

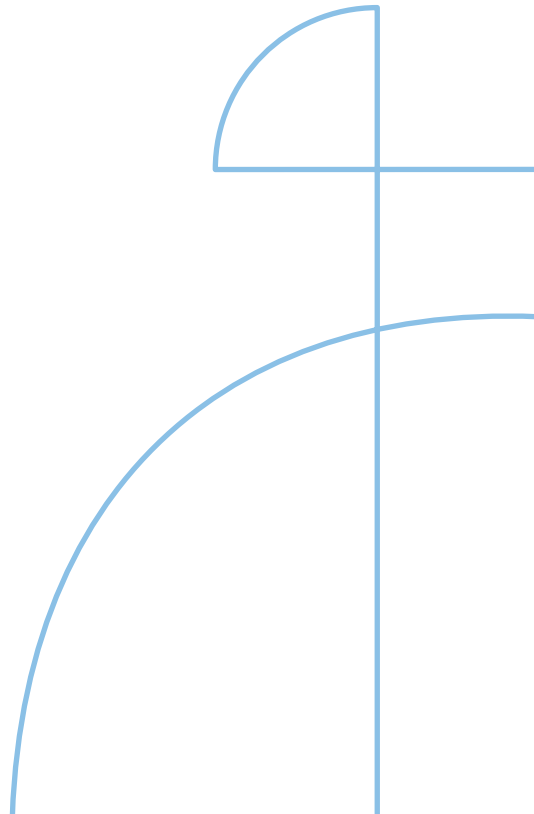


Doctoral Thesis in Information and Communication Technology

Cell-Free Massive MIMO Networks: Practical Aspects and Transmission Techniques for Radio Resource Optimization

MAHMOUD ZAHER

KTH ROYAL INSTITUTE OF TECHNOLOGY



Cell-Free Massive MIMO Networks: Practical Aspects and Transmission Techniques for Radio Resource Optimization

MAHMOUD ZAHER

Academic Dissertation which, with due permission of the KTH Royal Institute of Technology, is submitted for public defence for the Degree of Doctor of Philosophy on Friday, May 16, 2025 at 9:15 in Ka-Sal C, Electrum, Kistagången 16, 164 40 Kista.

Doctoral Thesis in Information and Communication Technology
KTH Royal Institute of Technology
Stockholm, Sweden 2025

© Mahmoud Zaher

ISBN 978-91-8106-250-2
TRITA-EECS-AVL-2025:43

Printed by: Universitetservice US-AB, Sweden 2025

Abstract

The increasing demand for wireless data traffic poses a significant challenge for current cellular networks, requiring each new technology generation to enhance network capacity and coverage, and spectral efficiency (SE) per connected device. Massive multiple-input multiple-output (MIMO) technology has emerged as a key component of 5G and leverages a large number of antennas at each access point (AP) to spatially multiplex many user equipments (UEs) over the same time-frequency resources. Looking beyond 5G, the new cell-free massive MIMO technology has gained considerable attention due to its ability to exploit spatial macro diversity and achieve higher interference resilience. Unlike traditional cellular networks, the cell-free architecture consists of a dense deployment of distributed APs that collaboratively serve UEs across a large coverage area without predefined cell boundaries. This architecture improves the mobile network coverage and aims to provide a more uniform quality of service throughout the network. However, the primary challenges of cell-free massive MIMO include the high computational complexity required for signal processing and the substantial fronthaul capacity needed for information exchange between APs. Moreover, another major challenge is handover management to cope with changing channel conditions and UE mobility; since in a cell-free network handover needs to consider how to dynamically evolve the serving set of APs to each UE, which is more complicated than in a cellular network where each UE is served by a single AP and handover means changing the serving AP.

In this doctoral thesis, we provide distributed solutions to research problems related to power allocation and mobility management to address some of the inherent challenges of the cell-free network architecture. Additionally, we introduce a new method for characterizing unknown interference in wireless networks. Moreover, we propose efficient optimization procedures in the context of multicast beamforming optimization and establish a novel method for rank reduction in conjunction with semidefinite relaxation (SDR).

For the problem related to power allocation, a distributed machine learning-based solution that provides a good trade-off between SE performance and applicability for implementation in large-scale networks is developed with reduced fronthaul requirements and computational complexity as compared to a centralized solution, where the power allocation for all APs is computed at a central processor. The solution is divided in a way that enables each AP, or group of APs, to separately decide on the power coefficients to the UEs based on the locally available information at the AP without exchanging information with the other APs, however, still attempting to achieve a network wide optimization objective.

Regarding mobility management, a new soft handover procedure is devised for updating the serving sets of APs and assigning pilot signals to each UE in a dynamic scenario considering UE mobility. The algorithm is tailored to reduce the required number of handovers per UE and changes in pilot assignment. Numerical results show that our proposed solution identifies the essential refinements since it can deliver comparable SE to the case when the AP-UE association is completely redone.

As for interference modeling, we developed a new Bayesian-based technique to model the distribution of the unknown interference arising from scheduling variations in neighbouring cells. The method is shown to provide accurate statistical modeling of the unknown interference power and an effective tool for robust rate allocation in the uplink with a guaranteed target outage performance. The method was later extended to account for the unknown interference of neighbouring clusters in a cell-free network architecture.

Many wireless communication applications require sending the same data to multiple UEs; for example, in streaming live events, distributing software updates, or training of federated learning models. Physical-layer multicasting presents an efficient transmission topology to exploit the beamforming capabilities at the transmitting nodes and broadcast nature of the wireless channel to satisfy the demand for the same content from several UEs. The uniform service quality and improved coverage of the cell-free network architecture are particularly suitable for this transmission topology. In this regard, we propose a novel successive elimination algorithm coupled with SDR to extract a near-global optimal rank-1 beamforming solution to the max-min fairness (MMF) multicast problem in a cell-free massive MIMO network. A specifically tailored optimization algorithm is then designed, leveraging the alternating direction method of multipliers (ADMM) and offering significant improvements in computational requirements.

Keywords: Cell-free massive MIMO, power allocation, sum-SE maximization, proportional fairness, spectral efficiency, machine learning, handover, cluster formation, pilot signal assignment, unknown interference, outage, multi-user MIMO, semidefinite relaxation, rank reduction, physical-layer multicasting, downlink beamforming, ADMM.

Sammanfattning

Den ökande efterfrågan på trådlös datatrafik utgör en betydande utmaning för dagens cellulära nätverk, vilket kräver att varje ny teknikgeneration förbättrar nätverkskapacitet och täckningen, samt spektraleffektivitet (SE) per uppkopplad enhet. Massiv MIMO-teknik (multiple-input multiple-output) har dykt upp som en viktig komponent i 5G då den använder ett stort antal antenner på varje accesspunkt (AP) för att rumsligt multiplexa många användarutrustningar (UE) över samma tidsfrekvensresurser. Om man ser bortom 5G så har den nya cellfria massiva MIMO-tekniken fått stor uppmärksamhet på grund av sin förmåga att utnyttja rumslik makrodiversitet och uppnå högre interferensmotståndskraft. Till skillnad från traditionella cellulära nätverk består den cellfria arkitekturen av en tät uppsättning av distribuerade AP:er som samarbetar för att betjäna UE:er över ett stort täckningsområde utan fördefinierade cellgränser. Denna arkitektur förbättrar mobilnätets täckning och syftar till att ge en mer enhetlig tjänstekvalitet i hela nätverket. De primära utmaningarna med cellfri massiv MIMO inkluderar den höga beräkningskomplexiteten som krävs för signalbehandling och den betydande fronthaul-kapacitet som behövs för informationsutbyte mellan AP:er. En annan stor utmaning är dessutom överräknings-hantering för att klara av ändrade kanalförhållanden och UE-mobilitet; eftersom överräkning i ett cellfritt nätverk måste överväga hur man dynamiskt förändrar den betjänande uppsättningen av AP:er till varje UE, vilket är mer komplicerat än i ett cellulärt nätverk där varje UE betjänas av en enda AP och överräkning endast innebär att byta ansvarig AP.

I den här doktorsavhandlingen presenterar vi distribuerade lösningar på forskningsproblem relaterade till effekreglering och mobilitetshantering för att hantera några av de inneboende utmaningarna med den cellfria nätverksarkitekturen. Dessutom introducerar vi en ny metod för att karakterisera okända störningar i trådlösa nätverk. Vi föreslår även effektiva optimeringsprocedurer för optimering av multicast-lobformning och etablerar en ny metod för rangreduktion i samband med semidefinit relaxering (SDR).

För problemen kopplade till effekreglering utvecklas en distribuerad maskinlärningsbaserad lösning som ger en bra avvägning mellan SE-prestanda och tillämpbarhet för implementering i storskaliga nätverk med minskade fronthaulkrav och beräkningskomplexitet jämfört med en centraliserad lösning, där effekregleringen för alla AP:er beräknas på en central processor. Lösningen är uppdelat på ett sätt som gör det möjligt för varje AP, eller grupp av AP:er, att separat besluta om effektkoefficienterna till UE:er baserat på den lokalt tillgängliga informationen vid AP:erna utan att utbyta information med de övriga AP:erna, men ändå försöka uppnå ett nätverksomfattande optimeringsmål.

När det gäller mobilitetshantering utformas en ny överräknings-procedur för uppdatering av de betjänande uppsättningarna av AP:er och tilldelas pilot-signaler till varje UE i ett dynamiskt scenario med hänsyn till UE-mobilitet. Algoritmen är skräddarsydd för att minska det nödvändiga antalet överräkningar per UE och förändringar i pilottilldelningen. Våra numeriska resultat visar att den föreslagna lösningen identifierar de väsentliga förfiningarna eftersom den kan leverera jämförbar SE som i fallet där AP-UE-associationen görs om helt.

När det gäller störningsmodellering utvecklade vi en ny Bayesiansk metod för att modellera fördelningen av de okända störningar som uppstår på grund av schemalägningsvariationer i närliggande celler. Metoden har visat sig ge korrekt statistisk modellering av störningseffekten och är ett effektivt verktyg för robust hastighetsallokering i upplänken med en garanterad avbrottsprestanda. Metoden utökades senare för att ta hänsyn till de okända störningarna från angränsande kluster i en cellfri nätverksarkitektur.

Många trådlösa kommunikationstillämpningar kräver att samma data skickas till flera UE:er, till exempel vid streaming av liveevenemang, distribution av programuppdateringar eller träning av federerade inlärningsmodeller. Multicasting på det fysiska lagret är en effektiv överföringsmetod som utnyttjar lobformningsförmågan vid AP:erna och spridningsförmågan hos den trådlösa kanalen för att tillfredsställa efterfrågan på samma innehåll från flera UE:er. Den enhetliga servicekvaliteten och förbättrade täckningen hos den cellfria nätverksarkitekturen är särskilt lämplig för denna överföringstopologi. I detta fall föreslår vi en ny gradvis elimineringsalgoritm kopplad till SDR för att extrahera en nästan globalt optimal rang-1-lobformningslösning för multicastproblemet med max-min-rättvisa i ett cellfritt massivt MIMO-nätverk. En skräddarsydd optimeringsalgoritm designas sedan för att utnyttja en metod som kallas alternating direction method of multipliers (ADMM) och erbjuder betydande förbättringar av beräkningsbehov.

Nyckelord: Cellfri massiv MIMO, effektregering, summa-SE-maximering, proportionell rättvisa, spektraleffektivitet, maskininläring, översäckning, klusterbildning, pilotsignalstildelning, okända störningar, avbrottsnivå, MIMO för flera användare, semidefinit relaxering, rangreduktion, multicasting på fysiska lagret, nedlänkslobforming, ADMM.

Acknowledgement

First and foremost, I thank God for giving me the ability and the strength to overcome many challenges in my life. I thank God for all his blessings to enable me to complete this work.

I would like to point a special thanks to my supervisor Emil Björnson for his guidance and support. I truly cannot think of a better form of supervision and working environment. I would also like to thank my co-supervisor Marina Petrova for her feedback and support. I would like to express my gratitude to the opponent and grading committee members for their constructive comments and discussion to help me improve my work.

To my wife, thanks for making my life easier, believing in me and supporting me. To my parents and sisters, thank you for always believing in me since I was little and standing by my side. I could not have finished this work without all of your support. Our bond gives me confidence that I can achieve anything.

Mahmoud Zaher
Stockholm, May 2025

Contents

List of Papers	ix
List of Figures	xi
List of Acronyms	xiv
1 Introduction	1
1.1 Cell-Free Massive MIMO	2
1.1.1 Motivation	2
1.2 Physical-Layer Multicasting	2
1.2.1 Motivation	3
1.3 Research Objectives	3
1.4 Thesis Contributions	4
1.5 Thesis Outline	6
2 Background on Mobile Networks	9
2.1 Cellular Networks	9
2.1.1 Network Architecture	9
2.1.2 Mobility Management	11
2.1.3 Power Allocation	12
2.2 Cell-Free Networks	13
2.2.1 Network Architecture	14
2.2.2 Building Blocks	14
2.2.3 User-Centric Cell-Free Networks	15
2.2.4 Operational Strategies	17
2.2.5 Mobility Management	19
2.2.6 Power Allocation	20
2.3 A Brief on Ergodic Capacity	21
2.3.1 Capacity of Multi-user MIMO Channels	21
2.3.2 Block Fading Model	24
2.3.3 Use-and-then-Forget Bounding Technique	25
3 Physical-Layer Multicasting	27
3.1 Unicast, Broadcast and Multicast Transmissions	27

3.1.1	Unicast	27
3.1.2	Broadcast	28
3.1.3	Multicast	29
3.2	Multicast Beamforming Optimization	30
3.2.1	QoS and MMF Problem Formulations	31
3.2.2	Solution Approaches	32
3.3	Cell-Free Massive MIMO Multicasting	34
4	Summary of Included Papers	37
4.1	Paper I - Learning-Based Downlink Power Allocation in Cell-Free Massive MIMO Systems	37
4.1.1	Introduction	37
4.1.2	Related Work	38
4.1.3	Contribution	38
4.1.4	Sum-SE and PF Optimization Objectives	39
4.1.5	Selected Results	40
4.2	Paper II - Soft Handover Procedures in mmWave Cell-Free Massive MIMO Networks	43
4.2.1	Introduction	43
4.2.2	Related Work	43
4.2.3	Contribution	43
4.2.4	Channel Model	44
4.2.5	Selected Results	45
4.3	Paper III - A Bayesian Approach to Characterize Unknown Interfer- ence Power in Wireless Networks	48
4.3.1	Introduction	48
4.3.2	Related Work	48
4.3.3	Contribution	48
4.3.4	Unknown Interference Power Distribution	49
4.3.5	Selected Results	49
4.4	Paper IV - Unknown Interference Modeling for Rate Adaptation in Cell-Free Massive MIMO Networks	52
4.4.1	Introduction	52
4.4.2	Related Work	52
4.4.3	Contribution	52
4.4.4	Unknown Interference Power Distribution	53
4.4.5	Selected Results	54
4.5	Paper V - Cell-Free Beamforming Design for Physical-Layer Multi- group Multicasting	57
4.5.1	Introduction	57
4.5.2	Related Work	57
4.5.3	Contribution	58
4.5.4	QoS and MMF Multicast Problems	59
4.5.5	Selected Results	60

4.6	Paper VI - Low-Complexity SDP-ADMM for Physical-Layer Multi-casting in Massive MIMO Systems	63
4.6.1	Introduction	63
4.6.2	Related Work	63
4.6.3	Contribution	63
4.6.4	SDP-ADMM: Efficient Multicast Beamforming	64
4.6.5	Selected Results	65
5	Conclusions and Future Work	67
5.1	Conclusions	67
5.2	Future Work	69
	References	71

List of Papers

1. ***Learning-Based Downlink Power Allocation in Cell-Free Massive MIMO Systems***
Mahmoud Zaher, Özlem Tuğfe Demir, Emil Björnson and Marina Petrova
in IEEE Transactions on Wireless Communications, vol. 22, no. 1, pp. 174-188, Jan. 2023.
2. ***Soft Handover Procedures in mmWave Cell-Free Massive MIMO Networks***
Mahmoud Zaher, Emil Björnson and Marina Petrova
in IEEE Transactions on Wireless Communications, vol. 23, no. 6, pp. 6124-6138, June 2024.
3. ***A Bayesian Approach to Characterize Unknown Interference Power in Wireless Networks***
Mahmoud Zaher, Emil Björnson and Marina Petrova
in IEEE Wireless Communications Letters, vol. 12, no. 8, pp. 1374-1378, Aug. 2023.
4. ***Unknown Interference Modeling for Rate Adaptation in Cell-Free Massive MIMO Networks***
Mahmoud Zaher, Emil Björnson and Marina Petrova
in Proc. IEEE Wireless Communications and Networking Conference (WCNC), 2024.
5. ***Cell-Free Beamforming Design for Physical Layer Multigroup Multicasting***
Mahmoud Zaher, Emil Björnson and Marina Petrova
under review in IEEE Transactions on Wireless Communications, 2024.
6. ***Low-Complexity SDP-ADMM for Physical-Layer Multicasting in Massive MIMO Systems***
Mahmoud Zaher and Emil Björnson
Accepted in 23rd International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt), 2025.

Contributions by the author not included in the thesis. These represent the conference versions of Papers 1, 2, 5, the journal version of Paper 6 which is to be submitted and co-authored conference and journal articles.

