venkatraman *et al.*, 2016).

## 2.3 Problem Formulation

We propose a novel formulation of the transceiver design problem for the QoS-constrained weighted EE maximization (Q-WEEM) problem under per-BS maximum power and per-UE minimum rate constraints:

$$\underset{\substack{\mathbf{w}_{u,s},\mathbf{m}_{u,s},\\ t_{u,s,n}}}{\text{maximize}} \quad \frac{\sum_{u=1}^{U} \gamma_u \left( \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} \right)}{\sum_{n=1}^{N} \sum_{u=1}^{U} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2 + \zeta} \tag{2.5a}$$

$$\text{subject to} \quad \xi_u \leq \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} \leq \frac{Q_u}{\Delta_{\text{tti}}}, \qquad \forall u, \tag{2.5b}$$

$$\sum_{n=1}^{N} \sum_{u \in \mathcal{U}_b} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2 \leq P_b, \qquad \forall b, \tag{2.5c}$$

$$t_{u,s,n} \leq \log_2(1 + \Gamma_{u,s,n}), \qquad \forall u, s, n, \tag{2.5d}$$

[1] Considering this packet arrival process and a given packet size distribution, we can determine the number of arriving bits, which will be used throughout the chapter.

where $\gamma_u > 0$ denotes the priority weight of UE $u$, $\xi_u$ models the rate requirement of UE $u$, $P_b$ is the power budget of BS $b$ and $\zeta$ accounts for the circuit power consumption. Thus, while the transmission power is used for data transmission, the circuit power is a constant quantity accounting for the dissipation in analog hardware, digital signal processing, backhaul signaling, and other overhead costs (such as cooling and power supply losses) (Björnson *et al.*, 2015). To avoid excessive allocation of the resources, the sum of bits transmitted to UE $u$ over all the sub-channels and streams cannot be higher than the amount of bits or number of backlogged bits available or waiting in the buffer $Q_u$, as modeled in (2.5b). Constraints (2.5c) control the power budget of each BS $b$. Finally, constraints (2.5d) model the rate of each UE $u$ on sub-channel $n$ over stream $s$ in a relaxed format, which will be used later to turn the optimization problem into a convex form.

The optimization variables are transmit beamforming vectors $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)} \in \mathbb{C}^{N_{\mathrm{T}}}$, receive beamforming vectors $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)} \in \mathbb{C}^{N_{\mathrm{R}}}$ and the number of transmitted bits over the $s$-th stream and $n$-th sub-channel of UE $u$, i.e., $t_{u,s,n}$. Thus, the formulated problem handles the transceiver design for global EE maximization subject to per-UE minimum rate requirements, while also considering finite-buffer traffic models. It is worth noting that problem (2.5) is solved for each time instant $i$.

In problem (2.5), the objective function is defined as the ratio between the achievable rate and the total power consumed by the system, which makes (2.5) a fractional programming problem. Fortunately, a mathematical framework that provides insights into this class of optimization problems can be found in (Dinkelbach, 1967; Charnes; Cooper, 1962; Schaible, 1974; Shen; Yu, 2018; Ródenas *et al.*, 1999; Jagannathan, 1966). However, in general, algorithms for computing the solution of fractional problems require the problem to have a convex optimization domain, which clearly does not occur in (2.5). Therefore, approximation based approaches are required for deriving a tractable and practical solution for problem (2.5).

## 2.4 Problem Reformulation and Centralized Solution

In this section a centralized solution is proposed. First, the original EE optimization problem (2.5) is reformulated in Section 2.4.1, then a method based on fractional programming is presented in Section 2.4.2 for solving the problem.

### *2.4.1 Problem Reformulation*

Due to the non-convexity of problem (2.5), we first exploit the relationship between the MSE and the achievable SINR, when minimum mean squared error (MMSE) receivers are employed (Shi *et al.*, 2011; Christensen *et al.*, 2008). Therefore, given the MMSE receiver assumption, we can take some advantage in reformulating our problem. Let $\epsilon_{u,s,n}$ be the MSE for UE $u$ and stream $s$ on sub-channel $n$, given by (Shi *et al.*, 2011; Christensen *et al.*, 2008):

$$\epsilon_{u,s,n} \triangleq \mathbb{E}\big[|\mathbf{w}_{u,s,n}^{\mathrm{H}}\mathbf{y}_{u,s,n} - x_{u,s,n}|^2\big] = \mathbb{E}\big[\mathbf{w}_{u,s,n}^{\mathrm{H}}\mathbf{y}_{u,s,n}\mathbf{y}_{u,s,n}^{\mathrm{H}}\mathbf{w}_{u,s,n}\big] - \mathbb{E}\big[\mathbf{w}_{u,s,n}^{\mathrm{H}}\mathbf{y}_{u,s,n}x_{u,s,n}^{*}\big] - \mathbb{E}\big[x_{u,s,n}\mathbf{y}_{u,s,n}^{\mathrm{H}}\mathbf{w}_{u,s,n}\big] + \underbrace{\mathbb{E}\big[x_{u,s,n}x_{u,s,n}^{*}\big]}_{=1}.$$

Substituting $\mathbf{y}_{u,s,n}$ from (2.1), we can write:

$$\begin{aligned}\epsilon_{u,s,n} &\triangleq \mathbb{E}\big[|\mathbf{w}_{u,s,n}^{\mathrm{H}}\mathbf{y}_{u,s,n} - x_{u,s,n}|^2\big] \\ &= |1 - \mathbf{w}_{u,s,n}^{\mathrm{H}}\mathbf{H}_{b_u,u,n}\mathbf{m}_{u,s,n}|^2 + \sigma^2\|\mathbf{w}_{u,s,n}\|^2 + \sum_{i=1}^{U}\sum_{\substack{j=1 \\ (i,j)\neq(u,s)}}^{S_{i,n}} |\mathbf{w}_{u,s,n}^{\mathrm{H}}\mathbf{H}_{b_i,u,n}\mathbf{m}_{i,j,n}|^2, \end{aligned} \tag{2.6}$$

and the MMSE receiver for stream $s$ of UE $u$ on sub-channel $n$ is given by (Shi *et al.*, 2011; Christensen *et al.*, 2008):

$$\mathbf{w}_{u,s,n} = \left(\sum_{i=1}^{U}\sum_{j=1}^{S_{i,n}} \mathbf{H}_{b_i,u,n}\mathbf{m}_{i,j,n}\mathbf{m}_{i,j,n}^{\mathrm{H}}\mathbf{H}_{b_i,u,n}^{\mathrm{H}} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{H}_{b_u,u,n}\mathbf{m}_{u,s,n}. \tag{2.7}$$

One can find the detailed derivation of (2.7) for a more general case with CSI imperfection in (Fodor *et al.*, 2015; Eraslan *et al.*, 2013), which can be easily adapted to the case of perfect CSI adopted herein. Considering that MMSE receivers are used, the following useful relation holds (Christensen *et al.*, 2008; Shi *et al.*, 2011):

$$\epsilon_{u,s,n}^{-1} = 1 + \Gamma_{u,s,n}. \tag{2.8}$$

Note that, for fixed receive beamformers, the MSE expression in (2.6) is a convex function in terms of the transmit beamforming vectors $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)}$.

By replacing (2.8) in (2.5), we can reformulate the original problem as

$$\underset{\substack{\epsilon_{u,s,n},\, t_{u,s,n},\\ \mathbf{m}_{u,s,n}}}{\text{maximize}} \quad \frac{\sum_{u=1}^{U} \gamma_u \left( \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} \right)}{\sum_{n=1}^{N} \sum_{u=1}^{U} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2 + \zeta} \tag{2.9a}$$

$$\text{subject to} \quad t_{u,s,n} \leq -\log_2(\epsilon_{u,s,n}), \quad \forall u, s, n, \tag{2.9b}$$

$$|1 - \mathbf{w}_{u,s,n}^{\mathrm{H}} \mathbf{H}_{b_u,u,n} \mathbf{m}_{u,s,n}|^2 + \sum_{i=1}^{U} \sum_{\substack{j=1,\\ (i,j)\neq(u,s)}}^{S_{u,n}} |\mathbf{w}_{u,s,n}^{\mathrm{H}} \mathbf{H}_{b_i,u,n} \mathbf{m}_{i,j,n}|^2 + \sigma^2 \|\mathbf{w}_{u,s,n}\|^2 \leq \epsilon_{u,s,n}, \quad \forall u, s, n, \tag{2.9c}$$

(2.5b) and (2.5c).

### *2.4.2 Centralized Solution*

In order to find a solution to problem (2.9), firstly note that we can rewrite it as

$$\underset{\substack{\epsilon_{u,s,n},\, t_{u,s,n},\\ \mathbf{m}_{u,s}}}{\text{maximize}} \quad \frac{\phi(\epsilon_{u,s,n},\, t_{u,s,n}, \mathbf{m}_{u,s,n})}{\psi(\epsilon_{u,s,n},\, t_{u,s,n}, \mathbf{m}_{u,s,n})} \tag{2.10a}$$

$$\text{subject to} \quad \{\epsilon_{u,s,n}, t_{u,s,n}, \mathbf{m}_{u,s,n}\}_{\forall(u,s,n)} \in \mathbf{\Omega}, \tag{2.10b}$$

where

$$\phi(\epsilon_{u,s,n},\, t_{u,s,n}, \mathbf{m}_{u,s,n}) = \sum_{u=1}^{U} \gamma_u \left( \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} \right), \tag{2.11a}$$

$$\psi(\epsilon_{u,s,n},\, t_{u,s,n}, \mathbf{m}_{u,s,n}) = \sum_{n=1}^{N} \sum_{u\in\mathcal{U}} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2 + \zeta, \tag{2.11b}$$

and $\mathbf{\Omega}$ is a nonempty and compact set, which includes the constraints (2.5b), (2.5c), (2.9b) and (2.9c). Furthermore, note that $\phi(\cdot)$ and $\psi(\cdot) > 0$ are continuous, differentiable and real-valued functions of $\{\epsilon_{u,s,n},\, t_{u,s,n}, \mathbf{m}_{u,s,n}\}_{\forall(u,s,n)} \in \mathbf{\Omega}$.

Fortunately, the optimal solution of problem (2.9) can be obtained by Dinkelbach's algorithm (Dinkelbach, 1967; Ródenas *et al.*, 1999), whose fundamental idea is to determine the root of a function $F(\cdot)$ in an equivalent parametric problem. This algorithm is based on a theorem by Jagannathan (Jagannathan, 1966) concerning the relationship between fractional and parametric programming, as stated in Theorem 1.

**Theorem 1** ((Dinkelbach, 1967; Ródenas *et al.*, 1999; Jagannathan, 1966)). *$\eta^\star$ is the optimal solution of problem* (2.9)*, i.e.,* $\eta^\star = \frac{\phi(\epsilon^\star_{u,s,n}, t^\star_{u,s,n}, \mathbf{m}^\star_{u,s,n})}{\psi(\epsilon^\star_{u,s,n}, t^\star_{u,s,n}, \mathbf{m}^\star_{u,s,n})} = \underset{\substack{\epsilon_{u,s,n},\, t_{u,s,n},\\ \mathbf{m}_{u,s,n}\in\boldsymbol{\Omega}}}{\textit{maximum}} \left\{ \frac{\phi(\epsilon_{u,s,n}, t_{u,s,n}, \mathbf{m}_{u,s,n})}{\psi(\epsilon_{u,s,n}, t_{u,s,n}, \mathbf{m}_{u,s,n})} \right\}$*, if and only if*

$$F(\eta^\star) = \underset{\substack{\epsilon_{u,s,n},\, t_{u,s,n},\\ \mathbf{m}_{u,s,n}\in\boldsymbol{\Omega}}}{\textit{maximum}} \Big\{ \phi(\epsilon_{u,s,n},\, t_{u,s,n}, \mathbf{m}_{u,s,n}) - \eta^\star \psi(\epsilon_{u,s,n},\, t_{u,s,n}, \mathbf{m}_{u,s,n}) \Big\} = 0, \tag{2.12}$$

where $F(\cdot)$ is an auxiliary function with parameter $\eta$ and with a unique root at $\eta^\star$.

*Proof.* The proof of Theorem 1 can be found in (Dinkelbach, 1967; Ródenas *et al.*, 1999; Jagannathan, 1966) and thus it is not duplicated here. As a result, solving problem (2.9) is equivalent to finding the unique root of the auxiliary function $F(\cdot)$. □

To find the unique root of $F(\cdot)$, we formulate the following subproblem parameterized in $\eta$:

$$\underset{\substack{\epsilon_{u,s,n},\, t_{u,s,n},\\ \mathbf{m}_{u,s}}}{\text{maximize}} \quad \sum_{u=1}^{U} \gamma_u \left( \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} \right) - \eta \left( \sum_{n=1}^{N} \sum_{u\in\mathcal{U}} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2 + \zeta \right) \tag{2.13a}$$

$$\text{subject to} \quad \text{(2.5b), (2.5c), (2.9b) and (2.9c).}$$

However, note that the optimization domain of problem (2.13) is still non-convex even for fixed receive beamformers, $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$, due to the set of constraints in (2.9b). Nevertheless, as an alternative, we can resort to the SCA approach to relax the constraints in (2.9b) by a sequence of convex subsets (Venkatraman *et al.*, 2016). This can be accomplished using the first order Taylor approximation around a fixed MSE point $\tilde{\epsilon}_{u,s,n}$ as

$$-\log_2\left(\tilde{\epsilon}_{u,s,n}\right) - \frac{\epsilon_{u,s,n} - \tilde{\epsilon}_{u,s,n}}{\tilde{\epsilon}_{u,s,n}\log\left(2\right)} \geq t_{u,s,n}. \tag{2.14}$$

Now, replacing (2.9b) by (2.14) in problem (2.13), we can formulate the following parametric subproblem:

$$\underset{\substack{\epsilon_{u,s,n},\, t_{u,s,n},\\ \mathbf{m}_{u,s}}}{\text{maximize}} \quad \sum_{u=1}^{U} \gamma_u \left( \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} \right) - \eta \left( \sum_{n=1}^{N} \sum_{u\in\mathcal{U}} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2 + \zeta \right) \tag{2.15a}$$

$$\text{subject to} \quad \text{(2.5b), (2.5c), (2.9c) and (2.14),}$$

which is convex for $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)}$, $\{\epsilon_{u,s,n}\}_{\forall(u,s,n)}$ and $\{t_{u,s,n}\}_{\forall(u,s,n)}$ when $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$ is kept fixed.

Thus, problem (2.15) can be handled via convex optimization to solve problem (2.9) based on Theorem 1. The idea is, therefore, to successively solve the parametric subproblem (2.15) until the value of $\eta$ converges, i.e., when $F(\eta) = 0$, or alternatively when $F(\eta) \leq \varepsilon$, where $\varepsilon > 0$ is the convergence tolerance having a very small value. In Dinkelbach's approach, the update of parameter $\eta$ is based on the application of Newton's method, given as

$$\begin{aligned}\eta^{(l+1)} &= \eta^{(l)} - \frac{F(\eta^{(l)})}{F'(\eta^{(l)})} \\ &= \eta^{(l)} - \frac{\phi(\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n}) - \eta^{(l)}\psi(\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n})}{-\psi(\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n})} \\ &= \frac{\phi(\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n})}{\psi(\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n})}, \end{aligned} \tag{2.16}$$

where $\{\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n}\}_{\forall(u,s,n)}$ are found by solving parametric subproblem (2.15) and $l$ represents the iterations in the *outer loop* of Algorithm 1. The sequence of $\eta$ converges with a superlinear convergence rate (Isheden; Fettweis, ; Isheden *et al.*, 2012b). However, it is worth mentioning that the global optimality of the achieved solution for problem (2.5) can no longer be guaranteed, due to the iterative linear approximation procedure used during the problem reformulation in Section 2.4.1. The interested reader is referred to (Boyd *et al.*, 2007; Marks; Wright, 1978) for more details about the properties of the SCA method. Moreover, observe that the MSE approximation point, $\tilde{\epsilon}_{u,s,n}$, is updated with the MSE value $\epsilon_{u,s,n}$ found in the preceding step. The complete algorithm for the centralized solution is depicted in Algorithm 1.

**Algorithm 1:** Centralized approach for solving problem (2.9) via fractional programming and SCA.

1: **Input:** $\{Q_u, \gamma_u, \xi_u, \mathbf{H}_{b_u,u,n}\}_{\forall(b,u,s,n)}$, and $\eta^{(0)}$;
2: **Output:** $\{\mathbf{m}_{u,s,n}, \mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$;
3: **Initialize:** $\{\mathbf{m}^{(0)}_{u,s,n},\ \tilde{\epsilon}^{(0)}_{u,s,n}\}_{\forall(u,s,n)}$ randomly, $l \leftarrow 0$;
4: **repeat**
5:  **repeat**
6:  Generate $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$ using (2.7);
7:  **repeat**
8:  Find $\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n}$ solving convex parametric subproblem (2.15) using $\eta^{(l)}$;
9:  Update MSE point, i.e., $\tilde{\epsilon}_{u,s,n} \leftarrow \epsilon^{\star}_{u,s,n}$;
10:  **until** Convergence has been achieved.
11:  **until** Convergence has been achieved.
12:  Compute the updated value of $\eta$, making $\eta^{(l+1)} \leftarrow \phi(\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n})/\psi(\epsilon^{\star}_{u,s,n},\, t^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n})$;
13:  $l \leftarrow l + 1$;
14: **until** Convergence has been achieved.

In essence, for a given $\eta$ value, the idea of the *inner loop* in lines 5-11 of Algorithm 1

is to perform multiple SCA steps for each fixed receive beamformer update until $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)}$, $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$, $\{\epsilon_{u,s,n}\}_{\forall(u,s,n)}$ and $\{t_{u,s,n}\}_{\forall(u,s,n)}$ achieve some convergence criterion, which can be a predefined maximum number of iterations or a criterion based on the progress achieved in each iteration of the *inner loop*. It is worth mentioning that the more iterations are executed, the better is the performed SCA. The *outer loop* in lines $4$-$14$ is responsible for updating $\eta$ at each iteration $l$ with the new values of the optimization variables according to Dinkelbach's approach. Again, the convergence criterion of the *outer loop* can be either a maximum number of iterations, or based on the change of the $F(\eta)$ value.

At last, due to the update method of $\eta$, its value is expected to converge quickly. Algorithm 1 requires that a central processing unit be available to collect the channel information from all BSs to all users. This information is acquired by the BSs sending their channel estimations to the central processing unit, which poses a huge burden on the backhaul links. Then, upon receiving this channel information, the central unit runs Algorithm 1 to compute the transmit and receive beamformers, which are then fed back to the BSs and users. Therefore, BSs and users are not required to perform any computation of the transmit and receive beamformers. The convergence analysis of Algorithm 1 is presented in **APPENDIX A**.

The computational complexity of Algorithm 1 is dominated by the matrix inversion in (2.7) and by solving the parametric subproblem (2.15). The remaining equations are only linear equations, whose contribution to the overall computational complexity can be neglected. Solving (2.7) requires $UN_R^3$ operations (Shi *et al.*, 2011), while problem (2.15) can be converted into a second-order cone programming (SOCP) form and solved with $(SN_T)^{3.5}$ operations (Boyd; Vandenberghe, 2004). Thus, the per-iteration and per-subchannel computational complexity of Algorithm 1 is given by $\mathcal{O}\left(UN_R^3 + \{SN_T\}^{3.5}\right)$.

## 2.5 Decentralized Solution and Signaling

Due to the amount of involved network nodes and user devices, acquiring global CSI is a difficult task, which renders the centralized solution impractical. Therefore, in Section 2.5.1, we propose a decentralized solution where the variables are computed in a distributed fashion among users and BSs. In addition, in order to enable the decentralized processing, we also propose a signaling strategy in Section 2.5.2.

### *2.5.1 Decentralized Solution*

The parametric optimization subproblem (2.15) is, in general, not decoupled among BSs because of the interference terms and rate constraints present in the transmit beamformer update process. To handle this issue, in this section we propose an alternative way to solve this problem using a dual decomposition method based on the Lagrangian relaxation of constraints (2.5b). As a result, the proposed relaxed formulation of parametric subproblem (2.15) is

$$\begin{aligned}\underset{\epsilon_{u,s,n},\, t_{u,s,n},\mathbf{m}_{u,s}}{\text{maximize}} \quad & \sum_{u=1}^{U}\gamma_u\left(\sum_{n=1}^{N}\sum_{s=1}^{S_{u,n}} t_{u,s,n}\right) - \eta\left(\sum_{n=1}^{N}\sum_{u\in\mathcal{U}}\sum_{s=1}^{S_{u,n}}\|\mathbf{m}_{u,s,n}\|^2+\zeta\right) \\ & -\sum_{\forall u}\alpha_u\left(-\sum_{n=1}^{N}\sum_{s=1}^{S_{u,n}} t_{u,s,n}+\xi_u\right) - \sum_{\forall u}\beta_u\left(\sum_{n=1}^{N}\sum_{s=1}^{S_{u,n}} t_{u,s,n} - \frac{Q_u}{\Delta_{\text{tti}}}\right) \qquad (2.17a)\\ \text{subject to} \quad & (2.5c),\ (2.9c) \text{ and } (2.14),\end{aligned}$$

where the dual variables $\{\alpha_u,\ \beta_u\}_{\forall u}$ are fixed, while solving the primal parametric subproblem (2.17). However, $\{\alpha_u,\ \beta_u\}_{\forall u}$ can be updated according to the possible violations of the corresponding constraints.

Thus, to solve problem (2.9) using parametric subproblem (2.17), we propose Algorithm 2, which is a decentralized solution based on Dinkelbach's approach and the KKT conditions. Basically, similarly to Algorithm 1, multiple consecutive SCA updates can be performed for each fixed receive beamformer update and fixed $\eta$. The signaling aspects of Algorithm 2 are discussed in Section 2.5.2. The convergence analysis of Algorithm 2 is presented in **APPENDIX A**.

By employing the KKT conditions, considering that $\{\delta_b\}_{\forall b}$ and $\{\lambda_{u,s,n}, \mu_{u,s,n}\}_{\forall(u,s,n)}$ are the dual variables corresponding to the constraints in (2.5c), (2.9c) and (2.14) of (2.17),

respectively, and assuming that (2.9c) and (2.14) are tight, we have:

$$\mathbf{m}_{u,s,n}^{(e)} = \left( \sum_{\forall i} \sum_{\forall j} \lambda_{i,j,n}^{(e-1)} \mathbf{H}_{b_u,j,n}^{\mathrm{H}} \mathbf{w}_{i,j,n} \mathbf{w}_{i,j,n}^{\mathrm{H}} \mathbf{H}_{b_u,j,n} + (\eta^{(l)} + \delta_b)\mathbf{I} \right)^{-1} \lambda_{u,s,n}^{(e-1)} \mathbf{H}_{b_u,u,n}^{\mathrm{H}} \mathbf{w}_{u,s,n}^{(e-1)}, \tag{2.18a}$$

$$\epsilon_{u,s,n}^{(e)} = |1 - \mathbf{w}_{u,s,n}^{\mathrm{H}} \mathbf{H}_{b_u,u,n} \mathbf{m}_{u,s,n}^{(e)}|^2 + \sigma_u^2 \|\mathbf{w}_{u,s,n}\|^2 + \sum_{i=1}^{U} \sum_{\substack{j=1, \\ (i,j)\neq(u,s)}}^{S_{u,n}} |\mathbf{w}_{u,s,n}^{\mathrm{H}} \mathbf{H}_{b_i,u,n} \mathbf{m}_{i,j,n}^{(e)}|^2, \tag{2.18b}$$

$$t_{u,s,n}^{(e)} = -\log_2 (\epsilon_{u,s,n}^{(e-1)}) - \frac{\epsilon_{u,s,n}^{(e)} - \epsilon_{u,s,n}^{(e-1)}}{\epsilon_{u,s,n}^{(e-1)} \log(2)}, \tag{2.18c}$$

$$\mu_{u,s,n}^{(e)} = \left[ (\gamma_u + \alpha_u^{(e)}) - \beta_u^{(e)} \right]^+, \tag{2.18d}$$

$$\lambda_{u,s,n}^{(e)} = \lambda_{u,s,n}^{(e-1)} + \theta^{(e)} \left( \frac{\mu_{u,s,n}}{\epsilon^{(e-1)} \log(2)} - \lambda_{u,s,n}^{(e-1)} \right), \tag{2.18e}$$

where $e$ and $l$ represent the iterations in the *inner loop* (lines 6-21) and *outer loop* (lines 5-25), respectively, of Algorithm 2. We give the Lagrangian function and the KKT conditions of (2.17) in **APPENDIX B**. The convergence criterion for the *inner loop* of Algorithm 2 can be a predefined maximum number of iterations (as stated in line 20) or a criterion based on the progress achieved in each iteration of the *inner loop*. Regarding the *outer loop* of Algorithm 2, the convergence criterion can be either a maximum number of iterations (as stated in line 25) or based on the change of the $\eta$ value.

Firstly, observe that the dual variables $\lambda_{u,s,n}^{(e)}$ and $\mu_{u,s,n}^{(e)}$ are interdependent in (2.18e) and, therefore, one of the dual variables is maintained fixed while the other is optimized. We assume that $\lambda_{u,s,n}^{(e)}$ is fixed to compute $\mu_{u,s,n}^{(e)}$. The dual variable $\lambda_{u,s,n}^{(e)}$, in its turn, is a point in the line segment between $\lambda_{u,s,n}^{(e)}$ and $\frac{\mu_{u,s,n}^{(e)}}{\epsilon^{(e-1)}\log(2)}$, determined by using a diminishing or fixed step size $\theta^{(e)} \in (0,\ 1)$. The choice of $\theta^{(e)}$ is system-dependent, and its value affects the convergence behavior, and controls the oscillations in the UEs' rates when negative (before projection) due to over-allocation.

From a practical point of view, we can observe that, for $t_{u,s,n}^{(e)}$ greater than $Q_u$, the corresponding dual variable $\mu_{u,s,n}^{(e)}$ can be negative and, due to the projection operator in (2.18d), it will be zero. Consequently, $\lambda_{u,s,n}^{(e)} < \lambda_{u,s,n}^{(e-1)}$, as it can be seen in (2.18e). Once $\lambda_{u,s,n}^{(e)}$ is reduced, the precoder weight in (2.18a) is lowered to make the rate $t_{u,s,n}^{(e)} < t_{u,s,n}^{(e-1)}$.

In the *inner loop* of Algorithm 2, the KKT expressions in (2.18) are solved in an

**Algorithm 2:** Decentralized approach for (2.9) via KKT conditions and fractional programming.

**Input:** $\{Q_u, \gamma_u, \xi_u, \mathbf{H}_{b_u,u,n}\}_{\forall(b,u,s,n)}$, and $\eta^{(0)}$;
**Output:** $\{\mathbf{m}_{u,s,n}, \mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$;
**Initialize:** $\{\mathbf{m}^{(0)}_{u,s,n},\ \tilde{\epsilon}^{(0)}_{u,s,n}\}_{\forall(u,s,n)}$ randomly, $\lambda^{(0)}_{u,s,n} \leftarrow 1$ and $l \leftarrow 0$, $e \leftarrow 1$;
BS: Use initial $\{\mathbf{m}^{(0)}_{u,s,n}\}_{\forall(u,s,n)}$ to transmit pilots;
**repeat**
  **repeat**
    UE: Generate $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$ using (2.7);
    **repeat**
      UE: Measure MSE $\epsilon^{(e)}_{u,s,n}$ as shown in (2.6);
      UE: Compute $t_{u,s,n}$ using (2.18c);
      UE: Update variable $\alpha_u$ using (2.19a);
      UE: Update variable $\beta_u$ using (2.19b);
      UE: Update variable $\mu_{u,s,n}$ from (2.18d);
      UE: Update the weights $\lambda_{u,s,n}$ from (2.18e);
      UE: Send $\lambda_{u,s,n}$ to BS using uplink signaling;
      BS: Exchange $\{\lambda_{u,s,n}\}_{\forall(u,s,n)}$ via backhaul link;
      BS: Solve $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)}$ from (2.18a);
      BS: Use $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)}$ to transmit pilots;
      $e \leftarrow e + 1$;
    **until** Convergence has been achieved or $e > E_{\max}$.
  **until** Convergence has been achieved.
  BS: Exchange the total transmit bits and consumed power via backhaul link;
  BS: Update $\eta^{(l+1)}$ using (2.20);
  $l \leftarrow l + 1$;
**until** Convergence has been achieved or $l > L_{\max}$.

iterative manner after a random initialization of $\{\mathbf{m}^{(0)}_{u,s,n},\ \tilde{\epsilon}^{(0)}_{u,s,n}\}_{\forall(u,s,n)}$, while the dual variables $\{\lambda^{(0)}_{u,s,n}\}_{\forall(u,s,n)}$ are initialized with ones, so that all UEs in the system have equal priority. Then, we begin by fixing the receive beamforming vectors to be the MMSE receive beamformers in (2.7). On the other hand, the transmit beamforming vectors in (2.18a) depend on the BS specific dual variable $\delta_b$ and $\eta^{(l)}$. In this case, $\delta_b$ should be chosen to meet the complementary slackness condition of the power budget constraint. Note that if the power constraint is not active when solving (2.18a) for $\delta_b = 0$, then the beamformers are optimal. Otherwise, the optimal value of $\delta_b$ can be obtained using one dimensional search techniques, e.g., bisection method, with respect to the power budget constraint (Shi *et al.*, 2011).

Once the MSE values $\{\epsilon^{(e)}_{u,s,n}\}$ are computed, we update the variables $\{t^{(e)}_{u,s,n}\}_{\forall(u,s,n)}$ using (2.18c). Meanwhile, $\{\mu^{(e)}_{u,s,n}\}_{\forall(u,s,n)}$ and $\{\lambda^{(e)}_{u,s,n}\}_{\forall(u,s,n)}$ are updated according to (2.18d) and (2.18e), respectively. Besides that, the SCA operating point is updated with the current MSE value.

Next, in the dual update, the rate demand weight factors $\{\alpha_u\}_{\forall u}$ and queue weight factors $\{\beta_u\}_{\forall u}$ follow, from their respective constraint violations, as

$$\alpha_u^{(e)} = \left[\alpha_u^{(e-1)} + \theta^{(e)}\left(-\sum_{n=1}^{N}\sum_{s=1}^{S_u} t_{u,s,n}^{(e)} + \xi_u\right)\right]^+, \tag{2.19a}$$

$$\beta_u^{(e)} = \left[\beta_u^{(e-1)} + \theta^{(e)}\left(\sum_{n=1}^{N}\sum_{s=1}^{S_u} t_{u,s,n}^{(e)} - \frac{Q_u}{\Delta_{\text{tti}}}\right)\right]^+, \tag{2.19b}$$

where $\theta^{(e)}$ is the step size of the current iteration. This also corresponds to a subgradient update of the dual variables in terms of (2.17) with the approximate MSE constraints (for more details, see (Bertsekas, 1999; Palomar; Mung Chiang, 2006)).

Finally, after the convergence of the values $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)}$ and $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$ for a given $\eta$, we update it in the *outer loop* as follows

$$\eta^{(l+1)} = \frac{\sum_{u=1}^{U}\gamma_u\left(\sum_{n=1}^{N}\sum_{s=1}^{S_{u,n}} t_{u,s,n}\right)}{\left(\sum_{n=1}^{N}\sum_{u=1}^{U}\sum_{s=1}^{S_{u,n}}\|\mathbf{m}_{u,s,n}\|^2 + \zeta\right)}. \tag{2.20}$$

In order to obtain practical distributed precoders design, we assume that each BS $b$ notifies all BSs the number of transmit bits and consumed power by means of backhaul signaling, so that $\eta^{(l+1)}$ can be computed locally at each BS.

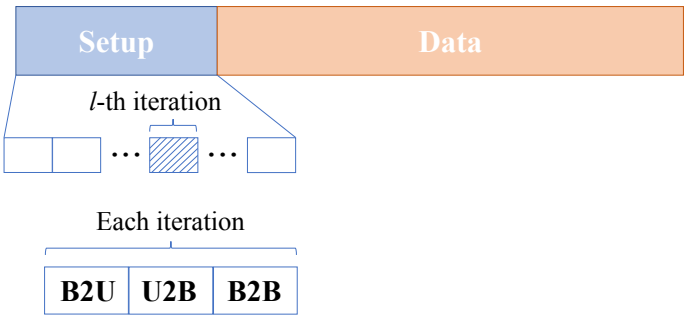
The computational complexity of Algorithm 2 is dominated by the matrix inversions in (2.7) and (2.18a), and the MSE computation in (2.6). Such operations are also required by the benchmarking algorithms in (Shi *et al.*, 2011; Kaleva *et al.*, 2016; Venkatraman *et al.*, 2016) used herein. Apart from the computations of (2.6), (2.7) and (2.18a), Algorithm 2 computes the value of other variables using linear equations. Therefore, the per-iteration and per-subchannel computational complexity of Algorithm 2 is known and given by $\mathcal{O}\left(U^2N_TN_R^2 + U^2N_T^2N_R + U^2N_T^3 + UN_R^3\right)$, which is within the capabilities of the hardware of current BSs and UE terminals considering a moderate number of transmit/receive antennas.

### *2.5.2 Signaling*

In this section, we analyze the main signaling aspects involved in the implementation of the proposed decentralized algorithm, which includes pilot and backhaul links. To this end, we extend the signaling scheme proposed in (Tölli *et al.*, 2019) to support Algorithm 2.

The proposed decentralized algorithm requires only local CSI knowledge. The computations are distributed among the BSs and UEs, which then exchange some limited

Figure 7 – Frame structure.




Source: Created by the author.

information using the adopted signaling scheme. Therefore, the decentralized algorithm does not require a central processing unit for its deployment, which can be very useful in scenarios where such a unit is not available. For the local processing that occurs at each node, the intermediate optimization variables are obtained using the proposed signaling scheme. In (2.18a), for example, note that the knowledge of the UEs' receive beamformers, corresponding interference channels and scalar variable $\eta$ are required. Thus, to acquire this information, during the execution of the algorithm, an over-the-air (bidirectional) signaling scheme is considered for the BS-to-UE (B2U) and UE-to-BS (U2B) communications. Meanwhile, the BS-to-BS (B2B) communication occurs via backhaul links.

Based on the above considerations and on the frame structure presented in (Tölli *et al.*, 2019), Fig. 7 shows the proposed frame structure. Note that the frame structure is split into two parts: setup and data. The first part is responsible for beamformer setup, while the second one is specifically used for data transmission. Both the over-the-air and the backhaul signaling occur in the beamformer setup phase. The over-the-air signaling is further divided into two phases. The forward pilot transmission from BS to the UEs, which corresponds to lines $4$ and $18$ of Algorithm 2, occurs in the first phase (denoted by B2U). The second phase (U2B) contains the backward signaling from each UE to its serving BS, which corresponds to line $15$ of Algorithm 2. Finally, the backhaul signaling for B2B communication, which corresponds to lines $16$ and $22$ of Algorithm 2, is used to share the $\{\lambda_{u,s,n}\}_{\forall(u,s,n)}$ weights, as well as the total transmit bits and consumed power of the BSs to compute the $\eta$ value.

As mentioned, there is some communication among the BSs via backhaul links, and some communication between BSs and UEs via over-the-air signaling. In practice, there may be some imperfection in the information acquired via the backhaul links due to delays, limited capacity, transmission errors or quantization error that might occur in those links. Regarding the

over-the-air information exchange, there may be some degree of imperfection in case transmission errors or delay occur. Furthermore, in the proposed scheme, the over-the-air signaling exchange used for CSI estimation may be impacted by pilot contamination, which may degrade the quality of the acquired CSI.

## 2.6 Performance Evaluation

This section evaluates the performance of the proposed solutions. Section 2.6.1 provides the main simulation assumptions and parameters, while the obtained results are shown and discussed in Sections 2.6.2, 2.6.3 and 2.6.4.

### *2.6.1 Simulation Assumptions*

We consider the downlink of multi-cell MIMO-OFDM scenarios with the following main parameters $\{B, U, U_b, N_{\mathrm{T}}, N_{\mathrm{R}}\} = \{4, 16, 4, 8, 2\}$. Each BS is located at the center of a hexagonal cell with radius of $250$ m and the UEs are uniformly distributed within the cell. The BS and UE heights are $25$ m and $1.5$ m, respectively. Unless otherwise stated, the power budget $P_b$ is set to $33$ dBm, the circuit power constant ($\zeta$) is set to $10$ W and the minimum data rate to $1$ bit/s/Hz. The 5G stochastic radio channel for dual mobility (5G-StoRM) channel model (Pessoa *et al.*, 2019; 3GPP, 2016) is used for all links considering the urban micro (UMi) scenario for a carrier frequency of $2$ GHz. The duration of 1 TTI ($\Delta_{\text{tti}}$) is $1$ ms. The results are obtained from $100$ Monte-Carlo simulations.

To benchmarking the proposed solution presented in Section 2.5, we consider three state-of-the-art solutions. For algorithms that were originally designed for full-buffer scenarios, we first consider the weighted minimum mean squared error (WMMSE) algorithm (Shi *et al.*, 2011), which solves a rate-unconstrained weighted sum-rate maximization problem. Besides, we also consider the algorithm proposed in (Kaleva *et al.*, 2016), which solves a weighted sum-rate maximization problem with minimum rate requirements, referred herein as Kaleva. Concerning finite-buffer approaches, we consider the solution from (Venkatraman *et al.*, 2016), referred to as joint space-frequency resource allocation (JSFRA) algorithm. Basically, the idea of JSFRA is to minimize the total number of remaining backlogged bits of the buffers at the BSs. Since the benchmarking algorithms are distributed solutions, we use the proposed decentralized solution in Algorithm 2, which is applicable to more practical situations and is hereafter referred to as

Q-WEEM.