7. ***Distributed DNN Power Allocation in Cell-Free Massive MIMO***
Mahmoud Zaher, Özlem Tuğfe Demir, Emil Björnson and Marina Petrova
in Proc. 55th Asilomar Conference on Signals, Systems and Computers, 2021.
8. ***Mobility Management in mmWave Cell-Free Massive MIMO Networks***
Mahmoud Zaher, Emil Björnson and Marina Petrova
in Proc. IEEE International Conference on Communications (ICC), 2023.
9. ***Near-Optimal Cell-Free Beamforming for Physical Layer Multigroup Multicasting***
Mahmoud Zaher, Emil Björnson and Marina Petrova
in Proc. IEEE Global Communications Conference (GLOBECOM), 2024.
10. ***Low-Complexity Optimal Beamforming for Physical-Layer Multicasting in Cell-Free Massive MIMO***
Mahmoud Zaher and Emil Björnson
to be submitted to IEEE Transactions on Wireless Communications, 2025.
11. ***Interference Cancellation in Full Duplex Uplink Multi-User MIMO DF Relaying System with Imperfect CSI***
Sarah Imam, **Mahmoud Zaher** and Ahmed El-Mahdy
in Proc. International Wireless Communications and Mobile Computing (IWCMC), 2021.
12. ***Joint Energy and Latency Optimization in Federated Learning over Cell-Free Massive MIMO Networks***
Afsaneh Mahmoudi, **Mahmoud Zaher** and Emil Björnson
in Proc. IEEE Wireless Communications and Networking Conference (WCNC), 2024.
13. ***Low-Latency and Energy-Efficient Federated Learning over Cell-Free Networks: A Trade-off Analysis***
Afsaneh Mahmoudi, **Mahmoud Zaher** and Emil Björnson
in IEEE Open Journal of the Communications Society, 2025.
14. ***On the Performance of Harmonic Backscattering for Zero-Energy Devices***
Paschalina Foti, Boules Mouris, Thiemo Voigt, **Mahmoud Zaher** and Mehrnaz Afshang
Accepted in IEEE Wireless Power Technology Conference and Expo (WPTCE), 2025.

List of Figures

2.1	Cellular network architecture.	10
2.2	Cell-free network architecture. The solid lines represent fronthaul links.	15
2.3	User-centric cell-free massive MIMO. Each UE is served by a subset of the APs.	16
4.1	Comparison of power allocation strategies.	41
4.2	Robustness against perturbations.	42
4.3	Site map for a European city with $K = 5$ moving UEs over a $400 T_i$ time interval.	46
4.4	CDF of the DL SE per UE with MR and RZF precoding, $K = 20$	47
4.5	CDF of the DL SE per UE with MR and RZF precoding, $K = 50$	47
4.6	Outage probability for different numbers of unknown interferers and different distances between desired UE and BS.	51
4.7	ϵ -outage SE with RZF.	51
4.8	Site map and UE distribution with $K_u = 100$ unknown interferers. Serving cluster is located inside the marked circle.	54
4.9	Outage probability for different numbers of unknown interferers.	55
4.10	ϵ -outage SE with RZF.	56
4.11	Average minimum SE for different numbers of UEs.	61
4.12	Average computational time in seconds.	62
4.13	CDF of the max-min SE for $K = 15$ UEs.	66

List of Acronyms

1G	First generation
4G	Fourth generation
5G	Fifth generation
6G	Sixth generation
ADMM	Alternating direction method of multipliers
AoD	Angle-of-departure
AP	Access point
BC	Broadcast Channel
BS	Base station
CDF	Cumulative distribution function
CDMA	Code division multiple access
C-RAN	Cloud-radio access network
CoMP-JT	Coordinated multipoint-joint transmission
CPU	Central processing unit
CSI	Channel state information
DL	Downlink
DNN	Deep neural network
EP	Edge processor
HetNet	Heterogeneous network
LMMSE	Linear Minimum mean-square error
LoS	Line-of-sight
MIMO	Multiple-input multiple-output
MMSE	Minimum mean-square error
MR	Maximum ratio
MU	Multi-user
PF	Proportional fairness
SDP	Semidefinite program
SDR	Semidefinite relaxation
SE	Spectral efficiency
SINR	Signal-to-interference-plus-noise ratio
SNR	Signal-to-noise ratio
TTT	Time-to-trigger
UatF	Use-and-then-forget
UC	User-centric
UDN	Ultra-dense network

UE	User equipment
UL	Uplink
ULA	Uniform linear array
WMMSE	Weighted minimum mean-square error
ZF	Zero-forcing

Chapter 1

Introduction

Wireless communication has revolutionized modern connectivity, enabling seamless data transmission across vast distances. Since the wireless traffic is continuously growing, each new mobile network generation must find new ways to multiplex more devices, improve network coverage and increase the spectral efficiency (SE) per device. The cornerstone technology of 5G is massive multiple-input multiple-output (MIMO) [15, 16]. Massive MIMO systems consist of access points (APs) equipped with a large number of active antennas to enable spatially multiplexing of a large number of user equipments (UEs) on the same time-frequency resource. In a conventional cellular network, the coverage area is divided into smaller geographic regions called cells, where each UE is usually served by only one AP that delivers the strongest signal to the UE's current location. This cellular architecture is designed specifically to promote efficient frequency reuse, however, suffers from several problems. An important problem is that the received signal power from an AP to a given UE rapidly decays with the propagation distance. As a consequence, the signal-to-noise ratio (SNR) perceived by the UEs may vary greatly depending on where they are within the coverage area. A common difference reaches about 40 dB between cell-center and cell-edge UEs [17]. In addition, cell-edge UEs are more likely to suffer from higher interference levels from neighbouring cells, thereby they may experience a significantly lower signal-to-interference-plus-noise-ratio (SINR) compared to cell-center UEs.

Another major concern in wireless networks is the prevalent problem of unknown interference. The unknown interference is an outcome of scheduling variations in neighbouring cells, which is hard to measure in practice due to the large signalling overhead incurred. This makes the task of rate allocation to guarantee a specific target outage a difficult problem for the serving node, resulting in frequent outage situations regardless of the employed rate. Inspired by the Bayesian radar detection problems [18, 19], that model the variance of a non-Gaussian spherically invariant random process, the unknown interference power can be equivalently modelled in wireless networks based on its statistical information.

1.1 Cell-Free Massive MIMO

Looking beyond 5G, cell-free massive MIMO (also known as distributed MIMO) is an emerging post-cellular architecture that is meant for rebuilding wireless networks to provide ubiquitous connectivity. The key characteristic is that a large number of distributed APs jointly serve the UEs within a given coverage area without creating artificial cell boundaries [17, 20, 21]. Cell-free massive MIMO has the ability to mitigate the aforementioned large SINR variations that may be experienced in a conventional cellular network through utilizing spatial macro-diversity. This allows for an improvement in network connectivity and energy efficiency.

The main challenges of building a cell-free network infrastructure lie in the computational complexity of signal processing and the huge fronthaul requirements for information exchange between the APs to allow for joint service [22–25]. Hence, it is desirable to build the network using limited-capacity fronthaul links between the APs and the central processing unit (CPU) [23, 24, 26, 27].

1.1.1 Motivation

Cell-free massive MIMO might become a foundation for 6G technologies owing to its ability to utilize macro-diversity and multi-user (MU) interference suppression, providing almost uniform service to the UEs [21, 28]. Inspired by the increasing attention on cell-free networks in both academia and the industry sector, the present work is pivoted on addressing some of the practical deployment problems in cell-free massive MIMO networks that may have been overlooked in the previous literature.

There are two main ways of operating a cell-free network: centralized and distributed operation [17]. Due to the aforementioned challenges, the distributed operation is adopted in this part of the thesis, where several functionalities are done at the AP level using the locally estimated channels. This enables reducing the fronthaul requirements for performing these functionalities. The focus is on providing distributed solutions to the power allocation and mobility management problems that become more complicated in a cell-free network, where each UE is connected to several APs at the same time, compared to a conventional cellular network with each UE being served by a single AP.

Afterwards, we tackle the problem of unknown interference from neighbouring cells, in a cellular massive MIMO system, which cannot be instantaneously measured at the serving AP. The setup is then extended to model the unknown interference of neighbouring service clusters in a cell-free massive MIMO network setting, which is particularly diverse and requires additional attention for devising an accurate interference model.

1.2 Physical-Layer Multicasting

In many wireless communication applications, there is a demand for the same data content by several UEs. Physical-layer multicasting represents an efficient alternative

to conventional unicast transmission when the same content is to be transmitted to a group of UEs simultaneously. It allows for serving numerous devices without exhausting network resources or the need for temporarily setting up extra network infrastructure in large events. For instance, mobile network operators can leverage it for video streaming of live events to multiple subscribers, providing a cost-effective alternative to traditional unicast-based content delivery. Other applications of multicasting include videoconferencing as well as broadcasting of machine learning models in federated learning applications, where multiple UEs need to receive the same data stream. Since every UE in a multicast group must be able to decode the data, the minimum SINR across the group determines what data rate can be utilized. Consequently, physical-layer multicasting is particularly well-suited for the cell-free network architecture. Emerging wireless communication applications requiring same content delivery to multiple UEs on cell-free networks can greatly benefit from this transmission topology.

1.2.1 Motivation

Physical-layer multicasting can support a wide range of applications in modern wireless communication networks. It provides an efficient scheme to transmit data to a group of UEs using a single transmission, thereby eliminating unnecessary co-channel interference and offering significant improvement in radio resource utilization [29]. Finding the optimal beamforming to boost the desired signal of several UEs simultaneously is a challenging task. The previous literature either relied on computationally expensive optimization procedures that can only guarantee convergence to a local optimal solution or on utilizing low-complexity well-known unicast precoding schemes with minimal adjustments to cope for the multicast problem, offering inferior performance. Moreover, recent works have suggested relatively lower-complexity solutions, however, still only guaranteeing convergence to a local optimum. To this end, we develop novel optimization procedures to efficiently find the near-global optimal multicast beamforming vectors. Further, we devise an ultra-low complexity heuristic algorithm that is specifically tailored for the multicast problem. In addition, we recognize the inherent compatibility between physical-layer multicasting and the cell-free massive MIMO architecture that can provide uniform service quality within the mobile network. For this reason, we keep the focus on the cell-free massive MIMO architecture to exemplify our solutions and to demonstrate that they are relevant in the 6G context.

1.3 Research Objectives

A primary objective of the thesis is to study practical deployment problems that may be inherent to the operation of the cell-free network architecture and propose efficient and effective solutions to these problems. Specifically, due to the aforementioned challenges of complexity, fronthaul requirements and mobility management, we formulate the following research questions:

RQ1: Can power allocation in cell-free massive MIMO networks be performed in a distributed manner, while still aiming to achieve a network-wide optimization objective?

We address the question considering imperfect channel state information (CSI), real-time processing constraints, the necessary input information and the robustness of the design.

RQ2: How can the pilot assignment and cluster formation decisions evolve due to UE mobility with a distributed solution that requires limited computational complexity and signalling overhead?

We address the question considering a mmWave ray-tracing based channel model, imperfect CSI, random UE mobility, spatially consistent shadow fading and distributed decision metrics.

Next, we shift the focus to the prevalent problem of unknown interference in wireless networks, which is then extended to consider unknown interference in cell-free networks. We formulate the following research question in this regard:

RQ3: How to perform rate allocation in the uplink to guarantee a target outage probability with incomplete knowledge of the interference statistics?

We address the question considering an MU-MIMO setup in both cellular and cell-free network settings, imperfect CSI, different combining schemes and the robustness of the design.

Afterwards, recognizing the alignment between physical-layer multicasting and the cell-free massive MIMO architecture, we tackle the challenging problem of beamforming design for multicast transmissions with a focus on cell-free networks. Specifically, we formulate the following research question:

RQ4: How to efficiently optimize beamforming configurations for multicast transmissions with respect to appropriate network utility maximization objectives?

We address the question considering quality-of-service (QoS) and max-min fair (MMF) optimization objectives, multiple multicast groups, perfect CSI and the computational requirements of the proposed solutions.

1.4 Thesis Contributions

The thesis is dedicated to analyzing and solving the research questions in the four topics listed earlier. Regarding RQ1, we have developed a learning-based solution to the power allocation problem in the downlink (DL) of a cell-free massive MIMO network. The solution features distributed or clustered feedforward deep neural networks (DNNs) that operate at the AP level or edge processor (EP) level which is responsible for a cluster of APs. The results of this study are published in the following papers:

- Mahmoud Zaher, Özlem Tuğfe Demir, Emil Björnson and Marina Petrova, “Distributed DNN Power Allocation in Cell-Free Massive MIMO,” *in Proc. 55th Asilomar Conference on Signals, Systems and Computers, 2021*.

- Mahmoud Zaher, Özlem Tuğfe Demir, Emil Björnson and Marina Petrova, “Learning-Based Downlink Power Allocation in Cell-Free Massive MIMO Systems,” in *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 174-188, Jan. 2023.

Own contributions: My contribution in the first study was to derive analytical expressions for the conventional optimization problems, the design and testing of the learning-based solutions, developing a simulation environment and writing the papers under the supervision of my advisors.

As for RQ2 concerning mobility management, a handover procedure for the update of the serving set of APs and assigned pilot sequence to each UE has been developed in a mobile scenario. The procedure is devised in order to minimize the number of *Master* AP handovers and changes in pilot assignment. The results of this work are published in the following papers:

- Mahmoud Zaher, Emil Björnson and Marina Petrova, “Mobility Management in mmWave Cell-Free Massive MIMO Networks,” in *Proc. IEEE International Conference on Communications (ICC)*, 2023.
- Mahmoud Zaher, Emil Björnson and Marina Petrova, “Soft Handover Procedures in mmWave Cell-Free Massive MIMO Networks,” in *IEEE Transactions on Wireless Communications*, vol. 23, no. 6, pp. 6124-6138, June 2024.

Own contributions: My contribution in the second study was to design the initial access and handover algorithms, the development of a novel Maneuvering Smooth Random Waypoint Mobility model for testing the proposed solutions, produce ray-tracing based channels in a real urban site map using [30] with some bugs corrected and enhance the configuration on my own, and write the papers under the supervision of my advisors.

For RQ3 about the problem of unknown interference in wireless networks, we developed an analytical framework to model the distribution of the unknown interference power arising from scheduling variations in neighbouring cells, which is hard to measure in practice due to the large signalling overhead incurred. After which a method for UL rate allocation with guaranteed target outage has been developed. The analytical derivation and rate allocation method is later expanded to consider unknown interference from neighbouring clusters in a cell-free massive MIMO network. The results of this work are published in the following papers:

- Mahmoud Zaher, Emil Björnson and Marina Petrova, “A Bayesian Approach to Characterize Unknown Interference Power in Wireless Networks,” in *IEEE Wireless Communications Letters*, vol. 12, no. 8, pp. 1374-1378, Aug. 2023.
- Mahmoud Zaher, Emil Björnson and Marina Petrova, “Unknown Interference Modeling for Rate Adaptation in Cell-Free Massive MIMO Networks,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2024.

Own contributions: My contribution in the third study was to derive the analytical expressions for the distribution of the unknown interference from neighbouring cells/clusters, the development of a method for robust rate allocation with guaranteed target outage, developing a simulation environment to test the proposed solution and writing the papers under the supervision of my advisors.

For RQ4, we devised a novel iterative optimization procedure coupled with semidefinite relaxation (SDR) to obtain a near-global optimal rank-1 beamforming solution to the QoS and MMF multicast problems. We further developed a low-complexity heuristic algorithm that provides an effective solution to the multicast beamforming optimization problem at a vastly reduced computational time. The proposed heuristic marks the first algorithm that is specifically tailored for the multicast problem with computational requirements in the milliseconds range. Afterwards, we derived a low-complexity alternating direction method of multipliers (ADMM) based optimization procedure to solve the problem to near-global optimality with reduced computational requirements. The results of this work are presented in the following papers:

- Mahmoud Zaher, Emil Björnson and Marina Petrova, “Near-Optimal Cell-Free Beamforming for Physical Layer Multigroup Multicasting,” *in Proc. IEEE Global Communications Conference (GLOBECOM), 2024.*
- Mahmoud Zaher, Emil Björnson and Marina Petrova, “Cell-Free Beamforming Design for Physical Layer Multigroup Multicasting,” *under review in IEEE Transactions on Wireless Communications, 2024.*
- Mahmoud Zaher and Emil Björnson, “Low-Complexity SDP-ADMM for Physical-Layer Multicasting in Massive MIMO Systems,” *Accepted in 23rd International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt), 2025.*
- Mahmoud Zaher and Emil Björnson, “Low-Complexity Beamforming for Physical-Layer Multicasting in Cell-Free Massive MIMO,” *to be submitted to IEEE Transactions on Wireless Communications, 2025.*

Own contributions: My contribution in the fourth study was to propose and formulate the research problem, devise the iterative optimization procedure and heuristic algorithm, derive a novel ADMM formulation for the proposed semidefinite program (SDP) that marks the first SDP-ADMM algorithm in the field, and write the papers under the supervision of my main advisor.

1.5 Thesis Outline

The rest of the thesis is organized as follows. Chapter 2 provides a background on cellular and cell-free network architectures as well as a brief on the ergodic capacity

expressions utilized throughout the thesis. Chapter 3 gives an overview of physical-layer multicasting and presents design criteria for beamforming optimization. In Chapter 4, a summary of the included papers in the thesis with a layout of the most relevant literature in each topic is presented. Chapter 5 concludes the thesis with the key takeaways and provides a brief description of possible future directions to extend the present work.

Chapter 2

Background on Mobile Networks

This chapter provides an overview of the cellular and cell-free network architectures and highlights the main differences between them that fall within the scope of the present work. We discuss the main challenges that are inherent to the cell-free network architecture. Afterwards, it paves the way with a brief analysis on the ergodic achievable rate expressions that are used throughout the thesis.

2.1 Cellular Networks

This section lays out the cellular network technology with a brief overview of the evolution between different generations. Cellular networks have been first deployed for commercial use since the first-generation (1G) launched in Tokyo, Japan in 1979. The 1G networks relied on analog transmissions, and supported only voice calls. With the advancements in digital technology, the wireless revolution began in the early 1990s with the first digital wireless mobile networks [31, 32]. Throughout the past four decades, digital wireless networks have evolved generation after generation with the ultimate aim of increasing the mobile network capacity and maximum number of devices that can be simultaneously served within a given area. Towards the evolution of mobile networks to the fifth generation (5G) around 2020, the mobile network coverage and data rates have reached far beyond the early generations. It has become an integral part of life, providing numerous services that are crucial to its users.

2.1.1 Network Architecture

For a cellular mobile network to provide coverage over a large geographical area, the area is divided into smaller areas that are referred to as *cells*. Each cell has a base station (BS) that sends and receives information to the UEs that are located within its coverage area. The coverage area can usually be determined by the places

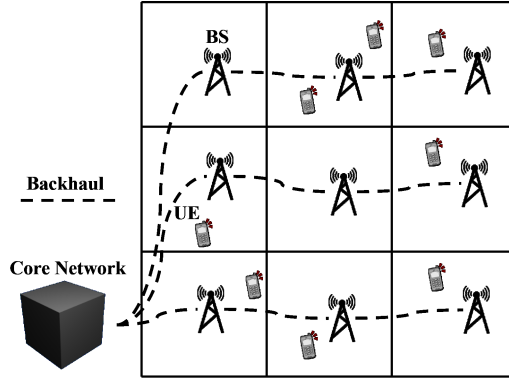


Figure 2.1: Cellular network architecture.

at which the BS has the highest received signal strength compared to other BSs in the network.

The wireless transmission range is determined by the propagation environment. Depending on the propagation environment, the received signal power at a given location from a BS may decay quadratically, or even faster, with the propagation distance. Consequently, this together with the transmission power of the BS, govern the coverage area that corresponds to a cell. The traditional mobile network infrastructure consists of a set of geographically distributed BSs that the UE can choose between. The BSs are typically deployed at elevated locations (e.g., at rooftops) to avoid having obstructions in the propagation path to many places in the area. The current deployment of mobile networks relies on the cellular architecture, which means that each UE connects to one BS; that is, the one with the highest received signal at the UE's current location. Figure 2.1 shows the basic infrastructure of a cellular network. The processing for the transmitted/received signals is usually done on dedicated hardware at each cell site. The square area around each BS marks the cell that the BS provides service to. Note that in reality the cells will not have such symmetric shape due to the existence of obstructions (such as buildings) that will affect the deployment locations of the BSs and cause irregularities in the shape of their coverage areas.

The infrastructure of contemporary cellular networks is divided into two main categories: the edge and the core network. The edge is responsible for the physical-layer communications with the UEs. It consists of BSs and other hardware units that are directly involved in that matter. On the other hand, the core network is responsible for routing of data packages and connection to the Internet, thereby it enables all the services requested by the UEs. The connections between the edge and the core are referred to as *backhaul* links and can either be fully wired (e.g., using fiber cables) or partially wireless (e.g., using fixed microwave links) [17].

2.1.2 Mobility Management

In a cellular network, each BS is responsible for providing service within a single cell in the entire network area. In order for the cellular network to support mobile devices, the concept of handover was introduced. In this section, we will discuss the key principles of the procedure and the two types of handover utilized in cellular networks.

Due to the existence of moving objects within the network area, the propagation environment is constantly being reshaped. This change results in a variation of the received powers from the different BSs that are in the vicinity of the UE. Naturally, a moving UE will experience even larger variations in signal strength from both the serving and neighbouring BSs as its location evolves with time. This is simply due to the fact that as the UE moves, its distances to the nearby BSs as well as the obstacles in the respective propagation paths change. For that purpose, the UE needs to keep monitoring the received power levels to the nearby BSs in addition to its current serving BS. If the power level from a BS exceeds that of the serving BS, then it is logical that the UE should start being serviced by that BS. This represents the intuition behind the concept of handover in cellular networks. In a handover event, the old cell that the UE belongs to is referred to as the *source* cell, whereas the new cell to which the UE “hands-off” is referred to as the *target* cell.

As stated earlier, the channel between a UE and the surrounding BSs is constantly developing due to variations in the UE’s location as well as the propagation environment. When a UE starts moving away from the coverage region of its current serving cell and enters the region of a neighbouring cell, the received signal strength from its serving BS is compared to that of the target cell. If the signal of the target cell is higher than that of the source cell, the handover should be initiated. It is likely that at the border between two cells, there should be some overlap of the coverage from the BSs in both cells. This is to avoid having uncovered regions within the network area. When the UE is at the border, the received signal strength from both BSs may be very similar. In order to avoid the so-called “ping-pong” handover, where the UE keeps on shifting between service from two cells due to small-scale channel variations, a hysteresis handover margin and time-to-trigger (TTT) has been utilized since the early cellular generations. The idea is that the target cell must keep a higher signal strength to the UE with some margin for a predetermined time interval in order for the handover event to be initiated.

In the beginning, the concept of *hard* handover was introduced to support mobility in cellular networks. This concept is commonly referred to as the *break-before-make*, since whenever a handover event is triggered, the UE first releases the resources allocated by the source cell, before the channel in the target cell is engaged. In order to improve cell-edge performance and reduce the connection drop rate, *soft* handover was introduced in code division multiple access (CDMA) systems. Conversely, this is referred to as the *make-before-break*. When a handover event is triggered, the UE is first allocated resources by the target cell, and is simultaneously served by the two BSs before the release of the resources allocated by the source BS.

In 5G cellular networks, dual/multiple connectivity is a new scheme that has been standardised to allow connection to multiple BSs on different component carriers, as outlined in [33]. It is utilized in heterogeneous networks (HetNets) to allow connecting to a macro BS (which may possibly be 4G) that utilizes sub-6 GHz frequencies, and simultaneously to one or more 5G small cells typically operating at mmWave frequencies. In this case, a UE is connected to the network through two distinct channels occupying a portion of both frequency bands. The connectivity to the lower frequency macro layer provides additional robustness to the inherently less robust high frequency small cell layer [33]. This is fundamentally different from a cell-free network that coherently combines the signals from multiple BSs that are transmitted on the same time-frequency resources. In the cell-free architecture that is described later in Section 2.2, we consider a single frequency band where the UEs are serviced and rely on spatial diversity (instead of frequency diversity) to offer a stable connection to every UE.

2.1.3 Power Allocation

Every BS in a cellular network is usually required to simultaneously provide service to multiple UEs residing within its coverage area. The BS must then have a mechanism to divide its resources on the different UEs. Among these resources is the power that the BS allocates for the transmission of the data symbols intended for each UE. A BS has a maximum power that it can utilize at any point in time; this is referred to as the power budget. The BS distributes this power on the active UEs depending on a number of factors. These could be the required data rate of the UE, the pathloss that the UE experiences with respect to other UEs and the needs of the other UEs at a given point in time. It is worth mentioning that the problem of power allocation in a conventional orthogonal access scheme, where each subcarrier has only a single active UE at a given point in time, is relatively simpler than the case of MU-MIMO where several UEs share the same time-frequency resources through spatial multiplexing. In the latter, the level of interference perceived by a given UE is directly related to the powers allocated by the BS to the other UEs, resulting in more complex relations between the SINRs of the different UEs. In this thesis, we focus on the more involved MU-MIMO case that offers significant gains to the capacity of mobile networks and have been utilized in its latest generations.

The problem of power allocation in cellular networks has been extensively studied in the last decades. Depending on the objective of the network operator, different utility functions can be used to divide the powers from the BSs to their respective UEs. In [15, 16, 34], several power allocation strategies have been listed to achieve a network-wide optimization objective. For example, the objective could be to maximize the minimum SE among all UEs, this is referred to as max-min fairness power allocation. This scheme focuses only on having a common SE for all the UEs, thereby it does not exploit the possible gains for the UEs having better channel conditions. Another extreme option is to maximize the sum-SE that the network can provide. Although it may sound tempting at first, such objective may lead to

large SE variations, i.e., poor fairness, among the UEs. This means that most of the power will be allocated to the UEs with better channel conditions, whereas others may be completely deprived from the network resources. A well-known trade-off between the two approaches listed above is the proportional fairness (PF) power allocation. The objective of this scheme is to maximize the product of the SINRs (or SEs) of the UEs. In this way, the utility balances between SE maximization for the UEs with better channel conditions, and not completely depriving less fortunate UEs from the network resources (as this would result in the product being zero),

Channel hardening refers to the phenomenon where the small-scale fading effects average out as the number of antennas increases, making the channel behave in a more deterministic manner. Particularly in massive MIMO networks, the channel hardening makes power allocation different than in single-antenna networks. The difference is that there is no need to adapt the transmit powers by each BS to small-scale fading variations, but only to the large-scale fading characteristics. In this case, the frequency with which new power allocation decisions need to be made is determined by the large-scale UE mobility and scheduling decisions in the network, meaning that it need not be done in every coherence interval. This distinctive characteristic of massive MIMO enables advanced power allocation schemes, that were previously considered too complex for practical implementation [16]. Nevertheless, the above mentioned schemes utilize fading coefficients from different parts of the network in order to maximize a network-wide objective. Such centralized solution might be difficult to implement in a cellular network where the power allocation decisions are done separately by each BS. In Section 2.2.6, we will further discuss possible centralized and distributed power allocation strategies in the context of cell-free massive MIMO networks.

It is worth noting that since in a cellular network each UE is only served by a single BS, the received desired signal power at the UE can be only influenced by the amount of power that the BS decides to allocate to that UE. Compared to a cell-free network, this limits the ability of the cellular network to manage the interference perceived by the UEs due to the reception of unintended transmissions that are meant for other UEs, which is particularly important in massive MU-MIMO networks that enable the multiplexing of several UEs on the same time-frequency resources. In the next section, we will detail some of the important building blocks of cell-free networks, clarifying their increased potential to manage the interference levels through power allocation.

2.2 Cell-Free Networks

Cell-Free massive MIMO has great potential to boost the minimum achievable rates in mobile networks. The cell-free network topology is based on cooperation between different APs to provide seamless connectivity to the UEs by eliminating the cell boundaries. Such a topology offers an increase in the SINR by associating each UE with multiple geographically distributed APs for coherent combination of the

transmitted/received signals. The connection to multiple distributed APs adds reliability to the network and achieves better spectral and energy efficiency.

This section provides an overview on cell-free massive MIMO network architecture and operational strategies.

2.2.1 Network Architecture

Going beyond the cellular paradigm, the key idea behind cell-free networks is to eliminate poor cell-edge performance that is existent in cellular networks. It can also enable denser network deployments through the cooperation between the APs. The difference between the cellular and cell-free infrastructures lies in the edge part of the network, whereas the core network does not encounter major differences. For the edge network, the cell-free topology incorporates more cooperation between the APs when compared to the cellular architecture. Such cooperation may exist in different levels, for example as described in [20]. Depending on the level of cooperation, the APs may be responsible for performing some tasks locally, if cooperation in those tasks is not very necessary. In any case, the cell-free terminology can be considered applicable if each UE is simultaneously served by a cluster of distributed APs that is particularly chosen for that UE. The role of each AP in the cluster and the amount of power allocated to the UE may vary depending on the importance of the AP, with respect to others in the serving cluster, to the UE.

Typically, a cell-free massive MIMO network is composed of L geographically distributed APs, each equipped with N antennas. The APs jointly serve K single antenna UEs that are arbitrarily distributed within a given coverage area. Each AP is connected via a fronthaul to a CPU, which is responsible for the cooperation between the APs conveying UL and DL data between them and other necessary signals. Throughout the thesis, we assume the connections are error-free. Figure 2.2 [1] depicts the network architecture. The CPU may be regarded as a logical entity, meaning that it may be physically located at one or more of the APs, or at a separate physical location. The CPU is connected to the core network through a backhaul link that is not shown in the figure, and which is used to facilitate different services that are requested by the UEs by providing connection to the internet or other service providers. The difference between cellular and cell-free networks lies in the infrastructure and joint signal processing at the APs, however, this does not impose any difference from the UE's perspective. It should be possible for the same UE to connect to both types of networks without upgrading its software [17].

2.2.2 Building Blocks

The cell-free network paradigm represents a combination of three different technologies. The first is the C-RAN (Cloud-Radio Access Network) which is a network architecture where a group of APs is connected to a single CPU that carries out most of the baseband processing for the APs. The synergy between the C-RAN architecture and the cell-free network deployment makes the cell-free architecture a

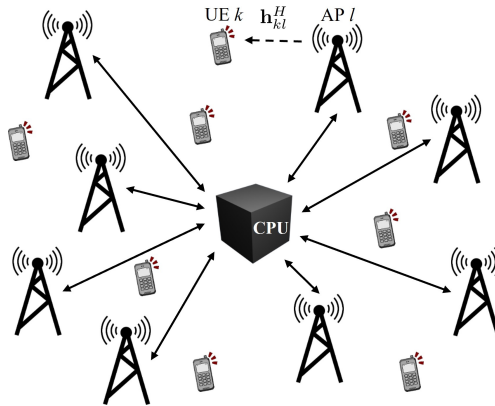


Figure 2.2: Cell-free network architecture. The solid lines represent fronthaul links.

natural application of that of the C-RAN. The second is the CoMP-JT (Coordinated Multipoint-Joint Transmission). This technology is the basis for the cooperation strategy between APs that is utilized in cell-free networks; such that APs jointly transmit the same signals to the UEs for coherent reception. The main difference is that CoMP-JT was originally introduced where only a selected number of UEs (usually those located at the cell-edge) are served by several BSs at the same time [35], whereas cell-free networks utilize the cooperation principles for all the UEs in the network. The key factor behind this is that CoMP required a lot of signalling between the BSs when it was first introduced, which have also resulted in achieving relatively small practical gains. Combining it with the C-RAN architecture makes it easier since most of the baseband processing is implemented at the same place; that is, the CPU. The third building block of cell-free networks lies in the massive MIMO processing literature. Since the signals are jointly transmitted/received by a large number of antennas, linear processing schemes that have been developed in the massive MIMO literature represent a perfect match for the joint processing of signals from the APs. These processing schemes represent a foundation for signal processing when the dimensionality of the transmitted/received signals becomes large, regardless of the network deployment architecture.

2.2.3 User-Centric Cell-Free Networks

For scalable network operation, it is unrealistic that all APs serve all UEs. In a conventional cellular setup, the cells are divided based on network-centric clustering where each AP represents a cell. The problem with this topology lies in the high inter-cell interference level that is achieved at the borders between cells. This results in low SINRs and corresponding low achievable SE. Alternatively, cell-free massive MIMO is based on user-centric (UC) clustering. UC clustering is a means

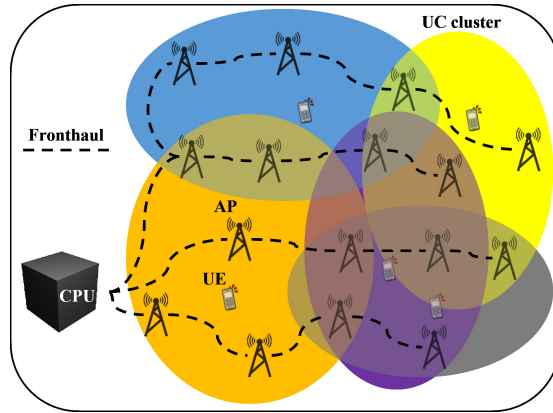


Figure 2.3: User-centric cell-free massive MIMO. Each UE is served by a subset of the APs.

of determining the serving set (or cluster) of APs to each UE such that the UE is always located at the center of the cluster. By putting the focus on the UEs, the serving clusters can be dynamically optimized for every UE's current location boosting the achievable SE.

Although the UC topology may achieve high SE gains, the design of such a dynamic clustering algorithm is a complex task. There are three main challenges for UC clustering: managing overlapping clusters, how to perform pilot assignment, and mobility management and cluster update. These challenges are inherent to the superior dynamic UC clustering.

As depicted in Figure 2.3 [2], since the serving cluster of APs is optimized for every UE, there may exist partial overlap between the UC clusters of the UEs. Each UE needs to be assigned a specific pilot sequence such that the serving cluster of APs can estimate the corresponding channel to that UE. UEs having overlapping AP clusters may then suffer from poor channel estimation due to severe pilot contamination if they are assigned the same pilot sequence. This necessitate that the decision on the serving clusters and pilot assignment needs to be jointly performed for every UE. In many possible situations, a UE may have partially overlapping AP cluster with several other UEs making this joint decision on pilot assignment and cluster formation a complex problem. All these challenges imply that the possible UC solutions incur large computational complexity and signalling overhead, which is one of the important aspects that we consider in Paper II when designing the UC clustering and handover procedure.

With regard to mobility management, since each UE is served by a cluster of APs, there exist more AP-UE associations in a cell-free network compared to a conventional cellular network. Naturally, as the UEs move in a given coverage area, the serving clusters and assigned pilots must be updated such that they keep suitable

for the UEs' current locations. The challenge here is how to decide on updating the clusters and pilots without the need to change everything for all UEs. In other words, one needs to keep the required changes in the network to a minimal level in order to limit the signalling overhead that is necessary for performing the updates.

2.2.4 Operational Strategies

A cell-free network can be primarily operated in two ways: centralized operation and distributed operation [17]. In the centralized operation, the CPU is responsible for performing the precoding and combining of signals in the DL and UL, respectively. The APs must then share all their information, including the pilot signals received by each AP, to the CPU to perform the channel estimation and precoding/combining. In this case, the APs may be viewed as relays that facilitate the communication between the CPU and the UEs. For the distributed case, no instantaneous CSI needs to be transmitted over the fronthaul, since each AP is responsible for locally performing the transmit precoding and receive combining based on the locally estimated channels. In this case, the CPU is only responsible for the data encoding/decoding.

There exist different levels of cooperation that may be implemented between the APs, with each having a slightly different functional split between the APs and the CPU. These may be referred to as semi-centralized implementations. The level of cooperation provides a trade-off between performance, and the required computational complexity and fronthaul signalling. For example, the APs may provide the CPU with only statistical CSI to enable it to optimize the combination of the data estimates from each AP in the UL, by giving higher weights to nearby APs or APs that experience less interference for a given UE.