Regarding the discussion of the results, besides the EE metric, we use other metrics for performance comparison. In this chapter we consider the buffer usage efficiency (BUE) metric, that aims to evaluate the performance of all solutions with respect to the efficiency in reducing the number of bits queued in the buffer. In addition, we also define the transmit power usage efficiency (PUE) metric, which takes into account the consumed transmit power and the maximum transmit power available at the BS. Mathematically, these metrics are expressed as

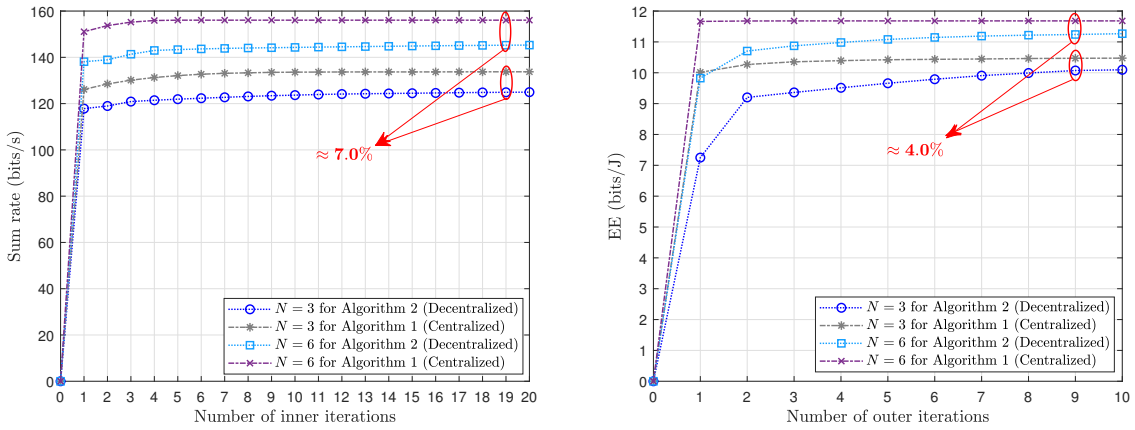
$$\text{BUE} = 1 - \frac{\sum_{u=1}^{U} \left[ \frac{Q_u}{\Delta_{\text{tti}}} - t_u \right]^+}{\sum_{u=1}^{U} \frac{Q_u}{\Delta_{\text{tti}}}}, \quad \text{and} \tag{2.21a}$$

$$\text{PUE} = 1 - \frac{\sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{u \in \mathcal{U}_b} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2}{B \times P_b}. \tag{2.21b}$$

It should be noticed that the BUE and PUE metrics are always between 0 (worst case) and 1 (best case).

### *2.6.2 Convergence Analysis and Performance Gap between Centralized and Decentralized Solutions*

Figure 8 – Number of inner and outer iterations for the proposed solution, considering different numbers of sub-channels ($N$).




(c) *Inner loop.* (d) *Outer loop.*

Source: Created by the author.

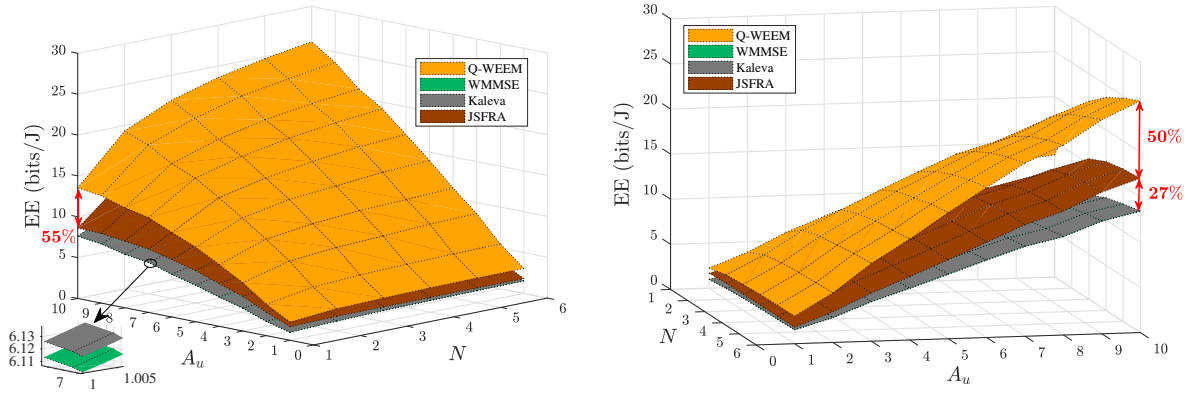
Our discussion starts with Fig. 8, which illustrates the evolution of the proposed iterative algorithms for two different values of the number of sub-channels in the system. An

important aspect of the first figure is to show the performance gap between the proposed centralized and decentralized solutions. Fig. 8c shows the convergence of the *inner loops*, i.e., lines 5-11 of Algorithm 1 and lines 6-21 of Algorithm 2. Meanwhile, Fig. 8d illustrates the convergence of the *outer loops* presented in lines 4-14 and 5-25 of Algorithms 1 and 2, respectively. Consequently, Fig. 8 shows the successive increases of EE (Fig. 8d) through the *outer loop* of each algorithm, which is responsible for updating the Dinkelbach parameter ($\eta$) after the respective convergence of their *inner loops* (Fig. 8c).

Moreover, note that because the EE update is based on Newton's method, the Dinkelbach parameter value converges quickly. The advantages of employing multiple sub-channels are also illustrated in Fig. 8, where the sum-rate and EE values increase significantly, as the number of system sub-channels ($N$) increases. After convergence is achieved in both figures, we highlight that the performance gap between centralized and decentralized solutions is of at most $7\%$. However, we emphasize that a higher performance of the centralized solution is expected due to its centralized nature, which allows its *inner loop* to solve an optimization problem with global CSI knowledge by employing convex optimization computational tools, e.g., CVXPY (Diamond; Boyd, 2016). Therefore, although the centralized approach provides a slightly higher performance, it also requires higher computational cost, besides the fact that it requires a central processing unit to be available.

### *2.6.3 Performance Comparison Under Perfect CSI Estimation*

Figure 9 – EE versus average number of packet arrivals ($A_u$) and number of sub-channels ($N$) for the Q-WEEM, WMMSE, Kaleva and JSFRA algorithms.




(c) Perspective 1. (d) Perspective 2.

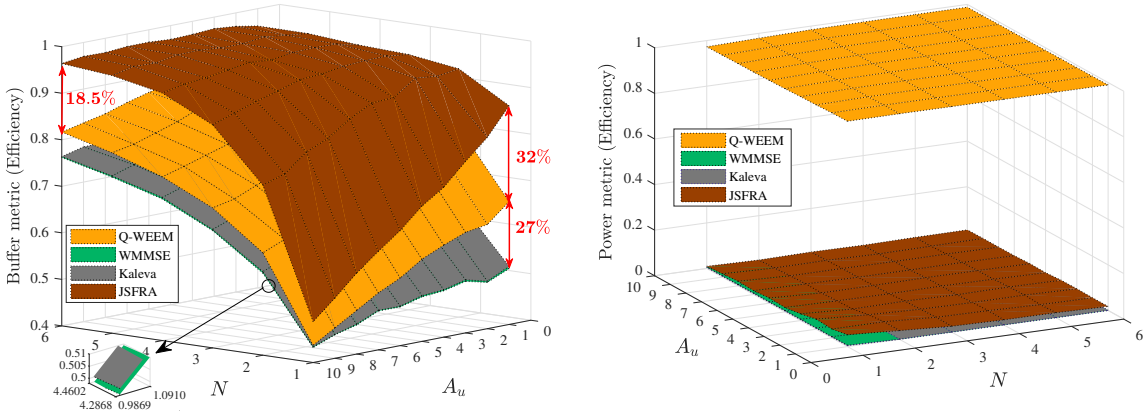
Source: Created by the author.

Since the decentralized solution achieves a performance close to that of the centralized solution with lower complexity, and since the comparison solutions are decentralized algorithms, hereafter we only use our decentralized approach (Q-WEEM).

Fig. 11 depicts the EE (in bits per Joule) performance achieved by the solutions Q-WEEM, WMMSE (Shi *et al.*, 2011), Kaleva (Kaleva *et al.*, 2016) and JSFRA (Venkatraman *et al.*, 2016), versus the average number of packet arrivals and the number of sub-channels. Firstly, as expected, we have a considerable EE increase for all solutions as both the number of sub-channels and the number of packet arrivals increase, with an emphasis on the performance of the Q-WEEM solution, which seeks to find a transceiver design providing the best trade-off between sum-rate and consumed energy in the system.

As a result, this leads to a higher level of EE in our proposed solution. Furthermore, note that it tends to increase its performance gap in relation to other solutions as $A_u$ and $N$ increase. It is also important to note that the worst performances in terms of EE are for the solutions WMMSE and Kaleva, while an intermediate performance is obtained by the JSFRA solution. Indeed, this can be explained by Figs. 10c and 10d. These figures compare the efficiency of emptying the buffer (BUE metric) and the power consumption (PUE metric) as the functions of the average number of packet arrivals ($A_u$) and the number of sub-channels ($N$).

Figure 10 – BUE and PUE metrics versus average number of packet arrivals ($A_u$) and number of sub-channels ($N$) for the Q-WEEM, WMMSE, Kaleva and JSFRA algorithms.



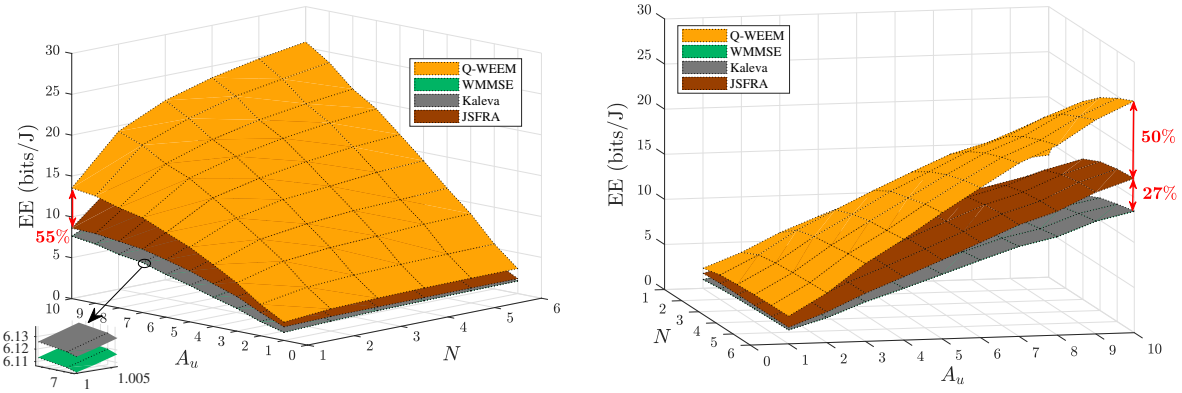

(c) BUE metric.

(d) PUE metric.

Source: Created by the author.

Fig. 10c shows that the JSFRA solution is the most efficient in minimizing the number of backlogged queuing bits or, in other words, it is the most efficient in the number of transmitted bits. This is highly relevant in finite-buffer scenarios, and it happens because

Figure 11 – EE versus average number of packet arrivals ($A_u$) and number of sub-channels ($N$) for the Q-WEEM, WMMSE, Kaleva and JSFRA algorithms.




(c) Perspective 1.

(d) Perspective 2.

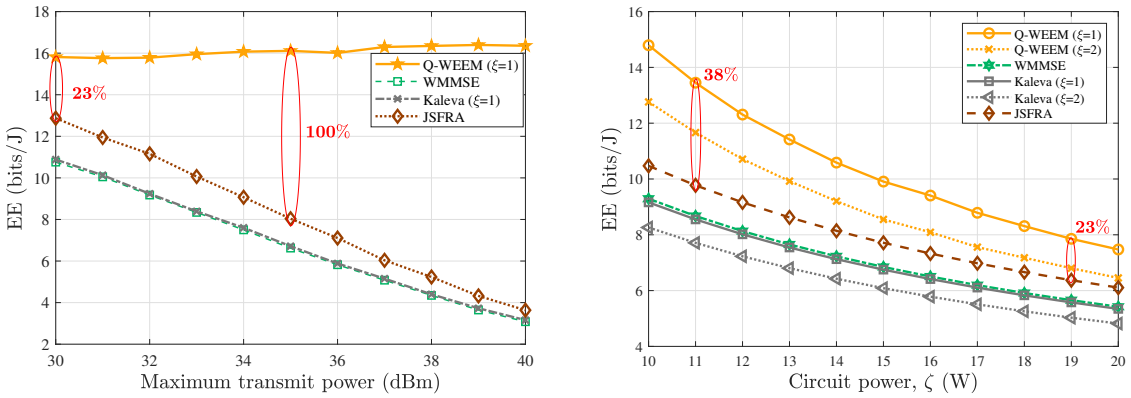
Source: Created by the author.

the JSFRA algorithm prioritizes UEs with a higher number of queued bits, before considering the UEs with a smaller number of backlogged bits. Due to this strategy, the JSFRA solution is clearly superior in minimizing the buffer size. On the other hand, the WMMSE and Kaleva solutions do not take into account the buffer size and, due to their greedy resource allocation policies, they tend to empty the queue of UEs with good channel conditions before considering the UEs with poor channels.

Although this may seem to be an efficient strategy to obtain more bits, the UEs with the best channel gains are not necessarily the same as the ones with the greatest number of backlogged bits. Therefore, the WMMSE and Kaleva solutions are less efficient in emptying the buffer, i.e., reducing the number of queued bits. Furthermore, note that both emptying the buffer and maximizing the sum rate are greedy strategies from the point of view of using the transmit power as shown in Fig. 10d. Indeed, in this figure, we can see that the power consumption of the JSFRA, WMMSE and Kaleva solutions are quite similar and, therefore, it is natural for the JSFRA solution to reach a greater EE than the achieved EE of the WMMSE and Kaleva solutions.

Note that although the Q-WEEM solution has a lower performance compared to JSFRA with respect to the total number of residual bits remaining in the buffer, the Q-WEEM solution outperforms the WMMSE and Kaleva solutions, since it considers the buffer size as a constraint. Moreover, another significant advantage of Q-WEEM can be observed in Fig. 10d, which shows a huge gain in terms of consumed power compared to all the considered solutions. Note that this is a direct consequence of maximizing the EE metric, which is the goal of our

Figure 12 – EE versus BS power and circuit power for the Q-WEEM, WMMSE, Kaleva and JSFRA algorithms, considering two QoS values.




(c) EE versus BS power.

(d) EE versus circuit power.

Source: Created by the author.

problem. Particularly, our power consumption model captures both the power consumption at the base station and the power consumption of the network infrastructure, which is a constant value (circuit power). Thereby, in order to find the best trade-off between the total number of transmitted bits and the consumed power, our solution allocates the available power more efficiently. Consequently, it yields the highest PUE metric. This undoubtedly contributes for Q-WEEM to achieve a high EE, as shown in Fig. 11.

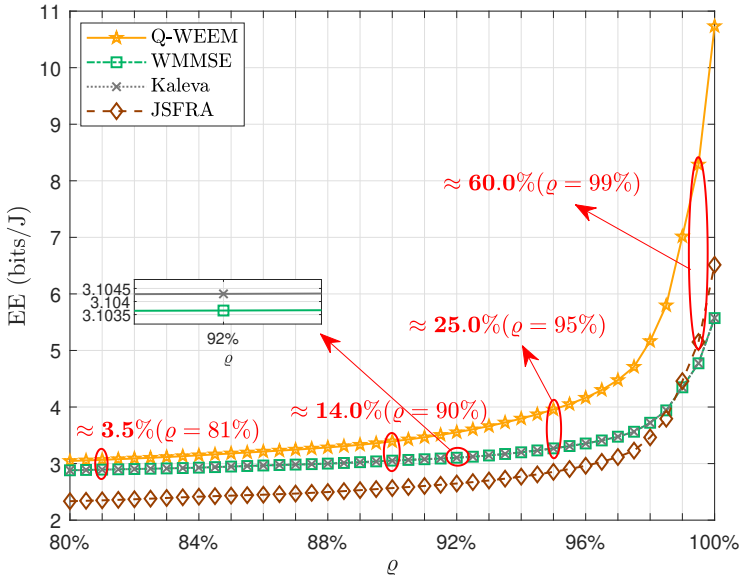
The dependence of EE on the maximum transmit power and the circuit power is shown in Fig. 12. This figure is particularly interesting because it shows that once a given QoS value is guaranteed, the Q-WEEM scheme tends to keep the EE value practically constant, regardless of the power budget available at the BS. In contrast, the other solutions tend to consume more and more power and, therefore, suffer expressive losses in terms of EE as more transmit power is available, as shown in Fig. 12c. The reason why the EE achieved by our solution remains almost constant when the total available power increases is the fact that the proposed scheme does not use more power to satisfy the minimum QoS requirements. Meanwhile, the benchmarking solutions increase the power consumption in order to maximize their objective functions.

Regarding Fig. 12d, we observe a different behavior, because all simulated algorithms are strongly affected by the $\zeta$ value, since it is the same for all solutions. Even so, we can observe that our solution provides a higher EE metric regardless of $\zeta$ value, because it uses the transmit power more efficiently, as discussed earlier. Furthermore, it is worth mentioning that, as $\zeta$ increases, the differences among the EE values for all algorithms are less significant. This

demonstrates that the improvement in EE also depends on a relative comparison between the total power available at the BS and the power consumed by the electronic circuits.

### *2.6.4 Performance Comparison Under Imperfect CSI Estimation*

Figure 13 – EE versus $\varrho$ for the Q-WEEM, WMMSE, Kaleva and JSFRA algorithms.




Source: Created by the author.

During the derivation of Algorithms 1 and 2, we assumed perfect knowledge of CSI at the transmitters and receivers. In order to analyze and compare the performance of the proposed distributed solution (i.e., Algorithm 2) with that of the benchmarking schemes, we model the CSI imperfection by assuming that the central unit estimates the channel using MMSE estimation. Therefore, the estimated channel matrix satisfies (Rusek *et al.*, 2013):

$$\hat{\mathbf{H}}_{b_i,s,n} = \varrho \mathbf{H}_{b_i,s,n} + \sqrt{1-\varrho^2}\mathbf{E}, \tag{2.22}$$

where $\mathbf{E} \in \mathbb{C}^{N_R \times N_T}$ is an error matrix with complex Gaussian independent and identically distributed (i.i.d.) entries with zero mean and unit variance. Meanwhile, $0 \leq \varrho \leq 1$ denotes the reliability of the estimate. Based on the model shown in (2.22), we perform simulations to evaluate the effects of imperfect CSI in the investigated algorithms.

The results are shown in Fig. 13, which depicts the EE metric for the solutions Q-WEEM, WMMSE, Kaleva and JSFRA for different values of $\varrho$, i.e., for various levels of

estimation reliability. An important conclusion based on this result is that the largest gains of the proposed solution over the benchmarking solutions are in regions where the reliability of the estimation is above $95\%$. This result indicates that our solution is superior compared with the benchmarks when the CSI quality is relatively high ($\varrho > 0.95$). However, when the value of $\varrho$ is below this threshold, the Q-WEEM scheme provides only slightly better performance than the benchmarking algorithms. This occurs because, similarly to the benchmarking solutions, Q-WEEM assumes close-to-perfect channel estimation. Improving the performance of the proposed solution with imperfect CSI is left for future works.

## 2.7 Chapter Summary

In this chapter, we developed a new framework for energy-efficient resource allocation in multi-cell MIMO-OFDM systems. The analyzed problem focused on maximizing the global energy efficiency of the system while satisfying constraints in terms of per-BS maximum power and per-UE minimum data rate, while considering a finite-buffer traffic model. This led to a non-convex and fractional optimization problem, which is hard to solve. To deal with this problem, we exploited the well-known MSE-SINR relation when using MMSE receivers. In this context, we first proposed a centralized iterative solution based on Dinkelbach's approach to fractional problems. Then, we also proposed an iterative decentralized solution with relaxed feasible initialization requirements, based on the dual decomposition and Lagrangian relaxation of the rate constraints. Both the centralized and the decentralized solutions are proven to converge to a KKT point of the formulated optimization problem.

Our simulation employed a realistic channel model with space, frequency and time correlations. Regarding the obtained results, the performance of the proposed approach was analyzed in terms of EE, number of backlogged bits in the buffer and consumed power for a variety of different system parameters. The proposed solution obtained gains of up to $55\%$ in terms of EE when varying the number of available frequency sub-channels, and gains of $38\%$ in terms of EE when the circuit power is appropriately taken into account. This was mainly accomplished by reducing the power consumption of the system, while still meeting the per-UE minimum rate demands. Finally, we also investigated the effect of imperfect CSI, and found that it can significantly affect the performance of the system in terms of EE. Further investigations to improve the performance of the proposed solution under imperfect CSI and considering other power consumption models are left for future work.

# 3 A GAME-THEORETIC DESIGN TO POWER CONTROL IN MASSIVE MIMO NETWORKS

Several previous works have proposed game-theoretic approaches to controlling the pilot and data power levels in the uplink of both single and multi-cell multiuser MIMO (MU-MIMO) systems. Unfortunately, the vast majority of existing works design these power control schemes under the assumption that the wireless channels between the mobile terminals and the serving base station are block fading. Meanwhile, several papers have shown that modeling fast fading channels as autoregressive (AR) processes with known or estimated state transition matrices give much more accurate results than those suggested by block fading models.

Thus, this chapter proposes a game-theoretic approach to controlling the uplink pilot and data power levels in a MU-MIMO system, in which the wireless channels are AR processes with mobile terminal-specific state transition matrices. We find that the proposed approach outperforms a classical cellular path-loss compensating fractional power control scheme and a game-theoretic power control scheme designed for block fading channels.

## 3.1 Introduction and Related Works

Many works demonstrated that the fundamental trade-off of sharing the available power resources between pilot and data symbols plays an important role on the performance of MU-MIMO systems (Fodor *et al.*, 2015; Zhao *et al.*, 2018; Zhao *et al.*, 2019). Particularly in (Fodor *et al.*, 2015), which focuses on minimizing the MSE with CSI errors, the authors derived the actual MMSE receiver as a function of the employed pilot-and-data power control (PDPC) and, thereby, demonstrated that significant gains are obtained when the number of BS antennas is large. Meanwhile, given that the sum-MSE minimization has a relatively close relation to the sum-rate maximization, this objective has become an important optimization metric and has been widely considered in recent works (Zhao *et al.*, 2018; Zhao *et al.*, 2019; Papazafeiropoulos *et al.*, 2021). In general, while (Papazafeiropoulos *et al.*, 2021) focused on the impact of hardware impairments, (Zhao *et al.*, 2018) and (Zhao *et al.*, 2019) were concerned to propose alternative optimization methods based on game theory to deal with non-convexity issues and allocating data power assuming a multi-user case.

However, the effect of channel aging was not considered in the above works. In this context, while a substantial body of literature exists on PDPC considering channel estimation errors in block fading channels, this analysis in MU-MIMO systems considering channel aging

models has been relatively unexplored and, therefore, it is the focus of this chapter. Existing works showed that exploiting the memoryful property of time-variant fast fading channels is relevant (Hijazi; Ros, 2010; Kashyap *et al.*, ; Fodor *et al.*, 2021). Specifically, when the channel can be modeled as an AR or autoregressive moving average (ARMA) process, its state can be well estimated when the correlation between subsequent channel realizations can be assumed to be known (Hijazi; Ros, 2010; Kashyap *et al.*, ; Fodor *et al.*, 2021). Recently, assuming a scenario where the channel is modeled as an AR process, (Fodor *et al.*, 2021) showed that the performance of the MMSE receiver can be further improved when the evolution of the channel in time is exploited via Kalman filter (KF). However, the results were limited to a single-user scenario and, thus, the impact of multi-user cases with interference and data power allocation algorithms were not investigated therein.

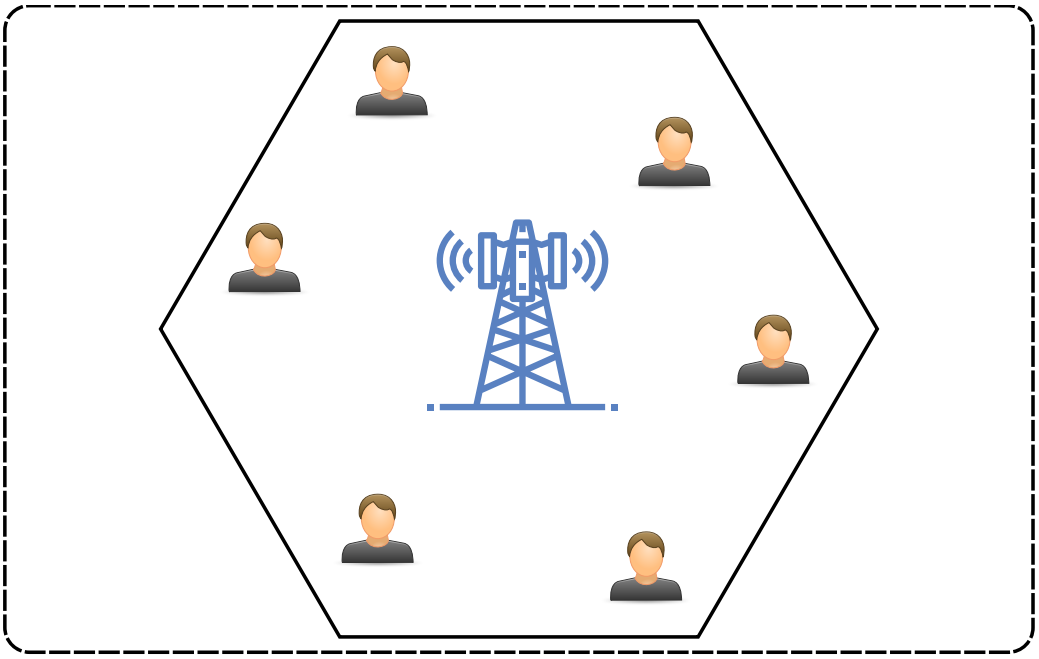
It can be intuitively expected that in multi-user scenarios a distributed power control scheme should take into account that different users may experience different channel conditions that correspond to different parameters of the AR fading process. However, this scenario is particularly challenging and an efficient solution is not trivial, as each user needs to acquire its own AR parameters and some system state information prior to computing its pilot and data powers within the power budget and without compromising the system performance.

### *3.1.1 Main Contributions*

In order to tackle these challenges, in this chapter we consider a MU-MIMO uplink system, assuming a more realistic and general situation where the channel is modeled as an AR(1) process. Specifically, our main contributions are:

1. We propose a game-based distributed power allocation algorithm where, initially, each user estimates its state transition matrix and, with the assistance of the BS, obtains some information about the system state using a novel signaling scheme. Then, each user iteratively finds a solution that minimizes its own MSE metric, while subject to interference and a limited power budget.
2. Using computational simulations, we show that the proposed game-based power control converges within a finite number of iterations, given a convergence threshold, and provides a performance close to the optimal approach while also outperforming conventional solutions from (Zhao *et al.*, 2019) and the uplink power allocation policy used by LTE (3GPP, 2021).

Figure 14 – A single-cell massive MIMO system with multiple users.



Source: Created by the author.

3. Finally, we present non-trivial new results in which we analyze the sensitivity of the proposed power control in more realistic scenarios where the channel autocorrelation coefficients of the channel are not perfectly known.

## 3.2 Network Model

We consider a single-cell massive MIMO system with multiple users according to Figure 14. Specifically, we consider the uplink of a cellular MU-MIMO system composed by a BS equipped with $N$ antennas and $K$ single-antenna UEs, which are grouped in the set $\mathcal{K}$, with $K \ll N$. Each UE employs an orthogonal pilot sequence, so that no interference between pilots is present in our system, i.e., $K < \tau_{\mathrm{p}}$, where $\tau_{\mathrm{p}}$ is the number of orthogonal pilot sequences. We also assume a comb type arrangement of the pilot symbols, in which the coherence bandwidth is composed of $\tau_{\mathrm{p}}$ subcarriers dedicated for pilot sequences and $\tau_{\mathrm{d}}$ subcarriers for data symbols, with $\tau_{\mathrm{p}}$ and $\tau_{\mathrm{d}}$ fixed. Moreover, each UE transmits at a constant total power, $P_{\mathrm{tot}}$, which is distributed unequally among the subcarriers, so that $\tau_{\mathrm{p}}\rho_k^{\mathrm{p}} + \tau_{\mathrm{d}}\rho_k = P_{\mathrm{tot}},\ \forall k \in \mathcal{K}$ must be ensured, where $\rho_k^{\mathrm{p}}$ and $\rho_k$ denote the pilot and data powers of UE $k$, respectively. Thus, let $\mathbf{s} = \left[s_1, \ldots, s_{\tau_{\mathrm{p}}}\right]^{\mathrm{T}} \in \mathbb{C}^{\tau_{\mathrm{p}} \times 1}$ be an orthogonal pilot sequence with $\left|s_i\right|^2 = 1$, for $i = 1, \ldots, \tau_{\mathrm{p}}$.

The $N \times \tau_{\mathrm{p}}$ matrix of the received pilot signal at the BS from UE $k \in \mathcal{K}$ is:

$$\mathbf{Y}_k^{\mathrm{p}} = \alpha_k \sqrt{\rho_k^{\mathrm{p}}} \mathbf{h}_k \mathbf{s}_k^{\mathrm{T}} + \mathbf{N}, \quad k \in \mathcal{K}, \tag{3.1}$$

where $\mathbf{s}_k \in \mathbb{C}^{\tau_{\mathrm{p}} \times 1}$ is the pilot sequence of UE $k$, $\alpha_k$ and $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ are the large and small scale fading between UE $k$ and the BS, respectively. Specifically, $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ is a circularly symmetric complex normal distributed column vector, and $\mathbf{N} \in \mathbb{C}^{N \times \tau_{\mathrm{p}}}$ is the spatially and temporally additive white Gaussian noise (AWGN) with variance $\sigma_{\mathrm{p}}^2$. We assume that in the consecutive CSI acquisition periods the complex channel vector follows a discrete time AR(1) process:

$$\mathbf{h}_k = \mathbf{h}_k(t) = \mathbf{A}_k \mathbf{h}_k(t-1) + \boldsymbol{\vartheta}_k(t) \in \mathbb{C}^{N \times 1}, \quad k \in \mathcal{K}, \tag{3.2}$$

where $\boldsymbol{\vartheta}_k(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_k)$ is a complex normal distributed noise vector with zero mean and covariance matrix $\mathbf{Q}_k$, which is identically and independently distributed in consecutive CSI acquisition periods. Meanwhile, $\mathbf{A}_k \in \mathbb{C}^{N \times N}$ is the transition matrix of the $\mathrm{AR}(1)$ process of UE $k$. We assume that $\mathbf{h}_k(t)$ is stationary, implying that its mean vector and covariance matrix are constant, i.e., $\mathbf{h}_k(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_k),\ \forall t$ and $k \in \mathcal{K}$, which according to (3.2) leads to $\mathbf{C}_k = \mathbf{A}_k \mathbf{C}_k \mathbf{A}_k^{\mathrm{H}} + \mathbf{Q}_k,\ k \in \mathcal{K}$.

The MU-MIMO received data signal at the BS at time $t$, $\mathbf{y}(t) \in \mathbb{C}^{N \times 1}$, can be written as:

$$\mathbf{y}(t) = \alpha_k \mathbf{h}_k(t) \sqrt{\rho_k} x_k(t) + \sum_{\forall i \neq k} \alpha_i \mathbf{h}_i(t) \sqrt{\rho_i} x_i(t) + \mathbf{n}_{\mathrm{d}}(t), \tag{3.3}$$

where $x_k(t)$ is the transmitted data symbol by UE $k \in \mathcal{K}$ and $\mathbf{n}_{\mathrm{d}}(t)$ is the noise on the received data signal with element-wise variance $\sigma_{\mathrm{d}}^2$. The BS uses the least squares (LS) estimator and, thus, the estimated channel of UE $k$ is given by

$$\hat{\mathbf{h}}_k = (\alpha_k \sqrt{\rho_k^{\mathrm{p}}})^{-1} \mathbf{Y}_k^{\mathrm{p}} \mathbf{s}_k^* (\mathbf{s}_k^{\mathrm{T}} \mathbf{s}_k^*)^{-1} = \mathbf{h}_k + \mathbf{w}_k,\ k \in \mathcal{K}, \tag{3.4}$$

where $\mathbf{w}_k = \left(\alpha_k \tau_{\mathrm{p}} \sqrt{\rho_k^{\mathrm{p}}}\right)^{-1} \mathbf{N} \mathbf{s}_k^*$. Also, $\hat{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k),\ k \in \mathcal{K}$, with

$$\mathbf{R}_k = \mathbb{E}\left\{\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^{\mathrm{H}}\right\} = \mathbf{C}_k + \sigma_{\mathrm{p}}^2 (\alpha_k^2 \tau_{\mathrm{p}} \rho_k^{\mathrm{p}})^{-1} \mathbf{I}_N, \quad k \in \mathcal{K}, \tag{3.5}$$

where $\mathbf{I}_N$ is the identity matrix of size $N$. The channel estimation error is defined as $\mathbf{w}_k = \hat{\mathbf{h}}_k - \mathbf{h}_k,\ k \in \mathcal{K}$, so that $\mathbf{w}_k \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Lambda}_k)$ with

$$\boldsymbol{\Lambda}_k = \sigma_{\mathrm{p}}^2 (\alpha_k^2 (P_{\mathrm{tot}} - \tau_{\mathrm{d}} \rho_k))^{-1} \mathbf{I}_N = q_k \mathbf{I}_N,\ k \in \mathcal{K}. \tag{3.6}$$

We also consider that the channel estimator at the BS uses KF, as an alternative to LS channel estimation, to predict the channel as: $\mathbf{h}_k^{\mathrm{f}}(t) = \mathbf{A}_k \mathbf{h}_k^{\mathrm{a}}(t-1),\ k \in \mathcal{K}$, where $\mathbf{h}_k^{\mathrm{a}}(t-1) = \mathbb{E}_{\mathbf{h}_k(t-1)}\left\{\mathbf{h}_k(t-1)\middle|\hat{\mathbf{h}}_k(t-1)\right\} = \mathbf{D}_k\hat{\mathbf{h}}_k(t-1)$ is the best estimate of $\mathbf{h}_k(t-1)$ based only on the current observation $\hat{\mathbf{h}}_k(t-1)$ at time instant $(t-1)$, with $\mathbf{D}_k = \mathbf{C}_k\mathbf{R}_k^{-1}$ (Fodor *et al.*, 2021).