In the first part of the thesis, we focus on the distributed operation in the DL of a cell-free massive MIMO network with standard unicast content delivery. As stated earlier, with such operation the precoding is done at each AP separately using the locally estimated channels. Let the channel between UE k , $k = 1, \dots, K$, and AP l , $l = 1, \dots, L$, be denoted by $\mathbf{h}_{kl} \in \mathbb{C}^{N \times 1}$ for a particular subcarrier at a given time instant. We consider maximum ratio (MR) and regularized zero-forcing (RZF) precoding. The precoding vector $\bar{\mathbf{w}}_{kl} \in \mathbb{C}^{N \times 1}$ intended for UE k that is computed by AP l can thus be written as

$$\bar{\mathbf{w}}_{kl} = \begin{cases} \hat{\mathbf{h}}_{kl} & \text{for MR,} \\ \left(\sum_{i \in \mathcal{D}_l} p_i \hat{\mathbf{h}}_{il} \hat{\mathbf{h}}_{il}^H + \sigma^2 \mathbf{I}_N \right)^{-1} p_k \hat{\mathbf{h}}_{kl} & \text{for RZF,} \end{cases} \quad (2.1)$$

where $\hat{\mathbf{h}}_{kl}$ represents the locally estimated channel at AP l . Logically, an AP is only required to estimate the channels of the UEs that it is currently serving. These precoding vectors are then scaled to satisfy the power constraints of the APs. The power allocation strategies for dividing the available power on the different UEs are discussed in Section 2.2.6

Local Channel Estimation

In order to perform channel estimation, each UE is assigned a pilot sequence of length τ_p from a set of τ_p mutually orthogonal pilot sequences. In the channel estimation phase, the UEs transmit their pilots on the same time-frequency resources. Let $t_k \in \{1, \dots, \tau_p\}$ denote the index of the pilot assigned to UE k . The received signal at AP l , which is intended for the estimation of the UEs' channels, $\mathbf{Y}_l^p \in \mathbb{C}^{N \times \tau_p}$ is computed as

$$\mathbf{Y}_l^p = \sum_{i=1}^K \sqrt{p_i} \mathbf{h}_{il} \boldsymbol{\theta}_{t_i}^T + \mathbf{N}_l, \quad (2.2)$$

where p_i is the transmit power of UE i , $\boldsymbol{\theta}_{t_i} \in \mathbb{C}^{\tau_p \times 1}$ denotes the pilot sequence assigned to UE i such that $\|\boldsymbol{\theta}_{t_i}\|^2 = \tau_p$, and $\mathbf{N}_l \in \mathbb{C}^{N \times \tau_p}$ represents the additive i.i.d. complex Gaussian noise at AP l during the entire channel estimation phase.

Different channel estimation techniques can be applied at the APs. We consider the optimal minimum mean square error (MMSE) channel estimation and the least square (LS) estimation. Although MMSE estimation is known to achieve better quality estimates than LS estimation, the advantage of the LS estimation is that it offers a lower computational complexity compared to the MMSE estimation and does not require knowledge of the statistical properties of the channels which may be cumbersome to obtain under mobility. Also, there may not exist a simple closed form MMSE solution for some channel models other than the correlated Rayleigh fading model that is adopted in this work.

After correlating the received signal at AP l with the normalized pilot t_k for the estimation of UE k channel, the signal $\mathbf{y}_{t_k l}^p \in \mathbb{C}^{N \times 1}$ is obtained as

$$\begin{aligned} \mathbf{y}_{t_k l}^p &= \mathbf{Y}_l^p \boldsymbol{\theta}_{t_k}^* / \sqrt{\tau_p} \\ &= \sum_{\substack{i=1 \\ t_i=t_k}}^K \sqrt{\tau_p p_i} \mathbf{h}_{il} + \mathbf{n}_{t_k l}, \end{aligned} \quad (2.3)$$

where $\mathbf{n}_{t_k l} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ represents the additive Gaussian noise vector at AP l . Throughout the thesis, we consider correlated Rayleigh channel fading. The MMSE estimate for the channel between UE k and AP l is thus estimated as [17]

$$\begin{aligned} \hat{\mathbf{h}}_{kl} &= \sqrt{\tau_p p_k} \mathbf{R}_{kl} \left(\sum_{\substack{i=1 \\ t_i=t_k}}^K \tau_p p_i \mathbf{R}_{il} + \sigma^2 \mathbf{I}_N \right)^{-1} \mathbf{y}_{t_k l}^p \\ &\sim \mathcal{N}_{\mathbb{C}} \left(\mathbf{0}, \tau_p p_k \mathbf{R}_{kl} \boldsymbol{\Psi}_{kl}^{-1} \mathbf{R}_{kl} \right) \end{aligned} \quad (2.4)$$

where $\mathbf{\Psi}_{kl} = \mathbb{E}\{\mathbf{y}_{t_{kl}}^p (\mathbf{y}_{t_{kl}}^p)^H\} = \sum_{i=1, t_i=t_k}^K \tau_p p_i \mathbf{R}_{il} + \sigma^2 \mathbf{I}_N$ represents the correlation matrix of the received pilot signal and \mathbf{R}_{kl} represents the spatial correlation matrix of \mathbf{h}_{kl} .

As for the LS estimator, the channel between UE k and AP l is estimated as

$$\hat{\mathbf{h}}_{kl} = \frac{\mathbf{Y}_l^p \boldsymbol{\theta}_{t_k}^*}{\tau_p \sqrt{p_k}} = \mathbf{h}_{kl} + \sum_{\substack{i=1 \\ t_i=t_k}}^K \sqrt{\frac{p_i}{p_k}} \mathbf{h}_{il} + \frac{\mathbf{n}_{t_{kl}}}{\sqrt{\tau_p p_k}}. \quad (2.5)$$

It is clear from (2.4) and (2.5) that the LS estimation offers reduced complexity compared to the optimal MMSE estimation, since it neither requires information about the spatial correlation matrices nor performing any matrix inversion for computing the channel estimates, at the expense of reduced quality of the acquired CSI.

2.2.5 Mobility Management

In Section 2.1.2, we have overlaid the basic concepts of handover in cellular networks where each AP is only responsible for the service of UEs within the boundaries of its cell. In this section, we discuss the main similarities and differences between cellular and cell-free networks, from the perspective of handover.

One of the main identifiers of cell-free networks is that a UE is served by a cluster of distributed APs. The idea is to eliminate poor cell-edge performance that is inherent to the cellular architecture through cooperation between the APs to jointly serve the UEs in a given coverage area. As a result, the cell-free architecture does not have the rigid boundaries of the cellular architecture. This calls for more advanced procedures to decide on how the serving cluster of APs and assigned pilot sequence to a UE is chosen at any given point in time. More importantly, considering UE mobility, the changes in the serving clusters and assigned pilots to the UEs should be kept to a minimal level, to limit the associated signalling overhead, while retaining the achievable SE of the network.

Similar to the case of cellular networks, undesired ping-pong handovers due to small-scale channel variations can be combated by a straightforward adoption of the TTT concept that is commonly used in the handover procedures of current cellular networks. Accordingly, TTT is not the focus of this work. We instead focus on handover procedures that are triggered by large-scale UE mobility, since pilot assignment and cluster update represent the primary difference between cell-free and conventional cellular networks, from a handover algorithm perspective [2].

The AP-UE associations rely on the average channel gains between the APs and UEs. The constantly developing UE locations due to mobility result in a change of the channel gains between the APs and UEs in the network, which should consequently change the AP-UE associations. Moreover, a limitation to the performance of a cell-free network is the level of pilot contamination, because it reduces the channel estimation quality at the APs and creates coherent interference [17]. Accordingly, the

main goal of a handover algorithm in a cell-free network should consider the following three primal factors: to reduce the number of handovers, necessary pilot assignment decisions and pilot contamination, while maintaining good SE performance for all the UEs in the network with distributed operation and reasonable computational complexity.

Again, we focus on distributed algorithms that require computations of local metrics at a given AP or UE with the local information measured at each of them, in order to reduce the required fronthaul signalling between the APs to perform the updates. Signalling is then only needed for the purpose of sending and receiving the decisions on the pilot assignment and cluster update between the relevant entities. On the other hand, centralized schemes rely on computing centralized metrics that include information from several locations within the network. This necessitates that several coefficients are required to be exchanged between the APs to first compute these metrics then communicate the resulting decisions, imposing increased signalling overhead. It is worth noting that the decisions need to be communicated at the handover rate, whereas communicating large-scale fading coefficients for computation of centralized metrics will be required at the rate of change of the coefficients which is expected to be much faster than the handover rate. In Paper II, we address this problem by devising a handover algorithm that works in a distributed manner such that a minimal number of pilot changes, which affects the update of the serving set of APs, is required for each UE.

2.2.6 Power Allocation

Another functionality that APs in a mobile network need to perform is power allocation. As stated in Section 2.1.3, in the case of a cellular network, there exist limited potential of the gains that may be achieved through employing an intelligent power allocation algorithm when compared to a cell-free network. In the cell-free topology, each UE is served by a cluster of APs. This entails that there exist more AP-UE links in the network increasing the amount of perceived interference at each UE. Moreover, the received desired signal power at any given UE may be achieved by different power allocations from the serving APs, i.e., the same received power level for the desired signal can be achieved through different combinations of transmitted power levels by the serving APs. Thus, for the same desired signal gain, the interference on other UEs at different parts of the network may be controlled by deciding on the power allocations from the serving APs to a given UE.

In a cell-free network, power allocation can be performed in a centralized or distributed manner. In the centralized case, the CPU is responsible for deciding on the allocated power from each AP to each UE. Although this may ultimately provide better performance, the metrics upon which the allocation is decided as well as the resulting power coefficients need to be communicated between the APs and the CPU. This imposes large fronthaul signalling requirements. On the other hand, a distributed implementation, where each AP is responsible for assigning its powers to the UEs it is serving using only the local information available at the AP,

does not require any fronthaul signalling for performing power allocation. Logically, the distributed operation may provide reduced performance due to the limited information available at a given AP about the channel gains and AP-UE links of the other APs in the network. To that end, this calls for developing an efficient distributed algorithm that has the ability to achieve relatively close performance to the centralized algorithms with reduced fronthaul requirements.

Centralized power allocation relies on maximizing a network-wide optimization objective [21, 37]. Among the widely adopted utility functions are the sum-SE and PF, which are presented later in Section 4.1.4. In a large scale network with many APs and UEs, the conventional approach of solving the optimization problems is computationally taxing and potentially violate real-time processing constraints. Another approach is to employ a learning-based solution to approximate the optimal powers [38–40]. With this approach, the optimization problem may be solved offline for training the learning model, and then the trained model can decide on the power allocations quickly without having to solve the optimization problem in real-time. However, the fronthaul signalling requirements remain for the centralized learning-based solutions as the learning model is still implemented at the CPU and utilizes information that needs to be gathered at the CPU from all over the network.

As for distributed power allocation in cell-free networks, previously these schemes relied on heuristic metrics, as in [28]. Such heuristic power allocation is usually based on intuitive choices about how the powers should be divided on the different UEs, yet they do not have any network-wide optimization goals. Since the network-wide optimizations utilize information that is gathered from different parts in the network, it is difficult to define an optimization procedure that can be performed in distributed manner. This entails that there is no obvious way of having a distributed implementation of the power allocation while still aiming to optimize a network-wide objective. In Paper I, we develop distributed/clustered learning-based procedures that can mimic centralized network-wide optimizations with a relatively small performance loss.

2.3 A Brief on Ergodic Capacity

The ergodic capacity is an important performance measure that defines reliable communication at the highest rate at which information can be sent over the channel with an arbitrarily small probability of error. This section provides the theoretical foundations behind the achievable SE expressions that are utilized throughout the thesis, and that represent lower bounds on the ergodic capacity.

2.3.1 Capacity of Multi-user MIMO Channels

Multi-user MIMO (MU-MIMO) systems play a fundamental role in modern wireless communications, where a BS equipped with multiple antennas serves multiple users simultaneously. The capacity region of the MU-MIMO DL channel, also known as

the broadcast channel (BC), assuming perfect channel knowledge is achieved using non-linear precoding such as Dirty Paper Coding (DPC).

Capacity Region with Dirty Paper Coding

Dirty Paper Coding is a non-linear precoding technique that pre-cancels known interference at the transmitter. Assuming perfect channel knowledge at the transmitter, the capacity region of the MIMO BC with M transmit antennas and K single-antenna UEs is given by [41]:

$$\mathcal{R} = \bigcup_{\substack{p_1, \dots, p_K \geq 0 \\ \sum_{k=1}^K p_k \leq P_{\max}}} \mathcal{R}_{p_1, \dots, p_K}^{\text{ul}}, \quad (2.6)$$

which is the union of virtual uplink regions with the powers p_1, \dots, p_K given by

$$\mathcal{R}_{p_1, \dots, p_K}^{\text{ul}} = \left\{ (R_1, \dots, R_K) : \sum_{k \in \mathcal{K}} R_k \leq R_{\mathcal{K}}^*, \forall \mathcal{K} \subseteq \{1, \dots, K\}, \text{ and } R_k \geq 0 \forall k \right\} \quad (2.7)$$

and

$$R_{\mathcal{K}}^* = \log_2 \det \left(\mathbf{I} + \sum_{k \in \mathcal{K}} \frac{p_k}{\sigma^2} \mathbf{h}_k \mathbf{h}_k^H \right), \quad (2.8)$$

where $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ is the channel vector between the transmitter and UE k , σ^2 is the Gaussian noise variance. p_k is the transmit power allocated to UE k such that $\sum_{k=1}^K p_k \leq P_{\max}$, where P_{\max} is the maximum DL power available at the transmitter. The optimal input distribution is Gaussian, and the capacity region can be achieved by different power allocation and successive encoding order across different UEs.

Achievable Rate with Linear Precoding

Although DPC achieves the capacity region, its practical implementation is complex. One of the main issues with DPC is the encoding delay which grows with the number of UEs, since encoding is done successively for the different UEs. Alternatively, linear precoding techniques are often employed to achieve suboptimal but computationally feasible solutions. Particularly when the transmitter has a relatively large number of antennas compared to the number of UEs, the loss in the achievable rates for the UEs due to restricting the precoding to be linear becomes small [41].

For linear precoding, the transmitted signal is given by:

$$\mathbf{x} = \mathbf{W}\mathbf{s}, \quad (2.9)$$

where $\mathbf{s} \sim \mathcal{CN}(0, \mathbf{I}_K)$ is the vector of transmitted symbols and $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_K]$ is the precoding matrix. The achievable rate for UE k for a particular (known)

realization of the channels \mathbf{h}_k , $\forall k$ is then given by:

$$R_k = \log_2 \left(1 + \frac{\rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{i=1, i \neq k} \rho_i |\mathbf{h}_k^H \mathbf{w}_i|^2 + \sigma^2} \right), \quad (2.10)$$

where ρ_k is the allocated power to UE k by the transmitter.

The DL rate region with linear processing is achieved with the optimal linear MMSE (LMMSE) precoding. The normalized precoding vector for UE k is then given by

$$\mathbf{w}_k = \frac{\bar{\mathbf{w}}_k}{\|\bar{\mathbf{w}}_k\|}, \quad (2.11)$$

where

$$\bar{\mathbf{w}}_k = \left(\sum_{i=1}^K p_i \mathbf{h}_i \mathbf{h}_i^H + \sigma^2 \mathbf{I}_M \right)^{-1} p_k \mathbf{h}_k. \quad (2.12)$$

Utilizing the LMMSE precoding and utilizing the UL-DL duality, an achievable rate for UE k is

$$R_k^{*\text{lin}} = \log_2 \left(1 + p_k \mathbf{h}_k \left(\sum_{i=1, i \neq k}^K p_i \mathbf{h}_i \mathbf{h}_i^H + \sigma^2 \mathbf{I}_M \right)^{-1} \mathbf{h}_k^H \right), \quad (2.13)$$

where p_k . The achievable rate region with linear precoding is then given by

$$\mathcal{R}^{\text{lin}} = \left\{ (R_1, \dots, R_K) : R_k = R_k^{*\text{lin}}, k = 1, \dots, K, \text{ for some } p_1, \dots, p_K \geq 0, \right. \\ \left. \text{and } \sum_{k=1}^K p_k \leq P_{\max} \right\}. \quad (2.14)$$

These linear precoding techniques offer a practical trade-off between performance and complexity, making them suitable for modern wireless systems where computational efficiency is essential.

Now, consider the same discrete memoryless channel with fast fading such that the channel realizations belong to an ergodic random process. If the channel realizations are known, then the ergodic channel capacity and ergodic achievable rates under linear processing are given as in (2.7) and (2.14), with the difference that the UE rates are averaged over the different channel realizations; that is, $R_\kappa = \mathbb{E} \{R_\kappa^*\}$ and $R_k = \mathbb{E} \{R_k^{*\text{lin}}\}$, respectively.

In practice, the transmitter will not have complete knowledge of the channel which makes the problem of defining the capacity region for MU-MIMO systems lack a simple closed-form expression. In the next subsections, we will present a

general method for obtaining achievable rates that represent a lower bound on the ergodic channel capacity, and that is independent of the utilized channel estimation technique and channel model.

2.3.2 Block Fading Model

The wireless channel is always changing due to the movement of objects and UEs in the wireless medium; but it can be treated as piecewise constant within a time interval called the coherence time. The channel coherence time is determined by the speed of movement and is typically a few milliseconds in mobile networks. It determines the time interval in which the channel can be approximated as a constant. Similarly, the channel coherence bandwidth defines the frequency interval in which the frequency response is approximately constant. It relies on the delay spread resulting from the path differences travelled by each path of the multi-path channel. The delay spread is defined as the time difference between the first and last propagation path.

Since there are many physical factors that affect both the coherence time and coherence bandwidth, it is hard to accurately determine their values in practice. By definition, the coherence time is the time after which substantial change in the channel is expected. As a rule of thumb, this time can be calculated by the time it takes a UE to move a quarter of the wavelength. Precisely, the coherence time can be approximated as [15]

$$T_c = \frac{\lambda}{4v_{\text{UE}}}, \quad (2.15)$$

where λ is the operating wavelength and v_{UE} is the UE speed. As for the coherence bandwidth, it can be considered as inversely proportional to the delay spread, where a suitable proportionality constant maybe chosen based on the propagation environment and application. A common approximation for the coherence bandwidth is

$$B_c = \frac{1}{2T_d}, \quad (2.16)$$

where T_d denotes the delay spread. By considering a sufficiently small time and frequency intervals, the concept of *block fading* has been introduced where the channels can be regarded as static and frequency-flat [16]. The time-frequency block with a time duration equal to the coherence time and a frequency width equal to the coherence bandwidth is referred to as the *channel coherence block*, denoted as $\tau_c = T_c B_c$. In this way, the channel between any two antennas in the network can be represented by a single complex-valued scalar within the coherence block.

Every UE has a different value of the coherence block depending on their speed and current location in the network. In order to have a common coherence block length, a practical solution is to fix the coherence block based on the worst-case

scenario that the network should support. In this thesis, we adopt the concept of coherence blocks in order to abstract away some technicalities about channel acquisition and focus on the algorithmic design of solutions to the problems stated earlier in Section 2.2.

2.3.3 Use-and-then-Forget Bounding Technique

The so-called use-and-then-forget (UatF) represents a lower bound on the ergodic channel capacity. The name comes from the fact that this method utilizes the channel estimates of the UEs in order to determine the precoding/combining vectors, as described in Section 2.2.4. However, it does not utilize this side information of the channel estimates in the decoding of the signals.

Let us consider the DL model of the cell-free network described in [1]. All UEs are assumed to be served by all the L APs on the same time-frequency resources. As a result, the received DL signal at UE k , $k = 1, \dots, K$, is given by

$$\begin{aligned} y_k^{dl} &= \sum_{l=1}^L \mathbf{h}_{kl}^H \sum_{i=1}^K \sqrt{\rho_{il}} \mathbf{w}_{il} s_i + n_k \\ &= \sum_{l=1}^L \sqrt{\rho_{kl}} \mathbf{h}_{kl}^H \mathbf{w}_{kl} s_k + \sum_{l=1}^L \sum_{\substack{i=1 \\ i \neq k}}^K \sqrt{\rho_{il}} \mathbf{h}_{kl}^H \mathbf{w}_{il} s_i + n_k, \end{aligned} \quad (2.17)$$

where $\rho_{il} \geq 0$ is the allocated DL power by AP l to UE i and $\mathbf{w}_{il} \in \mathbb{C}^{N \times 1}$ is the corresponding normalized precoding vector, such that $\|\mathbf{w}_{il}\|^2 = 1$. Moreover, s_i denotes the zero-mean signal intended for UE i and $n_k \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ represents the noise at UE k . In (2.17), the first term represents the signal intended for UE k and the second term is the inter-user interference due to the simultaneous transmissions by the APs to the UEs on the same time-frequency resources.

In the UatF method, only the part of the intended signal to UE k that is received over the average precoded channel, $\mathbb{E}\{\mathbf{h}_{kl}^H \mathbf{w}_{kl}\}$, is treated as the true desired signal. The part of the desired signal s_i received over the deviation from the mean value; that is $\mathbf{h}_{kl}^H \mathbf{w}_{kl} - \mathbb{E}\{\mathbf{h}_{kl}^H \mathbf{w}_{kl}\}$, has zero mean and can thus be treated as an uncorrelated noise signal in the detection. To elaborate, the received signal in (2.17) can be rewritten as

$$\begin{aligned} y_k^{dl} &= \underbrace{\sum_{l=1}^L \sqrt{\rho_{kl}} \mathbb{E}\{\mathbf{h}_{kl}^H \mathbf{w}_{kl}\} s_k}_{\text{Desired signal over average channel}} + \underbrace{\sum_{l=1}^L \sqrt{\rho_{kl}} \left(\mathbf{h}_{kl}^H \mathbf{w}_{kl} - \mathbb{E}\{\mathbf{h}_{kl}^H \mathbf{w}_{kl}\} \right) s_k}_{\text{Interference due to channel uncertainty}} \\ &+ \underbrace{\sum_{l=1}^L \sum_{\substack{i=1 \\ i \neq k}}^K \sqrt{\rho_{il}} \mathbf{h}_{kl}^H \mathbf{w}_{il} s_i}_{\text{Inter-user interference}} + \underbrace{n_k}_{\text{Noise}}. \end{aligned} \quad (2.18)$$

With this technique, the considered effective channel for the desired signal is deterministic. Employing the UatF capacity bounding method, a lower bound on the ergodic DL capacity over fast-fading channels with imperfect CSI in a cell-free system can be obtained. It is called an achievable SE and the SE expression for UE k is

$$\text{SE}_k = \frac{\tau_d}{\tau_c} \log_2(1 + \text{SINR}_k) \quad (2.19)$$

where τ_d denotes the portion of the coherence block that is dedicated for DL data transmission and SINR_k is the effective SINR of UE k . Using (2.18), the effective SINR is computed as

$$\text{SINR}_k = \frac{\rho_{kl} \left| \sum_{l=1}^L \mathbb{E} \{ \mathbf{h}_{kl}^H \mathbf{w}_{kl} \} \right|^2}{\sum_{i=1}^K \rho_{il} \mathbb{E} \left\{ \left| \sum_{l=1}^L \mathbf{h}_{kl}^H \mathbf{w}_{il} \right|^2 \right\} - \rho_{kl} \left| \sum_{l=1}^L \mathbb{E} \{ \mathbf{h}_{kl}^H \mathbf{w}_{kl} \} \right|^2 + \sigma^2} \quad (2.20)$$

Note that with the UatF technique, the numerator of the SINR contains only the desired signal over the average channel. Although this method may provide a looser bound on the capacity compared to other methods in the literature, it has the advantage of being applicable along with any channel estimation technique and fading model. In massive MIMO networks, the gap between the achievable SE computed using the UatF bound and the channel capacity becomes smaller as the channel hardens [16]; that is, $\frac{\sum_{l=1}^L \mathbf{h}_{kl}^H \mathbf{w}_{kl}}{NL} \approx \frac{\mathbb{E} \{ \sum_{l=1}^L \mathbf{h}_{kl}^H \mathbf{w}_{kl} \}}{NL}$. This along with the fact that the UatF concept provides more analytically tractable SE expressions has made it popular in massive MIMO networks. For cell-free massive MIMO networks, the level of channel hardening is expected to be less than their cellular counterpart, yet some level of channel hardening may still be achievable. In particular, the approximation above becomes tight, in general, when having a large number of antennas per AP and/or a large number of APs.

The variance of the inter-user interference in (2.20) might be changing rapidly due to scheduling variations that occur elsewhere in the network. We denote UEs that are being served by other BSs or AP clusters than that of a given UE k as unknown co-channel interferers. In practice, the large signalling overhead required prohibits keeping track of this rapidly changing interference. Particularly in the uplink, where the transmitting UEs' locations and antenna configurations cannot be predicted, identifying a suitable data rate is challenging. For that purpose, we develop a method for robust rate allocation with guaranteed predefined target outage for the uplink of a single cell massive MIMO system and a cell-free network in Papers III and IV, respectively.

Chapter 3

Physical-Layer Multicasting

Physical-layer multicasting presents an efficient transmission topology to exploit the beamforming capabilities at the transmitting nodes and broadcast nature of the wireless channel to satisfy the demand for the same content from several UEs. In multicasting, UEs are served using the same transmission and time-frequency resources, which avoids unnecessary co-channel interference between UEs requesting the same data. The difficulty is to find the suitable beamforming configuration that guarantees an acceptable minimum data rate, among the receiving UE group, to the multicast transmission.

It has received considerable attention since it can support various applications such as video streaming of live events to multiple subscribers over mobile networks, virtual reality, internet of things (IoT) applications, videoconferencing and broadcasting federated learning models where many UEs need to receive the same data content. A primary goal of multicasting is to improve energy efficiency and minimize unnecessary transmissions to reduce interference, thereby providing a remarkable improvement in radio resource utilization [29]. This makes the development of efficient multicast transmission techniques a vital ingredient in 6G networks.

3.1 Unicast, Broadcast and Multicast Transmissions

3.1.1 Unicast

Unicast transmission is a fundamental communication technique in wireless networks where a sender transmits data to a single, specific receiver. Unlike broadcast or multicast transmission, which deliver data to multiple UEs simultaneously, unicast transmission establishes a dedicated link between the transmitter and the intended recipient. This one-to-one communication model is widely used in applications such as voice calls, video streaming, and web browsing, where personalized data delivery is required.

One of the key advantages of unicast transmission is its ability to provide tailored communication with optimized resource allocation for the intended receiver.

Since each transmission is independent and dedicated to a single UE, network resources such as bandwidth and power can be dynamically adjusted based on QoS requirements and channel conditions of each receiver separately. This offers more flexibility for customized resource allocation to each of the individually served UEs. Despite its benefits, unicast transmission faces several challenges, particularly in high-demand wireless environments. The individualized nature of unicast communication leads to increased network congestion, especially in scenarios where a large number of UEs request high-bandwidth services.

In modern wireless communication systems, unicast transmission is implemented using various technologies, including Orthogonal Frequency Division Multiplexing (OFDM), MIMO, and adaptive modulation schemes [16]. These techniques enhance the robustness of the transmission by mitigating issues such as interference, fading, and congestion. With the evolution of wireless technologies towards 6G, unicast transmission continues to play a crucial role in enabling high-speed, low-latency communication. As wireless networks become more advanced, optimizing unicast transmission will remain a key focus for improving user experience and overall network performance.

3.1.2 Broadcast

Broadcasting involves sending the same data content to multiple UEs using the same transmission, typically without requiring interaction with them. This method can be used in applications such as video streaming of live events/news and software distribution/updates. By enabling large-scale dissemination of information, broadcasting represents an efficient solution for scenarios where the same data must reach multiple recipients simultaneously. In broadcasting, the same information can be transmitted to all UEs with no knowledge of their individual channel conditions. Moreover, the need for channel-specific beamforming is eliminated, streamlining the transmission process.

Since in broadcasting the APs do not target specific UEs, a reasonable strategy is to employ omnidirectional transmission to attempt at covering the entire area around each AP. This method operates without the need for CSI, making it particularly suitable for applications like emergency alerts or initial network synchronization where uniform coverage and reliability is critical, while limited information is available at the transmitting nodes.

Despite its simplicity, broadcasting suffers from several challenges. One major challenge is maintaining consistent desired signal quality for UEs experiencing different channel conditions. Effective power allocation is also essential to prevent coverage gaps and minimize interference if multiple broadcast transmissions occur simultaneously in the same region or in adjacent regions. A distinctive feature of broadcasting is the use of excess transmission power to ensure the weakest receivers get sufficient SNR. In addition, some broadcast transmissions may require special receiver equipment with high directional antennas pointed towards the transmitter for effective reception. Nevertheless, broadcasting remains a scalable and efficient

solution that requires no sophisticated CSI-based precoding and maintains broad accessibility.

3.1.3 Multicast

With the evolution of 6G networks, multicast transmission is becoming more advanced through the integration of technologies such as MIMO beamforming. These technologies help improve multicast efficiency and reliability, making it a crucial component of modern wireless communication systems. As the demand for high-quality content delivery grows, optimizing multicast transmission will remain essential for enhancing network performance and user experience.

Multicast transmission is a wireless communication technique where a sender transmits data to multiple selected receivers simultaneously. Unlike unicasting, which involves one-to-one communication, and broadcasting, which transmits data to all UEs within a network, multicasting provides an efficient way to deliver data to a selected group of UEs. A primary advantage of multicasting is its efficient use of network resources, since the same data is transmitted to multiple UEs simultaneously instead of having multiple unicast transmissions delivering identical content. This makes multicasting particularly beneficial in scenarios requiring large-scale content delivery, such as streaming live events, communicating federated learning models to multiple UEs or updating software across multiple devices [5].

Physical-layer multicasting and application-layer multicasting differ in their approach to data distribution within a network. Physical-layer multicasting operates at the lower layers of the network stack, leveraging techniques such as beamforming to transmit data efficiently to multiple UEs on the same time-frequency resources using a single transmission. This method reduces redundant transmissions, making it highly efficient when the same data content is required by many UEs within a network area. On the other hand, application-layer multicasting functions at the higher layers, using software-based approaches to replicate and distribute data through unicast connections or overlay networks. While it offers greater flexibility, it often results in higher consumption of network resources and increased latency due to duplicate transmissions. The choice between these two approaches depends on factors such as network scalability, efficiency requirements and infrastructure capabilities. In this thesis, we focus on physical-layer multicasting. Thereafter, whenever the term multicasting is used, it refers to the physical-layer based multicasting.

Through CSI-based precoding, multicasting can overcome the fundamental disadvantages of broadcasting, such as coping with different and varying channel conditions among the UEs. It allows for providing QoS guarantees to the targeted receivers, which is not possible with a simple broadcast transmission [42]. Moreover, it may offer higher energy efficiency since it avoids wasting transmission power in areas that do not contain the targeted UEs. These plausible gains come at the expense of increased system and computational complexity, which should be taken carefully into consideration when designing a large-scale data distribution platform.

Despite its advantages, multicast transmission in wireless networks faces several challenges. One major challenge is coping with varying channel conditions among different receivers, since a multicast transmission is limited by the channel conditions of all UEs receiving it. Compared to unicasting, where resources can be optimized for a single receiver, multicasting must consider the weakest link in the multicast group, which can lead to performance limitations. This calls for advanced beamforming optimization in order to boost the signal quality of all UEs receiving the multicast transmission. Another aspect is security concerns as multicast transmissions can be more susceptible to unauthorized access, requiring encryption and authentication mechanisms to safeguard the data.

3.2 Multicast Beamforming Optimization

Unlike conventional isotropic broadcasting, in multicasting, optimization of the transmitter beamforming can be done to enhance the received signal at the intended multicast receivers when CSI is available at the transmitter. With this technique, multicast beamforming optimization can adapt the precoding at the transmitting nodes to achieve different utility maximization objectives and provide QoS guarantees to the intended receivers. In this thesis, we focus on the MMF and QoS utility functions that represent suitable optimization objectives for the multicast problem. The reason is that since all UEs receiving the multicast transmission need to be able to decode the data, the minimum UE rate in a UE group constitutes the bottleneck for multicast transmissions [43].

As with unicast transmission, multicast beamforming has to take into account the following three components: 1) Interference mitigation, 2) Boosting the desired signal, and 3) Power allocation. With regard to interference mitigation, an advantage of multicasting over unicasting is that it avoids unnecessary co-channel interference within the group of UEs requesting the same data content. Other than this, the mitigation of interference from other multicast or unicast transmissions is done in a similar manner to the unicast case. The primary difference between unicast and multicast precoder design is that in multicasting, each precoder is responsible for boosting the desired signal gain for all UEs receiving the multicast transmission, which requires additional ministration in the precoder design. As such, the component that is responsible for boosting the desired signal of the entire multicast UE group should consist of a weighted sum of all channels within the group. This conforms with the optimal multicast beamforming structure derived in [44].

Consider a single transmitter equipped with N antennas and K single-antenna UEs, that are arbitrarily distributed in a large service area, served with a single multicast transmission. We assume a narrow-band channel, such that each channel realization is frequency-flat and quasi-static in time. The channel realizations are assumed to be available at the transmitter. The channel between UE k and the transmitter, normalized by the noise power, is denoted as $\mathbf{h}_k \in \mathbb{C}^N$. The received

signal at UE k is computed as

$$y_k^{\text{dl}} = \mathbf{h}_k^H \mathbf{w} s + n_k, \quad (3.1)$$

where s denotes the zero-mean unit-variance multicast signal intended for all UEs, \mathbf{w} represents the common multicast precoding vector, and $n_k \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ is the normalized Gaussian noise at UE k . As a result, an achievable SE of the k^{th} UE under the perfect CSI assumption can be evaluated as

$$\text{SE}_k^{\text{dl}} = \log_2 \left(1 + \text{SNR}_k^{\text{dl}} \right), \quad (3.2)$$

where the SNR of UE k is given by $\text{SNR}_k^{\text{dl}} = |\mathbf{h}_k^H \mathbf{w}|^2$.

3.2.1 QoS and MMF Problem Formulations

The objective of the QoS multicast problem is to minimize the total transmitter power, subject to some predefined SNR requirement at the intended multicast receivers. The QoS problem can thus be formulated as

$$\underset{\mathbf{w}}{\text{minimize}} \quad \|\mathbf{w}\|^2 \quad (3.3a)$$

$$\text{s.t.} \quad |\mathbf{h}_k^H \mathbf{w}|^2 \geq \gamma, \quad \forall k \in \{1, \dots, K\}, \quad (3.3b)$$

where γ represents the SNR constraint for all UEs receiving the multicast transmission. This problem is a non-convex quadratically constrained quadratic program (QCQP), which is NP-hard in general [42, 45–47]. The difficulty in solving this problem lies in the non-convex SNR constraints.

For the MMF problem, the objective is to maximize the minimum SNR among the multicast UEs subject to a total power constraint at the transmitter, which can be viewed as the inverse of the QoS problem. We assume that the multicast precoding vector satisfies a short-term power constraint, which means that the power constraint must be satisfied for each channel realization. Accordingly, $\|\mathbf{w}\|^2 \leq P_{\max}$, where P_{\max} represents the maximum transmitter power. The MMF multicast problem can then be formulated as

$$\underset{\mathbf{w}}{\text{maximize}} \quad \min_{\forall k \in \{1, \dots, K\}} \quad |\mathbf{h}_k^H \mathbf{w}|^2 \quad (3.4a)$$

$$\text{s.t.} \quad \|\mathbf{w}\|^2 \leq P_{\max}. \quad (3.4b)$$

By establishing the equivalence between the QoS and MMF problems, it can be deduced that the corresponding MMF problem is also NP-hard [42, 47]. In the next subsection, we present two powerful techniques to find approximate solutions to these non-convex problems.