Using the result of Lemma 2 from (Fodor *et al.*, 2021) and by considering a multi-user case where each user has an independent path loss value, we can derive the individual MSE of the UEs and use it to define important aspects of the proposed solution based on game theory such as the payoff function and Nash equilibrium. Using a generic receiver vector $\mathbf{G}_k(t)$ and assuming $\mathbb{E}\{x_k\} = 0$, $\mathbb{E}\{\mathbf{n}_{\mathrm{d}}\} = 0$, $\mathbb{E}\{x_k x_k^*\} = 1$ and $\mathbb{E}\{\mathbf{n}_{\mathrm{d}}\mathbf{n}_{\mathrm{d}}^{\mathrm{H}}\} = \sigma_{\mathrm{d}}^2\mathbf{I}_N$, the MSE of the estimated data symbols of UE $k$ is:

$$\mathrm{MSE}_k\big(\boldsymbol{\rho}, \mathbf{G}_k(t), \mathbf{h}_k(t)\big) = \mathbb{E}_{x_k,\mathbf{n}_{\mathrm{d}},\mathbf{h}_{i\neq k}}\left\{\left|\mathbf{G}_k(t)\mathbf{y}(t) - x_k(t)\right|^2\right\} =$$
$$\mathbf{G}_k(t)\boldsymbol{\Theta}_k(t)\mathbf{G}_k^{\mathrm{H}}(t) - \mathbf{G}_k(t)\boldsymbol{\gamma}_k(t) - \boldsymbol{\gamma}_k^{\mathrm{H}}(t)\mathbf{G}_k^{\mathrm{H}}(t) + 1, \tag{3.7}$$

where $\boldsymbol{\gamma}_k(t) = \alpha_k\sqrt{\rho_k}\mathbf{h}_k(t)$, $\boldsymbol{\Theta}_k(t) = \alpha_k^2\rho_k\mathbf{h}_k(t)\mathbf{h}_k^{\mathrm{H}}(t) + \sum_{\forall i\neq k}\alpha_i^2\rho_i\mathbf{C}_i + \sigma_{\mathrm{d}}^2\mathbf{I}_N$, and $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_K] \in \mathbb{R}^{1\times K}$ is the vector of data powers of all UEs. For convenience, we define $\boldsymbol{\rho}_{-(k)}$ as the vector of data powers of all UEs except of UE $k$. Depending on the context, we interchangeably use the notation $\boldsymbol{\rho}$ or $(\rho_k, \boldsymbol{\rho}_{-(k)})$. Although the expression in (3.7) is useful we need to derivate the MSE as a function of $\hat{\mathbf{h}}_k(t)$ and $\hat{\mathbf{h}}_k(t-1)$ and, thus, we define:

$$\mathrm{MSE}_k^{\mathrm{KF}}\big(\boldsymbol{\rho}, \mathbf{G}_k(t), \hat{\mathbf{h}}_k(t), \hat{\mathbf{h}}_k(t-1)\big) = \mathbb{E}_{\mathbf{h}_k(t)|\hat{\mathbf{h}}_k(t),\hat{\mathbf{h}}_k(t-1)}\mathrm{MSE}_k\big(\boldsymbol{\rho}, \mathbf{G}_k(t), \mathbf{h}_k(t)\big)$$
$$= \mathbf{G}_k(t)\boldsymbol{\Phi}_k(t)\mathbf{G}_k^{\mathrm{H}}(t) - \mathbf{G}_k(t)\boldsymbol{\Upsilon}_k(t) - \boldsymbol{\Upsilon}_k^{\mathrm{H}}(t)\mathbf{G}_k^{\mathrm{H}}(t) + 1, \tag{3.8}$$

where, according to (Fodor *et al.*, 2021, Lemma 2), we can easily see that

$$\boldsymbol{\Phi}_k(t) = \mathbb{E}_{\mathbf{h}_k(t)|\hat{\mathbf{h}}_k(t),\hat{\mathbf{h}}_k(t-1)}\boldsymbol{\Theta}_k(t) = \alpha_k^2\rho_k\mathbf{E}_k\zeta_k(t)\zeta_k^{\mathrm{H}}(t)\mathbf{E}_k^{\mathrm{H}} + \alpha_k^2\rho_k\mathbf{Z}_k - \alpha_k^2\rho_k\mathbf{C}_k + \xi + \sigma_{\mathrm{d}}^2\mathbf{I}_N, \tag{3.9}$$

$$\boldsymbol{\Upsilon}_k(t) = \mathbb{E}_{\mathbf{h}_k(t)|\hat{\mathbf{h}}_k(t),\hat{\mathbf{h}}_k(t-1)}\boldsymbol{\gamma}_k(t) = \alpha_k\sqrt{\rho_k}\mathbf{E}_k\zeta_k(t), \tag{3.10}$$

where $\xi = \sum_{\forall i}\alpha_i^2\rho_i\mathbf{C}_i$. Taking the derivative of (3.8) with respect to $\mathbf{G}_k(t)$ and making it equal to 0, we can find the optimal MMSE+KF receiver $\mathbf{G}_k^\star(t)$. Substituting $\mathbf{G}_k^\star(t)$ into (3.8), we obtain:

$$\min_{\mathbf{G}_k(t)}\mathrm{MSE}_k^{\mathrm{KF}}\big(\boldsymbol{\rho}, \mathbf{G}_k(t), \hat{\mathbf{h}}_k(t), \hat{\mathbf{h}}_k(t-1)\big) =$$
$$\mathrm{MSE}_k^{\mathrm{KF}}\big(\boldsymbol{\rho}, \hat{\mathbf{h}}_k(t), \hat{\mathbf{h}}_k(t-1)\big) = 1 - \boldsymbol{\Upsilon}_k^{\mathrm{H}}(t)\boldsymbol{\Phi}_k^{-1}(t)\boldsymbol{\Upsilon}_k(t). \tag{3.11}$$

From (3.11) and by assuming that the antennas are properly spaced and the channel coefficients at the different antennas are uncorrelated and identically distributed, i.e., $\mathbf{C}_k = c_k\mathbf{I}_N$, $\mathbf{A}_k = a_k\mathbf{I}_N$, $\mathbf{\Lambda}_k = q_k\mathbf{I}_N$ and $\mathbf{Q}_k = \theta_k\mathbf{I}_N$, where $\theta_k = (c_k - a_k c_k a_k^*)$, we can derive a closed-form expression for the unconditional MSE function according to (Fodor *et al.*, 2021, Section V).

## 3.3 Power Control Approaches for MIMO Systems

Standardized by 3GPP, the fractional power control (FPC) method is a power control mechanism for the uplink of the LTE system. Using parameters estimated by the UE and received from the BS, it specifies the data transmit power for each UE. In general, this approach is parameter-sensitive and its expression consists of two components: the first one is the *open loop* term, whose main objective is trying to compensate channel variations; the second component, called the *closed loop*, basically aims to adjust the open loop term (Baracca *et al.*, ). Commonly, the FPC strategy is simplified so that closed loop corrections are omitted and, thus, the expression of the data transmit power allocated by a given UE $k$ to a resource block simplifies to: $\rho_k = \min\{P_{\text{tot}}, P_{\text{o}} + \beta\alpha_k\}$, where $P_{\text{o}}$ is specific to each UE in the cell and $\beta$ is the fraction of path loss that will be compensated for.

As previously discussed, (Zhao *et al.*, 2018) and (Zhao *et al.*, 2019) proposed approaches for allocating data transmit power in MU-MIMO systems to deal with the problem of minimizing the sum of the MSEs assuming $\mathbf{A}_k = \mathbf{0}$. Therein, efficient non-cooperative games were proposed, where the users minimize their own MSE in a selfish manner to minimize the sum-MSE. Iterative algorithms based on game theory aim to find a Nash equilibrium as a solution, i.e., a strategy profile at which no player has any incentive for unilateral deviation, as shown in Definition 1:

**Definition 1.** *An $\epsilon$-Nash equilibrium of a game is achieved when the $\mathrm{MSE}_k(\boldsymbol{\rho})$ is such that for all UE $k$:*

$$\mathrm{MSE}_k(\boldsymbol{\rho}) \leq \mathrm{MSE}_k\big(\rho_k, \boldsymbol{\rho}_{-(k)}\big) + \epsilon, \ \forall \rho_k. \tag{3.12}$$

Differently from (Zhao *et al.*, 2018; Zhao *et al.*, 2019), in the next section we propose a power allocation policy modeled as a discrete and finite non-cooperative game that depends on the AR channel model parameters and therefore the impact of aging can be properly evaluated. Thus, our solution generalizes the algorithms designed for block fading channels. Moreover, the

proposed algorithm allows us to analyze the sensitivity of its performance with regard to model parameter estimation errors.

### 3.4 Problem Formulation and Decentralized Solution

In this section, we optimize the problem of PDPC in power-constrained MU-MIMO systems as

$$\underset{\boldsymbol{\rho}\in\boldsymbol{\Omega}}{\text{minimize}} \sum_{k\in\mathcal{K}} \gamma_k \mathrm{MSE}_k^{\mathrm{KF}}\left(\boldsymbol{\rho}\right), \tag{3.13}$$

where $\boldsymbol{\Omega} = \left\{\boldsymbol{\rho} \in \mathbb{R}_{+}^{*1\times K} \;\middle|\; \rho_k \in \left(0, P_{\text{tot}}/\tau_{\text{d}}\right), k \in \mathcal{K}\right\}$ and $\{\gamma_k\}_{\forall k} \in \mathbb{R}$ are fixed weights for each UE of the system.

In problem (3.13), the power level of a given UE affects not only its MSE but also the MSEs of the others. Indeed, each UE has the objective of minimizing its own MSE. However, a power choice completely independent of the system conditions can lead to a high interference scenario, resulting in an inefficient sum-MSE. Traditionally, in situations where conflicting interests are present, game theory can provide a mathematical basis for the analysis of interactive decision-making processes (MacKenzie; Silva, 2006). Given that each UE can set its power level without coordination with others, we can model problem (3.13) as a non-cooperative power control strategy game, i.e., problem (3.13) can be expressed as a tuple, $\mathcal{G} = \left\{\mathcal{K}, \boldsymbol{\Upsilon}, \{\mathrm{MSE}_k^{\mathrm{KF}}\left(\boldsymbol{\rho}\right)\}_{\forall k}\right\}$, where $\mathcal{K}$ is a finite set of all players (UEs) and $\boldsymbol{\Upsilon} = \mathcal{P}_1 \times \mathcal{P}_2 \times \cdots \times \mathcal{P}_K$ is the strategy space. For each player $k$, $\mathcal{P}_k = [\rho_{\text{min}}, \rho_{\text{max}}]$ is a discrete and finite set of strategies or actions available, with $\rho_{\text{min}} > 0$ and $\rho_{\text{max}} < P_{\text{tot}}/\tau_{\text{d}}$. Also, the $\mathrm{MSE}_k^{\mathrm{KF}}\left(\boldsymbol{\rho}\right)$: $\boldsymbol{\Upsilon} \to \mathbb{R}$, $k \in \mathcal{K}$, is a real-valued utility function where, for every player $k$, each profile of strategy or action $\boldsymbol{\rho} \in \boldsymbol{\Omega}$ is associated with a payoff, i.e., $\mathrm{MSE}_k^{\mathrm{KF}}\left(\boldsymbol{\rho}\right)$.

In order to develop a game-based solution to problem (3.13), we can exploit an important property of its payoff function. Since the $\mathrm{MSE}_k^{\mathrm{KF}}\left(\boldsymbol{\rho}\right)$ has a convex nature within the range $\mathcal{P}_k \subset \left(0, P_{\text{tot}}/\tau_{\text{d}}\right)$ (Fodor *et al.*, 2021), given the data power allocation $\boldsymbol{\rho}_{-(k)}$ of the other UEs, there exists an optimal unique value $\rho_k^{\star}$ that minimizes $\mathrm{MSE}_k^{\mathrm{KF}}\left(\boldsymbol{\rho}\right)$. We assume in our solution that when it is given the opportunity to set the power level, each UE will choose a *best response* to the actions of other players, i.e., $\rho_k^{\star}(\boldsymbol{\rho}_{-(k)}) = \textbf{arg min}_{\rho_k\in\mathcal{P}_k} \mathrm{MSE}_k^{\mathrm{KF}}(\rho_k, \boldsymbol{\rho}_{-(k)})$. Thereby, a Pareto efficient solution is expected.

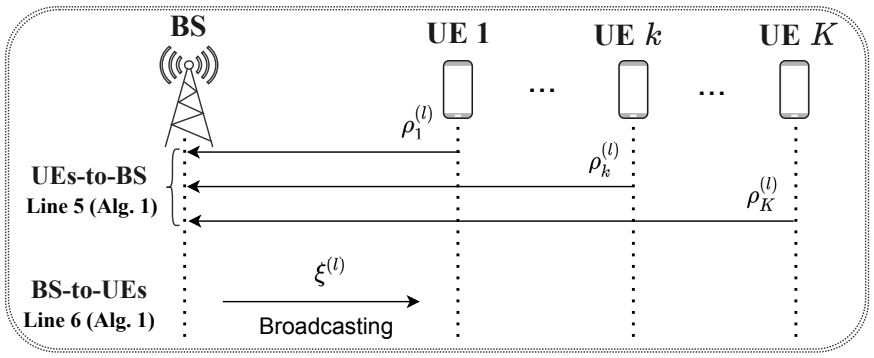
**Assumption 1.** *Considering fixed weights and given the current strategies of the other UEs, if each UE in the non-cooperative game $\mathcal{G}$ takes actions in order to minimize its MSE, then the solution obtained at the end of the game should be a good Pareto solution to problem* (3.13).

The complete procedure of our proposed solution to problem (3.13) is described in Algorithm 3, which can be used in both traditional and massive MU-MIMO scenarios. Initially, we consider that each UE $k$ can estimate its own AR parameter $a_k$ as shown in line 3 (Esfandiari *et al.*, 2020). Using the set $\mathcal{P}_k$, each UE can test a finite number of strategies and select the one that maximizes its payoff. In case of an improvement of the $\mathrm{MSE}_k^{\mathrm{KF}}(\boldsymbol{\rho})$, each UE updates its data power, otherwise the current data power value is kept. Note that the best response approach followed in our game is performed in a fully distributed way, and therefore, increasing $K$ does not provide a significant additional cost to obtain a solution. As a result, the computational complexity of Algorithm 3 does not depend on $K$ and is given by $|\mathcal{P}_k|\mathcal{O}(N)$. Through the definition of $\mathcal{G}$ and Algorithm 3, we are interested in computing a Nash equilibrium as a solution to problem (3.13), as shown in Definition 1. Although not every type of game has a Nash equilibrium as a solution, since the game $\mathcal{G}$ is finite in terms of players and possible strategies for each player, it has a Nash equilibrium in either mixed or pure strategies (MacKenzie; Silva, 2006). Moreover, as presented in Lemma 1, when the best response dynamic converges, it necessarily achieves a Nash equilibrium as a solution, as shown below:

**Algorithm 3:** Non-cooperative power control strategy game

1: **Input:** Initialize $\boldsymbol{\rho}^{(0)} \in \boldsymbol{\Omega}$, $l \leftarrow 0$, $\{\gamma_k\}_{\forall k}$ and $\epsilon > 0$;
2: **Output:** Data power vector $\boldsymbol{\rho}$;
3: Each UE $k$ estimates its own AR parameter $a_k$;
4: **loop**
5: UEs report the current data power to BS;
6: BS measures and broadcasts $\xi^{(l)}$;
7: Find the MSE function using (3.11);
8: Find $\rho_k^{\star} \leftarrow \mathbf{arg\ min}_{\rho_k \in \mathcal{P}_k} \mathrm{MSE}_k^{\mathrm{KF}}(\rho_k, \boldsymbol{\rho}_{-(k)}^{(l)}),\ k \in \mathcal{K}$;
9: **for** $k \in \mathcal{K}$ **do**
10: **if** $\mathrm{MSE}_k^{\mathrm{KF}}\left(\rho_k^{(l)}, \boldsymbol{\rho}_{-(k)}^{(l)}\right) - \mathrm{MSE}_k^{\mathrm{KF}}\left(\rho_k^{\star}, \boldsymbol{\rho}_{-(k)}^{(l)}\right) > \epsilon$ **then**
11: $\rho_k^{(l+1)} \leftarrow \rho_k^{\star}$;
12: **else**
13: $\rho_k^{(l+1)} \leftarrow \rho_k^{(l)}$;
14: **end if**
15: **end for** $l \leftarrow l + 1$;
16: **end loop when** $\rho_k^{(l)} = \rho_k^{(l-1)},\ k \in \mathcal{K}$;

Figure 15 – Signaling exchange scheme involved in Algorithm 3.




Source: Created by the author.

**Lemma 1.** *If the best reponse dynamic for game $\mathcal{G}$ converges to a particular strategy profile $\boldsymbol{\rho}^* \in \boldsymbol{\Omega}$, then, $\boldsymbol{\rho}^*$ must be a Nash equilibrium (MacKenzie; Silva, 2006).*

It is worth mentioning that the authors in (Zhao *et al.*, 2019) proved that Algorithm 3 converges to a Nash equilibrium for a particular case where $\mathbf{A}_k = \mathbf{0}$. In short, the authors were able to analytically find $\rho_k^\star(\boldsymbol{\rho}_{-(k)})$ and, by showing that $\rho_k^\star(\boldsymbol{\rho}_{-(k)})$ is within a limited range and is strictly monotonically increasing with respect to $\boldsymbol{\rho}_{-(k)}$, the convergence of their algorithm was proved by means of the monotone convergence theorem. However, obtaining an analytical expression for $\rho_k^\star(\cdot)$ when $\mathbf{A}_k \neq \mathbf{0}$ is equivalent to finding the root of a high-order polynomial, which is extremely challenging. Although it is intuitive to imagine that $\rho_k^\star(\cdot)$ is always strictly monotonically increasing in the data power of the other UEs in any interference-aware system, the formal convergence proof of Algorithm 3 when $\mathbf{A}_k \neq \mathbf{0}$ remains an open problem.

During the execution of Algorithm 3, an over-the-air signaling scheme must be considered for the UEs-to-BS and BS-to-UEs communications. For that, in line $5$, each UE initially reports to the BS its current data power value. Next, in line $6$, the BS broadcasts $\xi$ and $\boldsymbol{\Phi}_k(t)$ can be computed in (3.9) in order to obtain the MSE function. This signaling process allows the UEs to have some idea of the level of interference in the system to make a proper choice of data powers in line $8$. The proposed signaling scheme is depicted in Fig. 15.

## 3.5 Simulation Assumptions and Discussions

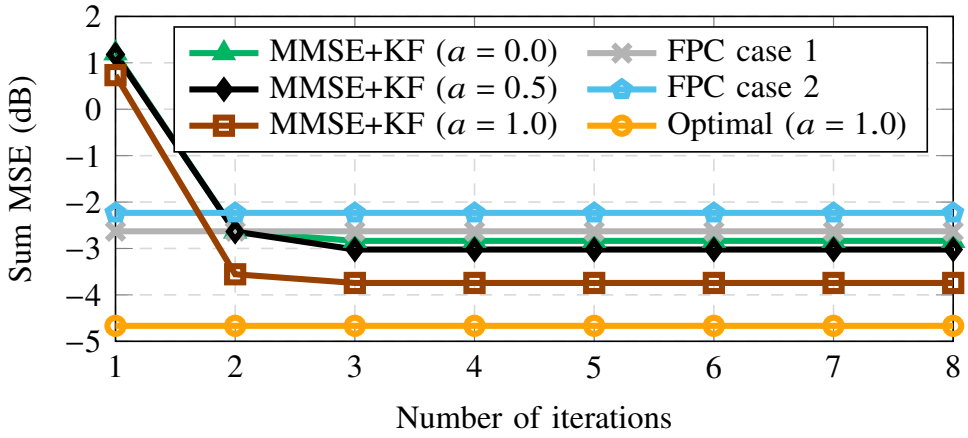
The setup considered herein is an extension of the scenario from (Fodor *et al.*, 2021) to multi-user scenarios. Unless otherwise stated, the values of $a_k = a, \forall k$ and are perfectly known. Also, $N = 100$, $K = 4$, and the path loss between the BS and UEs is uniformly

distributed within the range $[60, 75]$ dB. We adopt $\tau_{\mathrm{p}} = 1$, $\tau_{\mathrm{d}} = 11$, $\sigma_{\mathrm{p}}^2 = \sigma_{\mathrm{d}}^2 = 7.1659 \times 10^{-16}$ W, $P_{\mathrm{tot}} = 250$ mW and $\gamma_k = 1, k \in \mathcal{K}$. The results were obtained by running $500$ Monte Carlo simulations.

We use $5$ schemes for performance comparison: Algorithm 3 (marked as MMSE+KF game in the plots), the optimal solution obtained by exhaustive search (marked as Optimal), the naive solution in which the estimated channel is taken as if it were the actual channel (marked as Naive game) and the conventional solution where $\mathbf{A}_k = \mathbf{0}$ presented in (Zhao *et al.*, 2019) (marked as MMSE game). We also consider the fractional power control (marked as FPC), which uses the MMSE receiver from (Zhao *et al.*, 2019) that was designed for block fading channels. For that solution, we consider two different values for $P_{\mathrm{o}}$ and set the best $\beta$ for each of them, namely: case 1 with $P_{\mathrm{o}} = -45$ dBm, $\beta = 0.8$, and case 2 with $P_{\mathrm{o}} = -40$ dBm, $\beta = 0.7$. All results were obtained from an Intel Core i7-7500U computer with $2.90$ GHz and $8$ GB RAM.

### *3.5.1 Convergence Behavior*

Figure 16 – Convergence curves of Algorithm 3.




Source: Created by the author.

Fig. 16 shows the convergence curves of Algorithm 3 for three values of the AR parameter. The average per iteration execution time of Algorithm 3 is around $25$ ms[1]. We also show the FPC solutions and the optimal solution obtained via exhaustive search, with average execution times around $0.5$ ms and $135$ s, respectively. We emphasize the rapid convergence of Algorithm 3 that, as stated in Lemma 1, achieves a Nash equilibrium regardless of the choice of

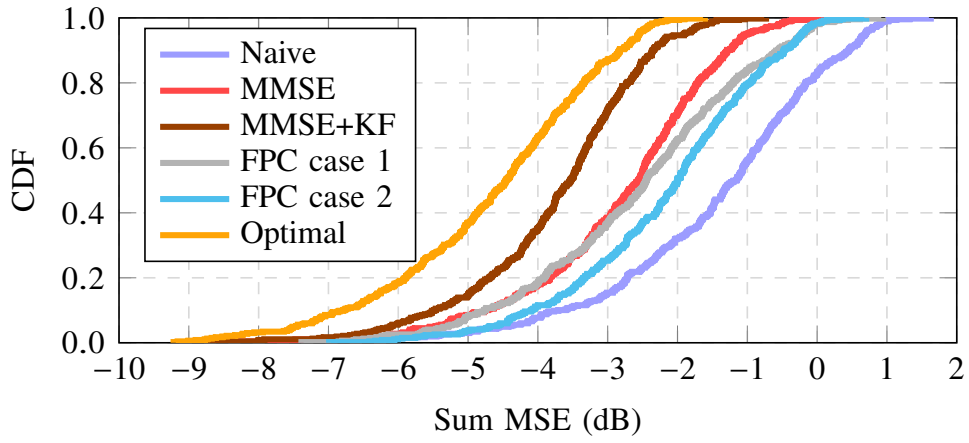
[1] For real systems the execution time of Algorithm 3 would be further improved by optimizations in the code, programming language and hardware.

the AR parameter, even though the Nash Equilibrium changes depending on the AR parameter value. Note that after the first iteration only marginal gains are obtained. Indeed, such a rapid convergence is interesting as it avoids excessive exchange of information between the BS and the UEs in real systems, which occurs in lines $5$ and $6$ of Algorithm 3. Also, it is worth noting that important gains are obtained as the $a$ value increases, i.e., as the memoryfull property of the fading channel is exploited. For $a = 1$, a gap smaller than 1 dB is reached in Nash equilibrium compared to the optimal solution.

### *3.5.2 Performance Comparison*

Taking into account the MMSE and Naive games as benchmarking, Fig. 17 depicts cumulative distribution function (CDF) curves of the sum-MSE. For the MMSE+KF game and the optimal solution we assume $a = 1$. In this figure, the performance gain at the $50$-th percentile of sum-MSE of the proposed solution compared to all other sub-optimal solutions is at least $1$ dB. The superior performance of the proposed solution is due to the fact that, differently from the other solutions, it can exploit channel aging models through the AR parameter. Particularly, the FPC solution does not directly deal with the interference generated by the UEs. Even though the proposed solution is a non-cooperative game, the power choice of a given UE $k$ is based on the $\xi$ value, which includes the current powers of all UEs.

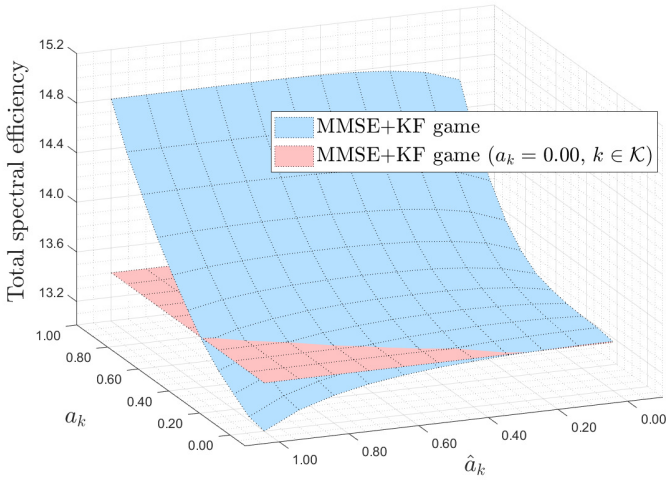
Figure 17 – CDF curves of the sum of MSEs.




Source: Created by the author.

Although previous results assume perfect knowledge of the AR parameter, in real scenarios the actual values $\{a_k\}_{\forall k}$ are unknown and must be estimated (Esfandiari *et al.*, 2020). In this context, Figs. 18 and 19 investigate how our solution deals with estimation errors. First,

Figure 18 – Total spectral efficiency for MMSE+KF solution versus the actual AR parameter ($a_k$) and its estimated value ($\hat{a}_k$).



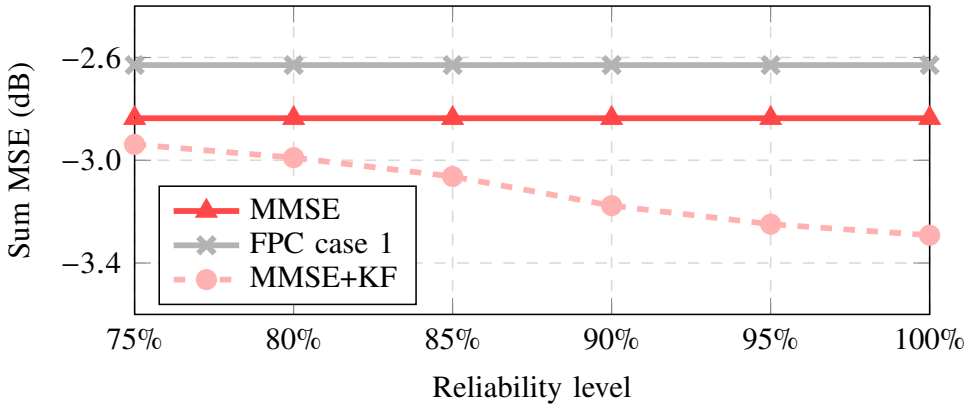

Source: Created by the author.

Fig. 18 shows the total spectral efficiency as a function of the actual AR parameter ($a_k$) and its estimated value ($\hat{a}_k$). For the red surface the channel is not correlated in time, i.e., $a_k = 0,\ \forall k$, while for the blue surface the UEs have the same value for the AR parameter but we vary it within $[0, 1]$. Interestingly, this result shows two important features of Algorithm 3. First, we can observe that the proposed solution is sensitive to the parameter $a_k$. This is interesting because, as $a_k$ varies in the system, important performance gains can be achieved compared to the case that assumes memoryless channel models (red surface).

The second aspect is that the proposed algorithm becomes robust to estimation errors as $a_k$ increases. Specifically, this last result is less trivial and can be well observed when $a_k = 1,\ \forall k$. In this situation, note that any estimate to $a_k$ leads to negligible performance losses. Instead, when $a_k = 0,\ \forall k$, a good estimate is considerably more important because a poor estimate for $a_k$ leads to lower performance compared to the most trivial case of the game, i.e., when it is designed for block fading channels.

Differently from Fig. 18, in Fig. 19 the UEs have independent values of the AR parameter chosen randomly between 0 and 1. We assume that $\hat{a}_k$ can be found with a given reliability level according to $\hat{a}_k = \varrho a_k + \sqrt{1-\varrho^2}u,\ k \in \mathcal{K}$, where $0 \leq \varrho \leq 1$ denotes the reliability level of the estimate and $u$ is a Gaussian variable with zero mean and unit variance. In Fig. 19 the dashed line shows the performance of Algorithm 3 in more realistic scenarios when it

Figure 19 – Sum-MSE for different reliability levels ($\varrho$).




Source: Created by the author.

is executed in the presence of estimation errors and with independent values of the AR parameter for each UE. We also depict the best curve of the LTE solution (i.e., case 1) and the MMSE game for benchmaking purposes.

Interestingly, even in the presence of estimation errors, the proposed solution outperforms the best benchmarking solutions evaluated herein. Furthermore, as expected, the performance of the proposed solution increases when the reliability of the estimation increases. However, note that, in general, the reliability level has little impact on the proposed allocation policy and only a small performance loss arises as $\varrho$ decreases.

## 3.6 Chapter Summary

Numerical results showed that, by exploiting the correlation structure present in real channels, the proposed solution outperforms the classical LTE FPC and game theory-based solutions designed for block fading channels. In scenarios where the knowledge of channel correlation coefficients is not perfect, it was shown that estimation errors have little impact on the performance of the proposed solution.

# 4 A GAME-THEORETIC DESIGN TO POWER CONTROL IN CELL-FREE NETWORKS

This chapter investigates cell-free massive multiple input multiple output systems with a particular focus on uplink power allocation. In these systems, uplink power control is highly non-trivial, since a single user terminal is associated with multiple intended receiving base stations. In addition, in cell-free systems, distributed power control schemes that address the inherent spectral and energy efficiency targets are desirable. By utilizing tools from game theory, we formulate our proposal as a non-cooperative game, and using the best-response dynamics, we obtain a distributed power control mechanism.

To ensure that this power control game converges to a Nash equilibrium, we apply the theory of potential games. Differently from existing game-based schemes, interestingly, our proposed potential function has a scalar parameter that controls the power usage of the users. Numerical results confirm that the proposed approach improves the use of the energy stored in the battery of user terminals and balances between spectral and energy efficiency.

## 4.1 Introduction and Related Works

Undoubtedly, radio resource management (RRM) is a major issue in the design of modern mobile networks. In interference-limited systems, for example, power control plays an indispensable role in managing interference, ensuring proper signal strength at the intended receivers and saving energy. Recognizing the scarcity of the energy resource and the growing worldwide energy concern, efficient power control solutions have definitely become a key requirement for the continued success of wireless systems (Miao *et al.*, 2012; Liu *et al.*, 2014; He *et al.*, 2017). Especially related to the uplink and given the ever-increasing growth in mobile subscriptions, an efficient power allocation strategy is important to reduce energy demands and battery consumption. By mitigating interference levels, power control has also the advantage of providing a more uniform throughput among users. Furthermore, an optimized energy consumption contributes to reducing environmental impacts, e.g., heat dissipation and electronic pollution (Miao, 2013).

However, the power management in cellular networks is a fairly complex problem. In general, efficiently controlling power usage with multiple interfering users may lead to non-polynomial time (NP)-hard problems, and in these cases obtaining optimal solutions is extremely challenging. Normally, within the power allocation framework, the main difficulties in finding

alternative solutions are the performance coupling among the users as well as their inherently selfish behaviors. Consequently, a good solution needs to deal with the interactions among several independent users with contrasting interests. In this context, game theory provides a natural framework for developing mechanisms when many individuals with conflicting interests interact. Therefore, it is a promising approach to study interactions among contending users in order to seek feasible and practically viable solutions (MacKenzie; Silva, 2006; Osborne, 2004). Indeed, there has been growing interest in adopting game-theoretic methods to propose alternative solutions in mobile communications, see, e.g., (Buzzi *et al.*, 2012; Xie *et al.*, 2014; Zhao *et al.*, 2018; Zhao *et al.*, 2019; Saraiva *et al.*, 2022; Myung *et al.*, 2022).

In (Buzzi *et al.*, 2012) and (Xie *et al.*, 2014), the authors focused on an uplink power control game-based solutions for orthogonal frequency division multiple access (OFDMA) systems and cognitive radio networks, respectively. To address the problem of minimizing the sum of the mean squared errors (MSEs), power control schemes for the uplink of massive multiuser MIMO (MU-MIMO) systems were proposed in (Zhao *et al.*, 2018; Zhao *et al.*, 2019; Saraiva *et al.*, 2022). More specifically, considering block fading channels, (Zhao *et al.*, 2018) and (Zhao *et al.*, 2019) relied on game theory to optimize the pilot-to-data power ratio assuming single and multi-cell cases, respectively. Likewise, a game-based approach to controlling the pilot and data power levels was presented in (Saraiva *et al.*, 2022) while considering more realistic auto-regressive channels.

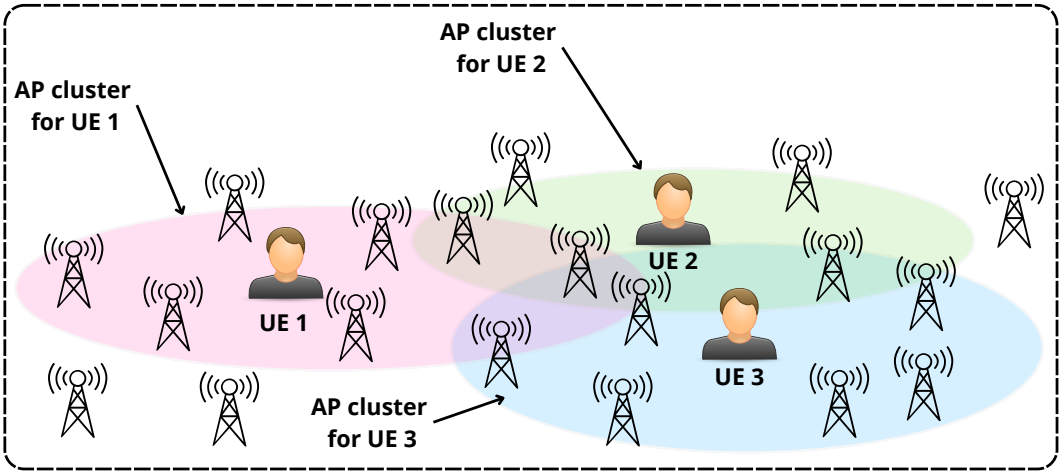
On the other hand, in more recent network architectures such as cell-free, game-theoretic approaches to radio resource allocation have not been widely explored in the literature, mainly related to the uplink. Recently, in (Myung *et al.*, 2022), a power control game was proposed for cell-free massive multiple input multiple output (MIMO) systems, but the authors focused on the downlink.

### *4.1.1 Main Contributions*

Inspired by the above discussion, this chapter considers the problem of power control for the uplink of cell-free massive MIMO systems. Specifically, our main contributions are:

1. Due to the inherent competitive nature of the multi-user and user-centric environment, we use a game-theoretic framework and model the problem as a strategic non-cooperative game, which can often provide feasible and convenient alternatives for a distributed implementation.

Figure 20 – A cell-free system with multiple single-antenna users.




Source: Created by the author.

2. We use novel payoff functions based on an adapted signal-to-interference ratio (SIR) expression. More importantly, different from existing works, our solution is designed as a parameterized potential game for which the existence and uniqueness of a Nash equilibrium is ensured.
3. Thereby, we show that the proposed power control achieves efficient solutions with respect to different network objectives such as sum-rate maximization, max-min fairness or power consumption minimization.

## 4.2 Network Model

We focus on a cell-free system with multiple single-antenna users according to Figure 20. Specifically, we consider a cell-free massive MIMO system consisting of $K$ single-antenna user equipments (UEs) and $L$ access points (APs) equipped with $N$ antennas grouped in the sets $\mathcal{K}$ and $\mathcal{L}$, respectively. The APs and UEs are deployed randomly in a wide area without boundaries. A central processing unit (CPU) connects with the APs via a backhaul network.

Particularly, we analyze a cell-free massive MIMO system operating in time division duplexing (TDD) mode with a pilot phase for channel estimation and a data transmission phase. Each coherence block is divided into $\tau_{\mathrm{p}}$ channel uses for uplink pilots, $\tau_{\mathrm{u}}$ for uplink data and $\tau_{\mathrm{d}}$ for downlink data such that $\tau_{\mathrm{c}} = \tau_{\mathrm{p}} + \tau_{\mathrm{u}} + \tau_{\mathrm{d}}$. The channel between AP $l$ and UE $k$ is denoted as $\mathbf{h}_{kl} \in \mathbb{C}^N$ and $\mathbf{h}_k = \left[\mathbf{h}_{k1}^{\mathrm{T}}, \ldots, \mathbf{h}_{kL}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{NL}$ is the collective channel from all APs. In each coherence block, an independent realization from a correlated Rayleigh fading distribution is drawn as $\mathbf{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_{kl})$, where $\mathbf{R}_{kl}$ is the spatial correlation matrix describing

the spatial property of the channel and $\beta_{kl} = \mathrm{tr}(\mathbf{R}_{kl})/N$ is the large-scale fading coefficient that describes pathloss and shadowing (Björnson; Sanguinetti, 2020; Chen *et al.*, 2021). The Gaussian distribution models the small-scale fading whereas the positive semi-definite correlation matrix $\mathbf{R}_{kl}$ describes the large-scale fading, including shadowing, pathloss, spatial channel correlation and antenna gains. Given that the APs are spatially distributed in the system, the channel vectors of different APs are independently distributed, i.e., $\mathbb{E}\left\{\mathbf{h}_{kl'}\left(\mathbf{h}_{kl}\right)^{\mathrm{H}}\right\} = \mathbf{0}$ when $l' \neq l$. The collective channel is distributed as $\mathbf{h}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_k)$, where $\mathbf{R}_k = \mathrm{diag}\left(\mathbf{R}_{k1}, \ldots, \mathbf{R}_{kL}\right) \in \mathbb{C}^{NL \times NL}$ is the block-diagonal spatial correlation matrix (Björnson; Sanguinetti, 2020).

We define the block-diagonal matrices $\mathbf{D}_k = \mathrm{diag}\left(\mathbf{D}_{k1}, \ldots, \mathbf{D}_{kL}\right) \in \mathbb{C}^{NL \times NL}$,$k \in \mathcal{K}$, where $\mathbf{D}_{il} \in \mathbb{C}^{N \times N}$, $i \in \mathcal{K}$ and $l \in \mathcal{L}$ is the set of diagonal matrices, determining which AP antennas may transmit to which UEs. More specifically, the $n$-th diagonal entry of $\mathbf{D}_{il}$ is 1 if the $n$-th antenna of AP $l$ is allowed to transmit and to decode signals from UE $k$, and 0 otherwise. Based on the definition of the set of matrices $\mathbf{D}_{il}$, we define a matrix $\mathbf{A} \in \mathbb{R}^{K \times L}$ specifying the AP selection, where $A_{k,l} = 1$ if AP $l$ is allowed to transmit and to decode signals from UE $k$, i.e., if $\mathrm{tr}\left(\mathbf{D}_{kl}\right) > 0$, and 0 otherwise. For the conciseness of mathematical descriptions, we denote by $\mathcal{M}_k = \{l \,|\, A_{k,l} = 1, l \in \mathcal{L}\}$ the subset of APs serving UE $k$. Meanwhile, $\mathcal{D}_l = \{k \,|\, A_{k,l} = 1, k \in \mathcal{K}\}$ is the subset of UEs served by AP $l$.

### *4.2.1 Channel Estimation*

We consider that there are $\tau_{\mathrm{p}}$ mutually orthogonal $\tau_{\mathrm{p}}$-length pilots, with $\tau_{\mathrm{p}}$ being a constant independent of $K$. Let $\mathcal{S}_t \subset \mathcal{K}$ be the subset of UEs assigned to pilot $t$. When the UEs in $\mathcal{S}_t$ transmit, the received signal $\mathbf{y}_{tl}^{\mathrm{pilot}} \in \mathbb{C}^N$ at AP $l$ is

$$\mathbf{y}_{tl}^{\mathrm{pilot}} = \sum_{i \in \mathcal{S}_t} \sqrt{\tau_{\mathrm{p}} \rho_i} \mathbf{h}_{il} + \mathbf{n}_{tl}, \tag{4.1}$$

where $\rho_i$ is the transmit power of UE $i$, $\tau_{\mathrm{p}}$ is the processing gain, and $\mathbf{n}_{tl} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ is the thermal noise. Note that, since we assume a massive access scenario with a large number of UEs, i.e., $K > \tau_{\mathrm{p}}$, several UEs share the same pilot as shown in (4.1), leading to pilot contamination.