3.2.2 Solution Approaches

The advancements in multicast beamforming research highlight the increasing importance of efficient transmission techniques for emerging wireless applications. Several methods have been proposed to approximate non-convex QCQPs. Primarily, two relaxation techniques were introduced to solve these non-convex QoS and MMF problems.

The first approach to solving these problems was proposed in [42]. The idea is to cast the problem as an equivalent SDP and then utilize SDR to solve the relaxed convex problem by dropping the non-convex rank constraint. By defining the notation $\mathbf{H}_k = \mathbf{h}_k \mathbf{h}_k^H$ and $\mathbf{W} = \mathbf{w} \mathbf{w}^H$, and utilizing the fact that $|\mathbf{h}_k^H \mathbf{w}|^2 = \text{tr}(\mathbf{h}_k \mathbf{h}_k^H \mathbf{w} \mathbf{w}^H)$, the SNR of UE k is equivalent to

$$\text{SNR}_k^{\text{dl}} = \text{tr}(\mathbf{H}_k \mathbf{W}). \quad (3.5)$$

This SNR formulation is suitable to construct the SDP of the QoS and MMF multicast problems. The QoS problem is then reformulated as

$$\underset{\mathbf{W}}{\text{minimize}} \quad \text{tr}(\mathbf{W}) \quad (3.6a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{H}_k \mathbf{W}) \geq \gamma, \quad \forall k \in \{1, \dots, K\}, \quad (3.6b)$$

$$\mathbf{W} \succeq \mathbf{0}, \quad \text{rank}(\mathbf{W}) = 1. \quad (3.6c)$$

By dropping the non-convex rank-1 constraint, the relaxed convex SDP problem can be solved using a standard solver.

For the single multicast group case considered here, a similar SDP formulation is possible to directly solve the MMF problem. The MMF problem can be reformulated as

$$\underset{\mathbf{W}}{\text{maximize}} \quad \min_{\forall k \in \{1, \dots, K\}} \text{tr}(\mathbf{H}_k \mathbf{W}) \quad (3.7a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{W}) \leq P_{\text{max}} \quad (3.7b)$$

$$\mathbf{W} \succeq \mathbf{0}, \quad \text{rank}(\mathbf{W}) = 1. \quad (3.7c)$$

Likewise, the non-convex rank-1 constraint is dropped to obtain the relaxed convex SDP problem that can be solved with a standard solver. Alternatively, the relation between the QoS and MMF problems can be exploited to solve the MMF problem by means of solving multiple instances of the QoS problem. The idea is based on the fact that the optimal objective values of the relaxed MMF and QoS problems are monotonically non-decreasing in the transmitter power budget and the SNR target, respectively [29, 48]. For that reason, a bisection search is performed over that SNR target such that the resulting required power is equal to the available power budget at the transmitter. This alternative solution is necessary in the case of multigroup multicasting, where the relaxed MMF problem reformulation is only quasi-convex; that is, the problem is only convex given a specific lower bound on the SNR target.

The SDR approach results in a solution matrix that, with high probability, violates the rank constraint of the original non-convex problem, and so provides an infeasible solution. If the matrix does have rank one, then its principal component represents the optimal solution to the original problem. Otherwise, the matrix provides a lower bound on the required power to meet the constraints, as one of the original problem's constraints has been removed. Post-processing of the output solution matrix is thus essential to find a feasible rank-1 solution. Advancements in optimization have introduced techniques for deriving high-quality solutions to the original problem. These methods rely on randomization, where the matrix is used to generate a set of candidate beamforming vectors, from which the solution that results in the highest objective value is selected. There are two main problems with this approach. The first problem is that the approximation quality deteriorates quickly as the system size grows large. The second problem is when considering multiple multicast groups, a power control optimization problem needs to be solved for each candidate set of beamforming vectors and a large number of candidate sets needs to be tested for a relatively good approximation quality leading to exorbitant computational complexity.

Recent work in the area of multicast beamforming optimization suggested utilizing the SCA technique to find a solution to the QoS and MMF problems iteratively. The main idea of this approach is to substitute the non-convex SNR constraints with tighter convex constraints, and solve the convex subproblem repeatedly until convergence to a local optimum of the original non-convex problem. First, the SNR of the k^{th} UE can be rewritten as

$$\text{SNR}_k^{\text{dl}} = |\mathbf{h}_k^H \mathbf{w}|^2 = \mathbf{w}^H \mathbf{H}_k \mathbf{w}. \quad (3.8)$$

Note that $\mathbf{H}_k, \forall k$ are rank-1 positive semidefinite matrices. Accordingly, for any arbitrary vector $\mathbf{v} \in \mathbb{C}^N$, the following result holds

$$(\mathbf{w} - \mathbf{v})^H \mathbf{H}_k (\mathbf{w} - \mathbf{v}) \geq 0, \quad \forall k \in \{1, \dots, K\}. \quad (3.9)$$

Reordering the terms, we have

$$\mathbf{w}^H \mathbf{H}_k \mathbf{w} \geq 2\Re(\mathbf{v}^H \mathbf{H}_k \mathbf{w}) - \mathbf{v}^H \mathbf{H}_k \mathbf{v}, \quad \forall k \in \{1, \dots, K\}. \quad (3.10)$$

Now, the non-convex SNR constraint $|\mathbf{h}_k^H \mathbf{w}|^2 \geq \gamma$ can be replaced by the stricter convex constraint given by

$$2\Re(\mathbf{v}^H \mathbf{H}_k \mathbf{w}) - \mathbf{v}^H \mathbf{H}_k \mathbf{v} \geq \gamma. \quad (3.11)$$

An approximation of the QoS problem can then be formulated as

$$\underset{\mathbf{w}}{\text{minimize}} \quad \|\mathbf{w}\|^2 \quad (3.12a)$$

$$\text{s.t.} \quad 2\Re(\mathbf{v}^H \mathbf{H}_k \mathbf{w}) - \mathbf{v}^H \mathbf{H}_k \mathbf{v} \geq \gamma, \quad \forall k \in \{1, \dots, K\}, \quad (3.12b)$$

which is a convex optimization problem given a specific instance of \mathbf{v} . Starting from an initial feasible solution, a sequence of the convex problem is solved such that the vector \mathbf{v} is set to the solution of the convex problem of the previous iteration. This approach guarantees convergence to a local optimum solution of the original non-convex QoS problem. A similar procedure can be developed to directly solve the MMF problem with this technique, where in each iteration the following convex approximation problem is solved:

$$\underset{\mathbf{w}}{\text{maximize}} \quad \min_{\forall k \in \{1, \dots, K\}} \quad 2\Re\left(\mathbf{v}^H \mathbf{H}_k \mathbf{w}\right) - \mathbf{v}^H \mathbf{H}_k \mathbf{v} \quad (3.13a)$$

$$\text{s.t.} \quad \|\mathbf{w}\|^2 \leq P_{\max}. \quad (3.13b)$$

As in the case of the QoS problem, by solving a sequence of the convex approximation problem, a local optimum solution can be achieved for the original non-convex MMF problem. This approach represents the state-of-the-art in multicast beamforming optimization, however, it generally suffers from two limitations. The first is that it can only guarantee convergence to a local optimum. The second is that it requires initialization with a feasible solution. A good initial feasible point can be challenging to obtain particularly for large systems. In Papers V and VI, we develop a novel optimization procedure that addresses these limitations and provides a reduced computational complexity compared to the state-of-the-art. Additionally, we propose a high-performing fast heuristic algorithm for the problem to bring the computational complexity down to the milliseconds range.

To emphasize the difference between the unicast and multicast beamforming optimization, we note that for the unicast case, the corresponding QoS and MMF problems will have all constraints met with equality at the optimum, even for multiple simultaneous unicast transmissions. On the other hand, it is very likely that most constraints will be over-satisfied for the multicast case [42]. This is particularly true for a cellular network where the common information multicast signal is transmitted from a single source. The main reason is that there exist large SNR differences between UEs belonging to the same multicast transmission. To exemplify this, consider two UEs with the same channel direction but different channel gains. In order to satisfy a certain received SNR constraint at the weaker UE, the UE with the stronger channel gain must over-satisfy that constraint. This will be different in the case of having multiple transmitting nodes to the multicast signal.

3.3 Cell-Free Massive MIMO Multicasting

Unlike traditional cellular systems, where transmission is coordinated by a single BS within a cell, cell-free massive MIMO employs a large number of distributed APs to jointly serve all UEs without cell boundaries. The distributed nature of APs enhances spatial diversity, which improves the likelihood of successful reception even under challenging channel conditions. In the context of multicasting, the APs

collaboratively transmit a common message to a group of UEs, such as software updates or federated learning model updates, by exploiting beamforming techniques. However, since UEs experience different channel gains, the system must optimize the beamforming weights to balance performance across all UEs while minimizing power consumption.

As stated earlier, the UE having the minimum SINR within a multicast group limits the rate that can be utilized for a multicast transmission, since every UE in the group must be able to decode the data. Moreover, a selling point for the cell-free architecture is its ability to provide uniform coverage throughout the network area. This is achieved since the available power for the multicast transmission is better distributed across different APs in the coverage area such that every UE will have a relatively good channel gain from a subset of the APs. In other words, the cell-free network architecture mitigates the large pathloss variations between the multicast UEs, that is inherent in the cellular architecture. As a result, it can be more likely that multicast QoS constraints are met with equality at the optimum due to spatial-macro diversity. This verifies the potential of cell-free massive MIMO to deliver uniform service within the mobile network, which can particularly enhance the performance of multicast transmissions.

We highlight that the solution methodology for the multicast problems presented in this chapter are not specific to the network architecture. However, we emphasize the synergy between the cell-free network architecture and multicast transmission. This can be described as the necessary condition for every UE within a multicast group to satisfy a certain SNR requirement can be more easily met with the uniform service quality that is achievable with the cell-free network architecture.

From another perspective, beamforming design can be more complicated in a cell-free massive MIMO network, as compared to a cellular massive MIMO network, for the same reason that adheres to the more uniform coverage of cell-free networks. That is, the difference in the average channel gains of the serving APs to a given UE makes the design of a precoder that aims to approach the QoS/MMF objective particularly challenging. On the one hand, each AP should try to equalize the desired signal gains of the multicast UEs. On the other hand, further away UEs should not be given more power by an AP, but rather be focused on by other closer by APs. This necessitates additional care especially in the design of heuristic precoding methods [5].

Chapter 4

Summary of Included Papers

In this chapter, we summarize the main findings of the four studies, stated in Section 1.4, carried out in this doctoral thesis. The detailed analyses can be found in the list of papers included at the beginning of the document.

4.1 Paper I - Learning-Based Downlink Power Allocation in Cell-Free Massive MIMO Systems

In a cell-free massive MIMO network, the different UEs are served simultaneously on the same time-frequency resource by several APs in the coverage area. For that reason, power allocation plays an important role in limiting MU interference and optimizing the network performance [16, 49]. The role of power allocation in a cell-free network can be thought of as a way to control the signal strengths for desired and interfering signals at different parts in the network. Since the desired signal is a combination of signals from different sources, achieving a desired signal level can be done through countless possible transmission levels from those sources. Some of which may be preferred over others in order to limit the interference at a given part of the network.

4.1.1 Introduction

In the network-wide power allocation schemes that are conventionally used in cell-free massive MIMO systems, the CPU is responsible to select the power allocated to each UE by each AP using the channel statistics that are shared from the APs. The rate with which new power allocation decisions need to be made and communicated to the APs is determined by the users' large-scale mobility and scheduling decisions in the network. Optimally, the system must be capable of making new power allocation decisions at the same frequency as the scheduling decisions can change, which happens at the subframe level (1 ms) in 5G. Large-scale UE mobility also requires changes in power allocation but this happens more slowly. Ideally, the

transmit powers should also be selected locally at each AP and this study develops efficient algorithms for that purpose.

4.1.2 Related Work

It was noticed in [37] that the network-wide DL max-min fairness problem is quasi-convex, and the authors solved it using general-purpose solvers. More efficient dedicated algorithms for network-wide DL max-min fairness and sum-SE maximization were developed in [21]. However, the computational complexities of the algorithms grow polynomially with the number of APs and UEs, thus resulting in unscalable solutions that potentially violate real-time processing constraints. An alternative solution to the power allocation problem is to employ the “learn to optimize” approach which makes use of the fact that DNNs can learn rich patterns and approximate complex function mappings [27, 38–40, 50]. This allows for a significant reduction in computational complexity compared to solving the optimization problems with traditional methods, resulting in a solution that is real-time implementable.

The authors in [51] develop a DNN for joint pilot and data power control for sum-SE maximization in the uplink (UL). In [49], a centralized deep convolutional neural network is designed to approximate the power coefficients for the max-min fairness objective in the cell-free massive MIMO DL. Moreover, [26] develops another centralized deep learning solution for the sum-SE maximizing power control in the UL of a cell-free system, taking into account the limited-capacity fronthaul links between the APs and the CPU. In [27], centralized and distributed DNNs were developed for the max-min fairness power allocation in a small cell-free massive MIMO setup. Especially in the distributed case, the results show that the approximation performance was not satisfactory, likely due to two main factors: 1) the inputs were not pre-processed based on prior domain knowledge of what information is essential for power allocation; and 2) the total transmit power from each AP was not approximated by the DNN, which is essential in the distributed case.

4.1.3 Contribution

We propose a distributed solution to the DL power allocation problem in a cell-free network. We first formulate the network-wide sum-SE and PF-based power allocation optimization problems. To solve these non-convex problems, we employ a weighted minimum mean square error (WMMSE) algorithm inspired by the previous work [21]. It is worth mentioning that PF-based power allocation was not considered in [21]. We take an ADMM approach to efficiently obtain the solution to the quadratically-constrained quadratic sub-problems of the WMMSE algorithms. This allows for generating a large amount of training data in a reasonable time. We then train feedforward DNNs to approximate the respective power coefficients. To the best of our knowledge, this is the first work to provide distributed DNN solutions

to the sum-SE and PF power allocation problems in the context of cell-free massive MIMO. The main contributions can be summarized as follows:

- We develop a fully distributed feedforward DNN for each AP to approximate the per-AP normalized power coefficients and the total transmitted power from the AP, using only local information available at the AP as input, however, utilizing the network-wide sum-SE and PF solutions in the training phase as the labeled output. Accordingly, each AP employs a DNN that implicitly learns the essential network structure and propagation environment. A heuristic closed-form power allocation, based on the LSF coefficients between the AP and the UEs, is used as the input to the DNN. The heuristic input is chosen to make the input/output relation of the DNN model less complex.
- We develop another distributed DNN that utilizes side information in the form of the ratio of the LSF coefficients from a given AP, to the LSF coefficients from all APs in the network for performance improvement. We further implement a clustered DNN where the LSF parameters of a small number of APs is fed to a common DNN to simultaneously learn the “optimal” power allocation of the cluster. This architecture provides better performance at the expense of increased fronthaul, compared to the fully distributed case.
- We perform a complexity analysis of the required processing and run-time for the proposed DNN models in comparison to the conventional methods of solving the power allocation optimization problems.

4.1.4 Sum-SE and PF Optimization Objectives

The sum-SE maximizing power allocation problem can be expressed as

$$\begin{aligned} & \underset{\{\mu_{kl}:\forall k,l\}}{\text{maximize}} && \sum_{k=1}^K \log_2(1 + \text{SINR}_k) \\ & \text{subject to} && \sum_{k=1}^K \mu_{kl}^2 \leq P_{\max}^{\text{dl}}, \quad l = 1, \dots, L, \end{aligned} \quad (4.1)$$

whereas the PF optimization problem is expressed as

$$\begin{aligned} & \underset{\{\mu_{kl}:\forall k,l\}}{\text{maximize}} && \sum_{k=1}^K \ln(\log_2(1 + \text{SINR}_k)) \\ & \text{subject to} && \sum_{k=1}^K \mu_{kl}^2 \leq P_{\max}^{\text{dl}}, \quad l = 1, \dots, L, \end{aligned} \quad (4.2)$$

with SINR_k being the signal-to-interference-plus-noise ratio (SINR) of the k^{th} UE, μ_{kl} is the square root of the allocated power to UE k by AP l and P_{\max}^{dl} is the per AP power budget.

Table 4.1: Network simulation parameters.

Area of interest (wrap around)	1000 m \times 1000 m
Bandwidth	20 MHz
Number of APs	$L = 16$
Number of UEs	$K = 20$
Number of antennas per AP	$N = 4$
Pathloss exponent	$\alpha = 3.76$
Per-AP maximum DL transmit power	$P_{\max}^{\text{dl}} = 1 \text{ W}$
UL transmit power	$p_i = 100 \text{ mW}$
UL/DL noise power	-94 dBm
Coherence block length	$\tau_c = 200$
Pilot sequence length	$\tau_p = 10$

4.1.5 Selected Results

In this section, we evaluate the performance of the proposed distributed learning-based power allocation solutions. The network simulation parameters are summarized in Table 4.1. In the following, the fully distributed DNN will be referred to as “DDNN”, the distributed DNN with side information as “DDNN-SI” and the clustered DNN as “CDNN”.

The estimated run-time of the DNN models is significantly lower than the conventional approach of solving the optimization problem using the WMMSE-ADMM algorithm. In Table 4.2, we recorded the average run-time, for 100 samples, of the WMMSE-ADMM algorithm and the DNN models for the processing schemes and optimization objectives presented earlier. It can be seen that the DNN models satisfy real-time processing constraints that mainly rely on scheduling decisions in the network, which can be in the order of 1 ms. Note that the recorded run-times for the learning-based solutions represent the total time taken by all the distributed/clustered DNNs implemented to support the entire cell-free network, i.e., the computational time for each DNN is less than 1 ms. It is clear that the CDNN provides the least (total) computational complexity by facilitating a common DNN for the APs in each cluster. We use the same platform, a 4 core Intel(R) Core i5-10310U CPU with 1.7 GHz base frequency and 4.4 GHz max turbo frequency. The programs are all written in Python 3.8.

To shed the light on the performance difference between the proposed deep learning solutions and practical LSF-based heuristic allocation strategies, as well as the gap to the WMMSE-ADMM benchmark, Figure 4.1 plots the CDF of the total DL SE for sum-SE maximization with RZF precoding. It can be seen that the proposed distributed and clustered DNNs provide a significant improvement over the state-of-the-art heuristic in [28]. Note that the DDNN-SI model achieves almost the same total DL SE as the CDNN model, and is thus omitted from the figure. Moreover, the average performance loss of the practical CDNN solution,

Table 4.2: Computational time in milliseconds for the proposed models.

Algorithm	Sum-SE		PF	
	MR	RZF	MR	RZF
ADMM	88.6	131.7	124.8	163.5
DDNN	8.8	9.2	8.6	9.0
DDNN-SI	9.7	9.8	9.5	9.7
CDNN	3.2	3.2	3.2	3.2

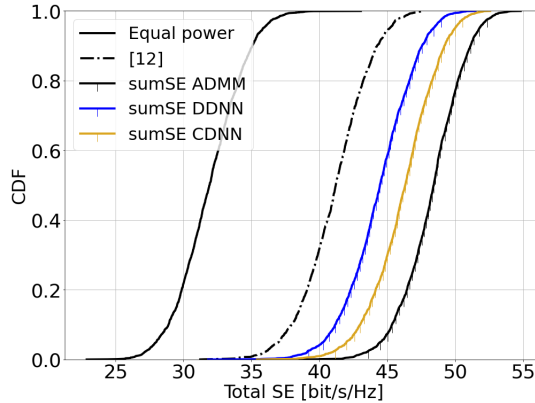


Figure 4.1: Comparison of power allocation strategies.

providing reduced fronthaul and run-time requirements, compared to the centralized WMMSE-ADMM algorithm is only about 5%. This result shows that our proposed learning-based solutions yield a good balance between system performance and feasibility for practical implementation.

Further, to demonstrate the robustness of the learning-based solutions, Figure 4.2 compares the CDF of the total DL SE with RZF for the original DDNN and CDNN models with perturbed LSF inputs. A truncated log-normal perturbation with a standard deviation of 1 dB, and truncated at 3 dB is chosen to model small variations in the propagation environment (for example due to vehicles and trees). Note that the DNNs are not retrained with the perturbed input in order to test their robustness against unknown variations that may occur in a practical cell-free network. It is clear that the input perturbation does not have a noticeable effect on the fully distributed DDNN model, whereas the clustered model shows a minor degradation in performance. The reason behind this is that the clustered model has a larger input vector including the LSF coefficients of the UEs to different APs, and so the perturbations here affect as well the inter-relations between the LSF coefficients of a given UE to the APs in the cluster.

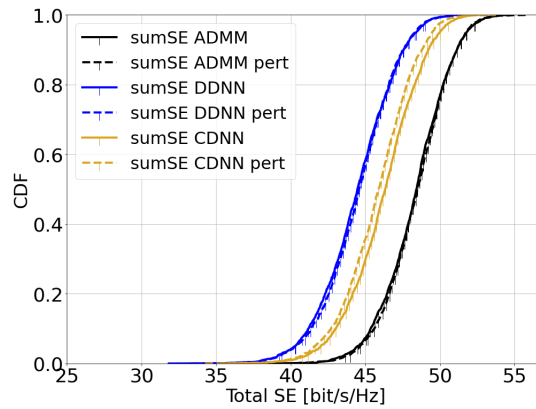


Figure 4.2: Robustness against perturbations.

4.2 Paper II - Soft Handover Procedures in mmWave Cell-Free Massive MIMO Networks

4.2.1 Introduction

In this study, we address UE mobility in the DL with imperfect CSI and pilot training. Aiming at extending traditional handover concepts to the challenging AP-UE association strategies of cell-free networks, distributed algorithms for joint pilot assignment and cluster formation are proposed in a dynamic environment considering UE mobility. The algorithms provide a systematic procedure for initial access and update of the serving set of APs and assigned pilot sequence to each UE. The principal goal is to limit the necessary number of AP and pilot changes, while limiting computational complexity. The pseudo-code for the proposed handover algorithms and mobility model are detailed in [2].

4.2.2 Related Work

To enable a scalable network operation, previous works have proposed UC clustering in which each UE is served by only a subset of the APs in the network [17, 25]. The general idea is to allow each AP to serve the UEs with the strongest estimated channels, and/or each UE to be served by APs that have the strongest channels to that UE.

The development of properly designed pilot assignment algorithms is pivotal to ensuring good performance in highly loaded networks [52]. Multiple matching algorithms to find the optimal pilot assignment either based on selected UC clusters or predetermined virtual clusters that a UE can be served by are presented in [52]. In [53], a dynamic pilot reuse scheme was developed such that a maximum of two UEs are allowed to share the same pilot sequence, under the condition that they have disjoint sets of serving APs. Pilot assignment schemes based on K-means clustering are proposed in [54]. However, these schemes rely on maximizing the minimum distance between UEs sharing the same pilots, thereby neglecting practical shadow fading effects. The effectiveness of such schemes is questionable in mmWave bands that are susceptible to line-of-sight (LOS) blockage.

While all the mentioned algorithms are developed and tested in static conditions, the impact of changing the serving set of APs for each UE due to mobility is not addressed. In [55], the first attempt for a handover algorithm in mmWave cell-free networks is presented based on stochastic channel modeling without channel estimation. Accordingly, only cluster formation and update were required with no developments to deal with pilot assignment/contamination.

4.2.3 Contribution

In this work, we address mobility management in the DL of a mmWave cell-free massive MIMO network. We consider fully digital beamforming based on imperfect

channel state information obtained from pilot training. We notice that reliability and handover are particularly important in mmWave networks, where APs have short range. An advantage of our proposed algorithm is that it does not require any centralized control allowing for a fully distributed operation, i.e., it is implemented at the AP level. To evaluate the performance of the proposed algorithm, we utilize ray-tracing software from [30] for mmWave operation to obtain realistic and spatially consistent channel modeling environment. The main contributions of the work can be summarized as follows:

- We propose a distributed algorithm for joint pilot assignment and cluster formation that identifies the initial AP cluster and assigned pilot sequence to each UE, where only UEs with disjoint AP clusters are allowed to share the same pilot sequence, to reduce the level of pilot contamination in the network.
- We propose a distributed algorithm for dynamic pilot assignment and cluster update, considering UE mobility, to extend traditional handover concepts to the more complicated AP-UE associations of the cell-free network topology. The algorithm is tailored to minimize the number of *Master* AP changes and necessary pilot assignment decisions, and to facilitate massive access to the network.
- We propose a modified pilot assignment strategy that takes into account the load distribution (i.e., used pilot sequences) of the possible serving APs that are in the vicinity of a given UE, showing improved performance, especially in highly loaded scenarios.
- We develop a *Maneuvering Smooth Random Waypoint Mobility* model to evaluate the performance of our proposed algorithms with realistic random UE mobility in a site map including buildings and other obstructions.

4.2.4 Channel Model

For scalable network operation, each UE is jointly served by a subset of the APs in the network. We assume a standard block fading channel model where the time-varying wideband channels are divided into time-frequency coherence blocks such that the channels are static and frequency-flat in each block [16]. The channels are updated following the update of the UE location due to mobility. A ray-tracing software [30] is utilized to provide realistic and spatially consistent channel realizations suitable for mmWave bands. The ray-tracer employs specular reflections upon the different objects in the site map to determine the path gains from AP l , $l = 1, \dots, L$, to each UE location. To create the channel coefficients in each coherence block, the path gains are combined such that the channel $\mathbf{h}_{kl} \in \mathbb{C}^{N \times 1}$ between UE k and AP l is

$$\mathbf{h}_{kl} = \sum_{i=1}^{N_{kl}} \alpha_i e^{-j\psi_i} \mathbf{a}(\phi_i), \quad (4.3)$$

where α_i represents the gain of path i , $\psi_i \sim \mathcal{U}[0, 2\pi)$, $\mathbf{a}(\phi_i)$ is the array response vector of AP l with ϕ_i being the angle of departure (AoD) of the i^{th} path, and N_{kl} corresponds to the number of paths from AP l to the location of UE k . We consider a 3D environment with a uniform linear array (ULA) horizontally placed at each AP with half-wavelength spacing between the antenna elements, such that

$$\mathbf{a}(\phi_i) = [1 \ e^{-j\pi \sin \phi_i} \ e^{-j2\pi \sin \phi_i} \ \dots \ e^{-j(N-1)\pi \sin \phi_i}]^T. \quad (4.4)$$

The average channel gain from an antenna at AP l to UE k is given by $\beta_{kl} = \sum_{i=1}^{N_{kl}} \alpha_i^2$.

4.2.5 Selected Results

In this section, we evaluate the performance of the proposed distributed pilot assignment and cluster update algorithm under UE mobility. We consider a cell-free network comprised of $L = 64$ APs deployed in an area of $600 \text{ m} \times 600 \text{ m}$. Only the inner $500 \text{ m} \times 500 \text{ m}$ are considered for generating the results, to reduce boundary effects. We assume K UEs are randomly and uniformly dropped within that area, with each UE being given a uniform random initial target location that lies within a distance of $d_{\text{seg}} \sim \mathcal{U}[50, 100] \text{ m}$ from the initial UE location. The simulation parameters are summarized in Table 4.3.

Figure 4.3 presents the considered site map with buildings/obstructions and $K = 5$ moving UEs over a duration of $400 T_i$ seconds. The results are averaged over 25 random initial UE locations such that for each random drop, UEs are allowed to move over a duration of 400 time intervals, i.e., $400 T_i$ seconds, according to the *Maneuvering Smooth Random Waypoint Mobility* devised in [2]. The APs are deployed at a height of 6 m (e.g., on utility poles and building facades), whereas UEs move along the same horizontal plane at a height of $h_{\text{UE}} = 1.5 \text{ m}$, with a constant speed of $v_{\text{UE}} = 1.5 \text{ m/s}$ representing handheld devices by pedestrians. The maximum cluster size is selected as $M_{\text{max}} = 5$ for the proposed algorithms which represents a suitable value for all simulated setups as shown in [2]. We use the output of the initial access algorithm in [2], denoted as “IA”, to determine the first pilot assignment and cluster formation for each UE. Hence, it serves as an initialization for the proposed pilot and cluster update algorithm. We adopt the low-complexity heuristic power allocation inspired by [1, Eq. 26], however, considering only the set of UEs served by a given AP. The power allocation scaling parameter v is chosen to be 0.5, which represents a good balance between total SE maximization and UE fairness.

Figures 4.4 and 4.5 compare the cumulative distribution function (CDF) of the SE per UE of the proposed dynamic algorithm with that of an ultra-dense network (UDN), where each UE is served by a single AP, under MR and RZF precoding for $K = 20$ and $K = 50$ moving UEs. We further show the CDF of the SE per UE that is achieved by running the IA algorithm in [2], at every time instance for $K = 20$ UEs. It is clear that the proposed dynamic algorithm achieves the same performance as the IA algorithm. This is remarkable since the proposed algorithm makes local

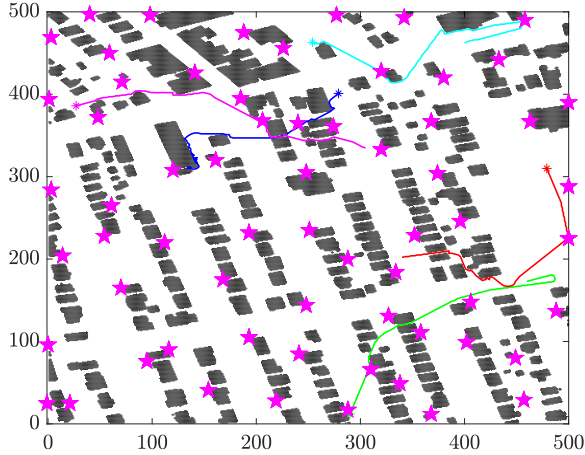


Figure 4.3: Site map for a European city with $K = 5$ moving UEs over a $400 T_i$ time interval.

Table 4.3: Network simulation parameters.

Area of interest	$500 \text{ m} \times 500 \text{ m}$
Carrier frequency	$f_c = 28 \text{ GHz}$
Bandwidth	500 MHz
Number of APs	$L = 64$
Number of UEs	$K \in [20, 60]$
Number of antennas per AP	$N = 8$
Per-AP maximum DL transmit power	$P_{\max}^{\text{dl}} = 1 \text{ W}$
UL transmit power	$p_i = 100 \text{ mW}$
UL/DL noise power	-80 dBm
Coherence block length	$\tau_c = 200$
Pilot sequence length	$\tau_p = 10$

refinements to the pilot assignment and cluster formation triggered by UE mobility, while the IA benchmark updates everything for all UEs at every time instance. Moreover, the proposed solution significantly outperforms the UDN architecture in terms of 95%-likely SE and median SE, for both MR and RZF precoding. The performance difference is seen to be larger for the case of RZF due to its ability to mitigate the added interference that appears in a cell-free network with more serving APs per UE than in the UDN. Comparing Figure 4.4 with Figure 4.5, we observe that the reduction in SE per UE with increasing K is smaller with the proposed solution compared to the UDN showing its superior capability to handle many UEs.

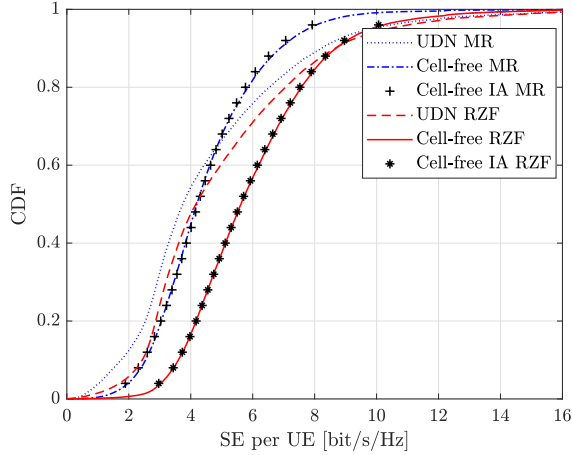


Figure 4.4: CDF of the DL SE per UE with MR and RZF precoding, $K = 20$.

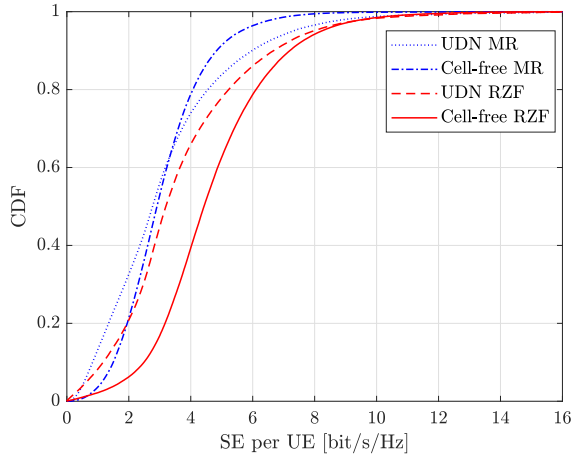


Figure 4.5: CDF of the DL SE per UE with MR and RZF precoding, $K = 50$.

Since master AP handovers are triggered mainly by large-scale UE mobility, we observed that the average number of master AP changes per UE remains fairly constant when varying the numbers of UEs K and maximum cluster sizes M_{\max} . This number is approximately 0.1 master AP handovers per second and is roughly the same as the average number of handovers per UE in the UDN system.

4.3 Paper III - A Bayesian Approach to Characterize Unknown Interference Power in Wireless Networks

4.3.1 Introduction

The existence of unknown interference is a prevalent problem in wireless communication systems. Specially for MU-MIMO networks where a large number of user equipments are served on the same time-frequency resources, the outage performance may be dominated by the unknown interference arising from neighbouring cells.

4.3.2 Related Work

Several works have focused on determining the signal-to-interference-plus-noise ratio (SINR) distribution in different MIMO setups. In [56], the cumulative distribution function (CDF) of the SINR is derived with equal power interferers, whereas [57] derives the SINR distribution in a Poisson field of interferers, both assuming MMSE combining and perfect CSI. In addition, [58, 59] approximate the SINR distribution with ZF and MMSE receivers assuming perfect CSI by a Gamma random variable for the equal power case and uniform interferers, respectively. More recent works [60, 61] derive an asymptotic approximation for the SINR distribution in a Poisson field of interferers with imperfect CSI by fitting a Gamma distribution to the co-channel interference in the UL and DL, respectively. In [62], the SINR distribution under uncorrelated Rayleigh fading, equal power and perfect CSI is presented for an MMSE MIMO system using regular patterns. A common assumption in these works is knowledge of the statistical properties of the underlying interference for computing the outage probability. In addition, numerous exponential family distributions have been proposed in Bayesian radar detection problems as in [18, 19], to model the variance of a non-Gaussian spherically invariant random process.

4.3.3 Contribution

Inspired by the Bayesian radar detection problems, we utilize an alternative approach to infer the type of distribution for the unknown interference power arising from UEs in neighbouring cells. Afterwards, we make use of observations of the total unknown interference power in a relatively large system to compute the sample mean and sample variance for this distribution that, under some practically feasible assumptions, converge to the true mean and variance, respectively. Further, we provide a tool for UL rate adaptation with guaranteed target outage performance based on our analytical work and validate the accuracy of our proposed method through Monte Carlo simulations.