For estimating the channels, the classic minimum mean squared error (MMSE) criterion has been recurrently employed in the literature. The MMSE estimate of $\mathbf{h}_{kl}$ for UE $k \in \mathcal{S}_t$ is $\hat{\mathbf{h}}_{kl} = \sqrt{\tau_{\mathrm{p}} \rho_k} \mathbf{R}_{kl} \boldsymbol{\Psi}_{tl}^{-1} \mathbf{y}_{tl}^{\mathrm{pilot}}$, where $\boldsymbol{\Psi}_{tl} = \sum_{i \in \mathcal{S}_t} \tau_{\mathrm{p}} \rho_i \mathbf{R}_{il} + \sigma^2 \mathbf{I}_N$ is the correlation matrix of (4.1). The estimated channel $\hat{\mathbf{h}}_{kl}$ and estimation error $\tilde{\mathbf{h}}_{kl} = \mathbf{h}_{kl} - \hat{\mathbf{h}}_{kl}$ are independent vectors distributed as $\hat{\mathbf{h}}_{kl} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{B}_{kl})$ and $\tilde{\mathbf{h}}_{kl} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C}_{kl})$, where $\mathbf{B}_{kl} = \mathbb{E}\left\{\hat{\mathbf{h}}_{kl} \hat{\mathbf{h}}_{kl}^{\mathrm{H}}\right\} =$

$\tau_{\mathrm{p}}\rho_k\mathbf{R}_{kl}\boldsymbol{\Psi}_{tl}^{-1}\mathbf{R}_{kl}$ and $\mathbf{C}_{kl} = \mathbb{E}\left\{\tilde{\mathbf{h}}_{kl}\tilde{\mathbf{h}}_{kl}^{\mathrm{H}}\right\} = \mathbf{R}_{kl} - \mathbf{B}_{kl}$.

### *4.2.2 Uplink Data Transmission*

During the uplink data transmission, AP $l$ receives the signal $\mathbf{y}_l \in \mathbb{C}^N$ from all UEs, as

$$\mathbf{y}_l = \sum_{k\in\mathcal{K}} \mathbf{h}_{kl} s_k + \mathbf{n}_l, \tag{4.2}$$

where $s_k \in \mathbb{C}$ is the signal transmitted from UE $k$ with power $\rho_k$ and $\mathbf{n}_l \sim \mathcal{NC}(\mathbf{0}, \sigma^2\mathbf{I}_N)$. However, since only a subset of APs take part in the signal detection, the estimate of $s_k$ is:

$$\hat{s}_k = \sum_{l\in\mathcal{L}} \mathbf{v}_{kl}^{\mathrm{H}}\mathbf{D}_{kl}\mathbf{y}_l = \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{h}_k s_k + \sum_{i\in\mathcal{K}\setminus\{k\}} \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{h}_i s_i + \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{n}, \tag{4.3}$$

where $\mathbf{v}_{kl} \in \mathbb{C}^N$ is a receive combining vector of AP $l$ for UE $k$, $\mathbf{v}_k = \left[\mathbf{v}_{k1}^{\mathrm{T}}, \ldots, \mathbf{v}_{kL}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{NL}$ denotes the collective of these combining vectors and $\mathbf{n} = \left[\mathbf{n}_1^{\mathrm{T}}, \ldots, \mathbf{n}_L^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{NL}$ collects all the noise vectors.

Preferably, for large-scale networks, it is more interesting to direct the main computational tasks to the APs in a distributed way and, thus, avoid overloading the CPU. Therefore, instead of sending $\{\mathbf{y}_{tl}^{\mathrm{pilot}}\}_{\forall t}$ and $\mathbf{y}_l$ to the CPU, each AP $l$ locally selects the combining vector $\mathbf{v}_{kl}$ and then it preprocesses its signal by computing local estimates of the data as $\hat{s}_{kl} = \mathbf{v}_{kl}^{\mathrm{H}}\mathbf{D}_{kl}\mathbf{y}_l$. Next, the local estimates of all APs that serve UE $k$ are sent to the CPU for final estimate of $s_k$, which is given by $\hat{s}_k = \sum_{l\in\mathcal{L}} \hat{s}_{kl}$. We utilize the *use-and-then-forget* bound to obtain the achievable spectral efficiency (SE).

**Lemma 2.** (Björnson; Sanguinetti, 2020; Chen *et al.*, 2021). *An achievable uplink SE for UE $k$ is*

$$\mathrm{SE}_k = \frac{\tau_{\mathrm{u}}}{\tau_{\mathrm{c}}} \log_2(1 + \mathrm{SINR}_k), \tag{4.4}$$

*where*

$$\mathrm{SINR}_k = \frac{\rho_k \left|\mathbb{E}\{\mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{h}_k\}\right|^2}{\sum\limits_{i\in\mathcal{K}} \rho_i \mathbb{E}\left\{\left|\mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{h}_i\right|^2\right\} - \rho_k\left|\mathbb{E}\{\mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{h}_k\}\right|^2 + \sigma^2\mathbb{E}\left\{\left|\left|\mathbf{D}_k\mathbf{v}_k\right|\right|^2\right\}}. \tag{4.5}$$

In general, any combining vector that depends on the local channel estimates and statistics can be used in the signal to interference-plus-noise ratio (SINR) expression (4.5), but the

expectations in it cannot be computed in closed form for any set of values $\{\mathbf{v}_{kl}\}_{\forall k,l}$. With simpler combining vector structures, such as maximum ratio combining (MRC), i.e., when $\mathbf{v}_{kl} = \hat{\mathbf{h}}_{kl}$, it is possible to obtain closed form expressions. Nevertheless, the performance of MRC is quite limited, and significant performance gains can be obtained when using combining vectors also with distributed structures but based on the MMSE criterion, such as local partial MMSE (LP-MMSE) combining (Björnson; Sanguinetti, 2020; Chen *et al.*, 2021), whose combining vector $\mathbf{v}_{kl}^{\mathrm{LP-MMSE}}$ is:

$$\mathbf{v}_{kl}^{\mathrm{LP-MMSE}} = \rho_k \left( \sum_{i \in \mathcal{D}_l} \rho_i \left( \hat{\mathbf{h}}_{il} \hat{\mathbf{h}}_{il}^{\mathrm{H}} + \mathbf{C}_{il} \right) + \sigma^2 \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_{kl}. \tag{4.6}$$

However, when using $\{\mathbf{v}_{kl}^{\mathrm{LP-MMSE}}\}_{\forall k,l}$ the SINR expression in (4.5) can only be computed via Monte Carlo simulations (Björnson; Sanguinetti, 2020).

## 4.3 Decentralized Solution

In this section, we describe the RRM employed in the network that consists of two parts. In the first part, we simply adopt the algorithm for joint initial access, pilot assignment, and cluster formation proposed in (Björnson; Sanguinetti, 2020, See Section V-A). Then, in the second part, differently from (Björnson; Sanguinetti, 2020), we pay special attention to power control and propose a game-theoretic model of the interactions among users assuming a distributed management framework, which is presented in details in the following subsections.

### *4.3.1 Game Theoretic Approach*

In the context of game theory, the players are considered as entities with the ability of observation and reaction. For the proposed game model, in order to mitigate potential interference levels in the uplink and obtain a suitable data power profile, the players are the users themselves. We define the proposed non-cooperative game $\mathcal{G}$ as $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{P}_k\}_{\forall k}, \{\mu_k(\rho_k, \boldsymbol{\rho}_{(-k)})\}_{\forall k}\right\}$, where $\mathcal{K}$ is the set of UEs, i.e., a finite set of players. For a given UE $k$, $\mathcal{P}_k = [\rho_{\min}, \rho_{\max}]$ is a finite set of available strategies or actions, where $\rho_{\min} > 0$ and $\rho_{\max} \leq P_{\max}$ with $P_{\max}$ being the maximum uplink data power. In the context of the proposed game, the data power value $\rho_k \in \mathcal{P}_k$ denotes the strategy chosen by UE $k$ and $\boldsymbol{\rho}_{(-k)}$ denotes the strategies of all the UEs other than UE $k$. Therefore, $\boldsymbol{\rho} = (\rho_k, \boldsymbol{\rho}_{(-k)}) = [\rho_1, \cdots, \rho_k, \cdots, \rho_K]^{\mathrm{T}} \in \mathbb{R}^K$ represents the profile of data powers of all UEs, i.e., a power allocation strategy for the system. Moreover, $\mu_k(\rho_k, \boldsymbol{\rho}_{(-k)})$: $\boldsymbol{\Upsilon} \rightarrow \mathbb{R}$,

is a real-valued utility/payoff function where, $\boldsymbol{\Upsilon} = \mathcal{P}_1 \times \mathcal{P}_2 \times \cdots \times \mathcal{P}_K$ is the strategy space. Note that, for every chosen strategy by UE $k$, the power profile $(\rho_k, \boldsymbol{\rho}_{(-k)})$ is associated with a payoff, i.e., $\mu_k(\rho_k, \boldsymbol{\rho}_{(-k)})$. Thus, the payoff function quantifies the preferences of each UE to a given action, provided the knowledge of others' actions.

Typically, a non-cooperative game is a procedure where each player will selfishly choose an action that improves its own utility function given the current strategies of the other players. Then, a key issue when designing a game is the choice of the payoff function. Specifically, for game $\mathcal{G}$ the independent actions of the UEs to set their power values should not only provide satisfactory individual solutions but should also mitigate potential interference levels in the network. Thereby, we design a payoff function that enables UEs to have lower data power levels while causing less interference in other UEs, given by:

$$\mu_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)}) = \gamma_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)}) + \lambda_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)}), \tag{4.7}$$

where

$$\gamma_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)}) = \frac{\sum\limits_{\forall i \neq k} \rho_i \Big( \sum\limits_{l \in \mathcal{M}_i} \beta_{i,l} \Big)^{\alpha}}{\rho_k \Big( \sum\limits_{l \in \mathcal{M}_k} \beta_{k,l} \Big)^{\alpha}}, \tag{4.8a}$$

$$\lambda_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)}) = \rho_k \Big( \sum_{l \in \mathcal{M}_k} \beta_{k,l} \Big)^{\alpha} \sum_{\forall i \neq k} \frac{1}{\rho_i \Big( \sum\limits_{l \in \mathcal{M}_i} \beta_{i,l} \Big)^{\alpha}}, \tag{4.8b}$$

and $\alpha$ is an input parameter of game $\mathcal{G}$.

Particularly, the term $\gamma_k(\alpha, \boldsymbol{\rho})$ is based on the reciprocal of the SIR expression shown in (Chen *et al.*, 2021, Section V, Equation (51)). First, assuming the particular case where $\{\lambda_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)})\}_{\forall k} = 0$ and given $\alpha$ and $\boldsymbol{\rho}_{(-k)}$ fixed, it is easy to see that the best strategy or action for each UE exists and it would be to minimize the payoff function in (4.7) by choosing the highest possible value for the data power. However, we add the term $\lambda_k(\alpha, \boldsymbol{\rho})$ to the payoff function $\mu_k(\alpha, \boldsymbol{\rho})$ in order to make the decision of the UEs non-trivial and especially less selfish. In a general case, i.e., even if $\{\lambda_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)})\}_{\forall k} \neq 0$, it is possible to obtain a single value that minimizes (4.7) as it is a convex function in $\mathcal{P}_k, \forall k \in \mathcal{K}$ (for more details, see **APPENDIX C**). By solving $\partial \mu_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)}) / \partial \rho_k = 0$, we can find the unique minimizer, $\rho_k^*$, of (4.7), as follows:

$$\rho_k^* = \sqrt{\left( \sum_{\forall i \neq k} \rho_i \Big( \sum_{l \in \mathcal{M}_i} \beta_{i,l} \Big)^{\alpha} \right) \left( \sum_{\forall i \neq k} \frac{\Big( \sum\limits_{l \in \mathcal{M}_k} \beta_{k,l} \Big)^{2\alpha}}{\rho_i \Big( \sum\limits_{l \in \mathcal{M}_i} \beta_{i,l} \Big)^{\alpha}} \right)^{-1}}. \tag{4.9}$$

From the point of view of the UEs, the term $\lambda_k(\alpha, \boldsymbol{\rho})$ in (4.7) represents a punishment for the UE who decides to excessively increase the value of the chosen data power. As a result, $\lambda_k(\alpha, \boldsymbol{\rho})$ can reduce interference levels in the system and avoid a greedy power allocation strategy, i.e., $\rho_k = P_{\max}$, $\forall k \in \mathcal{K}$.

In game theoretic approaches, we are interested in finding a Nash equilibrium as a solution. A Nash equilibrium is a strategy profile that satisfies the condition that no player can unilaterally improve its own payoff as shown in Definition 2:

**Definition 2.** *An $\epsilon$-Nash equilibrium of parameterized game $\mathcal{G}(\alpha)$ is achieved when the payoff function $\mu_k(\alpha, \boldsymbol{\rho})$ is such that for all UE $k$:*

$$\mu_k(\alpha, \boldsymbol{\rho}) \leq \mu_k\big(\alpha, \rho_k, \boldsymbol{\rho}_{-(k)}\big) + \epsilon, \ \ \rho_k \in \mathcal{P}_k. \tag{4.10}$$

However, games may have a large number of Nash equilibrium points or may not have any. Generally, finding or even characterizing the set of these equilibrium points in terms of existence or uniqueness is a difficult task. Fortunately, there is a particular case of non-cooperative games called *potential games* for which the existence and uniqueness of a Nash equilibrium is ensured (Scutari *et al.*, ). Basically, in a potential game the incentive of all players to change their actions can be expressed by a global payoff function called *potential function*. Mathematically, the proposed parameterized game $\mathcal{G}(\alpha)$ is a potential game if it complies with Definition 3 (MacKenzie; Silva, 2006).

**Definition 3.** *If the proposed parameterized game $\mathcal{G}(\alpha)$ is a potential game, then there exists a function $u : \boldsymbol{\Upsilon} \rightarrow \mathbb{R}$ such that $\forall k \in \mathcal{K}$ and $\forall \rho_k, \rho'_k \in \mathcal{P}_k$:*

$$u(\alpha, \rho'_k, \boldsymbol{\rho}_{(-k)}) - u(\alpha, \rho_k, \boldsymbol{\rho}_{(-i)}) = \mu_k(\alpha, \rho'_k, \boldsymbol{\rho}_{(-i)}) - \mu_k(\alpha, \rho_k, \boldsymbol{\rho}_{(-k)}). \tag{4.11}$$

*In this case, the function $u(\cdot)$ is called an exact potential function for the parameterized game $\mathcal{G}(\alpha)$* (MacKenzie; Silva, 2006; Osborne, 2004).

Note that based on Definition 3, $\mathcal{G}(\alpha)$ is a potential game if it is possible to define a potential function, i.e., an UE-independent function that measures the same amount of change or marginal payoff for any unilaterally deviating UE. By exploiting the definition of payoff function in (4.7), we can prove that $\mathcal{G}(\alpha)$ is a potential game by showing that it has an exact potential function as explained in the following result:

**Corolário 4.3.1.** *There is an exact potential function for the parameterized game $\mathcal{G}(\alpha)$ and it is given by:*

$$u(\alpha, \boldsymbol{\rho}) = \frac{1}{2} \sum_{k \in \mathcal{K}} \Delta\mu_k(\alpha, \boldsymbol{\rho}). \tag{4.12}$$

*Proof.* This can be demonstrated with a relatively simple sequence of steps. Let $\rho_{\tilde{k}}$, $\rho'_{\tilde{k}} \in \mathcal{P}_{\tilde{k}}$ be two different and arbitrary data power values for a generic UE $\tilde{k}$. Suppose that UE $\tilde{k}$ changes its data power from $\rho_{\tilde{k}}$ to $\rho'_{\tilde{k}}$, then the change of its payoff function is: $\Delta\mu_{\tilde{k}} = \mu_{\tilde{k}}(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) - \mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})$. Moreover, the change of $u(\alpha, \boldsymbol{\rho})$ is: $\Delta u = u(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) - u(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})$. By developing the expressions $\Delta\mu_{\tilde{k}}$ and $\Delta u$, it is possible to show that: $\Delta\mu_{\tilde{k}} = \Delta u$, and thus $u(\alpha, \boldsymbol{\rho})$ is an exact potential function for the parameterized game $\mathcal{G}(\alpha)$ (for more details, see **APPENDIX D**). □

### *4.3.2 Proposed Iterative Algorithm*

In order to develop a potential game-based approach to address the problem of power control for the uplink of cell-free massive MIMO systems, we propose a procedure where the power allocation is updated every iteration for each UE until reaching convergence. However, before specifically discussing this procedure, in order to achieve a practical implementation we define the following vector $\boldsymbol{\xi} \in \mathbb{R}^K$: $\boldsymbol{\xi} = [\xi_1, \cdots, \xi_k, \cdots, \xi_K]^{\mathrm{T}}$, where $\xi_k = \rho_k \Big( \sum_{l \in \mathcal{M}_k} \beta_{k,l} \Big)^{\alpha}$, $k \in \mathcal{K}$. Note that using the definition of $\boldsymbol{\xi}$, we can express the payoff function for each UE in terms of $\boldsymbol{\xi}$ only, as follows:

$$\mu_k(\boldsymbol{\xi}) = \frac{\sum\limits_{\forall i \neq k} \xi_i}{\xi_k} \; + \; \xi_k \sum_{\forall i \neq k} \frac{1}{\xi_i}. \tag{4.13}$$

The complete procedure for the proposed power allocation strategy (PAS) based on game theory is described in Algorithm 4. First, in line 1, we can define $\boldsymbol{\rho}^{(0)}$ using any naive power allocation strategy, e.g., $\rho_k^{(0)} = P_{\max}/n$, $k \in \mathcal{K}$, where $n \in \mathbb{R}_+^*$. Moreover, the input parameter $\alpha$ must also be initialized, which is an important variable for the game as will be discussed in Session 4.4.

In the *outer loop* (lines 3-15) the following procedure is repeated: after information exchange between the UEs and the Master AP, each UE will independently choose a *best response* to the actions of the other UEs according to line 7. Then, in case of an improvement of the payoff function, each UE updates its data power, otherwise the current data power value is kept. Analogously to the best response approach in line 7, the power update also occurs

**Algorithm 4:** Game-based power allocation strategy (Game-PAS)

1: **Input:** Initialize $\boldsymbol{\rho}^{(0)}$, $\alpha \in \mathbb{R}$, $l \leftarrow 0$, and $\epsilon > 0$;
2: **Output:** Data power vector $\boldsymbol{\rho}$;
3: **loop**
4: UEs report the current data power to Master AP;
5: Master AP measures and broadcasts $\boldsymbol{\xi}^{(l)}$;
6: Find the payoff function using $\boldsymbol{\xi}^{(l)}$ and (4.13);
7: Find $\rho_k^\star$ making $\rho_k^\star \leftarrow \min\{\rho_k^*, P_{\text{max}}\}$, where $\rho_k^*$ is defined in (4.9);
8: **for** $k \in \mathcal{K}$ **do**
9: **if** $\mu_k\big(\rho_k^{(l)}, \boldsymbol{\rho}_{-(k)}^{(l)}\big) - \mu_k\big(\rho_k^\star, \boldsymbol{\rho}_{-(k)}^{(l)}\big) > \epsilon$ **then**
10: $\rho_k^{(l+1)} \leftarrow \rho_k^\star$;
11: **else**
12: $\rho_k^{(l+1)} \leftarrow \rho_k^{(l)}$;
13: **end if**
14: **end for** $l \leftarrow l + 1$;
15: **end loop when** $\rho_k^{(l)} = \rho_k^{(l-1)},\ k \in \mathcal{K}$;

individually, i.e., in the *inner loop* (lines 8-14) the power update for each UE is performed in a distributed way and thus it does not depend on $K$. Finally, Algorithm 4 ends when no UE can improve its payoff by unilateral deviation (cf. Definition 2).

### *4.3.3 Signaling and Convergence Analysis*

During the execution of the proposed algorithm, an over-the-air signaling scheme must be considered for the two-way communication between the UEs and the Master AP, as performed in the lines 4 and 5 of Algorithm 4. Initially, each UE reports to the Master AP its current data power value. Next, the Master AP measures and then broadcasts $\boldsymbol{\xi}$. This is how each user receives the power allocation from other users.

Regarding the convergence of Algorithm 4, all finite potential games have the finite improvement path property (MacKenzie; Silva, 2006). Consequently, if every improvement path is finite and the best response approach provides an improvement of at least $\epsilon$, then it must necessarily converge (MacKenzie; Silva, 2006, See Chapter 5, Theorem 19).

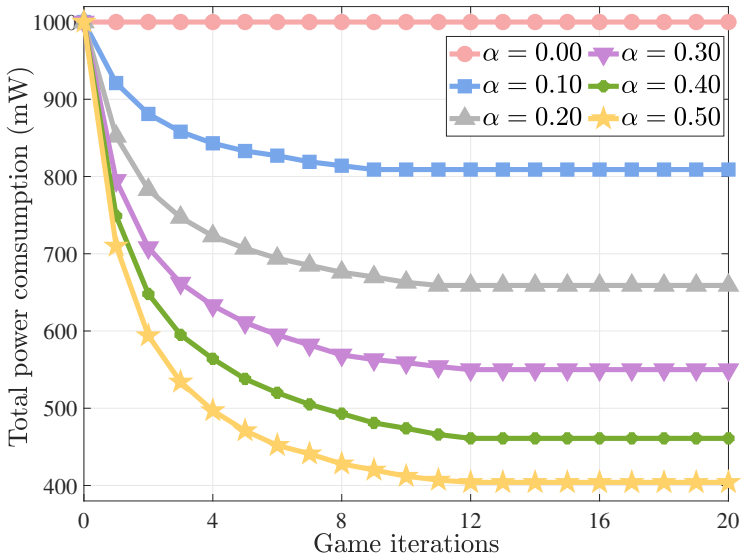
## 4.4 Simulation Results and Discussion

In this section, we conduct simulations relying on the setup introduced in (Björnson; Sanguinetti, 2020). In summary, the APs and UEs are independently and uniformly distributed in a $2$ km $\times$ $2$ km square. We apply the wrap-around technique to approximate an infinitely large network. Moreover, we assume that $\tau_{\text{c}} = 200$, $\tau_{\text{p}} = 10$, $\tau_{\text{u}} = 190$, $P_{\text{max}} = 100$ mW and a $20$ MHz bandwidth.

### *4.4.1 Convergence Behavior*

Fig. 21 shows the evolution of total power consumption of the UEs for different values of $\alpha$ versus the iterations of the game. For this result, we assume that initially all UEs transmit at their maximum power, i.e., $\rho_k^{(0)} = P_{\max}$, $k \in \mathcal{K}$. Thus, the total power usage at the beginning of the game is $K \cdot P_{\max} = 1000$ mW. In particular, when $\alpha = 0$ note that no UE has an incentive to change the initial power allocation strategy and in that case the convergence of Algorithm 4 is immediate.

Figure 21 – Impact of $\alpha$ on the total power consumption versus iterations of the game assuming $L = 100$, $N = 4$ and $K = 10$.




Source: Created by the author.

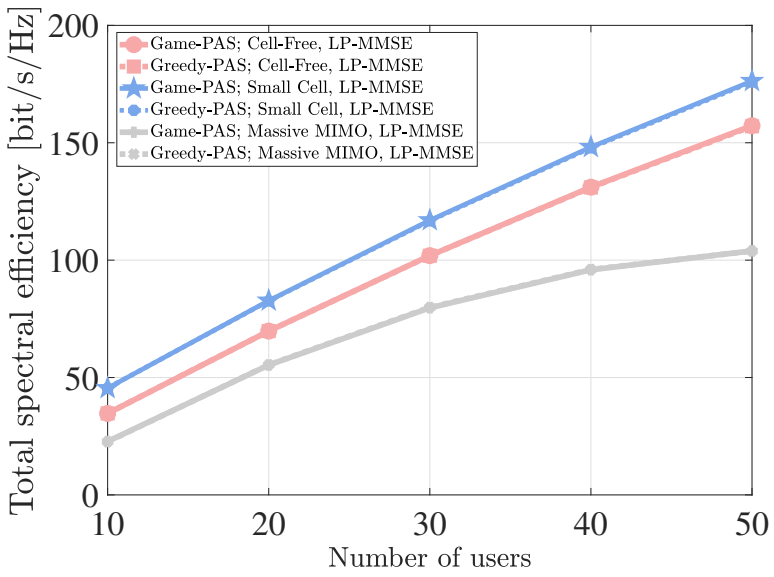
On the other hand, in a more general case when $\alpha \neq 0$, the convergence behavior is different. As $\alpha$ increases, the strategy followed by the UEs converges to situations where the power expenditure is decreasing, i.e., the UEs are turned to a low power mode. As a result, this allows to improve the use of the energy stored in the battery and it shows that, interestingly, $\alpha$ has a direct impact on energy-saving. Obviously, as the values of $\alpha$ vary, we also obtain different data rates for the UEs and, therefore, non-trivial solutions especially in terms of energy efficiency (EE) defined as bit/J can be obtained.

### *4.4.2 Performance Comparison*

In this section, the performance evaluation of the proposed power control is evaluated based on two aspects. First, we show the performance of Algorithm 4 in three different scenarios, namely, cell-free (discussed in Section 4.2), small cell and massive MIMO systems. This is interesting as it shows the good adaptability of game theory-based approaches to various frameworks. In each scenario, we also consider a baseline scheme, in which each UE transmits at full power, i.e., we use the greedy power allocation strategy (Greedy-PAS) as benchmarking. It has been shown in the literature that this power allocation strategy can provide good SE and fairness (Björnson; Sanguinetti, 2020).

Furthermore, we consider three different metrics: the total system SE, the minimum SE, and the total system EE, which is the sum of the EEs of each UE, defined as the ratio between the SE and the corresponding consumed power. Finally, our power control (Game-PAS) is performed for different $\alpha$ within the range $[0, 2]$ and the best performance for each metric is depicted.

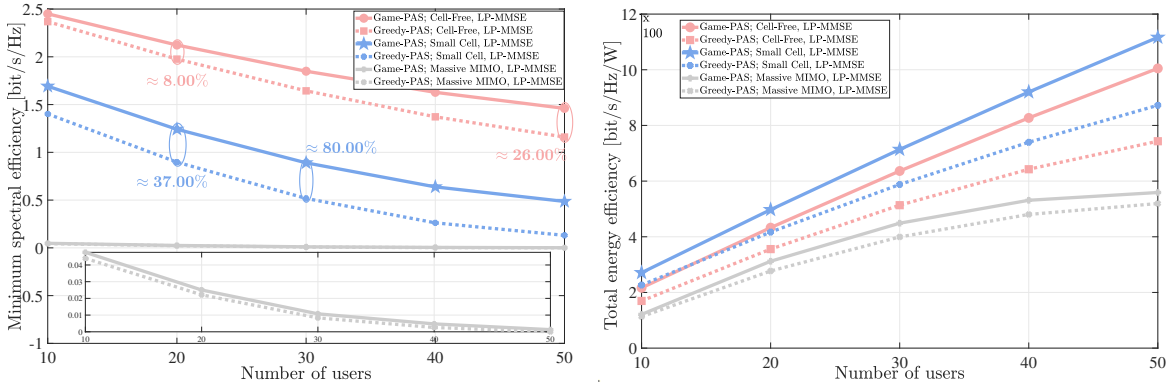
Figure 22 – Total spectral efficiency versus number of users assuming $L = 100, N = 4$ for the cell-free/small cell setups and $L = 4, N = 100$ for the massive MIMO case.




Source: Created by the author.

Fig. 22 plots the total spectral efficiency versus the number of users. Specifically in

Figure 23 – Minimum spectral efficiency and total energy efficiency versus number of users assuming $L = 100, N = 4$ for the cell-free/small cell setups and $L = 4, N = 100$ for the massive MIMO case.




(c) min-SE versus number of users.

(d) Sum-EE versus number of users.

Source: Created by the author.

this metric, the performance obtained by the proposed and baseline solutions are the same in all simulated setups. Basically, it means that from the point of view of total SE, and due to its simpler implementation, the Greedy-PAS solution has a better trade-off between performance and computational cost and is, therefore, the best option.

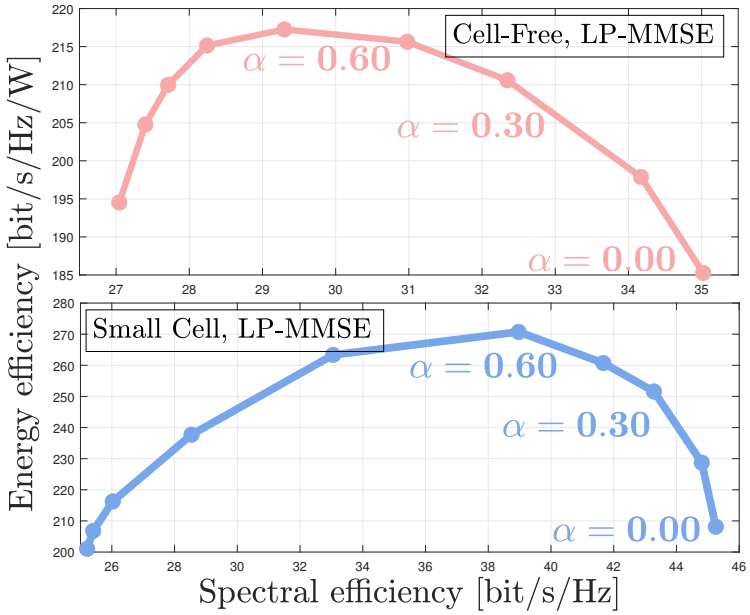
On the other hand, for the cell-free and small cell setups, significant performance gains in terms of minimum SE can be achieved using the proposed power control, as shown in Fig. 23c. Moreover, we highlight that the gains tend to increase as the number of UEs increases. For the cell-free case, for example, we have average percentage gains around $8\%$ and $26\%$ when $K = 20$ and $K = 50$, respectively. Note that even more expressive gains of the proposed solution are obtained for the small cell case. In general, under interference-limited environments, as the number of UEs in the system increases, the power control problem becomes more relevant. However, trivial power allocation strategies usually neglect the impact of increasing UEs and, consequently, are ineffective in mitigating network interference by means of power control.

Finally, we plot the EE in Fig. 23d. First, we highlight that the impact of $\alpha$ on the power usage shown in Fig. 21 has a direct effect in achieving enhanced EEs. Further, note that similarly to the minimum SE metric, the total EE performance gains also increase as $K$ increases. This is particularly interesting as increasing the number of UEs in the network can rapidly lead to a growing concern with excessive energy demand, especially for the Greedy-PAS solution. At this point, energy efficient solutions are important and a more robust power allocation strategy such as the Game-PAS is critical to reduce the energy cost per transmitted bit and to improve the

greenness of wireless systems.

### *4.4.3 Trade-off between EE and SE*

Figure 24 – Trade-off curve between EE and SE for the cell-free and small cell setups assuming $L = 100$, $N = 4$ and $K = 10$.




Source: Created by the author.

It is well-known that EE and SE are conflicting objectives and there exists an inherent trade-off between them. Thus, in the context of the proposed solution, it is interesting to show the impact of parameter $\alpha$ on the EE-SE trade-off. Fig. 24 presents the achieved SE and EE with different values of $\alpha$ for the cell-free and small cell setups. From the point of view of maximizing the SE, when $\alpha = 0.00$, as discussed in Fig. 21, the Game-PAS solution is equivalent to the Greedy-PAS solution and, in this case, the systems achieve high SE.

However, as $\alpha$ increases, the EE is gradually improved until reaching a maximum value when $\alpha = 0.60$. Also, for other values of $\alpha$, different solutions for EE and SE can be obtained. Therefore, Fig. 24 demonstrates that the proposed solution is efficient in achieving a flexible trade-off between EE and SE. For example, for small cells, when $\alpha = 0.30$, the EE metric has a gain around $20\%$ with a small cost in terms of SE.

## 4.5 Chapter Summary

In this chapter, we proposed a distributed game-theoretic method for power control in the uplink of cell-free systems. Simulations indicate that the proposed solution achieves significant performance gains in terms of minimum SE floor and power consumption with an improved EE. Moreover, by varying the $\alpha$, we showed that it is possible to achieve different solutions for EE and SE. Hence, the proposed solution simplifies the process of joint optimization of these metrics and allows to obtain useful trade-offs between EE and SE.

# 5 INTEGRATING AERIAL AND GROUND USERS: POWER CONTROL IN CELL-FREE NETWORKS

This chapter addresses the power control problem in cell-free uplink networks, containing both aerial users and traditional terrestrial users in the same network. In particular, we study the data power control optimization problem to maximize the weighted sum of the SEs, where each network user must satisfy its minimum SE demand. By adjusting the SE contribution weights for terrestrial and aerial users, we selectively prioritize certain user groups. Such model enhances current connectivity by addressing the needs and utilization behaviors of different user groups.

In this context, we first introduce a centralized solution as the upper bound of solving the problem, which is mainly based on convex optimization. Then, we propose two decentralized solutions to solve the problem. The first solution focuses on coordinate descent techniques and still benefits from the support of convex optimization. On the other hand, the second solution applies distributed deep reinforcement learning, using sophisticated artificial intelligence to create adaptive, self-improving systems in complex settings.

Moreover, in the context of our learning-based approach, we outline practical signaling considerations for a decentralized model. In our model, each user acts as an independent agent and tries to find the best policy locally, while minimizing the need for information exchange. In our analysis, we explore the implications of weighing the performance of ground and aerial users in terms of SE. Finally, the numerical analysis reveals that the performance of our decentralized solutions closely matches that of the centralized and convex approach.

## 5.1 Introduction

The main challenge in efficiently integrating ground user equipments (GUEs) with uncrewed aerial vehicles (UAVs) into current mobile network architectures lies in the differences in mobility patterns and channel quality characteristics of the two types of users. In this context, GUEs typically exhibit relatively predictable mobility and operate within well-defined network coverage areas, whereas UAVs, with their three-dimensional mobility and varying altitudes, encounter more complex and dynamically changing signal propagation environments. Moreover, by considering UAVs as a novel category of UE, there is a notable enhancement in the line of sight (LOS) connectivity between aerial UE and terrestrial base stations (BSs). This advancement, however, leads to a substantial elevation in system interference, necessitating the development of

innovative strategies (Fotouhi *et al.*, 2019; Tentu *et al.*, 2022).

Even with these constraints, over the past years the usage of UAVs in cellular networks has become more and more relevant, which is credited to their potential of improving connectivity, flexibility, instantaneity throughout the numerous areas of application in society. For example, in the area of logistics, drones bring groundbreaking approaches to delivering packages and managing supply chains, particularly valuable in isolated locations or densely populated urban areas. Environmental monitoring has also unlocked massive advancements with UAVs, which have made it possible to perform an expanded and more detailed data collection for awareness on climate change, pollution, and animal protection (Bai *et al.*, 2023).

### *5.1.1 Related Works*

Considering that cell-free network architectures are distinguished by their extensive array of distributed antennas, which collectively serve users within the same frequency spectrum, and are fundamentally engineered with a user-centric philosophy, the integration of UAVs alongside GUEs can be accomplished with remarkable efficiency in this type of system (D'Andrea *et al.*, ; D'Andrea *et al.*, 2020; Tentu *et al.*, 2022; Li *et al.*, 2023; Elwekeil *et al.*, 2023). In particular, in (Tentu *et al.*, 2022) a closed-form SE formula is developed for a cell-free massive MIMO system experiencing hardware impairments, incorporating both UAVs and GUEs operating over channels that exhibit a combination of spatially-correlated Rician and Rayleigh fading. Following this, the authors employ a novel block quadratic transformation method to optimize the non-convex global energy efficiency and SE metrics. In their works, both (D'Andrea *et al.*, ) and (D'Andrea *et al.*, 2020), the authors introduce power control strategies tailored for environments where UAVs and GUEs coexist. They demonstrate that the cell-free architecture outperforms conventional cellular networks in terms of efficiency and effectiveness.

More recently, (Li *et al.*, 2023) and (Elwekeil *et al.*, 2023) have put forward innovative strategies for the allocation of radio resources that efficiently cater to both GUEs and UAVs in cell-free systems. In (Li *et al.*, 2023), the authors introduce a novel algorithm for pilot allocation that is based on grouping users, complemented by an approach to enhance system performance through optimization of data power. Conversely, in (Elwekeil *et al.*, 2023), the focus shifts to ultra-reliable low-latency communication applications, which are crucial for a variety of critical applications. These include remote surgery, autonomous vehicle navigation, and, notably, UAV communications, underscoring their importance in advancing these technologically

demanding areas (Hassan *et al.*, 2021).

Therefore, particularly in the context of resource allocation, cell-free networks have emerged as a fundamental element for enabling the effective integration of GUEs with UAVs. This efficient integration is essential not only for harnessing the vast potential applications that UAVs can offer but also for ensuring that such advancements do not detract from the performance and experience of traditional terrestrial network users.

Alongside suggesting strategies for resource allocation to effectively manage the simultaneous presence of UAVs and GUEs within a cell-free system, a notable focus of the studies in (D'Andrea *et al.*, ; D'Andrea *et al.*, 2020; Tentu *et al.*, 2022; Li *et al.*, 2023; Elwekeil *et al.*, 2023) is the employment of conventional optimization techniques to realize these strategies. Nevertheless, the dependency of traditional optimization approaches on the SINR presents a significant hurdle. These approaches usually require a precise mathematical formulation of the SINR, which markedly restricts their practical application. Moreover, even with a closed expression for SINR at our disposal, the pursuit of more efficient solutions via convex optimization may depend on the application of advanced strategies.

In this context, we emphasize the approach taken by work (Tentu *et al.*, 2022), which incorporates fractional programming theory alongside quadratic transformations to deal with the global energy efficiency of the system. Additionally, in (Elwekeil *et al.*, 2023), we spotlight the application of precise and concave inequalities to more accurately estimate users' SE metrics. Nevertheless, it should be recognized that even these strategies, which facilitate an equivalent reformulation of the original problem or, at the very least, provide good approximations, may not always be feasible or easy to implement.

Regarding this matter, as outlined in (Björnson; Sanguinetti, 2020), modifying the beamforming architecture — specifically by implementing more sophisticated and robust beamforming strategies beyond conjugate beamforming — can result in the SE being represented through non-closed-form mathematical expressions. Consequently, in these scenarios, the task of optimizing network resources, like transmission power, using traditional methods becomes significantly more complex. This shift towards more complex beamforming structures challenges traditional optimization paradigms, necessitating innovative approaches that can accommodate the intricacies of these advanced beamforming techniques without relying on explicit mathematical expressions of SE for optimization.

Moreover, in future wireless networks, such as the 6th generation (6G), the challenges

we will encounter in enhancing performance, scalability, and efficiency are set to become exceedingly complex, reaching levels never seen before. Indeed, convex optimization methods have been the backbone of network optimization due to their mathematical tractability and the guarantee of finding global optimal solutions in specific problem domains. Yet, as mobile networks evolve, the problems we encounter often deviate from the neatly defined, convex domains. For example, in multi-UAV scenarios, the interplay between UAV trajectory planning and resource allocation, coupled with the large volume of devices, significantly complicates network design. This complexity renders traditional methods, such as convex optimization and dynamic programming, inadequate for addressing these challenges (Zeng *et al.*, 2016; Zeng; Zhang, 2017). Additionally, issues such as non-linearity, model uncertainty, and the large-scale of variables in 6G networks introduce complexities that challenge the boundaries of classical optimization methods.

In this scenario, artificial intelligence (AI) stands out as a key enabler for future wireless networks. Its effectiveness in managing large-scale challenges within wireless systems is well-established (Samir *et al.*, 2021). In this context, reinforcement learning represents an important area of study within machine learning, having profoundly influenced the advancement of AI over the past two decades (Luong *et al.*, 2019). Moreover, the integration of deep neural networks has elevated traditional reinforcement learning to new heights, paving the way for the emergence of powerful tools like deep reinforcement learning (DRL).

According to (Bai *et al.*, 2023), DRL has been thoroughly researched and successfully applied in four key big domains within the area of multi-UAV communications, including *resource allocation for wireless connectivity*. This domain can be further broken down into more specific sub-domains, each with unique characteristics that also facilitate the effective application of DRL. In case of resource allocation, it can be divided into three main categories: spectrum allocation and power control, wireless power transfer (WPT), and caching strategies. Within this first sub-domain, numerous significant and recent studies have delved into the application of learning techniques to tackle resource allocation issues in mobile networks, see, e.g., (Braga *et al.*, 2023; Shi *et al.*, 2021; Lee *et al.*, 2022; Zhong *et al.*, 2022; Nguyen *et al.*, 2022; Liu *et al.*, 2022). Recently, the work in (Braga *et al.*, 2023) presents a nearly optimal solution utilizing DRL to tackle the complex issue of simultaneously allocating pilot power and data in cell-free networks in scenarios without UAVs.

On the other hand, to reduce the interference generated by UAVs, the authors in (Shi

*et al.*, 2021) suggest a downlink/uplink decoupled access strategy for cellular networks that include UAV communications. This approach separates the control and data links of UAVs and the uplinks and downlinks of users across different serving BSs and operating frequencies. They introduce a deep Q-network (DQN)-based method aimed at enhancing the system's energy efficiency. A DQN-based method is also applied to power control in (Lee *et al.*, 2022; Zhong *et al.*, 2022). Employing a different learning strategy, (Nguyen *et al.*, 2022) proposes multi-UAV networks supported by a reflective intelligent surface (RIS) panel to enhance the network performance. To optimize the energy efficiency of the networks in question, focusing on both the transmission energy of the UAV and the RIS phase coefficients matrix, the authors implement efficient learning solutions, including the utilization of deep deterministic policy gradient and, for a more optimized solution, the robust proximal policy optimization (PPO) algorithms.

In the context of WPT, in (Liu *et al.*, 2022), the authors develop a strategy that simultaneously fine-tunes UAV flight paths and schedules wireless energy distribution to reduce the average delay experienced by internet of things (IoT) devices. UAVs are deployed to recharge IoT devices through WPT, allowing these devices to then transmit data back in the uplink using the energy they have received. Thus, despite the highly complex computational nature of the formulated problem, the implementation of the proposed DRL approach achieves near-optimal performance.

### *5.1.2 Main Contributions*

In this chapter, we consider resource allocation to optimize cell-free uplink networks given the concurrent service of GUEs and UAVs. Additionally, our study aims to identify efficient decentralized power control solutions for data transmission. The contributions of the current study are as follows:

- We formulate the joint optimization problem to improve the overall system efficiency by controlling data power. The novel perspective of our design is that it considers the weighted sum-SE for the GUEs and the UAVs. In addition, we also maintain the minimum SE requirement for all users to ensure fair service.
- By fine-tuning a weighted sum-SE, our methodology leverages an adjustable input weight to strategically allocate power resources, enabling the prioritization of either UAVs or GUEs based on operational needs. In an environment where GUEs and UAVs are utilized simultaneously, possessing a weight term to adjust power allocation among users is critical.

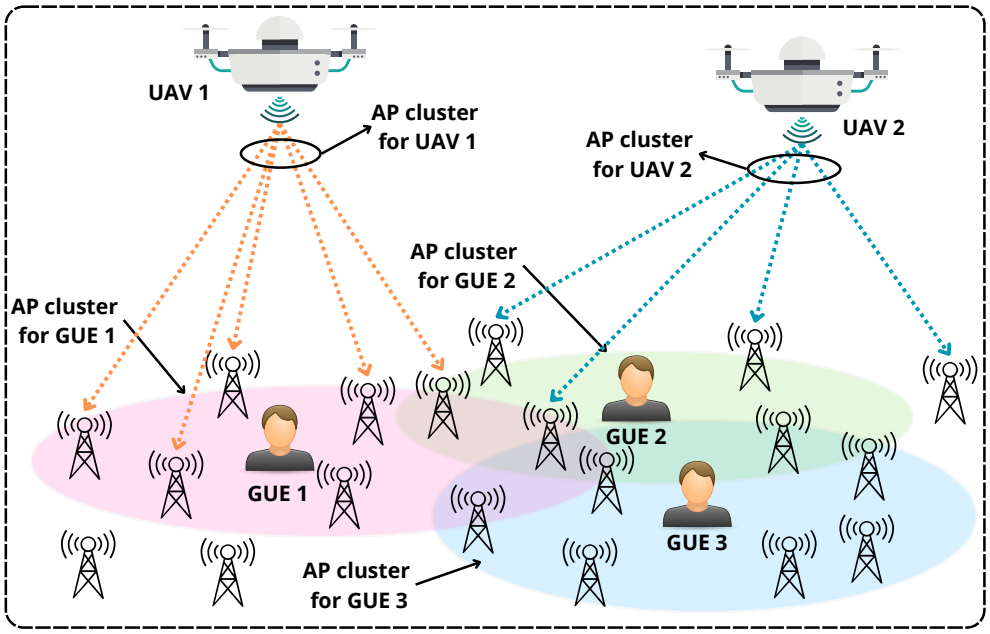
This approach plays a significant role in reducing the massive interference threats caused by UAVs on GUEs.

- To establish a practical upper bound benchmark, we develop a centralized solution based on convex optimization. This framework utilizes concave and tight approximations of SE expressions, which have been widely established in the literature, to enhance the accuracy and effectiveness of our solution. We then establish, based on the derived expressions, an effective decentralized solution. This solution is used to make our initial framework more scalable in light of a growing user population.
- In order to further expand our contributions by offering a solution that integrates high levels of flexibility and adaptability, we set forth a distributed DRL-based approach. This integrates a number of agents and considerably promotes a decentralized approach during the training phases, fully harnessing the capabilities of the state-of-the-art PPO algorithms. Using DRL helps achieve a more dynamic system capable of responding to fluctuating network patterns and lays the foundation of an architecture that can address, in a scalable manner, the challenges of power control that will be faced as the number of users continues to rise steadily. In addition, incorporating this approach grants our system the ability to continuously learn and adapt, which means it can offer effective solutions in conditions where closed-form SE expressions are unobtainable, and open problems arise, or situations where deriving convex SE expressions is impractical.

## 5.2 Network Model

Unlike the previous chapter, we now explore a cell-free system where GUEs and UAVs coexist as users within the same network, as illustrated in Fig. 25. Specifically, our focus is on an uplink cell-free network composed of $K$ users, each equipped with a single antenna, and $L$ APs where every AP is equipped with a uniform linear array containing $N$ antennas. These APs are seamlessly interconnected with edge-cloud processors, known as CPUs, through a network structure employing ideal fronthaul connections to facilitate efficient data exchange and processing. Beyond catering to the conventional GUE, our network extends its service capabilities to include aerial entities, specifically UAVs. We consider that there exists a set of APs denoted by $\mathcal{A}$ with $|\mathcal{A}| = L$, and a comprehensive set $\mathcal{K}$ encompassing all users. In this context, the aggregate quantity of users supported by the network is defined as $K = K_{\mathrm{g}} + K_{\mathrm{u}}$, where $K_{\mathrm{g}}$ corresponds to the complete count of GUEs and $K_{\mathrm{u}}$ signifies the overall number of

Figure 25 – A cell-free scenario where GUEs and UAVs operate together.




Source: Created by the author.

UAVs. As a result, the user group $\mathcal{K}$ is defined as the union of sets $\mathcal{G}$ and $\mathcal{U}$, where $\mathcal{G}$ represents the set of GUEs and $\mathcal{U}$ denotes the set of UAVs, i.e., $\mathcal{K} = \mathcal{G} \cup \mathcal{U}$.

In this scenario, we adhere to the user-AP association strategy commonly referenced in existing literature (Elwekeil *et al.*, 2023; D'Andrea *et al.*, 2020). Specifically, each user within the network is allocated service by a selected subset of the total APs available, this selection is determined based on the large-scale fading coefficients. This approach ensures that the allocation of APs to users improves the efficiency of the network's coverage and signal strength, taking into account the unique propagation characteristics and geographic distribution of both the users and the APs. Particularly, the collection $\{\mathcal{A}_k\}_{\forall k}$ defines the *cluster formation* of all users $k$, where $\mathcal{A}_k$ is the set containing the APs serving user $k$. Thus, the service provided to each user $k$ is facilitated by a specific subset of APs, denoted as $\mathcal{A}_k$ with $|\mathcal{A}_k| = M \leq L$ indicating the quantity of APs serving user $k$. Consequently, the notation $\mathcal{K}_a$ is utilized to denote the group of users that are serviced by AP $a$. In this context, our approach to the clustering issue is straightforward and uncomplicated, i.e., set $\mathcal{A}_k$ comprises the $M$ APs that exhibit the highest large-scale gain coefficients for user $k$.

### *5.2.1 TDD Transmission*

Consistent with the standards found in the cell-free recent literature (Li *et al.*, 2023; Elwekeil *et al.*, 2023; Antonioli *et al.*, 2022; Saraiva *et al.*, ), whether or not UAVs are involved, we have chosen to utilize TDD transmission. In this framework, the entire communication process unfolds within a coherence interval comprising $\tau_{\mathrm{c}}$ time-frequency samples. This interval is segmented into three distinct sub-intervals. The initial sub-interval, spanning a length of $\tau_{\mathrm{p}} < \tau_{\mathrm{c}}$, is dedicated to uplink channel estimation. During this phase, users transmit a pilot sequence, enabling every AP in the network to conduct channel estimation activities. The second sub-interval, also of length $\tau_{\mathrm{d}} < \tau_{\mathrm{c}}$, is allocated for downlink data transmission, during which each AP sends data to the users it supports. Lastly, the third sub-interval, extending for a duration of $\tau_{\mathrm{u}} < \tau_{\mathrm{c}}$, is designated for uplink data transmission, where each user sends its uplink data. The durations of the previously mentioned sub-intervals should be chosen to ensure that $\tau_{\mathrm{p}} + \tau_{\mathrm{u}} + \tau_{\mathrm{d}} = \tau_{\mathrm{c}}$. Furthermore, in this work, we operate under the assumption that sub-intervals of equal length are utilized for both uplink and downlink data transmissions.

### *5.2.2 Channel Model*

The channel vector connecting user $k$ to AP $a$ within the designated coherence interval is represented by $\mathbf{g}_{k,a}$. For the UAVs, we take into account that the small-scale fading adheres to a Rician distribution. This perspective views the channel as being composed of several paths, among which there is exclusively one LOS path. This characteristic is crucial, as the presence of a LOS component significantly influences the channel's behavior.

Conversely, in the case of GUEs, the small-scale fading is considered to follow a Rayleigh distribution. This distribution is acknowledged as a specific scenario of Rician fading where the Rician $R$-factor equals zero. This distinction underscores the contrasting propagation environments encountered by UAVs and GUEs, with the former experiencing conditions conducive to LOS communications and the latter predominantly influenced by non-LOS components.

Thus, $\mathbf{g}_{k,a}$ can be expressed as in (Elwekeil *et al.*, 2023; D'Andrea *et al.*, ; D'Andrea *et al.*, 2020; Li *et al.*, 2023):

$$\mathbf{g}_{k,a} = \sqrt{\frac{\beta_{k,a}}{R_{k,a}+1}} \left( \sqrt{R_{k,a}} \mathrm{e}^{j\vartheta_{k,a}} \mathbf{a}(\theta_{k,a}) + \mathbf{h}_{k,a} \right), \tag{5.1}$$

where $\beta_{k,a}$ characterizes the scalar coefficient modeling the channel path-loss and shadowing effects between the $k$-th user and the $a$-th AP, and $\mathbf{h}_{k,a} \in \mathbb{C}^N$ is the vector of independent and identically distributed (i.i.d.) small scale fading coefficients between user $k$ and AP $a$, i.e., $\mathbf{h}_{k,a} \sim \mathcal{CN}(0, \mathbf{I}_N)$, where $\mathbf{I}_N$ is the $N \times N$ identity matrix. Moreover, $R_{k,a}$ denotes the Ricean $R$-factor, which quantifies the dominance of the LOS component in the received signal between user $k$ and AP $a$. Denoting the random phase offset for the direct path, $\vartheta_{k,a}$ follows a uniform distribution in $[0, 2\pi]$, and $\mathbf{a}(\theta_{k,a}) \in \mathbb{C}^N$ represents the steering vector for the angle $\theta_{k,a}$, which characterizes the direct path between AP $a$ and user $k$.

Therefore, as highlighted before, a key distinction arises from the UAVs' typical higher LOS probability with multiple APs. This factor often results in UAVs exhibiting superior channel quality compared to GUEs. Additionally, the pathloss model for UAVs diverges significantly from that of GUEs. This variation further underscores the distinct nature of the channel characteristics between UAVs and GUEs. For GUEs, the channel model is constructed based on (3GPP, 2017c), whereas the specifications from (3GPP, 2017a) are used for UAVs.

### *5.2.3 Channel Estimation*

In this section, it is presumed that each AP employs linear MMSE channel estimation (D'Andrea *et al.*, ; D'Andrea *et al.*, 2020; Elwekeil *et al.*, 2023). In this context, the transmitted signal from user $k$ during the training phase can be represented by $\sqrt{\rho_k^{\text{pilot}}}\boldsymbol{\phi}_k^{\text{H}}$, where $\rho_k^{\text{pilot}}$ denotes the pilot power employed by the $k$-th user and $\boldsymbol{\phi}_k \in \mathbb{C}^{\tau_{\text{p}}}$ is the pilot sequence sent by user $k$ such that $||\boldsymbol{\phi}_k||^2 = 1$. Meanwhile, the corresponding component arriving at the AP $a$ is $\sqrt{\rho_k^{\text{pilot}}}\mathbf{g}_{k,a}\boldsymbol{\phi}_k^{\text{H}}$. Therefore, the cumulative received pilot signal at AP $a$ $\mathbf{Y}_a \in \mathbb{C}^{N \times \tau_{\text{p}}}$ is:

$$\mathbf{Y}_a = \sum_{k \in \mathcal{K}} \sqrt{\rho_k^{\text{pilot}}}\mathbf{g}_{k,a}\boldsymbol{\phi}_k^{\text{H}} + \mathbf{W}_a, \tag{5.2}$$

where $\mathbf{W}_a \in \mathbb{C}^{N \times \tau_{\text{p}}}$ contains the thermal noise contribution at the $a$-th AP, with i.i.d. $\mathcal{CN}(0, \sigma^2)$ as entries.

Subsequently, the AP $a$ can apply the pilot sequence used by a specific user $\tilde{k} \in \mathcal{K}$ to multiply with $\mathbf{Y}_a$, facilitating the calculation of the ensuing statistic:

$$\begin{aligned}\hat{\mathbf{y}}_{\tilde{k},a} &= \mathbf{Y}_a\boldsymbol{\phi}_{\tilde{k}} \\ &= \sqrt{\rho_{\tilde{k}}^{\text{pilot}}}\mathbf{g}_{\tilde{k},a} + \sum_{k \in \mathcal{K}\setminus\tilde{k}} \sqrt{\rho_k^{\text{pilot}}}\mathbf{g}_{k,a}\boldsymbol{\phi}_k^{\text{H}}\boldsymbol{\phi}_{\tilde{k}} + \mathbf{W}_a\boldsymbol{\phi}_{\tilde{k}}.\end{aligned} \tag{5.3}$$

Assuming knowledge of $\{\beta_{k,a}, R_{k,a}, \mathbf{a}(\theta_{k,a})\}_{\forall(k,a)}$, the linear MMSE channel estimate from user $\tilde{k}$ to AP $a$ in the coherence interval of interest, $\hat{\mathbf{g}}_{\tilde{k},a}$, can be calculated based on $\hat{\mathbf{y}}_{\tilde{k},a}$ as follows:

$$\hat{\mathbf{g}}_{\tilde{k},a} = \mathbf{D}_{\tilde{k},a}\hat{\mathbf{y}}_{\tilde{k},a}, \tag{5.4}$$

where

$$\mathbf{D}_{\tilde{k},a} = \sqrt{\rho_{\tilde{k}}^{\text{pilot}}}\mathbf{G}_{\tilde{k},a}\mathbf{B}_{\tilde{k},a}^{-1} \in \mathbb{C}^{N\times N}, \tag{5.5a}$$

$$\mathbf{G}_{\tilde{k},a} = \frac{\beta_{\tilde{k},a}}{R_{\tilde{k},a}+1}\left(R_{\tilde{k},a}\mathbf{a}(\theta_{\tilde{k},a})\mathbf{a}^{\text{H}}(\theta_{\tilde{k},a}) + \mathbf{I}_N\right), \tag{5.5b}$$

$$\mathbf{B}_{\tilde{k},a} = \sum_{i\in\mathcal{K}} \rho_{\tilde{k}}^{\text{pilot}}\mathbf{G}_{i,a}\left|\boldsymbol{\phi}_i^{\text{H}}\boldsymbol{\phi}_{\tilde{k}}\right|^2 + \sigma^2\mathbf{I}_N. \tag{5.5c}$$

As recently discussed in (Elwekeil *et al.*, 2023), it is important to note that power control strategies do not always require access to instantaneous channel state information (CSI). In particular, the matrix $\{\mathbf{D}_{k,a}\}_{\forall(k,a)}$ values can remain consistent across multiple successive coherence intervals. Consequently, the power control methodologies suggested within this context are designed to operate without the necessity of updating with the estimated channels for each coherence interval.

### *5.2.4 Uplink Data Transmission*

In the described scenario, during the uplink phase, both GUEs and UAVs are sending data to the APs within the network. The data signal $\mathbf{y}_a \in \mathbb{C}^N$ received at AP $a$ is:

$$\mathbf{y}_a = \sum_{k\in\mathcal{K}} \sqrt{\rho_k}\mathbf{g}_{k,a}s_k + \mathbf{w}_a, \tag{5.6}$$

where $\rho_k$ and $s_k$ are the uplink transmit power and the data symbol of user $k$, respectively, and $\mathbf{w}_a \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N) \in \mathbb{C}^N$ the additive white Gaussian noise (AWGN) vector, where $\sigma^2$ is the corresponding noise variance. In this context, (5.6) can be rewritten as follows:

$$\mathbf{y}_a = \sqrt{\rho_{\tilde{k}}}\mathbf{g}_{\tilde{k},a}s_{\tilde{k}} + \sum_{k\in\mathcal{K}\setminus\tilde{k}} \sqrt{\rho_k}\mathbf{g}_{k,a}s_k + \mathbf{w}_a, \tag{5.7}$$

where the first term represents the desired signal of user $\tilde{k}$ at AP $a$, while the second term corresponds to the interference affecting user $\tilde{k}$'s signal. Every AP $a$ decodes the received vector for only the set of its served users, i.e., $\tilde{k} \in \mathcal{K}_a$, by multiplying the received signal in (5.7) by the

uplink receive combining vector, $\boldsymbol{v}_{k,a}$, to obtain an estimate of the corresponding uplink data symbol, $\hat{s}_{\tilde{k},a}$, based on the local knowledge of AP $a$ as follows:

$$\hat{s}_{\tilde{k},a} = \boldsymbol{v}_{\tilde{k},a}^{\mathrm{H}} \mathbf{y}_a = \sqrt{\rho_{\tilde{k}}} \boldsymbol{v}_{\tilde{k},a}^{\mathrm{H}} \mathbf{g}_{\tilde{k},a} s_{\tilde{k}} + \sum_{k \in \mathcal{K} \setminus \tilde{k}} \sqrt{\rho_k} \boldsymbol{v}_{\tilde{k},a}^{\mathrm{H}} \mathbf{g}_{k,a} s_k + \boldsymbol{v}_{\tilde{k},a}^{\mathrm{H}} \mathbf{w}_a, \quad \tilde{k} \in \mathcal{K}_a. \tag{5.8}$$

Subsequently, each AP $a$ transmits the estimated value of $\hat{s}_{\tilde{k},a}$, $\tilde{k} \in \mathcal{K}_a$ to the CPU. This allows the CPU to compute the definitive estimate of the uplink symbol $\hat{s}_{\tilde{k}}$ for each user $\tilde{k} \in \mathcal{K}$ in the following manner:

$$\hat{s}_{\tilde{k}} = \sum_{a \in \mathcal{A}_{\tilde{k}}} \hat{s}_{\tilde{k},a} = \sqrt{\rho_{\tilde{k}}} \sum_{a \in \mathcal{A}_{\tilde{k}}} \boldsymbol{v}_{\tilde{k},a}^{\mathrm{H}} \mathbf{g}_{\tilde{k},a} s_{\tilde{k}} + \sum_{k \in \mathcal{K} \setminus \tilde{k}} \sqrt{\rho_k} \sum_{a \in \mathcal{A}_{\tilde{k}}} \boldsymbol{v}_{\tilde{k},a}^{\mathrm{H}} \mathbf{g}_{k,a} s_k + \sum_{a \in \mathcal{A}_{\tilde{k}}} \boldsymbol{v}_{\tilde{k},a}^{\mathrm{H}} \mathbf{w}_a, \quad \tilde{k} \in \mathcal{K}, \tag{5.9}$$

where the first term denotes the aggregate desired signal received at the CPU, the second term illustrates the total interference experienced by the CPU, and the last term encapsulates the overall AWGN present at the CPU.

In our work, we specifically opt for *conjugate beamforming*, i.e., $\boldsymbol{v}_{k,a} = \hat{\mathbf{g}}_{k,a}^{\mathrm{H}}$, $k \in \mathcal{K}$, $a \in \mathcal{A}$, attracted by its low computational complexity and its distributed implementation nature. While acknowledging that its performance might not reach the heights achievable with other combining methods, its prevalent use in the literature underscores its practicality and efficiency (Björnson; Sanguinetti, 2020), providing a valuable balance between performance and operational simplicity.

Moreover, the application of conjugate beamforming facilitates the derivation of a closed-form expression for the user's SINR. This characteristic is particularly advantageous for employing conventional optimization tools to effectively manage interference levels using this metric. It also simplifies the process of analyzing and optimizing network performance, making it more straightforward to implement strategies that enhance the overall quality of service (QoS). In this scenario, the definitive SINR expression for user $k \in \mathcal{K}$, utilizing conjugate beamforming, is presented as follows (Elwekeil *et al.*, 2023; D'Andrea *et al.*, 2020):

$$\mathrm{SINR}_k = \frac{\rho_k \left| \sum\limits_{a \in \mathcal{A}_k} \hat{\mathbf{g}}_{k,a}^{\mathrm{H}} \mathbf{g}_{k,a} \right|^2}{\sum\limits_{j \in \mathcal{K} \setminus k} \rho_j \left| \sum\limits_{a \in \mathcal{A}_k} \hat{\mathbf{g}}_{k,a}^{\mathrm{H}} \mathbf{g}_{j,a} \right|^2 + \sigma^2 \sum\limits_{a \in \mathcal{A}_k} \left|\left| \hat{\mathbf{g}}_{k,a} \right|\right|^2}. \tag{5.10}$$

Hence, in accordance with the classical Shannon capacity formula, the SE attained by a particular user $k \in \mathcal{K}$ is determined as follows:

$$\mathrm{SE}_k = \frac{\tau_{\mathrm{u}}}{\tau_{\mathrm{c}}} \log_2 \left(1 + \mathrm{SINR}_k\right). \tag{5.11}$$

Note that the expression in (5.11) is applicable to both GUEs and UAVs. However, as discussed, the channel quality between these two user types can differ significantly, resulting in very different SINR values, even with equal power allocation, for example. Therefore, applying different weights to the SE expressions can be important for effectively managing interference and compensating for the substantial differences in channel quality.

## 5.3 Problem Formulation

This section discusses creating a power control optimization framework designed for cell-free uplink data transmission, aiming to boost network efficiency by maximizing the weighted sum of SEs. This method enables adjustable performance balance between GUEs and UAVs based on system needs or priorities. This optimization problem, however, is not without constraints. It is crucial that each user or device within the network adheres to predetermined minimum SE standards, ensuring that the optimization process does not disproportionately favor certain users over others.

The mathematical formulation of this problem, which encapsulates the weighted sum of SEs alongside the minimum SE requirements, is presented as follows:

$$\max_{\boldsymbol{\rho}} \ \lambda \sum_{k \in \mathcal{G}} \mathrm{SE}_k\left(\boldsymbol{\rho}\right) + \left(1 - \lambda\right) \sum_{k \in \mathcal{U}} \mathrm{SE}_k\left(\boldsymbol{\rho}\right) \tag{5.12a}$$

subject to

$$\mathrm{SE}_k\left(\boldsymbol{\rho}\right) \geq \mathrm{SE}_{\min}^{k}, \quad k \in \mathcal{K}, \tag{5.12b}$$

$$0 \leq \rho_k \leq P_{\max}^{k}, \quad k \in \mathcal{K}, \tag{5.12c}$$

where $\lambda$ is designed as an input parameter. In detail, the incorporation of the weight parameter $\lambda$ into the problem formulation enables a versatile optimization strategy capable of adjusting the emphasis on network performance between GUEs and UAVs according to their operational significance.

Notably, such a framework is needed due to the two devices' roles being distinct and the difference in how much of an impact optimal performance would have in different

situations. In an emergency situation, the UAVs could be given preferential importance to enable quick gathering of information, while in urban areas, GUEs are prioritized to ensure the wide coverage of their devices. By tuning the weight parameter $\lambda$, whenever necessary, the best support optimization for the GUEs and UAVs is offered, thus contributing to the efficiency of the network and the user experience.

Meanwhile, by considering minimum SE constraints for each user in (5.12b), this optimization ensures every user receives its desired QoS. Also, it promotes fairness and ensures that no user group is unfairly disadvantaged. By avoiding marginalizing any user, it guarantees there is no significant service drop for any group. By being inclusive, it improves the network experience for all users and maintains the integrity and efficiency of the network by accurately addressing the diverse needs of its users.

Moreover, both GUEs and UAVs are required to ensure that their operations are performed within strict transmission power limitations. More precisely, the transmission power of both the GUEs and UAVs is limited between the values of zero and a fixed maximum threshold labeled as $P_{\max}^{k}, \forall k \in \mathcal{K}$. This is denoted in the corresponding constraints (5.12c).

## 5.4 Centralized Solution

In this part, our focus is on developing a centralized strategy for addressing the power control problem (5.12). However, despite the weight $\lambda$ being a specified input parameter, problem (5.12) exhibits a complex structure that is neither concave nor convex, presenting a significant hurdle in applying standard optimization techniques directly.

To address this challenge, we adopt a methodology recently introduced in (Elwekeil *et al.*, 2023), which centers around the concept of constructing a concave and tight lower bound for the SE expression. This approach leverages the principles of convex optimization by transforming the SE expression into a form that is both tractable and amenable to rigorous analysis. By establishing a concave lower boundary, the methodology not only simplifies the complexity of the original problem but also ensures that the solutions derived are both efficient and closely aligned with the optimal performance metrics.

In detail, to identify a concave lower bound for the SE expression, we employ the following inequality, proven in (Sheng *et al.*, 2018), for all $x > 0$, $y > 0$, and $\bar{x} > 0$, $\bar{y} > 0$:

$$\ln\left(1+\frac{x}{y}\right) \geq \ln\left(1+\frac{\bar{x}}{\bar{y}}\right) + \frac{\bar{x}}{\bar{y}}\left(2\frac{\sqrt{x}}{\sqrt{\bar{x}}} - \frac{x+y}{\bar{x}+\bar{y}} - 1\right), \tag{5.13}$$

where we define for SE in (5.11) the following terms:

$$x = \rho_k q_k,\ y = \sum_{j \in \mathcal{K} \setminus k} \rho_j u_{k,j} + t_k, \tag{5.14}$$

where

$$q_k = \left| \sum_{a \in \mathcal{A}_k} \hat{\mathbf{g}}_{k,a}^{\mathrm{H}} \mathbf{g}_{k,a} \right|^2,\ u_{k,j} = \left| \sum_{a \in \mathcal{A}_k} \hat{\mathbf{g}}_{k,a}^{\mathrm{H}} \mathbf{g}_{j,a} \right|^2, \text{ and } t_k = \sigma^2 \sum_{a \in \mathcal{A}_k} \left|\left| \hat{\mathbf{g}}_{k,a} \right|\right|^2. \tag{5.15}$$

Then, the SE expression can be approximated to a local tight concave lower bound $\widetilde{\mathrm{SE}}$ as

$$\mathrm{SE}_k(\boldsymbol{\rho}) \geq \widetilde{\mathrm{SE}}_k(\boldsymbol{\rho}) = \frac{\tau_{\mathrm{u}}}{\tau_{\mathrm{c}} \times \ln 2} \left( \ln(1 + \eta_k) + \eta_k \mu_k \right), \tag{5.16}$$

where

$$\eta_k = \frac{\rho_k^{(l-1)} q_k}{\sum_{j \in \mathcal{K} \setminus k} \rho_j^{(l-1)} u_{k,j} + t_k}, \tag{5.17a}$$

$$\mu_k = \left( 2 \frac{\sqrt{\rho_k}}{\sqrt{\rho_k^{(l-1)}}} - \frac{\rho_k q_k + \sum_{j \in \mathcal{K} \setminus k} \rho_j u_{k,j} + t_k}{\rho_k^{(l-1)} q_k + \sum_{j \in \mathcal{K} \setminus k} \rho_j^{(l-1)} u_{k,j} + t_k} - 1 \right). \tag{5.17b}$$

To arrive at an effective solution for the power control problem (5.12), the strategy involves substituting the original SE expressions with their concave approximations, i.e., $\widetilde{\mathrm{SE}}$. These approximated SE expressions, by virtue of their concavity, lend themselves to more tractable analysis and optimization. Following this substitution, we harness the power of successive convex optimization (SCO) techniques. This approach allows us to tackle the problem iteratively, solving a series of successively approximated convex problems. With each iteration, we refine our solution, progressively moving closer to an optimal power control. This iterative process continues until we achieve convergence, ensuring that the final solution is both efficient and closely aligned with the optimal power profile. This outlined process is detailed in Algorithm 5.

**Algorithm 5:** SCO-based Power Iterative Algorithm (SCOPIA)

1: **Input:** $\{\mathcal{A}_k\}_{\forall k}$, feasible point $\boldsymbol{\rho}^{(0)}$, $l \leftarrow 1$;
2: **Output:** Data power vector $\boldsymbol{\rho}$;
3: **loop**
4: Utilize $\{\widetilde{\mathrm{SE}}_k\}_{\forall k}$ and $\boldsymbol{\rho}^{(l-1)}$ in problem (5.12);
5: Solve the obtained problem to obtain $\boldsymbol{\rho}^{\star}$;
6: Set $\boldsymbol{\rho}^{(l)} \leftarrow \boldsymbol{\rho}^{\star}$, and $l \leftarrow l + 1$;
7: **end loop**

In that algorithm, line 1 initializes the values of $\{\mathcal{A}_k\}_{\forall k}$. Additionally, we begin with an initial feasible point $\boldsymbol{\rho}^{(0)}$. In practice, we find this starting point by generating it randomly, ensuring it adheres to set power constraints. We then repeatedly update this point: by solving a convex optimization problem using the power control vector from the previous iteration as input, we obtain a new vector for the next iteration. The convergence of the algorithm is assured by its design. Starting from a feasible point, $\boldsymbol{\rho}^{(0)}$, each iteration yields a better objective function value. Due to the problem's concave nature, the algorithm inevitably finds a local optimum. Thus, by iteratively solving approximated convex problems, we achieve locally optimal solutions. In practical terms, the described *loop* (lines 3-7) should continue iterating until: $\left|\left|\boldsymbol{\rho}^{(l)} - \boldsymbol{\rho}^{(l-1)}\right|\right|^2 / \left|\left|\boldsymbol{\rho}^{(l)}\right|\right|^2 \leq \epsilon$, where $\epsilon$ is an acceptable tolerance for the convergence of $\boldsymbol{\rho}$.

### *5.4.1 Key Features of this Proposed Solution*

In general, centralized and convex solutions deliver optimal or near-optimal performance, making them excellent benchmarks. However, given its centralized framework, it is important to note that the computational complexity of the proposed SCOPIA solution is directly tied to the system's user count, which compromises its scalability as the number of users approaches infinity, i.e., $K \to \infty$.

Moreover, the SCOPIA solution depends on meeting the QoS requirements for all users. Indeed, this can be a problem, as if even a small group of users fails to meet these conditions, the entire system experiences an outage event, and therefore no feasible solution is obtained.

Finally, it is important to note that the SCOPIA procedure does not precisely address the initially proposed problem in (5.12). This is due to the non-convex nature of the original problem. Therefore, it was only possible to efficiently solve problem (5.12) through convex optimization after applying concave approximations to the SE in (5.16). Specifically, for problem (5.12), as demonstrated in (Elwekeil *et al.*, 2023), the expressions in (5.16) effectively approximate the SEs for concave expressions, making them suitable for manipulation through convex optimization with the SCOPIA procedure.

However, finding effective approximations that can transform a non-convex problem into a convex one is not always feasible. In other words, convex optimization tools may inherently involve mathematical manipulations that can be challenging to derive or implement.

## 5.5 Decentralized Solutions

In this section, we introduce two decentralized power control optimization strategies, aiming to achieve scalable solutions that effectively accommodate an increasing number of users. First, in Section 5.5.1, our discussion encompasses the utilization of coordinate descent methods as a foundation for developing a robust decentralized solution. Furthermore, in Section 5.5.2, we delve into a learning approach, wherein multiple agents undergo training in a decentralized fashion.

### *5.5.1 Successive Convex Optimization and Coordinate Descent Methods for Decentralized Power Control*

In general, coordinate descent methods work by breaking down the optimization problem into smaller subproblems, where each subproblem focuses on optimizing a single parameter or a block of parameters while keeping the others fixed (Wright, 2015). This method can be particularly effective for large-scale optimization problems because it simplifies the optimization process and can be parallelized.

Indeed, by delving into the previously outlined concave SE expressions, we have the opportunity to integrate the principles of coordinate descent methods into problem (5.12). Hence, from the perspective of a typical user $\tilde{k} \in \mathcal{K}$, their personal power optimization objectives can be effectively realized by tackling the specified power control subproblem:

$$\max_{\rho_{\tilde{k}}} \ \lambda \sum_{k \in \mathcal{G}} \widetilde{\mathrm{SE}}_k \left( \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})} \right) + (1-\lambda) \sum_{k \in \mathcal{U}} \widetilde{\mathrm{SE}}_k \left( \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})} \right) \tag{5.18a}$$

$$\text{subject to } \widetilde{\mathrm{SE}}_{\tilde{k}} \left( \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})} \right) \geq \mathrm{SE}_{\min}^{\tilde{k}}, \tag{5.18b}$$

$$0 \leq \rho_{\tilde{k}} \leq P_{\max}^{\tilde{k}}, \tag{5.18c}$$

where the vector $\boldsymbol{\rho}_{(-\tilde{k})}$ represents the power levels of all users excluding user $\tilde{k}$. For problem (5.18), $\boldsymbol{\rho}_{(-\tilde{k})}$ remains constant, allowing for the exclusive optimization of user $\tilde{k}$'s power. Moreover, since subproblem (5.18) focuses on individual power control, it is essential that constraints (5.18b) and (5.18c) are satisfied exclusively for user $\tilde{k}$. Therefore, to develop a comprehensive power control strategy for the entire network, we utilize the individual power control subproblem (5.18) as a foundational component and introduce Algorithm 6, which is basically a distributed and fail-safe version of SCOPIA.

**Algorithm 6:** Distributed and Fail-Safe SCOPIA (DFS-SCOPIA)

1: **Input:** $\{\mathcal{A}_k\}_{\forall k}$, feasible point $\boldsymbol{\rho}^{(0)}$, $l \leftarrow 1$;
2: **Output:** Data power vector $\boldsymbol{\rho}$;
3: **loop**
4: **try:**
5: Use $\boldsymbol{\rho}_{(-\tilde{k})}^{(l-1)}$ and solve problem (5.18) to obtain $\rho_{\tilde{k}}^{\star}, \forall \tilde{k}$;
6: $\rho_{\tilde{k}}^{(l)} \leftarrow \rho_{\tilde{k}}^{\star}$;
7: **catch:**
8: $\rho_{\tilde{k}}^{(l)} \leftarrow P_{\max}^{\tilde{k}}$;
9: **end**
10: Signaling exchange among network nodes; $l \leftarrow l + 1$;
11: **end loop**

In particular, the DFS-SCOPIA procedure is a distributed solution strategy proposed as an alternative to addressing problem (5.12), designed to facilitate decentralized execution while maintaining the core principles and objectives of the original methodology. In this sense, for this alternative proposal, we in start in line 1 by establishing an initial power allocation vector for the network. Subsequently, we address the power of each user independently, solving for it while maintaining the power levels of other users constant, through the dedicated power control subproblem (5.18).

However, optimizing transmission power for each user while meeting QoS requirements is difficult due to the relationship between minimum rate needs, network interference, and user power levels. This challenge is especially pronounced in the early stages of Algorithm 6, where changes in power levels and shifting network conditions can cause unpredictable interference fluctuations.

#### *5.5.1.1 Key Features of this Proposed Solution*

In general Algorithm 6 can remain effective, progressing towards a solution even if a user's QoS demands are not initially met. If optimizing power control does not satisfy a user's minimum requirements, the algorithm defaults to assigning the maximum allowed power to that user. This strategy maintains network stability by avoiding algorithmic stalls and demonstrates a commitment to meeting user needs as closely as possible under existing conditions. Therefore, this method protects the network's functionality and aims to maximize user satisfaction within the scenario's limitations.

Furthermore, this strategy is meaningful for another reason. As discussed before, in

the initial iterations, a user may not actually reach their demanded QoS levels; at the same time, however, the very nature of iterations, i.e., signaling changes in multiple rounds of iterations, makes users' powers seem to stabilize. Indeed, this stabilization can significantly enhance QoS adherence as iterations progress. It is crucial to recognize, however, that this outcome is not guaranteed; the algorithm may actually complete without having guaranteed that every user meets their minimum SE requirements. Despite this, it is generally anticipated that a significant proportion of users will achieve satisfaction after many iterations.

Nevertheless, the signaling exchange in line $10$ can be complex. Specifically, the power optimization depends on the SE of each user in the network, i.e., in order for the power optimization to be decentralized, one must know the requisite elements to formulate the objective function of problem (5.18). Moreover, similar to the SCOPIA procedure, this proposed approach also relies on a closed and convex expression for the SE. In other words, the DFS-SCOPIA solution relies on a precisely well-defined structure for the optimization problem at hand, including both the objective function and the set of constraints.

### *5.5.2 Application of Distributed Deep Reinforcement Learning Methods for Managing Power Control*

In practice, there are several approaches based on DRL. Due to the continuous optimization space in the uplink power control problem (5.12), we propose a power control algorithm under the category of *actor-critic*.

#### *5.5.2.1 Definition of Agents and Actions*

In this context, we propose an approach to DRL involving multiple agents, where each user can be an individual agent. Therefore, there are $K$ agents in this approach. Moreover, we employ PPO agents, which support actors and critics and use recurrent deep neural networks as function approximator (Schulman *et al.*, 2017). In general, PPO aims to address the stability and efficiency problems encountered in earlier approaches like trust region policy optimization (TRPO) (Schulman *et al.*, 2017) by simplifying and improving the optimization process (Feriani; Hossain, 2021). Moreover, at present, PPO is the most broadly utilized policy gradient technique (Bai *et al.*, 2023). Fundamentally, PPO agents utilize a stochastic policy that is parameterized. In the case of continuous action spaces, this policy is implemented using a continuous Gaussian actor. The Gaussian actor takes an observation as input and generates a

random action as output by sampling from a Gaussian probability distribution.

To estimate the policy and value function, a PPO agent maintains two function approximators:

- $\pi\left(a|s;\phi_\pi\right)$, i.e., the actor, with parameters $\phi_\pi$, outputs the conditional probability of taking each action $a$ when in state $s$.
- $\mu\left(s;\phi_\mu\right)$, i.e., the critic, with parameters $\phi_\mu$, takes observation $s$ and returns the corresponding expectation of the discounted long-term reward.

In practice, the idea of this proposal is that agents can adjust the transmission power level through distributed power control, where the action $a \in \mathbb{R}$ of each actor is simply a continuous scalar value, and is chosen by taking into account the limitation of user $k$ maximum transmit power, that is: $0 \leq a \leq P_{\text{max}}^k, \forall k \in \mathcal{K}$. More specifically, at each time step, every agent can choose to increase, decrease, or maintain their power level, resulting in a power profile for problem (5.12).

To construct a function approximator that can be employed as a stochastic actor in a reinforcement learning agent operating in a continuous action space, a neural network is needed with two output layers: one for mean values and the other for standard deviations in each action dimension. In our case, where the actor selects a scalar for transmit power, the output includes one value for the mean and another for standard deviation. The standard deviation must be nonnegative, then we use a softplus or ReLU layer. Meanwhile, the mean value should be within the action range, so a scaling layer is used to adjust it to the desired range, i.e., $[0, P_{\text{max}}^k], \forall k \in \mathcal{K}$.

Additionally, to represent the parametrized value function in the critic, we define a neural network whose architecture consists of one input layer, which accepts the data from the observation channel, and one output layer that passes a scalar value. Critic's output is a scalar value that represents the expected discounted sum of future long-term rewards.

#### *5.5.2.2 Definition of Observations and Rewards*

Prior to exploring the intricacies of our reward function in the multi-agent learning framework, it is premised that every agent garners the same rewards and shares consistent environmental observations. This stems from the collective goal shared among all agents: to uncover a feasible solution that concurrently addresses the objectives of problem (5.12) in light of a specified input parameter $\lambda$.

Therefore, at a given time step $t$ when each agent $k$ selects an action $0 \leq a_k^{(t)} \leq P_{\text{max}}^k$,

**Algorithm 7:** Reward Optimization for SE and QoS (ROSE-Q)

**Require:** $\{\mathrm{SE}_k\}_{\forall k}, \{\mathrm{SE}^k_{\min}\}_{\forall k}, \gamma \leftarrow 0$ and $\lambda$;
1: $r_1 \leftarrow \lambda \sum_{k\in\mathcal{G}} \mathrm{SE}_k + (1-\lambda)\sum_{k\in\mathcal{U}} \mathrm{SE}_k$;
2: $r_2 \leftarrow \left(\frac{\sum_{k\in\mathcal{U}} \mathrm{SE}_k}{\sum_{k\in\mathcal{G}} \mathrm{SE}_k}\right)^{1-2\lambda}$;
3: **for** $k \in \mathcal{K}$ **do**
4: **if** $\mathrm{SE}_k < \mathrm{SE}^k_{\min}$ **then**
5: $\gamma \leftarrow \gamma + (\mathrm{SE}_k - \mathrm{SE}^k_{\min})$;
6: **end if**
7: **end for**
8: $r_3 \leftarrow \beta^\gamma$, $\beta > 1$;
9: $r \leftarrow \log(r_1 \times r_2 \times r_3)$; {reward function.}
10: **return** $r$;

representing their chosen transmit power level within the allowable maximum $P^k_{\max}, \forall k \in \mathcal{K}$, this leads to the formulation of a transmit power vector $\mathbf{a}^{(t)} \in \mathbb{R}^K$ for the entire network.

In the cell-free uplink environment we are looking at, all agents affect each other due to mutual interference, meaning that the transmission power decision of any agent $k$ influences the whole network. This interconnectedness requires agents to fully understand the network's SE levels before making decisions. Thus, we stipulate the observation or state $\mathbf{s} \in \mathbb{R}^K$ for any given user as follows:

$$\mathbf{s}^{(t)}_k = \left\{\mathrm{SE}_k(\mathbf{a}^{(t)})\right\}_{\forall k}, \quad k \in \mathcal{K}. \tag{5.19}$$

The reward signal $r^{(t)}$ is carefully constructed with three main components aimed at guiding agents towards optimal task performance and closer to desired outcomes. These components evaluate agent actions on efficiency, adherence to operational constraints, and impact on network performance. This reward function design encompasses the problem's various facets, motivating agents to not only improve their results but also contribute to the network's collective objectives. Below is the formulation of the reward function:

$$r^{(t)} = \log\left(r_1 \times r_2 \times r_3\right), \tag{5.20}$$

where components $r_1$, $r_2$, and $r_3$ are delineated in lines $1$, $2$, and $8$ of ROSE-Q procedure (Algorithm 7), respectively. In detail, component $r_1$, being the primary objective of problem (5.12), naturally plays an important role in the agents' reward function. Such a component ensures that the agents' actions are aligned with the overarching goals of problem (5.12).

On the other hand, component $r_2$ proves to be invaluable for achieving a customized or biased power control solution within the network, especially when the weight $\lambda$ leans heavily towards the extremes of $0$ or $1$. In such cases, component $r_2$ enables agents to more swiftly

learn the intended effect of input parameter $\lambda$. For instance, with $\lambda$ set to $0$, UAVs are given precedence. Here, the reward function not only aims to optimize component $r_1$, which represents the problem's primary objective, but it also seeks to maximize the ratio of the total SEs of UAVs to that of the GUEs. Consequently, the reward function increases as the sum-SE of UAVs increases, or conversely, as the sum-SE of GUEs diminishes. A parallel rationale applies when $\lambda = 1$. It is important to recognize that in scenarios where there is no explicit prioritization — specifically, when $\lambda = 0.5$ — component $r_2$ becomes functionally useless and does not influence the reward function.

Lastly, we introduce component $r_3$, which is a critical element in the system that enforces minimum SE requirements. This component imposes penalties on agents whenever they opt for power control solutions that fail to comply with the stringent requirements outlined in constraints set (5.12b). It is important to emphasize that variable $\gamma$ is set to zero when all users' needs are fully met, resulting in $r_3 = 1$. In this situation, there is no alteration to the agents' reward function. On the other hand, as the number of users not meeting the minimum SE threshold increases, variable $\gamma$ becomes increasingly negative. This decline directly influences the reduction of users' rewards, effectively serving as a form of punishment. Therefore, in such circumstances, component $r_3$ operates within a fractional range of $0 < r_3 < 1$, dynamically adjusting the rewards based on users' satisfaction with respect to the SE criteria.

#### *5.5.2.3 Training Algorithm*

In general, the concept behind PPO focuses on enhancing the stability of policy training. This is achieved by constraining the modifications applied to the policy during each training epoch, aiming to prevent excessively large updates to the policy. This is due to two main reasons. Firstly, empirical evidence suggests that smaller, incremental policy updates throughout the training process are more likely to converge to an optimal solution. Secondly, overly large adjustments in policy updates risk causing a drastic decline in performance, metaphorically causing the policy to "fall off the cliff". This can result in prolonged recovery times, or in some cases, the inability to recover at all.

The algorithm utilized for training the PPO agents is detailed in Algorithm 8. Regarding the described algorithm, we can begin by initializing the actors and critics with random parameters. The *loop* between steps $3$ to $5$ is executed numerous times, with each iteration being referred to as an *episode*. During each episode, $N$ experiences are generated, each structured as

a $4$-element tuple. This $4$-tuple comprises the current system observation, the selected action, the subsequent observation, and, naturally, the reward, i.e., $(\mathbf{s}^{(t)}, a^{(t)}, \mathbf{s}^{(t+1)}, r^{(t+1)})$, where $r^{(t+1)}$ is obtained from ROSE-Q procedure (Algorithm 7). The experiences generated are important for determining *advantage function* $D$, and *return function* $G$.

While there are some methods to achieve this, in our approach we employ the *generalized advantage estimator*, commonly referred to as GAE, to ascertain the estimates of these functions, as illustrated in step $4$. Mathematically, the advantage function for a time step $t$ is given by the sum (Schulman *et al.*, 2015):

$$D^{(t)} = \sum_{i=0}^{\infty} (\tilde{\gamma}\tilde{\lambda})^i \delta^{(t+i)}, \tag{5.21}$$

where $\delta^{(t+i)} = r^{(t+i)} + \tilde{\gamma}\mu\left(\mathbf{s}^{(t+i+1)}; \phi_\mu\right) - \mu\left(\mathbf{s}^{(t+i)}; \phi_\mu\right)$ is the temporal difference error. Meanwhile, $\tilde{\gamma}$ and $\tilde{\lambda}$ are the discount and the GAE factors, respectively. In our particular case, (5.21) is tailored to accommodate $N$ experiences per episode, as outlined in Algorithm 8.

Specifically, the GAE method employs the advantage function $D$ to compute the return function $G$. This, in its turn, facilitates the update of the critic parameters through the minimization of the loss function $\mathcal{L}_{\text{critic}}$ in step $5$. Mathematically, $\mathcal{L}_{\text{critic}}$ is given by:

$$\mathcal{L}_{\text{critic}} = \mathbb{E}\left\{\left(G - \mu\left(\mathbf{s}; \phi_\mu\right)\right)^2\right\}. \tag{5.22}$$

In practical terms, the critic parameters are updated by minimizing the loss $\mathcal{L}_{\text{critic}}$ across all sampled mini-batch data according to Algorithm 8 (step $5$). While the concept of updating the critic parameters is straightforward, the process of refining the policy is significantly more complex. In the case of PPO, policy updates are made with caution. This involves assessing the degree of change between the current and previous policies through a ratio of their probabilities. We then apply clipping to this ratio within a specified range, $[1 - \tilde{\epsilon}, 1 + \tilde{\epsilon}]$, effectively limiting how much the current policy can deviate from its predecessor (Schulman *et al.*, 2017). This mechanism justifies the "proximal" aspect of *proximal policy optimization*, as it limits or reduces the incentive for drastic shifts from the previous policy.

Hence, PPO employs a novel objective function known as the *clipped surrogate objective* (Schulman *et al.*, 2017). This function is designed to keep policy updates within a tight boundary, utilizing a clipping technique to ensure that changes to the policy are modest, as illustrated below:

$$\mathcal{L}_{\text{actor}} = \mathbb{E}\left\{\left(-\min\left(q(\phi_\pi)D, c(\phi_\pi)D\right) + \mathcal{W}\right)\right\}, \tag{5.23}$$

where

$$q\left(\phi_\pi\right) = \frac{\pi\left(a\middle|\mathbf{s};\phi_\pi\right)}{\pi\left(a\middle|\mathbf{s};\phi_\pi^{\text{old}}\right)}, \tag{5.24a}$$

$$c(\phi_\pi) = \max\left(\min\left(q(\phi_\pi), 1+\tilde{\epsilon}\right), 1-\tilde{\epsilon}\right), \tag{5.24b}$$

$$\mathcal{W} = w\mathcal{E}\left(\phi_\pi, \mathbf{s}\right), \tag{5.24c}$$

where $\pi\left(a\middle|\mathbf{s};\phi_\pi\right)$ and $\pi\left(a\middle|\mathbf{s};\phi_\pi^{\text{old}}\right)$ are the probabilities of taking action $a$ when in state $\mathbf{s}$, given the updated policy parameters $\phi_\pi$, and given the previous policy parameters $\phi_\pi^{\text{old}}$, respectively. Moreover, $\tilde{\epsilon}$ is the clip factor, $\mathcal{E}(\cdot)$ is the entropy loss, and $w$ is the entropy loss weight.

In more practical terms, function (5.23) is designed to avoid updates that are excessively large and potentially detrimental to the weights. Specifically, $q\left(\phi_\pi\right)$ represents the ratio of probabilities between the current policy and the previous one. Thus, if $q\left(\phi_\pi\right) > 1$, the action $a$ at state $\mathbf{s}$ is more likely in the current policy than the old policy. Instead, if $q\left(\phi_\pi\right)$ is between $0$ and $1$, the action is less likely for the current policy than for the old one. Note that this probability ratio offers a straightforward method to gauge the divergence between the old and current policy. However, if the action taken is significantly more likely under our current policy compared to the previous one, the probability ratio $q\left(\phi_\pi\right)$ will be very different from $1$, yielding a large step on that term and generating a policy update that is too aggressive. As a result, it becomes necessary to regulate this objective function by imposing penalties on changes that result in a ratio significantly deviating from $1$.

In this context, by employing a clip to the ratio $q\left(\phi_\pi\right)$ using $c(\phi_\pi)$, we prevent excessively large updates to the policy, as it restricts the current policy ($\pi\left(a\middle|\mathbf{s};\phi_\pi\right)$) from diverging too much from the previous one, i.e., $\pi\left(a\middle|\mathbf{s};\phi_\pi^{\text{old}}\right)$. Therefore, to derive function (5.23), we use both the clipped and unclipped objectives, and then select the lesser of the two as the final objective, effectively establishing a lower bound (or a pessimistic estimate) of the unclipped objective.

Moreover, it should be emphasized that (5.23) includes an entropy term ($w\mathcal{E}(\cdot)$). More specifically, the entropy term gauges the unpredictability of an agent's actions, reflecting its uncertainty. Incorporating this term helps adjusting the agent's inclination to explore. A higher entropy suggests more uncertainty in decisions, and by maximizing entropy, we encourage the agent to explore more. Indeed, this can be beneficial in complex settings to avoid regions of local

**Algorithm 8:** PPO Agents Training Algorithm (PPO-ATA)

1: Initialize the actor $\pi\left(a|\mathbf{s};\phi_\pi\right)$ and the critic $\mu\left(\mathbf{s};\phi_\mu\right)$ with parameters $\phi_\pi$ and $\phi_\mu$, respectively;
2: **loop**
3: Produce $N$ experiences adhering to the current policy, as follows:

$$\left(\mathbf{s}^{(0)}, a^{(0)}, \mathbf{s}^{(1)}, r^{(1)}\right), \ldots, \left(\mathbf{s}^{(N-1)}, a^{(N-1)}, \mathbf{s}^{(N)}, r^{(N)}\right).$$

4: For each episode step $t$, compute: $D^{(t)} = \sum_{i=t}^{N-1} (\tilde{\gamma}\tilde{\lambda})^{i-t}\delta^{(i)}$. Next, compute the return: $G^{(t)} = D^{(t)} + \mu\left(\mathbf{s}^{(t)};\phi_\mu\right)$.
5: Learn from mini-batches of experiences over multiple epochs.
- Sample a random mini-batch data set of size $M$.
- Update the critic parameters by minimizing $\mathcal{L}_{\text{critic}}$:

$$\mathcal{L}_{\text{critic}} = \frac{1}{2M}\sum_{i=1}^{M}\left(G^{(i)} - \mu\left(\mathbf{s}^{(i)};\phi_\mu\right)\right)^2.$$

- Update the actor parameters by minimizing $\mathcal{L}_{\text{actor}}$:

$$\mathcal{L}_{\text{actor}} = \frac{1}{M}\sum_{i=1}^{M}\left(-\min\left(q^{(i)}(\phi_\pi)D^{(i)}, c^{(i)}(\phi_\pi)D^{(i)}\right) + \mathcal{W}^{(i)}\right).$$

6: **end loop**

optima. For a continuous action space, the agent utilizes:

$$\mathcal{E}^{(i)}\left(\phi_\pi, \mathbf{s}^{(i)}\right) = \frac{1}{2}\sum_{a=1}^{A}\text{In}\left(2\pi e\sigma_{a,i}^2\right), \tag{5.25}$$

where $A$ is the number of continuous actions output by the actor, and $\sigma_{a,i}$ is the standard deviation for action $a$ when in state $\mathbf{s}^{(i)}$ following the current policy.

In practical terms, the actor parameters are updated by minimizing the loss function described in (5.23), across all sampled mini-batch data, as outlined in step 5 of Algorithm 8. Typically, to minimize these loss functions, we employ techniques such as the traditional *stochastic gradient descent* or the *Adam optimization algorithm* (Kingma; Ba, 2017).

Hence, in step 5 the core learning process unfolds. Drawing upon random samples from the mini-batch dataset, the algorithm refines the parameters of both the actor and the critic, aiming to minimize their respective loss functions to achieve optimal performance. Essentially, this step is where the system learns from experience, iteratively improving its performance to achieve the best possible results.

#### *5.5.2.4 Key Features of this Proposed Solution*

In this proposed solution, each agent collects its experiences independently during episodes, learning and adapting from their specific interactions with the environment free from

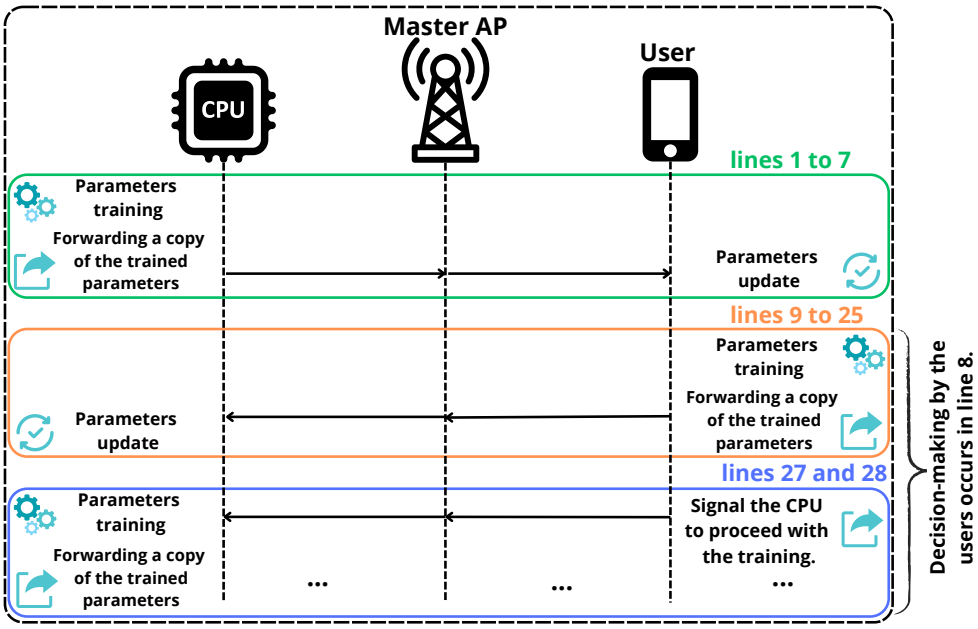
**Algorithm 9:** Periodic Learning and Adjustment Technique

1: **for** $i = 1, \ldots, T_{\text{CPU}}$ **do**
2: CPU: Allocate $K$ processing cores and in each of them run the PPO-ATA algorithm to simulate $K$ distributed agents. To construct the state and reward, it is essential to utilize $\{\text{SE}_k\}_{\forall k}$, which are readily accessible at this network layer and can be disseminated through backhaul links.
3: **end for**
4: CPU: After $T_{\text{CPU}}$ time steps, stop the training.
5: CPU: Send a copy of parameters $\phi_\pi$,$\phi_\mu$ to the master AP of each user.
6: Master AP: Send the parameters to corresponding users.
7: UE: Update the parameters $\phi_\pi$,$\phi_\mu$.
8: UE: After $\tilde{T}$ time steps, indicate to the master AP whether to proceed with training or not using a flag $f$.
9: **if** $f == 1$ **then**
10: **for** $i = 1, \ldots, T_{\text{User}}$ **do**
11: **for** $n = 1, \ldots, N$ **do**
12: Master AP: Broadcast the current state $\mathbf{s}$ to the users.
13: UE: Send $\rho_k$ to the master AP.
14: Master AP: Send $\{\rho_k\}_{\forall k}$ to the CPU.
15: CPU: Compute $\{\text{SE}_k\}_{\forall k}$ and $r$, and then send it to the master AP.
16: Master AP: Broadcast the next state $\mathbf{s}'$ and reward $r$ to the users.
17: UE: Create and store the experience $(\mathbf{s}, \rho_k, \mathbf{s}', r)$.
18: **end for**
19: UE: Implement steps 4 and 5 from the PPO-ATA algorithm.
20: **end for**
21: UE: After $T_{\text{User}}$ time steps, stop the training.
22: UE: Send a copy of parameters $\phi_\pi$, $\phi_\mu$ to the master AP.
23: Master AP: Send the parameters to the CPU.
24: CPU: Update the parameters $\phi_\pi$,$\phi_\mu$.
25: **Return to line 8**.
26: **else**
27: Master AP: Signal the CPU to proceed with the training.
28: **Return to line 1**.
29: **end if**

other agents' influence. Thus, parameters $\phi_\pi$ and $\phi_\mu$ are customized for each user, and distributed training allows our learning-based power control method to scale efficiently with user numbers. Moreover, in our learning-driven method, even when it is impossible to satisfy all users, we continue to find reasonable solutions without interruptions. This process is similar to the strategy discussed in Algorithm 6, highlighting our method's ability to adapt and achieve progress despite challenges.

Now, a significant benefit of our learning methodology, compared to the methods outlined in Algorithms 5 and 6, lies in its flexibility/adaptability. Differently from the previous approaches, our strategy herein is not constrained by the requirement for a convex formulation of SE to optimize transmission powers. Furthermore, it does not necessarily depend on a closed SE expression. One reason is that agents only need SE values for their state and reward functions, and those do not have to be derived from expressions, as they can be empirically collected. This aspect is particularly adequate for a dynamic environment, as directly computing an SE

Figure 26 – Signaling scheme for Algorithm 9, highlighting an overview of the three different blocks that occur in this algorithm.




Source: Created by the author.

expression might be remarkably difficult. Furthermore, signaling exchange among agents is simple, it basically depends on the SE values, allowing a broadcast transmission approach due to the uniform state and the reward shared between all agents.

Learning-based methods often struggle with prolonged agent training times. In this context, solely using mobile terminals for complex problem-solving in mobile networks is impractical due to limited user computational power. On the other hand, relying entirely on the network's superior computing capabilities for training becomes inefficient with the addition of new users, requiring more and more processing cores and potentially leading to bottlenecks.

To address this challenge, we consider a hybrid approach to optimize the use of computational resources. In this model, the initial intensive training phase is handled by the network's powerful computing centers. After this phase, the trained parameters are distributed to users, allowing training to continue on user devices in a decentralized fashion. Therefore, once the parameters of the agents have been sufficiently trained, the network then distributes parameters to the users, enabling the continuation of the training process in a decentralized manner on the users' devices.

This concept is more thoroughly explained in Algorithm 9. In this algorithm, in line 2, the CPU can orchestrate the training phase of the $K$ agents at the start, having the ability to distribute the $K$ agents to other independent processing cores also in several CPUs within the

network. This enables the utilization of advanced computational resources; for example, the use of specialized tools for machine learning like *graphics processing units* (i.e., GPUs) that reside within the network infrastructure can be used to improve or adjust parameters for each user. The environment's response to an agent's action, including the next state and reward, is easily distributed among agents via the backhaul link. This training typically spans $T_{\text{CPU}}$ time steps. But, the duration of $T_{\text{CPU}}$ can fluctuate, increasing or decreasing based on the availability of processing units.

Regardless, these processing units should not be permanently engaged, i.e., after $T_{\text{CPU}}$ time steps, training is halted according to line $4$. Subsequently, each pair of parameters $\phi_\pi$, $\phi_\mu$ is dispatched to the corresponding user, as detailed in lines $5$, $6$, and $7$. Thus, equipped with these parameters, users can make action decisions with practically no computational overhead, since the parameters have been pre-trained. In line $8$, after $\tilde{T}$ time steps, should there be a need for parameter adjustments due to new priorities or substantial changes in the channel, users can opt to either proceed with additional training on their own or alert the CPU to manage a recalibration of parameters $\phi_\pi$, $\phi_\mu$ over a new training period.

Hence, if users choose to proceed with the training, it must be carried out within the $T_{\text{User}}$ time step, following the sequence of instructions outlined from line $10$ to line $25$. Alternatively, users notify the master AP, which then communicates with the CPU to proceed the training, as depicted in line $27$. Following this training period, the aforementioned cycle is reiterated.

To better illustrate this approach, Fig. 26 illustrates the overarching concept of Algorithm 9, focusing on the signaling exchange among the CPU, master APs, and users. Within this figure, we categorize the primary procedural flows during the information exchange among network nodes into three distinct blocks. In this context, the initial green block involves training the parameters $\phi_\pi$, $\phi_\mu$ on the CPU and then forwarding them to the users. The orange and blue blocks are influenced by the users' decision to either conduct training to adjust the parameters $\phi_\pi$, $\phi_\mu$ autonomously (orange block, case $f = 1$), or to request that this training process be repeated on the CPU (blue block, case $f \neq 1$).

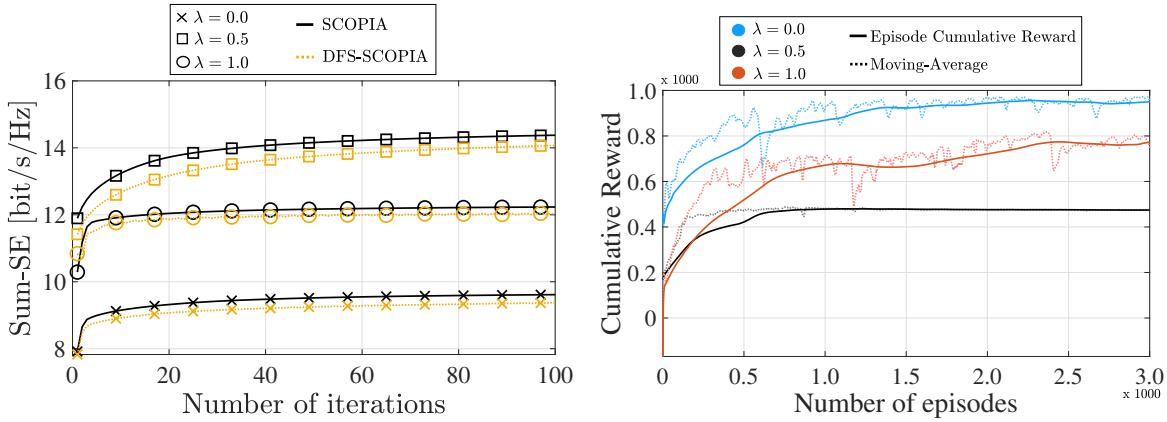
In general, the methodology described above provides a continuous mechanism for improving and adjusting the model. Using the superior computing powers of the network to occasionally enrich or change agent training when traffic volume is low allows the system to evolve in parallel with network conditions, users' actions, and traffic load.

## 5.6 Simulation Results

Our simulation scenario features a cell-free network model with $40$ APs and $8$ users, where half are designated as GUEs and the other half as UAVs (Zheng *et al.*, 2024). Our proposed learning solution features agents designed in accordance with the methodologies outlined in Section 5.5.2, augmented by two hidden layers, each equipped with $256$ neurons. The minimum SE for GUEs is $0.30$ bits/s/Hz, while for UAVs, it is $0.60$ bits/s/Hz. The maximum power for each user is $150$ mW. Each episode consists of $250$ experiences. Moreover, the agent learning rate is set to $25 \times 10^{-4}$, with a discount factor of $0.995$. The GAE factor is $0.95$. There are $3$ epochs, with a mini-batch size of $256$, a clip factor of $0.2$, and an entropy loss weight of $0.01$.

### *5.6.1 Convergence Behavior*

Figure 27 – Convergence curves for solutions based on convex optimization (SCOPIA, a centralized solution, and DFS-SCOPIA, a decentralized solution) and learning curves of agents across various $\lambda$ values.




(c) *SCOPIA and DFS-SCOPIA.* (d) *PPO-ATA.*

Source: Created by the author.

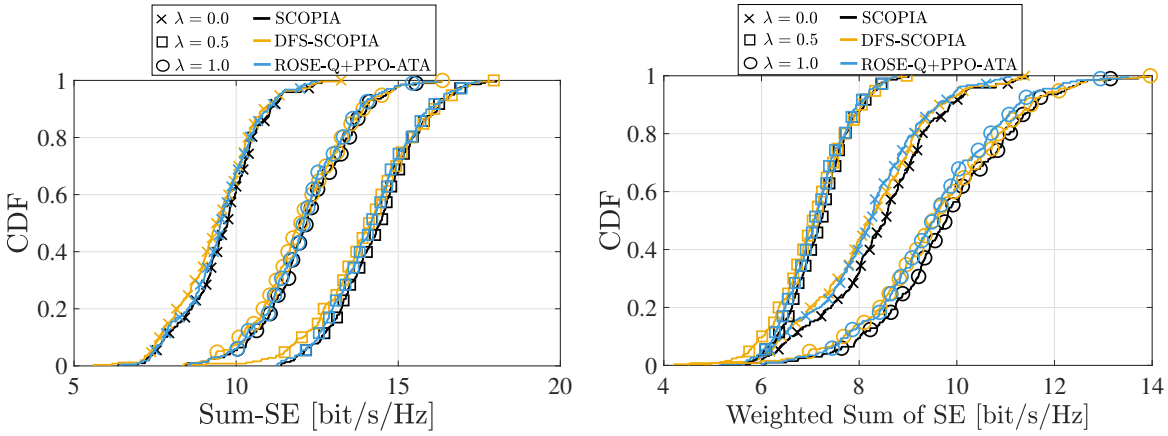
In Fig. 27c, we present the convergence curves, illustrating the sum-SE achieved by the solutions based on convex optimization, specifically SCOPIA and DFS-SCOPIA. This analysis incorporates various values of $\lambda$, which significantly influences the sum-SE. It is important to note that at $\lambda = 0.5$, the objective shifts towards maximizing the traditional sum-SE, resulting in the highest value of this metric. Note that the DFS-SCOPIA solution demonstrates a slight decrease in performance compared to the SCOPIA solution. Such a result is justified by the costs of distributing the SCOPIA solution. In general, the distribution can obstruct users from

effectively finding the optimal power allocation. Nonetheless, it is crucial to point out that the slight performance dip is compensated by the considerable scalability advantages of distributed solutions, particularly in scenarios or environments with many users.

In Fig. 27d, we illustrate the convergence of the proposed solution employing our learning-based approach. In general, in such solutions, convergence is defined by the accumulation of rewards over time. As agents gather maximum possible rewards, it is anticipated that they achieve the optimal policy, signifying the convergence in the learning process of the agents. Herein, we also demonstrate the learning convergence across varying values of $\lambda$. In this result, we emphasize that our solution for $\lambda = 0.5$ achieves convergence rapidly, in fewer than $1000$ episodes. Nevertheless, when prioritizing GUEs or UAVs exclusively (i.e., when $\lambda = 0.0$ or $\lambda = 1.0$), the difficulty in identifying efficient solutions increases considerably. In such scenarios, a greater number of episodes is required to develop an effective solution.

### *5.6.2 Performance Comparison*

Figure 28 – CDF curves for the sum-SE and for the weighted sum-SE, comparing SCOPIA, DFS-SCOPIA and ROSE-Q+PPO-ATA solutions across various $\lambda$ values.



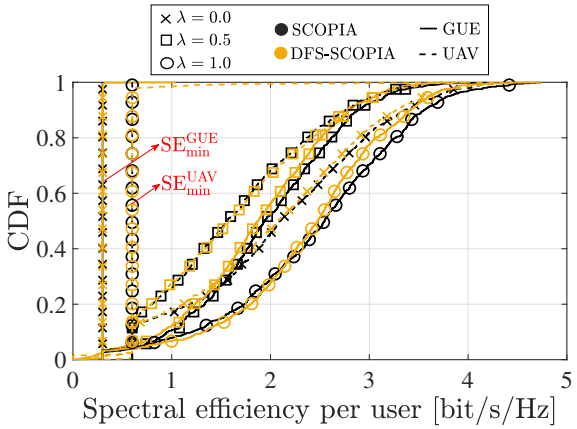

(c) *Sum-SE metric.* (d) *Weighted Sum-SE metric.*

Source: Created by the author.

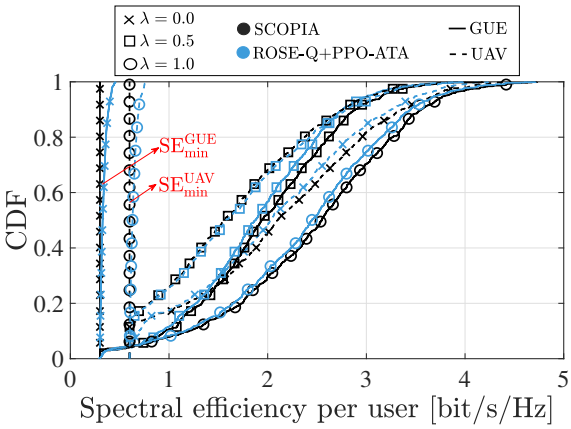
In Figs. 28c and 28d, we present the cumulative distribution functions (CDFs) for objective function (5.12a) and for the sum-SE, across three distinct values of $\lambda$, respectively. From these figures, we can see that our proposed distributed solutions demonstrate comparable performance to the centralized SCOPIA solution in both the weighted sum-SE (objective function) and the unweighted sum-SE aspect of problem (5.12).

In Figs. 29c and 29d, we more clearly illustrate the influence of $\lambda$ on the SE per user for both GUEs and UAVs. Additionally, we distinctly present the performance of our proposed decentralized solutions in comparison to the centralized approach. Particularly about these results, note that on the extreme ends of the $\lambda$ values, i.e., with $\lambda$ equaling $0$ or $1$, the system's operational focus shifts dramatically. When $\lambda = 1$, the emphasis is on maximizing the sum-SE of GUEs. This focus allows GUEs to effectively utilize network resources, thereby improving their throughput and QoS. Concurrently, UAV transition to a mode of operation that minimizes power consumption, adhering to just the essential SE requirements. On the other hand, with $\lambda = 0$, UAVs receive the utmost priority in terms of sum-SE. This prioritization enables UAVs to fully leverage the system's capabilities, enhancing their service quality and supporting high-demand applications. However, this leads GUEs to adopt a minimalistic approach toward energy use, striving to meet their SE needs with the least amount of power.

Figure 29 – CDF curves for SE per user of GUEs and UAVs, comparing SCOPIA, DFS-SCOPIA and ROSE-Q+PPO-ATA solutions solutions across various $\lambda$ values.




(c) *SCOPIA and DFS-SCOPIA.*

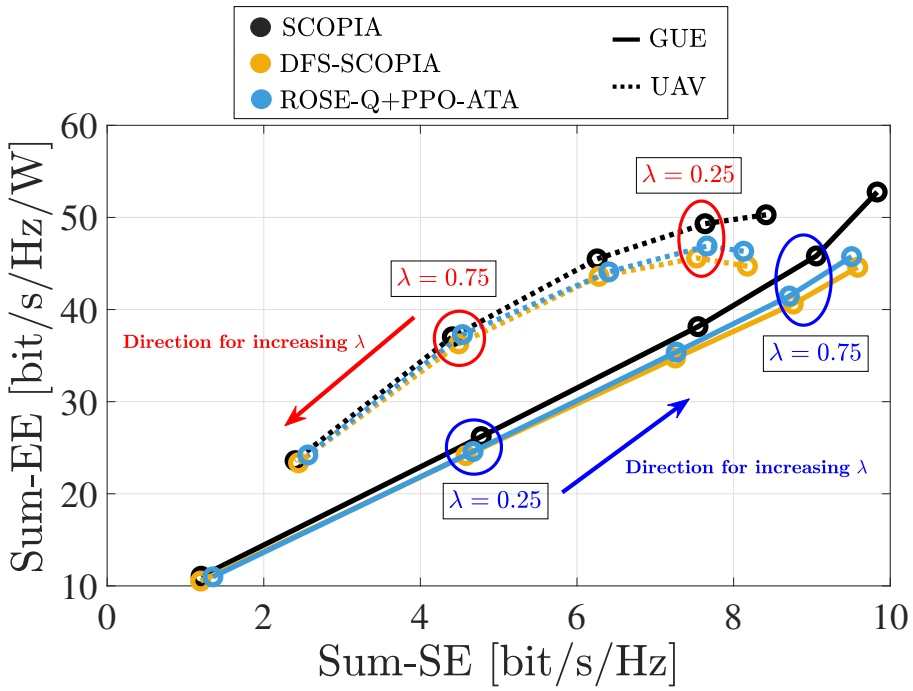



(d) *SCOPIA and ROSE-Q+PPO-ATA.*

Source: Created by the author.

It is important to emphasize that even when $\lambda = 0$, it does not signify complete priority for UAVs, as their allocated power seldom reaches the maximum available. If it were possible for them to utilize maximum power, UAVs would exhibit a significantly superior SE profile compared to GUEs due to their enhanced quality. However, the improved channel quality of UAVs also renders them potential sources of interference for GUEs. Therefore, even when $\lambda = 0$, UAV power levels are meticulously regulated to ensure that GUEs can meet their minimum SE requirements.

Figure 30 – Trade-off curves between Sum-EE and Sum-SE for SCOPIA, DFS-SCOPIA, and ROSE-Q+PPO-ATA solutions.




Source: Created by the author.

Hence, at these extreme values of $\lambda$, where the focus shifts towards prioritizing one group of users while only minimally meeting the needs of another, efficient power control becomes essential. In these situations, achieving such efficiency is unlikely through simplistic or trivial solutions. In fact, even when we have $\lambda = 0.5$ and problem (5.12) towards enhancing the classical sum-SE, it does not make simple strategies like *full power* and *fractional power control* effective. In general, these simpler approaches do not align with the core philosophy of problem (5.12), rendering a comparison with them uninteresting. This is not only due to their inability to address minimum SE constraints but also because they remain indifferent to the priority levels set by parameter $\lambda$.

In these results, we emphasize that our learning-based solution sustains performance comparable to the SCOPIA solution, irrespective of the prioritization levels governed by parameter $\lambda$, and it is achieved while adhering to the QoS constraints. This demonstrates our solution's capability to balance network priorities efficiently within stringent operational guidelines.

Although the objective function in (5.12a) primarily quantifies SE, it is crucial to recognize that altering the parameter $\lambda$ significantly influences users' EE as well. To demonstrate it, Fig. 30 illustrates the trade-off between sum-EE and sum-SE for both GUEs and UAVs as we

vary $\lambda$. Several observations can be made based on these results. First, it is essential to note the similar performance of the solution based on learning and the DFS-SCOPIA. Additionally, it is worth noting that $\lambda$ significantly influences the EE metric. Moreover, the UAVs with equivalent total SE achieve a higher total EE than GUEs, which results from the much better link quality experienced by UAVs mediated through the implicit LOS condition in the UAV channels.

Therefore, adjusting the value of $\lambda$ from $0$ to $1$ significantly impacts the SE and EE of both GUEs and UAVs, leading to substantial changes in their operation and performance. This adjustment serves as an important control mechanism, alternating its focus between GUEs and UAVs as needed. Finally, it is important to mention in this regard that while $\lambda$ changes in response to the priorities shift, our novel decentralized solutions ensure a level of performance that is almost the same as that of the centralized and convex methodologies, while also providing improved adaptability and scalability.

## 5.7 Chapter Summary

In this chapter, we explored the problem of managing data power in cell-free systems where GUEs and UAVs operate together. We proposed both centralized and decentralized solutions, utilizing convex optimization and DRL techniques. We highlight our learning-based strategy that successfully managed power allocation for GUEs and UAVs, addressing the distinct priority levels in terms of sum-SE between these two groups.

Additionally, considering the adaptable and easily modifiable characteristics of DRL-based approches, our proposed solution can be deployed in a distributed manner, utilizing a straightforward signaling protocol, which allows the investigation of more efficient combining structures, even though it does not require a closed-form or convex expression for SE. Indeed, investigating learning-based solutions that utilize varied combining structures is an attractive direction for further research. This exploration may also incorporate more realistic aspects, including mobility and time-correlated channels, along with objectives specifically tailored to the context of green communications.

# 6 CONCLUSIONS AND FUTURE WORKS

The main objective of this thesis was to investigate different strategies and solutions for power control in centralized and distributed MIMO networks. In this context, Chapter 1 provided an overview of centralized and distributed MIMO architectures and highlighted the significance of efficient RRM solutions in mobile networks.

Centralized uplink and downlink MIMO networks were studied in Chapter 2 and Chapter 3. More specifically, Chapter 2 investigated the transceiver design problem in MIMO interference broadcast channel (IBC) networks, focusing on maximizing global EE while adhering to QoS constraints related to data rates for all users. This study assumed a non-full buffer traffic model and employed a more realistic channel model that incorporates temporal and spatial correlation. To address this issue, we proposed both centralized and decentralized solutions based on the theory of fractional programming. In the results, we compared our solutions with various benchmark solutions and demonstrated that our proposed methods not only ensure efficient energy use but also can effectively reduce the size of the queue of bits in the buffer.

Meanwhile, Chapter 3 examined an uplink massive MIMO network, considering a temporally correlated channel modeled by a first-order autoregressive process. To effectively utilize this channel model, we designed a game theory-based decentralized solution that optimizes data and pilot power to minimize the sum of MSEs. Moreover, by decoupling the number of users from the computational complexity of our proposed solution, it has the potential to remain feasible regardless of the number of users. In the results, we compared our solutions with various benchmark solutions and demonstrated that our approach remains efficient even in scenarios where the parameter estimates of our autoregressive channel model contain errors.

Decentralized uplink MIMO networks were studied in Chapter 4 and Chapter 5. More specifically, Chapter 4 proposed a decentralized solution, categorized under potential games, for controlling data power in cell-free networks. In this solution, we also designed a method to decentralize the optimization of data power at the user level, resulting in a potentially scalable approach as well. The results demonstrated that our solution effectively achieves various network objectives, such as max-min fairness and SE, while also achieving an interesting trade-off between EE and SE.

Finally, Chapter 5 addressed the challenge of efficiently integrating GUEs and UAVs into cell-free networks. In this context, we investigated the weighted sum-SE problem with QoS constraints. This problem enabled the prioritization of power resources between GUEs and UAVs,

making it particularly relevant in scenarios where these users coexist. To deal with this problem, initially, we proposed centralized and decentralized solutions using convex optimization theory. In order to develop more flexible and adaptable frameworks, we then introduced a deep learning-based solution for power control, categorized under distributed actor-critic methods. Moreover, our proposed decentralized solutions can also adeptly accommodate the network's increasing user base. In the results, we demonstrated that our proposed decentralized solutions achieve comparable performance to centralized and convex approaches across various prioritization levels between GUEs and UAVs.

**Future Works**

Considering the context of MIMO networks, there are some general research directions which are potential topics for future investigation. Some of these directions are described below:

- Given the context of 5th generation (5G) networks, incorporating aspects of ultra-reliable and low-latency communications (URLLC) into the RRM solutions discussed in this thesis presents an intriguing area for future research.
- One aspect that was overlooked in the studies of this thesis is hardware impairments. In practice, accounting for these impairments is crucial, as they can significantly impact SE and EE metrics, among other performance indicators.
- In our studies of cell-free networks, we did not address AP selection solutions. Therefore, in the future, we can develop solutions that simultaneously tackle AP selection and power control.

Particularly in the context of UAV communications, future studies could explore the following perspectives:

- Studies on UAVs serving as aerial base stations to provide temporary coverage in areas with insufficient infrastructure, such as during emergencies or in rural locations, are indeed crucial.
- In the context of *network planning*, determining the optimal positions for UAVs to maximize coverage, minimize latency, and enhance network capacity are interesting studies. Moreover, designing efficient flight paths to maintain connectivity while minimizing energy consumption and avoiding obstacles may also be relevant research directions.
- In the context of *energy harvesting*, we have the interest of utilizing solar panels or other

energy harvesting methods to extend UAV operation time as well as developing protocols that minimize energy consumption without compromising performance.

- In the context of *interference management*, investigating techniques to mitigate interference between UAVs and GUEs remains a key and persisting research challenge within UAV communications. Specifically, this includes adaptive beamforming, dynamic frequency allocation, and collaborative interference management strategies.
- In the context of *mobility management*, is it possible to study advanced mobility management techniques to handle the dynamic nature of UAV-based networks. Specifically, this includes seamless handover protocols, mobility prediction algorithms, and dynamic resource allocation to ensure continuous and reliable connectivity.
- Finally, in the context of *QoS and QoE optimization*, developing QoS and quality of experience (QoE) optimization techniques for UAV-based communication networks to meet diverse user requirements, and investigating adaptive QoS frameworks that dynamically adjust to varying network conditions and user demands represent intriguing areas for future investigation.

Obviously, this is just a small sample of what can be studied in the future. More important than listing all possible future studies is the understanding that there will always be new discoveries and avenues for research. As Carl Sagan (Spangenburg; Moser, 2004) aptly said, "*Somewhere, something incredible is waiting to be known*." This essence of exploration and discovery is what makes science and research truly fascinating.

# BIBLIOGRAPHY

3GPP. *TR 38.900 v14.2.0: Study on Channel Model For Frequency Spectrum above 6 GHz*. Sophia Antipolis, France, 2016.

3GPP. *TR 36.777 V0.5.0; Technical Specification Group Radio Access Network; Study on Enhanced LTE Support for Aerial Vehicles (Release 15)*. Sophia Antipolis, France, 2017.

3GPP. *TR 36.814 v9.2.0: Further Advancements For E-UTRA Physical Layer Aspects*. Sophia Antipolis, France, 2017.

3GPP. *TR 36.814 V9.2.0; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements For E-UTRA Physical Layer Aspects (Release 9)*. Sophia Antipolis, France, 2017.

3GPP. *TS 38.213 V16.6.0; NR; Physical layer procedures for control (Release 16)*. Sophia Antipolis, France, 2021.

Akra, A.; Assaad, M. Energy efficient transmit beamforming under queueing stability constraints. In: **Proceedings of the IEEE Wireless Communications and Networking Conference**. Doha, Qatar: IEEE. Apr. 03-06, 2016.

Albreem, M. A.; Juntti, M.; Shahabuddin, S. Massive MIMO detection techniques: A survey. **IEEE Communications Surveys & Tutorials**, IEEE, v. 21, n. 4, p. 3109–3132, 2019.

Ameigeiras, P.; Wang, Y.; Navarro-Ortiz, J.; Mogensen, P. E.; Lopez-Soler, J. M. Traffic models impact on OFDMA scheduling design. **EURASIP Journal on Wireless Communications and Networking**, SpringerOpen, v. 2012, n. 1, p. 61, 2012.

Antonioli, R. P.; Braga, I. M.; Fodor, G.; Silva, Y. C. B.; de Almeida, A. L. F.; Freitas, W. C. On the energy efficiency of cell-free systems with limited fronthauls: Is coherent transmission always the best alternative? **IEEE Transactions on Wireless Communications**, IEEE, v. 21, n. 10, p. 8729–8743, 2022.

Antonioli, R. P.; Fodor, G.; Soldati, P.; F. Maciel, T. User scheduling for sum-rate maximization under minimum rate constraints for the MIMO IBC. **IEEE Wireless Communications Letters**, IEEE, v. 8, n. 6, p. 1591–1595, 2019.

Asplund, H. **Advanced Antenna Systems For 5G Network Deployments: Bridging the Gap Between Theory and Practice**. Amsterdam: Academic Press, Elsevier, 2020. ISBN 9780128223864.

Bai, Y.; Zhao, H.; Zhang, X.; Chang, Z.; Jäntti, R.; Yang, K. Toward autonomous multi-UAV wireless network: A survey of reinforcement learning-based approaches. **IEEE Communications Surveys & Tutorials**, IEEE, v. 25, n. 4, p. 3038–3067, 2023.

Baracca, P.; Giordano, L. G.; Rodriguez, A. G.; Geraci, G.; López-Pérez, D. Downlink performance of uplink fractional power control in 5G massive MIMO systems. In: **Proceedings of the IEEE Globecom Workshops**. Abu Dhabi, United Arab Emirates: IEEE. 09-13 Dec., 2018.

Basturk, I.; Chen, Y. Energy efficiency for MISO-OFDMA-based user-relay assisted cellular networks. **IEEE Systems Journal**, IEEE, v. 14, n. 4, p. 1–10, 2020.

Bertsekas, D. P. **Nonlinear Programming**. 2nd. ed. Belmont, MA, USA: Athena scientific, 1999. ISBN 9781886529007.

Björnson, E.; Sanguinetti, L. Scalable cell-free massive MIMO systems. **IEEE Transactions on Communications**, IEEE, v. 68, n. 7, p. 4247–4261, 2020.

Björnson, E.; Sanguinetti, L.; Hoydis, J.; Debbah, M. Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer? **IEEE Transactions on Wireless Communications**, IEEE, v. 14, n. 6, p. 3059–3075, 2015.

Boyd, S.; Vandenberghe, L. **Convex Optimization**. Cambridge, U.K.: Cambridge University Press, 2004. ISBN 9780521833783.

Boyd, S.; Xiao, L.; Mutapic, A.; Mattingley, J. **Sequencial Convex Programming - Notes for EE364b, Stanford University**. 2007. Disponível em: <https://see.stanford.edu/materials/lsocoee364b/Syllabus.pdf>.

Braga, I. M.; Antonioli, R. P.; Fodor, G.; Silva, Y. C. B.; Freitas, W. C. Decentralized joint pilot and data power control based on deep reinforcement learning for the uplink of cell-free systems. **IEEE Transactions on Vehicular Technology**, IEEE, v. 72, n. 1, p. 957–972, 2023.

Busari, S. A.; Huq, K. M. S.; Felfel, G.; Rodriguez, J. Adaptive resource allocation for energy-efficient millimeter-wave massive MIMO networks. In: **Proceedings of the IEEE Global Communications Conference (GLOBECOM)**. Abu Dhabi, United Arab Emirates: IEEE. 09-13 Dec., 2018.

Buzzi, S.; Colavolpe, G.; Saturnino, D.; Zappone, A. Potential games for energy-efficient power control and subcarrier allocation in uplink multicell OFDMA systems. **IEEE Journal of Selected Topics in Signal Processing**, IEEE, v. 6, n. 2, p. 89–103, 2012.

Calabrese, F. D.; Wang, L.; Ghadimi, E.; Peters, G.; Hanzo, L.; Soldati, P. Learning radio resource management in RANs: Framework, opportunities, and challenges. **IEEE Communications Magazine**, IEEE, v. 56, n. 9, p. 138–145, 2018.

Cavalcante, E. O.; Fodor, G.; Silva, Y. C. B.; Freitas, W. C. Distributed beamforming in dynamic TDD MIMO networks with BS to BS interference constraints. **IEEE Wireless Communications Letters**, IEEE, v. 7, n. 5, p. 788–791, 2018.

Cavalcante, E. O.; Fodor, G.; Silva, Y. C. B.; Freitas, W. C. Bidirectional sum-power minimization beamforming in dynamic TDD MIMO networks. **IEEE Transactions on Vehicular Technology**, IEEE, v. 68, n. 10, p. 9988–10002, 2019.

Chang, Z.; Wang, Z.; Guo, X.; Han, Z.; Ristaniemi, T. Energy-efficient resource allocation for wireless powered massive MIMO system with imperfect CSI. **IEEE Transactions on Green Communications and Networking**, v. 1, n. 2, p. 121–130, June 2017. ISSN 2473-2400.

Charnes, A.; Cooper, W. W. On nonlinear fractional programming. **Naval Research Logistics Quarterly**, Wiley-Blackwell, v. 9, n. 3–4, p. 181–186, 1962.

Chen, S.; Zhang, J.; Björnson, E.; Zhang, J.; Ai, B. Structured massive access for scalable cell-free massive MIMO systems. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 39, n. 4, p. 1086–1100, 2021.

Chien, T. V.; Björnson, E.; Larsson, E. G. Optimal design of energy-efficient cell-free massive MIMO: Joint power allocation and load balancing. In: **Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. Barcelona, Spain: IEEE. 04-08 May, 2020.

Christensen, S. S.; Agarwal, R.; Carvalho, E. D.; Cioffi, J. M. Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design. **IEEE Transactions on Wireless Communications**, IEEE, v. 7, n. 12, p. 4792–4799, 2008.

D'Andrea, C.; Rodriguez, A. G.; Geraci, G.; Giordano, L. G.; Buzzi, S. Cell-free massive MIMO for UAV communications. In: **Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops)**. Shanghai, China: IEEE. 20-24 May, 2019.

D'Andrea, C.; Rodriguez, A. G.; Geraci, G.; Giordano, L. G.; Buzzi, S. Analysis of UAV communications in cell-free massive MIMO systems. **IEEE Open Journal of the Communications Society**, IEEE, v. 1, p. 133–147, 2020.

Diamond, S.; Boyd, S. CVXPY: A Python-embedded modeling language for convex optimization. **Journal of Machine Learning Research**, v. 17, n. 83, p. 1–5, 2016.

Dinkelbach, W. On nonlinear fractional programming. **Management Science**, INFORMS, v. 133, n. 7, p. 492–498, 1967.

Dong, A.; Mao, R.; Yu, J.; Wang, Y.; Huang, B. Power control for SISO interference channel networks based on successive convex approximation. **EURASIP Journal on Wireless Communications and Networking**, SpringerOpen / Research Square, v. 2020, p. 1–15, 2020.

Elhoushy, S.; Ibrahim, M.; Hamouda, W. Cell-free massive MIMO: A survey. **IEEE Communications Surveys & Tutorials**, IEEE, v. 24, n. 1, p. 492–523, 2022.

Elwekeil, M.; Zappone, A.; Buzzi, S. Power control in cell-free massive MIMO networks for UAVs URLLC under the finite blocklength regime. **IEEE Transactions on Communications**, IEEE, v. 71, n. 2, p. 1126–1140, 2023.

Eraslan, E.; Daneshrad, B.; Lou, C.-Y. Performance indicator for MIMO MMSE receivers in the presence of channel estimation error. **IEEE Wireless Communications Letters**, IEEE, v. 2, n. 2, p. 211–214, 2013.

Ericsson. **Ericsson Mobility Report**. [S.l.], 2019. Disponível em: <https://www.ericsson.com/en/reports-and-papers/mobility-report>.

Esfandiari, M.; Vorobyov, S. A.; Karimi, M. New estimation methods for autoregressive process in the presence of white observation noise. **Signal Processing**, Elsevier, v. 171, p. 107480, 2020.

Feriani, A.; Hossain, E. Single and multi-agent deep reinforcement learning for AI-enabled wireless networks: A tutorial. **IEEE Communications Surveys & Tutorials**, IEEE, v. 23, n. 2, p. 1226–1252, 2021.

Fodor, G.; Fodor, S.; Telek, M. Performance analysis of a linear MMSE receiver in time-variant Rayleigh fading channels. **IEEE Transactions on Communications**, IEEE, v. 69, n. 6, p. 4098–4112, 2021.

Fodor, G.; Marco, P. D.; Telek, M. On minimizing the MSE in the presence of channel state information errors. **IEEE Communications Letters**, IEEE, v. 19, n. 9, p. 1604–1607, set. 2015.

Folkman, J.; Shapiro, N. On the continuity of the minimum set of a continuous function. **Journal of Mathematical Analysis and Applications**, Elsevier, v. 17, n. 3, p. 519–548, 1967.

Fotouhi, A.; Qiang, H.; Ding, M.; Hassan, M.; Giordano, L. G.; Garcia-Rodriguez, A.; Yuan, J. Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges. **IEEE Communications Surveys & Tutorials**, IEEE, v. 21, n. 4, p. 3417–3442, 2019.

Geraci, G.; Garcia-Rodriguez, A.; Galati Giordano, L.; López-Pérez, D.; Björnson, E. Understanding UAV cellular communications: From existing networks to massive MIMO. **IEEE Access**, IEEE, v. 6, p. 67853–67865, 2018.

Gogoi, P.; Sarma, A.; Borah, R.; Saikia, B. On the evolution of downlink physical layer in multi-antenna 3GPP LTE / LTE-A: A review. In: **Proceedings of the International Symposium on Advanced Computing and Communication (ISACC)**. Silchar, India: IEEE. 14-15 Sept., 2015.

Gong, S.; Wang, S.; Xing, C.; Ma, S.; Quek, T. Q. S. Robust superimposed training optimization for UAV assisted communication systems. **IEEE Transactions on Wireless Communications**, IEEE, v. 19, n. 3, p. 1704–1721, 2020.

Han, Z. **Game Theory in Wireless and Communication Networks: Theory, Models, and Applications**. Cambridge, U.K.: Cambridge University Press, 2012. ISBN 978-1107012083.

Hassan, B.; Baig, S.; Asif, M. Key technologies for ultra-reliable and low-latency communication in 6G. **IEEE Communications Standards Magazine**, IEEE, v. 5, n. 2, p. 106–113, 2021.

Hayat, S.; Yanmaz, E.; Muzaffar, R. Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint. **IEEE Communications Surveys & Tutorials**, IEEE, v. 18, n. 4, p. 2624–2661, 2016.

He, A.; Wang, L.; Chen, Y.; Wong, K.; Elkashlan, M. Spectral and energy efficiency of uplink D2D underlaid massive MIMO cellular networks. **IEEE Transactions on Communications**, IEEE, v. 65, n. 9, p. 3780–3793, 2017.

He, S.; Huang, Y.; Jin, S.; Yang, L. Coordinated beamforming for energy efficient transmission in multicell multiuser systems. **IEEE Transactions on Communications**, IEEE, v. 61, n. 12, p. 4961–4971, Dec. 2013.

Hijazi, H.; Ros, L. Joint data QR-detection and Kalman estimation for OFDM time-varying Rayleigh channel complex gains. **IEEE Transactions on Communications**, IEEE, v. 58, n. 1, p. 170–178, 2010.

Huang, R.; Guo, M.; Gu, C.; He, S.; Chen, J.; Sun, M. Toward scalable and efficient hierarchical deep reinforcement learning for 5G RAN slicing. **IEEE Transactions on Green Communications and Networking**, IEEE, v. 7, n. 4, p. 2153–2162, 2023.

Huq, K. M. S.; Mumtaz, S.; Bachmatiuk, J.; Rodriguez, J.; Wang, X.; Aguiar, R. L. Green HetNet CoMP: Energy efficiency analysis and optimization. **IEEE Transactions on Vehicular Technology**, IEEE, v. 64, n. 10, p. 4670–4683, 2015.

I, C.-L.; Rowell, C.; Han, S.; Xu, Z.; Li, G.; Pan, Z. Toward green and soft: A 5G perspective. **IEEE Communications Magazine**, IEEE, v. 52, n. 2, p. 66–73, 2014.

Iliev, T. B.; Ivanova, E. P.; Stoyanov, I. S.; Mihaylov, G. Y.; Beloev, I. H. Artificial intelligence in wireless communications - evolution towards 6G mobile networks. In: **Proceedings of the International Convention on Information, Communication and Electronic Technology**. Opatija, Croatia: MIPRO. 24-28 May, 2021.

Imoize, A. L.; Obakhena, H. I.; Anyasi, F. I.; Sur, S. N. A review of energy efficiency and power control schemes in ultra-dense cell-free massive MIMO systems for sustainable 6G wireless communication. **Sustainability**, MDPI, v. 14, n. 17, p. 11100, 2022.

Isheden, C.; Chong, Z.; Jorswieck, E.; Fettweis, G. Framework for link-level energy efficiency optimization with informed transmitter. **IEEE Transactions on Wireless Communications**, IEEE, v. 11, n. 8, p. 2946–2957, 2012.

Isheden, C.; Chong, Z.; Jorswieck, E.; Fettweis, G. Framework for link-level energy efficiency optimization with informed transmitter. **IEEE Transactions on Wireless Communications**, IEEE, v. 11, n. 8, p. 2946–2957, 2012.

Isheden, C.; Fettweis, G. P. Energy-efficient link adaptation on parallel channels. In: **Proceedings of the European Signal Processing Conference (EUSIPCO)**. Barcelona, Spain: IEEE. Aug. 29 to Sept. 2, 2011.

Jagannathan, R. On some properties of programming problems in parametric form pertaining to fractional programming. **Management Science**, INFORMS, v. 12, n. 7, p. 609–615, 1966.

Joshi, S. K.; Weeraddana, P. C.; Codreanu, M.; Latva-aho, M. Weighted sum-rate maximization for MISO downlink cellular networks via branch and bound. **IEEE Transactions on Signal Processing**, IEEE, v. 60, n. 4, p. 2090–2095, 2012.

Jungnickel, V.; Jaeckel, S.; Thiele, L.; Jiang, L.; Kruger, U.; Brylka, A.; von Helmolt, C. Capacity measurements in a cooperative MIMO network. **IEEE Transactions on Vehicular Technology**, IEEE, v. 58, n. 5, p. 2392–2405, 2009.

Kaleva, J.; Tölli, A.; Juntti, M. Decentralized sum rate maximization with QoS constraints for interfering broadcast channel via successive convex approximation. **IEEE Transactions on Signal Processing**, IEEE, v. 64, n. 11, p. 2788–2802, 2016.

Kashyap, S.; Mollen, C.; Björnson, E.; Larsson, E. G. Performance analysis of (TDD) massive MIMO with Kalman channel prediction. In: **Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. New Orleans, LA, USA: IEEE. 05-09 Mar., 2017.

Kingma, D. P.; Ba, J. **Adam: A Method for Stochastic Optimization**. [S.l.]: Cornell University Library, 2017. ArXiv preprint arXiv:1412.6980.

Kolawole, O. Y.; Biswas, S.; Singh, K.; Ratnarajah, T. Transceiver design for energy-efficiency maximization in mmWave MIMO IoT networks. **IEEE Transactions on Green Communications and Networking**, IEEE, v. 4, n. 1, p. 109–123, 2020.

Lakshminarayana, S.; Assaad, M.; Debbah, M. Energy efficient design in MIMO multi-cell systems with time average QoS constraints. In: **Proceedings of the IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)**. Darmstadt, Germany: IEEE. 16-19 Jun., 2013.

Larsson, E. G.; Edfors, O.; Tufvesson, F.; Marzetta, T. L. Massive MIMO for next generation wireless systems. **IEEE Communications Magazine**, IEEE, v. 52, n. 2, p. 186–195, 2014.

Lee, D.; Sun, Y. G.; Kim, S. H.; Kim, J.-H.; Shin, Y.; Kim, D. I.; Kim, J. Y. Multi-agent reinforcement learning-based resource allocation scheme for UAV-assisted internet of remote things systems. **IEEE Access**, IEEE, v. 11, p. 53155–53164, 2023.

Lee, S.; Yu, H.; Lee, H. Multiagent Q-learning-based multi-UAV wireless networks for maximizing energy efficiency: Deployment and power control strategy design. **IEEE Internet of Things Journal**, IEEE, v. 9, n. 9, p. 6434–6442, 2022.

Li, S.; Li, C.; Zhao, F.; Zhou, H. User group-based pilot allocation and data power optimization in cell-free massive MIMO for coexistence of UAVs and ground users. **Physical Communication**, Elsevier, v. 61, p. 102180, 2023.

Li, Y. **Deep Reinforcement Learning: An Overview**. [S.l.]: Cornell University Library, 2017. ArXiv preprint arXiv:1701.07274.

Li, Y.; Tian, Y.; Yang, C. Energy-efficient coordinated beamforming with individual data rate constraints. In: **Proceedings of the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications**. Hilton London Metropole, London, United Kingdom: IEEE. 8–11 Sept., 2013.

Lillicrap, T. P.; Hunt, J. J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous control with deep reinforcement learning. **arXiv preprint arXiv:1509.02971**, Cornell University Library, 2015.

Liu, F.; Zheng, K.; Xiang, W.; Zhao, H. Design and performance analysis of an energy-efficient uplink carrier aggregation scheme. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 32, n. 2, p. 197–207, 2014.

Liu, L.; Xiong, K.; Cao, J.; Lu, Y.; Fan, P.; Letaief, K. B. Average AoI minimization in UAV-assisted data collection with RF wireless power transfer: A deep reinforcement learning scheme. **IEEE Internet of Things Journal**, IEEE, v. 9, n. 7, p. 5216–5228, 2022.

Liu, X.; Lu, P. Solving nonconvex optimal control problems by convex optimization. **Journal of Guidance, Control, and Dynamics**, American Institute of Aeronautics and Astronautics, v. 37, n. 3, p. 750–765, 2014.

Lu, L.; Li, G. Y.; Swindlehurst, A. L.; Ashikhmin, A.; Zhang, R. An overview of massive MIMO: Benefits and challenges. **IEEE Journal of Selected Topics in Signal Processing**, IEEE, v. 8, n. 5, p. 742–758, 2014.

Luo, Z.-Q.; Yu, W. An introduction to convex optimization for communications and signal processing. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 24, n. 8, p. 1426–1438, 2006.

Luong, N. C.; Hoang, D. T.; Gong, S.; Niyato, D.; Wang, P.; Liang, Y.-C.; Kim, D. I. Applications of deep reinforcement learning in communications and networking: A survey. **IEEE Communications Surveys & Tutorials**, IEEE, v. 21, n. 4, p. 3133–3174, 2019.

MacKenzie, A.; Silva, L. **Game Theory For Wireless Engineers**. 1st. ed. 1537 Fourth Street, San Rafael, CA, USA: Morgan & Claypool Publishers, 2006. 1–86 p. (Synthesis Lectures on Communications, 1). ISBN 978-1598290165.

Mahapatra, R.; Nijsure, Y.; Kaddoum, G.; Ul Hassan, N.; Yuen, C. Energy efficiency tradeoff mechanism towards wireless green communication: A survey. **IEEE Communications Surveys & Tutorials**, IEEE, v. 18, n. 1, p. 686–705, 2016.

Marks, B. R.; Wright, G. P. Technical note — a general inner approximation algorithm for nonconvex mathematical programs. **Operations Research**, INFORMS, v. 26, n. 4, p. 681–683, 1978.

Marzetta, T. L. Massive MIMO: An introduction. **Bell Labs Technical Journal**, Nokia Bell Labs, v. 20, p. 11–22, 2015.

Miao, G. Energy-efficient uplink multi-user MIMO. **IEEE Transactions on Wireless Communications**, IEEE, v. 12, n. 5, p. 2302–2313, 2013.

Miao, G.; Himayat, N.; Li, G. Y.; Talwar, S. Low-complexity energy-efficient scheduling for uplink OFDMA. **IEEE Transactions on Communications**, IEEE, v. 60, n. 1, p. 112–120, 2012.

Misilmani, H. M. E.; El-Hajj, A. M. Massive MIMO design for 5G networks: An overview on alternative antenna configurations and channel model challenges. In: **Proceedings of the International Conference on High Performance Computing and Simulation (HPCS)**. Grand Hotel Savoia, Genoa, Italy: IEEE. 17–21 Jul., 2017.

Mkiramweni, M. E.; Yang, C.; Li, J.; Zhang, W. A survey of game theory in unmanned aerial vehicles communications. **IEEE Communications Surveys & Tutorials**, IEEE, v. 21, n. 4, p. 3386–3416, 2019.

Mohammadi, M.; Mobini, Z.; Ngo, H. Q.; Matthaiou, M. **Next Generation Multiple Access with Cell-Free Massive MIMO**. [S.l.]: https://arxiv.org/abs/2408.14598, 2023. TechRxiv preprint, available at https://arxiv.org/abs/2408.14598.

Myung, J.; Kim, K.; Ko, Y.-J. Transmit power control game with a sum power constraint in cell-free massive MIMO downlink. **IEEE Wireless Communications Letters**, IEEE, v. 11, n. 3, p. 632–635, 2022.

Nagib, A. M.; Abou-Zeid, H.; Hassanein, H. S. Safe and accelerated deep reinforcement learning-based O-RAN slicing: A hybrid transfer learning approach. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 42, n. 2, p. 310–325, 2024.

Ngo, H. Q.; Ashikhmin, A.; Yang, H.; Larsson, E. G.; Marzetta, T. L. Cell-free massive MIMO versus small cells. **IEEE Transactions on Wireless Communications**, IEEE, v. 16, n. 3, p. 1834–1850, 2017.

Ngo, H. Q.; Tran, L.; Duong, T. Q.; Matthaiou, M.; Larsson, E. G. On the total energy efficiency of cell-free massive MIMO. **IEEE Transactions on Green Communications and Networking**, IEEE, v. 2, n. 1, p. 25–39, 2018.

Nguyen, K. K.; Khosravirad, S. R.; da Costa, D. B.; Nguyen, L. D.; Duong, T. Q. Reconfigurable intelligent surface-assisted multi-UAV networks: Efficient resource allocation with deep reinforcement learning. **IEEE Journal of Selected Topics in Signal Processing**, IEEE, v. 16, n. 3, p. 358–368, 2022.

Oguejiofor, O.; Zhang, L. Global optimization of weighted sum-rate for downlink heterogeneous cellular networks. In: **Proceedings of the International Conference on Telecommunications (ICT)**. Makedonia Palace Hotel, Thessaloniki, Greece: IEEE. 16–18 May, 2016.

Osborne, M. An introduction to game theory. **Oxford University Press**, Oxford University Press, Oxford, United Kingdom, 2004.

Osborne, M. J.; Rubinstein, A. **A Course in Game Theory**. 1st. ed. Cambridge, Massachusetts, USA: MIT Press, 1994. ISBN 978-0262650403.

Palomar, D. P.; Mung Chiang. A tutorial on decomposition methods for network utility maximization. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 24, n. 8, p. 1439–1451, 2006.

Pan, Z.; Liu, J.; Shimamoto, S. Traffic-aware energy optimizing strategies for multi-cell coordinated green cellular networks. **IEEE Transactions on Green Communications and Networking**, IEEE, v. 2, n. 2, p. 418–431, 2018.

Papazafeiropoulos, A.; Bjornson, E.; Kourtessis, P.; Chatzinotas, S.; Senior, J. M. Scalable cell-free massive MIMO systems: Impact of hardware impairments. **IEEE Transactions on Vehicular Technology**, IEEE, v. 70, n. 10, p. 9701–9715, 2021.

Pengfei Xia; Shengli Zhou; Giannakis, G. B. Adaptive MIMO-OFDM based on partial channel state information. **IEEE Transactions on Signal Processing**, IEEE, v. 52, n. 1, p. 202–213, 2004.

Pennanen, H.; Tölli, A.; Kaleva, J.; Komulainen, P.; Latva-aho, M. Decentralized linear transceiver design and signaling strategies for sum power minimization in multi-cell MIMO systems. **IEEE Transactions on Signal Processing**, IEEE, v. 64, n. 7, p. 1729–1743, 2016.

Pessoa, A. M.; Guerreiro, I. M.; Silva, C. F. M. E.; Maciel, T. F.; Sousa, D. A.; Moreira, D. C.; Cavalcanti, F. R. P. A stochastic channel model with dual mobility for 5G massive networks. **IEEE Access**, IEEE, v. 7, p. 149971–149987, 2019.

Razaviyayn, M. **Successive Convex Approximation: Analysis and Applications**. Tese (Doutorado) — University of Minnesota, Minneapolis, Minnesota, USA, 2014.

Razaviyayn, M.; Baligh, M.; Callard, A.; Luo, Z. Joint user grouping and transceiver design in a MIMO interfering broadcast channel. **IEEE Transactions on Signal Processing**, IEEE, v. 62, n. 1, p. 85–94, 2014.

Razaviyayn, M.; Hong, M.; Luo, Z.-Q.; Pang, J.-S. Parallel successive convex approximation for nonsmooth nonconvex optimization. In: **Proceedings of the Advances in Neural Information Processing Systems 27 (NeurIPS)**. Montreal, Canada: MIT Press. 8–13 Dec., 2014.

Ródenas, R. G.; López, M. L.; Verastegui, D. Extensions of Dinkelbach's algorithm for solving non-linear fractional programming problems. **TOP: An Official Journal of the Spanish Society of Statistics and Operations Research**, Springer, v. 7, n. 1, p. 33–70, 1999.

Rusek, F.; Persson, D.; Lau, B. K.; Larsson, E. G.; Marzetta, T. L.; Edfors, O.; Tufvesson, F. Scaling up MIMO: Opportunities and challenges with very large arrays. **IEEE Signal Processing Magazine**, IEEE, v. 30, n. 1, p. 40–60, 2013.

Samir, M.; Ebrahimi, D.; Assi, C.; Sharafeddine, S.; Ghrayeb, A. Leveraging UAVs for coverage in cell-free vehicular networks: A deep reinforcement learning approach. **IEEE Transactions on Mobile Computing**, IEEE, v. 20, n. 9, p. 2835–2847, 2021.

Saraiva, J. V.; Antonioli, R.; Fodor, G.; Braga, I.; Freitas, W.; Silva, Y. A network-assisted game-theoretic design to power control in autoregressive fading channels. **IEEE Communications Letters**, IEEE, v. 26, n. 7, p. 1663–1667, abr. 2022.

Saraiva, J. V.; Antonioli, R. P.; Fodor, G.; Freitas, W. C.; Silva, Y. C. B. A distributed game-theoretic solution for power management in the uplink of cell-free systems. In: **IEEE Globecom Workshops (GC Wkshps)**. Rio de Janeiro, Brazil: IEEE. p. 1084–1089. 4–8 Dec., 2022.

Saraiva, J. V.; Braga Jr., I. M.; Monteiro, V. F.; Lima, F. R. M.; Maciel, T. F.; Freitas Jr., W. C.; Cavalcanti, F. R. P. Deep reinforcement learning for QoS-constrained resource allocation in multiservice networks. **Journal of Communication and Information Systems**, SBrT, v. 35, n. 1, p. 66–76, 2020.

Saraiva, J. V.; Lima, F. R. M.; Maciel, T. F.; Cavalcanti, F. R. P. Relay selection, subcarrier pairing and power allocation for energy efficiency and QoS guarantees. In: **Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)**. Barcelona, Spain: IEEE. 15–18 Apr., 2018.

Saxena, N.; Sahu, B. J. R.; Han, Y. S. Traffic-aware energy optimization in green LTE cellular systems. **IEEE Communications Letters**, IEEE, v. 18, n. 1, p. 38–41, 2014.

Schaible, S. Parameter-free convex equivalent and dual programs of fractional programming problems. **Zeitschrift für Operations Research**, Springer, v. 18, n. 5, p. 187–196, 1974.

Schulman, J.; Levine, S.; Moritz, P.; Jordan, M. I.; Abbeel, P. **Trust Region Policy Optimization**. 2017. <https://arxiv.org/abs/1502.05477>.

Schulman, J.; Moritz, P.; Levine, S.; Jordan, M.; Abbeel, P. **High-Dimensional Continuous Control Using Generalized Advantage Estimation**. 2015. <https://arxiv.org/abs/1506.02438>.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. **Proximal Policy Optimization Algorithms**. 2017. <https://arxiv.org/abs/1707.06347>.

Scutari, G.; Barbarossa, S.; Palomar, D. Potential games: A framework for vector power control problems with coupled constraints. In: **Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)**. Toulouse, France: IEEE. 14–19 May, 2006.

Scutari, G.; Palomar, D. P.; Facchinei, F.; Pang, J.-S. Convex optimization, game theory, and variational inequality theory. **IEEE Signal Processing Magazine**, IEEE, v. 27, n. 3, p. 35–49, 2010.

Seong, K.; Narasimhan, R.; Cioffi, J. M. Queue proportional scheduling via geometric programming in fading broadcast channels. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 24, n. 8, p. 1593–1602, 2006.

Shen, K.; Yu, W. Fractional programming for communication systems—part I: Power control and beamforming. **IEEE Transactions on Signal Processing**, IEEE, v. 66, n. 10, p. 2616–2630, 2018.

Sheng, Z.; Tuan, H. D.; Nasir, A. A.; Duong, T. Q.; Poor, H. V. Power allocation for energy efficiency and secrecy of wireless interference networks. **IEEE Transactions on Wireless Communications**, IEEE, v. 17, n. 6, p. 3737–3751, 2018.

Shi, Q.; Razaviyayn, M.; Luo, Z.; He, C. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel. **IEEE Transactions on Signal Processing**, IEEE, v. 59, n. 9, p. 4331–4340, 2011.

Shi, Y.; Hamdan, M. Q.; Alsusa, E.; Hamdi, K. A.; Baidas, M. W. A decoupled access scheme with reinforcement learning power control for cellular-enabled UAVs. **IEEE Internet of Things Journal**, IEEE, v. 8, n. 24, p. 17261–17274, 2021.

Silvagni, M.; Tonoli, A.; Zenerino, E.; Chiaberge, M. Multipurpose UAV for search and rescue operations in mountain avalanche events. **Geomatics, Natural Hazards and Risk**, Taylor & Francis, v. 8, n. 1, p. 18–33, 2017.

Spangenburg, R.; Moser, D. **Carl Sagan: A Biography**. Westport, CT, USA: Greenwood Press, 2004.

Sutton, R. S.; Barto, A. G. **Reinforcement Learning: An Introduction**. 1st. ed. Cambridge, MA, USA: MIT Press, 1998.

Tentu, V.; Sharma, E.; Amudala, D. N.; Budhiraja, R. UAV-enabled hardware-impaired spatially correlated cell-free massive MIMO systems: Analysis and energy efficiency optimization. **IEEE Transactions on Communications**, IEEE, v. 70, n. 4, p. 2722–2741, 2022.

Tölli, A.; Ghauch, H.; Kaleva, J.; Komulainen, P.; Bengtsson, M.; Skoglund, M.; Honig, M.; Lahetkangas, E.; Tiirola, E.; Pajukoski, K. Distributed coordinated transmission with forward-backward training for 5G radio access. **IEEE Communications Magazine**, IEEE, v. 57, n. 1, p. 58–64, 2019.

Tsouros, D. C.; Bibi, S.; Sarigiannidis, P. G. A review on UAV-based applications for precision agriculture. **Information**, MDPI, v. 10, n. 11, 2019.

Venkatraman, G.; Tölli, A.; Juntti, M.; Tran, L. Traffic aware resource allocation schemes for multi-cell MIMO-OFDM systems. **IEEE Transactions on Signal Processing**, IEEE, v. 64, n. 11, p. 2730–2745, 2016.

Venturino, L.; Zappone, A.; Risi, C.; Buzzi, S. Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination. **IEEE Transactions on Wireless Communications**, IEEE, v. 14, n. 1, p. 1–14, 2015.

Wang, S.; Jiang, F.; Zhang, B.; Ma, R.; Hao, Q. Development of UAV-based target tracking and recognition systems. **IEEE Transactions on Intelligent Transportation Systems**, IEEE, v. 21, n. 8, p. 3409–3422, 2020.

Weeraddana, P. C.; Codreanu, M.; Ephremides, M. Latva-aho and A. Weighted sum-rate maximization for a set of interfering links via branch and bound. **IEEE Transactions on Signal Processing**, IEEE, v. 59, n. 8, p. 3977–3996, 2011.

Wright, S. J. Coordinate descent algorithms. **Mathematical Programming**, Springer, v. 151, n. 1, p. 3–34, 2015.

Xie, X.; Yang, H.; Vasilakos, A. V.; He, L. Fair power control using game theory with pricing scheme in cognitive radio networks. **Journal of Communications and Networks**, IEEE, v. 16, n. 2, p. 183–192, 2014.

Yang, Y.; Pesavento, M.; Chatzinotas, S.; Ottersten, B. Energy efficiency optimization in MIMO interference channels: A successive pseudoconvex approximation approach. **IEEE Transactions on Signal Processing**, IEEE, v. 67, n. 15, p. 4107–4121, 2019.

Ye, X.; Zhou, Q.; Fu, L. Deep reinforcement learning-based scheduling for NR-U/WiGig coexistence in unlicensed mmWave bands. **IEEE Transactions on Wireless Communications**, IEEE, v. 23, n. 1, p. 58–73, 2024.

Zangwill, W. I. Convergence conditions for nonlinear programming algorithms. **Management Science**, INFORMS, v. 16, n. 1, p. 1–13, 1969.

Zappone, A.; Jorswieck, E. *et al.* Energy efficiency in wireless networks via fractional programming theory. **Foundations and Trends® in Communications and Information Theory**, Now Publishers, Inc., v. 11, n. 3-4, p. 185–396, 2015.

Zappone, A.; Jorswieck, E. A. Energy-efficient resource allocation in future wireless networks by sequential fractional programming. **Digital Signal Processing**, Elsevier, v. 60, p. 324–337, 2017.

Zeng, Y.; Zhang, R. Energy-efficient UAV communication with trajectory optimization. **IEEE Transactions on Wireless Communications**, IEEE, v. 16, n. 6, p. 3747–3760, 2017.

Zeng, Y.; Zhang, R.; Lim, T. J. Throughput maximization for UAV-enabled mobile relaying systems. **IEEE Transactions on Communications**, IEEE, v. 64, n. 12, p. 4983–4996, 2016.

Zhang, C.; Li, Z.; He, C.; Wang, K.; Pan, C. Deep reinforcement learning based trajectory design and resource allocation for UAV-assisted communications. **IEEE Communications Letters**, v. 27, n. 9, p. 2398–2402, 2023.

Zhang, S.; Zhang, H.; Di, B.; Song, L. Cellular UAV-to-X communications: Design and optimization for multi-UAV networks. **IEEE Transactions on Wireless Communications**, v. 18, n. 2, p. 1346–1359, 2019.

Zhang, Y. J.; Letaief, K. B. Optimizing power and resource management for multiuser MIMO/OFDM systems. In: **Proceedings of the IEEE Global Communications Conference (GLOBECOM)**. San Francisco, CA, USA: IEEE. v. 1. 1–5 Dec., 2003.

Zhao, P.; Fodor, G.; Dán, G.; Telek, M. A game theoretic approach to setting the pilot power ratio in multi-user MIMO systems. **IEEE Transactions on Communications**, IEEE, v. 66, n. 3, p. 999–1012, 2018.

Zhao, P.; Fodor, G.; Dán, G.; Telek, M. A game theoretic approach to uplink pilot and data power control in multi-cell MU MIMO systems. **IEEE Transactions on Vehicular Technology**, IEEE, v. 68, n. 9, p. 8707–8720, 2019.

Zhao, Y.; Niemegeers, I. G.; De Groot, S. H. Power allocation in cell-free massive MIMO: A deep learning method. **IEEE Access**, IEEE, v. 8, p. 87185–87200, 2020.

Zheng, J.; Zhang, J.; Du, H.; Niyato, D.; Ai, B.; Debbah, M.; Letaief, K. B. Mobile cell-free massive MIMO: Challenges, solutions, and future directions. **IEEE Wireless Communications**, IEEE, p. 1–8, 2024.

Zheng, K.; Zhao, L.; Mei, J.; Shao, B.; Xiang, W.; Hanzo, L. Survey of large-scale MIMO systems. **IEEE Communications Surveys & Tutorials**, IEEE, v. 17, n. 3, p. 1738–1760, 2015.

Zhong, R.; Liu, X.; Liu, Y.; Chen, Y. Multi-agent reinforcement learning in NOMA-aided UAV networks for cellular offloading. **IEEE Transactions on Wireless Communications**, IEEE, v. 21, n. 3, p. 1498–1512, 2022.

# APPENDIX A – CONVERGENCE ANALYSIS OF CENTRALIZED AND DECENTRALIZED ALGORITHMS

In this appendix, we present the convergence analysis for Algorithms 1 and 2 described in Sections 2.4.2 and 2.5.1, respectively. Our convergence analysis is split into two parts: we first analyze the convergence of the *outer loops* and then the convergence of the *inner loops* of Algorithms 1 and 2.

To prove the convergence of the *outer loops* of Algorithms 1 and 2, we need the following result stated in Proposition 1. We remark that the convergence proof for the *outer loop* of both Algorithms 1 and 2 is exactly equal as both *outer loops* are concerned with the convergence of the $\eta$ variable. Therefore, we only present such a proof once.

**Proposition 1.** *$F(\eta)$ defined in* (2.12) *is strictly monotonic decreasing with respect to $\eta$. Besides, for any feasible $\tilde{\eta}$ value, we have $F(\tilde{\eta}) \geq 0$.*

*Proof.* The proof of Proposition 1 can be found in (Dinkelbach, 1967). Thus, it is not reproduced here. □

To complete the convergence analysis of the *outer loops* of Algorithms 1 and 2, consider the value of $\eta$ in the $l$-th iteration, i.e., $\eta^{(l)}$. By Proposition 1, $F(\eta^{(l)}) \geq 0$ but according to Theorem 1, $F(\eta^{(l)}) = 0$ if and only if $\eta^{(l)}$ is the optimal solution of problem (2.9). Since $\eta^{(l)}$ is not necessarily the optimal solution of (2.9), we have $F(\eta^{(l)}) > 0$. Thus, if $\mathbf{x}^{(l)}$ is the global solution of the parametric subproblem (2.13) for $\eta^{(l)}$, then $\eta^{(l+1)} = \frac{\phi(\mathbf{x}^{(l)})}{\psi(\mathbf{x}^{(l)})}$, and $F(\eta^{(l)}) = \phi(\mathbf{x}^{(l)}) - \eta^{(l)}\psi(\mathbf{x}^{(l)}) = \eta^{(l+1)}\psi(\mathbf{x}^{(l)}) - \eta^{(l)}\psi(\mathbf{x}^{(l)}) = \psi(\mathbf{x}^{(l)})\left[\eta^{(l+1)} - \eta^{(l)}\right]$. Therefore, $\eta^{(l+1)} > \eta^{(l)}$ since necessarily $\psi(\mathbf{x}^{(l)}) > 0$. Consequently, given that $\eta$ is always increasing and $F(\cdot)$ is strictly monotonic decreasing with respect to $\eta$, we can see that as long as the number of iterations is large enough, we have $F(\eta^{\star}) = 0$, i.e., the *outer loops* of Algorithms 1 and 2 converge.

In the second part of our convergence analysis, we need to show that the sequences generated in the *inner loops* of Algorithms 1 and 2 converge to a unique solution and that this solution is a Karush-Kuhn-Tucker (KKT) point of problem (2.13). The second part of our convergence analysis is based on (Kaleva *et al.*, 2016) and extended to the proposed algorithms.

**Remark.** *Algorithms 1 and 2 can perform a sufficient number of iterations of the primal/dual decomposition and successive convex approximation (SCA) updates in the inner loops of Algo-*

*rithms 1 and 2. This assumption is needed to guarantee the monotonicity of the global objective function.*

We highlight that considering Remark A, the convergence proof for the *inner loop* of both Algorithms 1 and 2 becomes exactly equal. Thus, we only present the convergence proof for the *inner loop* once. To begin with the proof of the *inner loops* of Algorithms 1 and 2, we start by analyzing the feasible set of problem (2.13). It is demonstrated in (Kaleva *et al.*, 2016; Venkatraman *et al.*, 2016) that the feasible set of (2.13) is compact. This allows us to state that the update iterations of the variables $\{\mathbf{m}_{u,s,n}\}_{\forall(u,s,n)}$, $\{\mathbf{w}_{u,s,n}\}_{\forall(u,s,n)}$, $\{\epsilon_{u,s,n}\}_{\forall(u,s,n)}$ and $\{t_{u,s,n}\}_{\forall(u,s,n)}$ in the *inner loops* of Algorithms 1 and 2 can be represented using infimal maps (Kaleva *et al.*, 2016). In fact, since all optimization variables belong to compact regions, the infimal maps modeling the updates of the optimization variables are closed point-to-set maps (Folkman; Shapiro, 1967).

Next, we prove two propositions related to the convergence of the *inner loops* of Algorithms 1 and 2. Consequently, these propositions address the parametric subproblem (2.13), which is used to solve the problem (2.9) according to Theorem 1. As it will be seen later, this is important to complete our convergence analysis.

**Proposition 2.** *For a given fixed $\eta$ value, the objective function of parametric subproblem* (2.13) *is monotonic and converges with the inner loop of Algorithm 1.*

*Proof.* For a given $\eta$ value in the *inner loop* of Algorithm 1, the objective of the parametric subproblem (2.13) is strictly monotonic increasing. To show that, first, note that due to the power constraints, the objective function of (2.13) is clearly bounded. Moreover, note that before the SCA iterations, the transmit beamformers and $\eta$ variable are fixed, i.e., only the MMSE receivers are updated in the *inner loop*. It is well-known that the MMSE receive beamformers maximize the per-stream SINRs (Kaleva *et al.*, 2016; Pennanen *et al.*, 2016), i.e., they maximize the rate for each user. As a result, we necessarily have an increase in the rate and consequently an increase in the objective function of the parametric subproblem (2.13). During the update of the transmit beamformers, when the receive beamformers and $\eta$ variable are fixed, the monotonicity is ensured by the fact that an SCA subproblem is either the solution to the original problem or monotonically improve its objective, as shown in (Marks; Wright, 1978).

We can extend the monotonicity of the *inner loop* of Algorithm 1 to the *inner loop* of Algorithm 2 considering the assumption of a sufficient number of iterations of the primal/dual

decomposition and SCA updates. Therefore, since the objective function of (2.13) is bounded and monotonically increasing, we have that it necessarily converges with the *inner loops* of Algorithms 1 and 2. □

Given that the *inner loops* in Algorithms 1 and 2 can be modeled as closed infimal maps and are monotonic with respect to the objective of the parametric subproblem (2.13) for a fixed $\eta$ value, we can conclude that, according to the general convergence theorem presented in (Zangwill, 1969), the sequence of iterates produced by the *inner loops* of Algorithms 1 and 2 has at least one accumulation point and each accumulation point is a generalized fixed point, if the sequence of beamformers converges. However, we can make even stronger considerations and show that the *inner loops* of Algorithms 1 and 2 converge to a unique solution for all fixed points.

For the centralized algorithm, at any fixed point, the optimal transmit beamformers and the variables $\{\epsilon_{u,s,n}, t_{u,s,n}\}_{\forall(u,s,n)}$ are uniquely defined when the parametric subproblem (2.13) is solved via optimization. With respect to the decentralized algorithm, at any fixed point, the transmit beamformer in (2.18a) can be computed using a uniquely defined generalized inverse operation, such as the Moore-Penrose pseudoinverse (Kaleva *et al.*, 2016). Nevertheless, note that many viable fixed points are possible, i.e., a single fixed point cannot be guaranteed, since there is an SINR equivalence for different complex beamformers with some different phase rotation.

**Proposition 3.** *For a fixed $\eta$ value, any fixed point $\{\mathbf{w}^{\star}_{u,s,n}, \mathbf{m}^{\star}_{u,s,n}, t^{\star}_{u,s,n}, \epsilon^{\star}_{u,s,n}\}_{\forall(u,s,n)}$ of the inner loop of Algorithm 1 is a KKT point of problem* (2.13).

*Proof.* Based on the theorem shown in (Marks; Wright, 1978), we should show that the SCA algorithm in the *inner loop* of Algorithm 1 either stops at a KKT point or the limit of any convergent sequence is a KKT point with a slight difference due to the extra step involving the receive beamformer updates. To show this, firstly, note that the primal and dual constraints always hold for problem (2.9) since the convex approximation is only applied for constraints (2.14). Thus, only the MSE constraints still need to be analyzed.

We can approximate the constraints (2.9b) in each iteration. Let $g(\epsilon_{u,s,n}, \tilde{\epsilon}_{u,s,n})$ be the first-order Taylor approximation used to approximate $\log_2(\epsilon_{u,s,n})$ around a fixed MSE point in (2.14). Thus:

$$g(\epsilon_{u,s,n}, \tilde{\epsilon}_{u,s,n}) = \log_2(\tilde{\epsilon}_{u,s,n}) + \frac{\epsilon_{u,s,n} - \tilde{\epsilon}_{u,s,n}}{\tilde{\epsilon}_{u,s,n}\log(2)} + t_{u,s,n}, \tag{A.1}$$

where we can clearly notice that $g(\epsilon_{u,s,n}, \tilde{\epsilon}_{u,s,n})$ is a differentiable convex function and, furthermore, from the convergence to a fixed point it satisfies the following properties:

1. $g(\epsilon_{u,s,n}, \tilde{\epsilon}_{u,s,n}) \geq \log_2(\epsilon_{u,s,n})$
2. $g(\epsilon^*_{u,s,n}, \epsilon^*_{u,s,n}) = \log_2(\epsilon^*_{u,s,n})$
3. $\frac{\partial\, g(\epsilon^*_{u,s,n}, \epsilon^*_{u,s,n})}{\partial \epsilon_{u,s,n}} = \frac{\partial \log_2(\epsilon^*_{u,s,n})}{\partial \epsilon_{u,s,n}}$.

By (2.) and (3.) and according to the Theorem 1 presented in (Marks; Wright, 1978), the limit of any convergent sequence generated by the *inner loop* of Algorithm 1 is a KKT point of problem (2.9). This conclusion is also valid for the decentralized Algorithm 2 considering the assumption of a sufficient number of primal/dual iterations. □

Finally, our convergence analysis needs to show that the *outer loop* of Algorithm 1 stops in a KKT point of problem (2.9). For this, let $\eta^\star$ be the value of $\eta$ when the *outer loop* of Algorithm 1 converges. Thus, for $\eta^\star$ and as a consequence of Propositions 2 and 3, the solution $\{\mathbf{w}^\star_{u,s,n}, \mathbf{m}^\star_{u,s,n}, t^\star_{u,s,n}, \epsilon^\star_{u,s,n}\}_{\forall(u,s,n)}$ obtained in the *inner loop* of Algorithm 1 is a fixed point and also a KKT point of parametric subproblem (2.13). However, by Theorem 1, for $\eta^\star$, the optimality conditions to the parametric subproblem (2.13) and problem (2.9) are the same and, therefore, $\{\mathbf{w}^\star_{u,s,n}, \mathbf{m}^\star_{u,s,n}, t^\star_{u,s,n}, \epsilon^\star_{u,s,n}\}_{\forall(u,s,n)}$ is also a fixed point and a KKT point of problem (2.9). The last step involves the equivalence between a KKT point of problem (2.9) and (2.5), which follows from (Shi *et al.*, 2011).

# APPENDIX B – LAGRANGIAN FUNCTION AND KKT CONDITIONS FOR THE OPTIMIZATION PROBLEM

In the following, we present the Lagrangian function and KKT conditions of problem (2.17). First, the Lagrangian of (2.17) is given by

$$\begin{aligned}\mathcal{L}(\mathbf{w}_{u,s,n}, \mathbf{m}_{u,s,n}, t_{u,s,n}, \epsilon_{u,s,n}) = &- \sum_{u=1}^{U} \gamma_u \left( \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} \right) + \\ \eta \left( \sum_{n=1}^{N} \sum_{u \in \mathcal{U}} \sum_{s=1}^{S_{u,n}} \|\mathbf{m}_{u,s,n}\|^2 + \zeta \right) &+ \sum_{u=1}^{U} \alpha_u \left( - \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} + \xi_u \right) \\ &+ \sum_{u=1}^{U} \beta_u \left( \sum_{n=1}^{N} \sum_{s=1}^{S_{u,n}} t_{u,s,n} - \frac{Q_u}{\Delta_{\text{tti}}} \right) \\ + \sum_{n=1}^{N} \sum_{u=1}^{U} \sum_{s=1}^{S_{u,n}} &\left[ \mu_{u,s,n} \left( \log_2 \left( \tilde{\epsilon}_{u,s,n} \right) + \frac{\epsilon_{u,s,n} - \tilde{\epsilon}_{u,s,n}}{\tilde{\epsilon}_{u,s,n} \log(2)} + t_{u,s,n} \right) \right] \\ &+ \sum_{\forall b} \delta_b \left( \sum_{n=1}^{N} \sum_{u \in \mathcal{U}_b} \sum_{s=1}^{S_u} \|\mathbf{m}_{u,s,n}\|^2 - P_b \right) \\ + \sum_{n=1}^{N} \sum_{u=1}^{U} \sum_{s=1}^{S_{u,n}} &\left[ \lambda_{u,s,n} \left( |1 - \mathbf{w}_{u,s,n}^{\mathrm{H}} \mathbf{H}_{b_u,u,n} \mathbf{m}_{u,s,n}|^2 + \right. \right. \\ \left. \left. \sum_{i=1}^{U} \sum_{\substack{j=1, \\ (i,j) \neq (u,s)}}^{S_u} |\mathbf{w}_{u,s,n}^{\mathrm{H}} \mathbf{H}_{b_i,u,n} \mathbf{m}_{i,j,n}|^2 + \sigma_u^2 \|\mathbf{w}_{u,s,n}\|^2 - \epsilon_{u,s,n} \right) \right], \end{aligned} \tag{B.1}$$

in which $\{\delta_b\}_{\forall b}$ and $\{\lambda_{u,s,n}, \mu_{u,s,n}\}_{\forall (u,s,n)}$ are the dual variables corresponding to the constraints defined in (2.5c), (2.9c) and (2.14) in problem (2.17), respectively. In addition to the primal feasibility constraints and complementary slackness conditions, the KKT conditions are:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{m}_{u,s,n}} = (\eta + \delta_b) \mathbf{m}_{u,s,n} + \sum_{i=1}^{U} \sum_{j=1}^{S_{i,n}} \lambda_{i,j,n} \mathbf{H}_{b_u,j,n}^{\mathrm{H}} \mathbf{w}_{i,j,n} \mathbf{w}_{i,j,n}^{\mathrm{H}} \mathbf{H}_{b_u,j,n} \mathbf{m}_{u,s,n} - \lambda_{u,s,n} \mathbf{H}_{b_u,u,n}^{\mathrm{H}} \mathbf{w}_{u,s,n} = 0, \tag{B.2a}$$

$$\frac{\partial \mathcal{L}}{\partial t_{u,s,n}} = -\gamma_u - \alpha_u + \beta_u + \mu_{u,s,n} = 0, \tag{B.2b}$$

$$\frac{\partial \mathcal{L}}{\partial \epsilon_{u,s,n}} = \frac{\mu_{u,s,n}}{\tilde{\epsilon}_{u,s,n} \log_{(2)}} - \lambda_{u,s,n} = 0, \tag{B.2c}$$

$$\lambda_{u,s,n}, \mu_{u,s,n}, \delta_u \geq 0. \tag{B.2d}$$

# APPENDIX C – PROOF OF CONVEXITY OF PAYOFF FUNCTION

We can prove the convexity of the payoff function of a generic UE $\tilde{k}$, $\mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})$, by noticing that the second derivate of it is larger than zero over $\rho_{\tilde{k}} \in \mathcal{P}_k$. First, for $\alpha$ and $\boldsymbol{\rho}_{(-\tilde{k})}$ fixed, we have:

$$\frac{\partial^2 \mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})}{\partial \rho_{\tilde{k}}^2} = \frac{\partial^2 \gamma_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})}{\partial \rho_{\tilde{k}}^2} + \frac{\partial^2 \lambda_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})}{\partial \rho_{\tilde{k}}^2}. \tag{C.1}$$

In addition, by the definition of $\{\gamma_k(\cdot), \lambda_k(\cdot)\}_{\forall k}$ in (4.8), it is easy to see that:

$$\frac{\partial \lambda_k(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-k)})}{\partial \rho_k} = \Big( \sum_{l \in \mathcal{M}_k} \beta_{k,l} \Big)^{\alpha} \sum_{\forall i \neq k} \frac{1}{\rho_i \Big( \sum\limits_{l \in \mathcal{M}_i} \beta_{i,l} \Big)^{\alpha}}, \tag{C.2a}$$

$$\frac{\partial \gamma_k(\alpha, \rho_k, \boldsymbol{\rho}_{(k)})}{\partial \rho_k} = -\frac{\sum\limits_{\forall i \neq k} \rho_i \Big( \sum\limits_{l \in \mathcal{M}_i} \beta_{i,l} \Big)^{\alpha}}{\rho_k^2 \Big( \sum\limits_{l \in \mathcal{M}_k} \beta_{k,l} \Big)^{\alpha}}. \tag{C.2b}$$

Therefore:

$$\frac{\partial^2 \mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})}{\partial \rho_{\tilde{k}}^2} = \frac{\partial^2 \gamma_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})}{\partial \rho_{\tilde{k}}^2} = 2 \frac{\sum\limits_{\forall i \neq k} \rho_i \Big( \sum\limits_{l \in \mathcal{M}_i} \beta_{i,l} \Big)^{\alpha}}{\rho_k^3 \Big( \sum\limits_{l \in \mathcal{M}_k} \beta_{k,l} \Big)^{\alpha}}. \tag{C.3}$$

Given that $\{\rho_k, \beta_{k,l}\}_{\forall k,l} > 0$, then $\frac{\partial^2 \mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})}{\partial \rho_{\tilde{k}}^2} > 0$. Therefore, the payoff function $\mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})$ is convex over $\rho_{\tilde{k}} \in \mathcal{P}_{\tilde{k}}$ and, consequently, there exists a unique $\rho_{\tilde{k}}^{\star} \in \mathcal{P}_{\tilde{k}}$ that minimizes $\mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})$. Note that we can obtain $\rho_{\tilde{k}}^{\star}$ by solving $\frac{\partial \mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})}{\partial \rho_{\tilde{k}}} = 0$.

## APPENDIX D – PROOF THAT THE GAME $\mathcal{G}$ IS A POTENTIAL GAME

Let $\rho_{\tilde{k}}$, $\rho'_{\tilde{k}} \in \mathcal{P}_{\tilde{k}}$ be two different and arbitrary data power values for a generic UE $\tilde{k}$. Suppose UE $\tilde{k}$ changes its data power from $\rho_{\tilde{k}}$ to $\rho'_{\tilde{k}}$, then the change of its payoff function is:

$$\Delta\mu_{\tilde{k}} = \mu_{\tilde{k}}(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) - \mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}). \tag{D.1}$$

Moreover, the change of the function $u(\alpha, \boldsymbol{\rho})$ is:

$$\Delta u = u(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) - u(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}). \tag{D.2}$$

However, if $u(\alpha, \boldsymbol{\rho})$ in (4.12) is a potential function of our parameterized game $\mathcal{G}(\alpha)$, then $\Delta u = \Delta\mu_{\tilde{k}}$ according to Definition 3. To prove it, we can rewrite $\Delta u$ as follows:

$$\Delta u = \frac{1}{2}\left(\sum_{k\in\mathcal{K}} \mu_k(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) - \sum_{k\in\mathcal{K}} \mu_k(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})})\right). \tag{D.3}$$

Moreover, note that the terms present in (D.3) can also be rewritten separately as shown below

$$\begin{aligned}\sum_{k\in\mathcal{K}} \mu_k(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) &= \mu_1(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) + \cdots \\ &\quad + \mu_{\tilde{k}}(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) + \cdots + \mu_K(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) \\ &= \sum_{k\in\mathcal{K}} \gamma_k(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) + \sum_{k\in\mathcal{K}} \lambda_k(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}).\end{aligned} \tag{D.4a}$$

$$\begin{aligned}\sum_{k\in\mathcal{K}} \mu_k(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) &= \mu_1(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) + \cdots \\ &\quad + \mu_{\tilde{k}}(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) + \cdots + \mu_K(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) \\ &= \sum_{k\in\mathcal{K}} \gamma_k(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) + \sum_{k\in\mathcal{K}} \lambda_k(\alpha, \rho_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}).\end{aligned} \tag{D.4b}$$

Now, we separately highlight the terms of (D.4a), as follows:

$$\begin{aligned}\sum_{k\in\mathcal{K}} \gamma_k(\alpha, \rho'_{\tilde{k}}, \boldsymbol{\rho}_{(-\tilde{k})}) &= \left[\frac{\rho_2\Big(\sum\limits_{l\in\mathcal{M}_2}\beta_{2,l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}} + \cdots + \frac{\rho'_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}} + \cdots + \frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}\right] \\ &+ \cdots + \left[\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}} + \cdots + \frac{\rho'_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}} + \cdots + \frac{\rho_{K-1}\Big(\sum\limits_{l\in\mathcal{M}_{K-1}}\beta_{K-1,l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}\right]\end{aligned} \tag{D.5a}$$

$$\sum_{k\in\mathcal{K}}\lambda_k(\alpha,\rho'_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})=\left[\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho_2\Big(\sum\limits_{l\in\mathcal{M}_2}\beta_{2,l}\Big)^{\alpha}}+\cdots+\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho'_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}+\cdots+\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}\right]$$
$$+\cdots+\left[\frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}+\cdots+\frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho'_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}+\cdots+\frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho_{K-1}\Big(\sum\limits_{l\in\mathcal{M}_{K-1}}\beta_{K-1,l}\Big)^{\alpha}}\right]. \tag{D.5b}$$

Similarly for the terms of (D.4b), $\sum_{k\in\mathcal{K}}\gamma_k(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})$ and $\sum_{k\in\mathcal{K}}\lambda_k(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})$, we have:

$$\sum_{k\in\mathcal{K}}\gamma_k(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})=\left[\frac{\rho_2\Big(\sum\limits_{l\in\mathcal{M}_2}\beta_{2,l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}+\cdots+\frac{\rho_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}+\cdots+\frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}\right]$$
$$+\cdots+\left[\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}+\cdots+\frac{\rho_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}+\cdots+\frac{\rho_{K-1}\Big(\sum\limits_{l\in\mathcal{M}_{K-1}}\beta_{K-1,l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}\right] \tag{D.6a}$$

$$\sum_{k\in\mathcal{K}}\lambda_k(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})=\left[\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho_2\Big(\sum\limits_{l\in\mathcal{M}_2}\beta_{2,l}\Big)^{\alpha}}+\cdots+\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}+\cdots+\frac{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}\right]$$
$$+\cdots+\left[\frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho_1\Big(\sum\limits_{l\in\mathcal{M}_1}\beta_{1,l}\Big)^{\alpha}}+\cdots+\frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}+\cdots+\frac{\rho_K\Big(\sum\limits_{l\in\mathcal{M}_K}\beta_{K,l}\Big)^{\alpha}}{\rho_{K-1}\Big(\sum\limits_{l\in\mathcal{M}_{K-1}}\beta_{K-1,l}\Big)^{\alpha}}\right]. \tag{D.6b}$$

Given that the change of data power occurs specifically for UE $\tilde{k}$, the functions $\sum_{k\in\mathcal{K}}\mu_k(\alpha,\rho'_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})$ and $\sum_{k\in\mathcal{K}}\mu_k(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})$ have several terms in common and thereby

when we compute the difference between them and rearrange the terms, we have:

$$\begin{aligned}\sum_{k\in\mathcal{K}}\mu_k(\alpha,\rho'_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})-\sum_{k\in\mathcal{K}}\mu_k(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})}) &= 2\left[\frac{\sum\limits_{\forall i\neq\tilde{k}}\rho_i\Big(\sum\limits_{l\in\mathcal{M}_i}\beta_{i,l}\Big)^{\alpha}}{\rho'_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}+\sum_{\forall i\neq\tilde{k}}\frac{\rho'_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}{\rho_i\Big(\sum\limits_{l\in\mathcal{M}_i}\beta_{i,l}\Big)^{\alpha}}\right]\\ &\quad-2\left[\frac{\sum\limits_{\forall i\neq\tilde{k}}\rho_i\Big(\sum\limits_{l\in\mathcal{M}_i}\beta_{i,l}\Big)^{\alpha}}{\rho_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}+\sum_{\forall i\neq\tilde{k}}\frac{\rho_{\tilde{k}}\Big(\sum\limits_{l\in\mathcal{M}_{\tilde{k}}}\beta_{\tilde{k},l}\Big)^{\alpha}}{\rho_i\Big(\sum\limits_{l\in\mathcal{M}_i}\beta_{i,l}\Big)^{\alpha}}\right]\\ &= 2\left(\mu_{\tilde{k}}(\alpha,\rho'_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})-\mu_{\tilde{k}}(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})\right)=2\Delta\mu_{\tilde{k}}.\end{aligned} \tag{D.7}$$

Moreover, according to (D.3), we have:

$$\begin{aligned}\Delta u &= \frac{1}{2}\left(\sum_{k\in\mathcal{K}}\mu_k(\alpha,\rho'_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})-\sum_{k\in\mathcal{K}}\mu_k(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})\right)\\ &= \frac{1}{2}2\left(\mu_{\tilde{k}}(\alpha,\rho'_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})-\mu_{\tilde{k}}(\alpha,\rho_{\tilde{k}},\boldsymbol{\rho}_{(-\tilde{k})})\right)=\Delta\mu_{\tilde{k}}.\end{aligned} \tag{D.8}$$

Therefore, $u(\alpha,\boldsymbol{\rho})$ is an exact potential function for the parameterized game $\mathcal{G}(\alpha)$.