4.3.4 Unknown Interference Power Distribution

The main focus of this work is to characterize the distribution of the total unknown interference power coming from UEs in neighbouring cells. In practice, the unknown interference power from adjacent cells is not known at the desired cell.

In a large system, the number of unknown interferers in the vicinity of the desired cell is expected to be large. The total unknown interference represented by the sum of independent interference terms converges, by the central limit theorem, to a Gaussian distribution [56]. In the following, we derive the distribution of the unknown variance for a Gaussian random variable and utilize the result to characterize the distribution of the unknown interference power.

Theorem 1. *The marginal posterior distribution of the unknown variance of a random variable $X \sim \mathcal{N}(\mu_u, \sigma_u^2)$ with known mean has an Inverse-Gamma distribution with a pdf given by*

$$f(\sigma_u^2|x) = \frac{\zeta^{\frac{1}{2}}}{\Gamma(\frac{1}{2})} \left(\frac{1}{\sigma_u^2}\right)^{\frac{3}{2}} e^{-\frac{\zeta}{\sigma_u^2}}, \quad (4.5)$$

where $\zeta = \frac{(x-\mu_u)^2}{2}$ and $\Gamma(\cdot)$ is the gamma function.

proof: refer to [3].

In practice, the statistics of the total unknown co-channel interferers' power can be estimated by the desired cell. By observing a sufficiently large sample of these statistics and feasibly assuming that the statistics of the unknown interferers' power vary slowly compared to that interval, the sample mean and sample variance of the unknown interference power converge to the true mean and variance of the distribution, respectively [60]. Using the result of Theorem 1, the unknown interference power distribution is given by the following.

Corollary 1. *The unknown interference power has an Inverse-Gamma distribution with a pdf given by*

$$f_{\text{IUI}_k^u}(x) = \frac{\beta^{\frac{3}{2}}}{\Gamma(\alpha)} \left(\frac{1}{x}\right)^{\alpha+1} e^{-\frac{\beta}{x}}, \quad (4.6)$$

$$\alpha = \left(\frac{\mu^2}{v}\right) + 2, \quad \beta = \left(\frac{\mu^2}{v} + 1\right) \times \mu, \quad (4.7)$$

where α and β denote the shape and scale parameters, respectively, and μ and v represent the estimated mean and variance for the total unknown interference power.

4.3.5 Selected Results

In this section, we corroborate the derived analytical results by Monte Carlo simulations and evaluate the UL SE that can be achieved while satisfying different target outage probabilities. We consider a half-wavelength spaced uniform linear array of $N = 16$ antennas at the desired BS located at the origin. The locations of the desired UE and known interferers are assumed fixed throughout the simulations

Table 4.4: Network simulation parameters.

Bandwidth	20 MHz
Number of BS antennas	$N = 16$
Pathloss exponent	$\alpha = 3.76$
UL transmit power	$p_i = 100$ mW
UL noise power	-94 dBm
Coherence block length	$\tau_c = 200$
Pilot sequence length	$\tau_p = 10$

since we focus on the SINR variations caused by the inter-cell interference that is unknown at the serving BS. We employ two different simulations; the first considers a distance of $r = 100$ m whereas the second considers $r = 200$ m between the desired UE and serving BS. In both cases, $K_n = 5$ known interferers are dropped, each on a circle of radius $\{60, 100, 140, 180, 220\}$ m and having a fixed angle to the serving BS that is chosen in a uniformly random manner. In each simulation instance, K_u unknown interferers are randomly and uniformly dropped on a disk in the radius range of $[250, 500]$ m. The network simulation parameters are summarized in Table 4.4. We utilize the 3GPP Urban Microcell model for generating large-scale fading coefficients with correlated shadowing among the UEs as given in [20, Eq. (37)].

Figure 4.6 shows the CDF of the SINR at the desired UE with both the analytical model and Monte Carlo simulation results for different numbers of unknown interferers and different distances between the desired UE and serving BS, denoted by r . It is clear that the analytical performance matches tightly with the exact numerical evaluation for all simulated scenarios. The SEs achieved with MR combining are about the same for different numbers of unknown interferers in neighbouring cells, thus, we only show the case of $K_u = 20$ unknown interferers. The reason is that MR ignores interference and then the denominator of the SINR is dominated by interference from the known co-channel interferers in the desired cell. Alternatively, when employing RZF, the residual known interference is seen to be weaker than the unknown interference, resulting in CDF curves with larger variations. Further, a noticeable change in SINR is observed with changing the number of unknown interferers, showing that the unknown interference may significantly impact the outage performance.

To demonstrate the effectiveness of our analytical result, Figure 4.7 plots the ϵ -outage SE with RZF for different numbers of unknown interferers and different distances between the desired UE and serving BS. A baseline scheme, based on the classic idea of having a fixed fade margin is shown for comparison. The baseline utilizes no knowledge of the unknown interference for rate adaptation but instead divides the effective SINR of the desired UE, excluding the term IUI_k^u , by a fixed margin m to compensate for the unknown interference from UEs in neighbouring cells. Note that for $K_u \in \{30, 40\}$, $r = 200$ m, and $m = 3.10$, the resulting outage probability $\epsilon > 0.3$; hence, it is not shown in the figure. It can be seen that the resulting outage probability varies significantly with the number of unknown

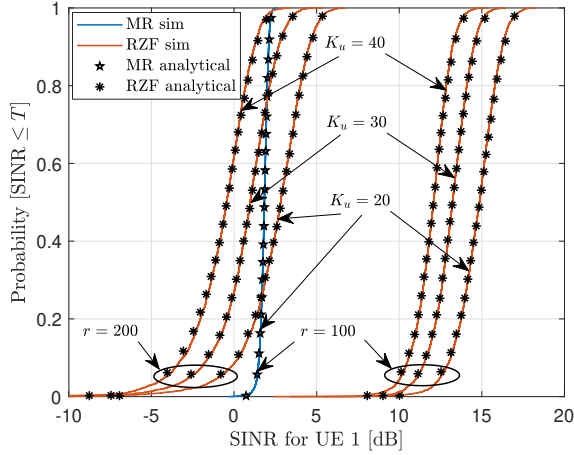


Figure 4.6: Outage probability for different numbers of unknown interferers and different distances between desired UE and BS.

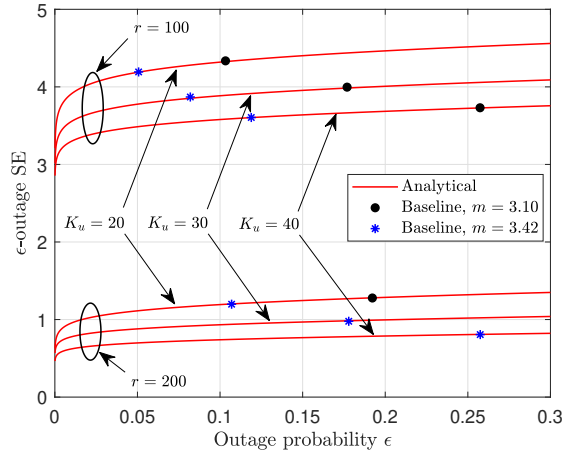


Figure 4.7: ϵ -outage SE with RZF.

interferers K_u and the distance r between the desired UE and BS. On the other hand, our proposed analytical model is able to determine the correct SE that results in any target predefined outage probability; thus, providing an effective tool for UL rate adaptation with guaranteed target outage performance.

4.4 Paper IV - Unknown Interference Modeling for Rate Adaptation in Cell-Free Massive MIMO Networks

4.4.1 Introduction

Wireless communication is continuously evolving to cope with the increasing traffic demands in mobile networks. Recently, three leading concepts have been widely investigated to boost the capacity of mobile networks: achieving higher SE, allocating more bandwidth, and network densification [2]. Cell-free massive MIMO is a combination of cooperative principles with massive MIMO processing to allow for more network densification. It poses as a potential candidate for 6G mobile communications due to its ability to utilize macro diversity and suppress MU interference to provide uniform service to the UEs [17]. Although the combination of cell-free massive MIMO and network densification can offer high data rates, the strength of co-channel interference from neighboring AP clusters can vary rapidly due to scheduling decisions and cause outages [3]. Massive signalling is required to measure instantaneous interference levels; thus, it is usually not done in practice [60].

4.4.2 Related Work

The common approach in prior works is either knowledge of the statistical properties of the underlying interference for computing the outage probability, which leads to expressions that are only valid for a particular set of assumptions; for example, precoding/combining and interference channel distribution; or fitting a single Gamma approximation to the SINR or interference which lacks a clear theoretical motivation and does not necessarily work well for all network configurations.

In addition to the literature that have focused on determining the SINR distribution in different cellular MIMO setups stated in Section 4.3.2. For the cell-free setup, the initial works [63, 64] have similarly fitted a Gamma distribution for each of the desired signal and interference to approximate the resulting SINR distribution.

4.4.3 Contribution

Our primal work [3] utilizes a Bayesian method to infer the distribution of the unknown interference power in a single-cell setup. This work addresses the more challenging cell-free massive MIMO scenario where the combined unknown interference from the APs perceived at the CPU negatively affects the SINR. Observations of the unknown interference power at each AP in the serving cluster are utilized separately to compute the distribution parameters. We derive an expression for the CDF of the total unknown interference power based on the characteristic functions of the unknown interference power at each AP. We verify the tightness of the analytical expressions numerically and provide a tool for UL rate adaptation with guaranteed target outage.

4.4.4 Unknown Interference Power Distribution

In practice, the unknown interference power from adjacent AP clusters, i.e., UEs not served by the same AP cluster as the desired UE, is not known in the desired UE's serving cluster, since it can change rapidly due to user mobility and scheduling decisions in the neighboring clusters [3, 60]. In our previous work [3], we have derived the distribution of the unknown interference power at a single AP as an Inverse-Gamma distribution. A typical AP-UE association strategy, as shown for pre-defined AP clusters in Figure 4.8, results in the strongest contributors to the unknown interference being different for the different APs. This results in the cross-covariance of the unknown interference being negligible. In other words, the covariance matrix for the total unknown interference power of the cluster will be diagonal with the unknown interference powers at the APs as the diagonal elements. Accordingly, the total unknown interference power can be approximated by $IUI_k^u \approx \sum_{l=1}^{L_s} |a_{kl}|^2 IUI_{kl}^u$, where IUI_{kl}^u corresponds to the unknown interference power at AP l and a_{kl} is the weight assigned by the CPU to the local signal estimate of AP l for UE k . The total unknown interference power IUI_k^u then comprises a weighted sum of the unknown interference powers experienced at each AP of the serving cluster with L_s APs.

Building on our previous result for the single-cell scenario, then for a dense network, the component terms $\{IUI_{kl}^u : l = 1, \dots, L_s\}$ can be modelled as independent non-identically distributed Inverse-Gamma random variables. We denote the shape and scale parameters for the l^{th} component by α_l and β_l , respectively, and so we have $IUI_{kl}^u \sim \text{IG}(\alpha_l, \beta_l)$. We stress that this approach is particularly useful as different unknown interferers can be modeled separately at each serving AP, regardless of the utilized clustering method. For computing the parameters of the Inverse-Gamma random variables, the statistics of the unknown co-channel interferers' power can be estimated separately by each AP in the serving cluster in a distributed fashion.

In the following, we derive the CDF of the total unknown interference power at the CPU as a weighted sum of independent non-identically distributed Inverse-Gamma random variables.

Theorem 1: *The CDF of the total unknown interference power is given by*

$$F_{IUI_k^u}(x) \approx \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \text{Im} \left(\frac{e^{-jtx} \phi(t)}{t} \right) dt, \quad (4.8)$$

$$\phi(t) = \prod_{l=1}^{L_s} \phi_l(|a_{kl}|^2 t), \quad (4.9)$$

$$\phi_l(t) = \frac{2(-j\beta_l t)^{\alpha_l/2}}{\Gamma(\alpha_l)} K_{\alpha_l} \left(2(-j\beta_l t)^{1/2} \right), \quad (4.10)$$

$$\alpha_l = \left(\frac{\mu_l^2}{v_l} \right) + 2, \quad \beta_l = \left(\frac{\mu_l^2}{v_l} + 1 \right) \mu_l, \quad (4.11)$$

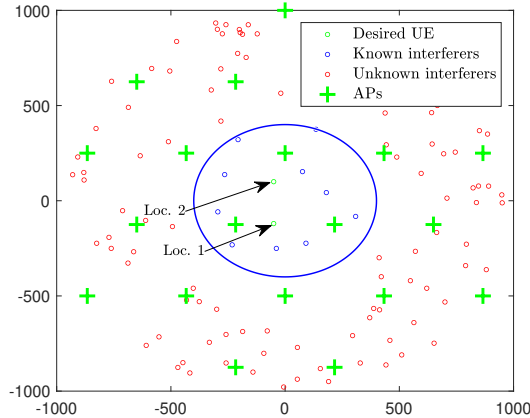


Figure 4.8: Site map and UE distribution with $K_u = 100$ unknown interferers. Serving cluster is located inside the marked circle.

where $\phi(t)$ represents the characteristic function of the total unknown interference power at the CPU, $\phi_l(t)$ denotes the characteristic function of the Inverse-Gamma random variable representing the unknown interference power at the l^{th} AP, $K_\alpha(\cdot)$ denotes the modified Bessel function of the second kind, and μ_l and v_l represent the estimated mean and variance for the unknown interference power at AP l .

Proof: The proof is given in [4].

4.4.5 Selected Results

In this section, we validate the derived distribution for the total unknown interference power by Monte Carlo simulations and determine the achievable UL SE that satisfies different target outage probabilities. Each AP is equipped with a half-wavelength spaced uniform linear array of $N = 16$ antennas. We consider a serving AP cluster of $L_s = 3$ APs and the first tier of 6 neighboring clusters comprising 18 APs. Since the focus of the paper is on the SINR variations caused by the inter-cluster interference that is unknown at the serving AP cluster, the locations of the desired UE and known interferers are assumed fixed throughout the simulations. The simulation setup is depicted in Figure 4.8. We employ two different simulations with the desired UE locations marked in Figure 4.8. In both cases, $K_n = 10$ known interferers are at fixed locations that are chosen in a uniformly random manner within a circle of radius 400 m, where the serving AP cluster of the desired UE is the $L_s = 3$ APs inside the marked circle. In each simulation step, K_u unknown interferers are randomly and uniformly dropped on a disk with the radius range of [450, 1000] m. The network simulation parameters are summarized in Table 4.5.

Figure 4.9 plots the CDF of the SINR at the desired UE with both the analytical model and Monte Carlo simulation results with different numbers of unknown

Table 4.5: Network simulation parameters.

Bandwidth	20 MHz
Number of serving APs	$L_s = 3$
Total number of APs	$L = 21$
Number of antennas per AP	$N = 16$
Number of known interferers	$K_n = 10$
Pathloss exponent	$\alpha = 3.67$
UL transmit power	$p_i = 100$ mW
UL noise power	-94 dBm
Coherence block length	$\tau_c = 200$
Pilot sequence length	$\tau_p = 10$

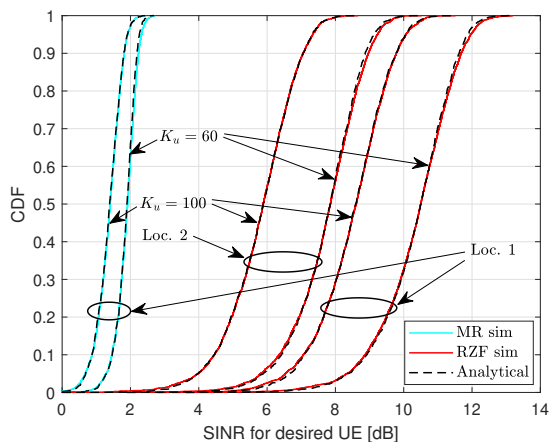


Figure 4.9: Outage probability for different numbers of unknown interferers.

interferers and different desired UE locations. The Fourier integral in (4.8) is approximated by a Riemann sum and computed efficiently using the Fast Fourier Transform (FFT) algorithm. We can see that the analytical performance tightly matches with the exact simulations for all simulated scenarios. We focus on RZF combining where the residual known interference is seen to be weaker than the unknown interference, which means that the unknown interference will be the dominant factor in the SINR's denominator. This results in a significant difference between the SINR curves when changing the number of unknown interferers for RZF, as well as larger variations within each curve caused by the unknown interference, which may potentially lead to outages. For the case of MR, the denominator of the SINR is dominated by the known interference from UEs belonging to the same serving AP cluster as the desired UE since it does not attempt to suppress the known interference. This results in lower SINRs in general with smaller SINR variations due to the unknown interference.

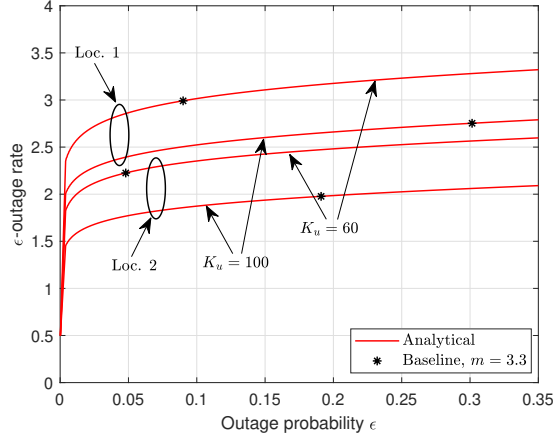


Figure 4.10: ϵ -outage SE with RZF.

In Figure 4.10, we demonstrate how our analytical results can be used for robust rate adaptation. The figure plots the ϵ -outage UL SE with RZF for different numbers of unknown interferers and different desired UE locations. We develop a baseline scheme for comparison, based on the classic idea of having a fixed fade margin. The baseline utilizes no knowledge of the unknown interference for rate adaptation. Alternatively, it divides the effective SINR of the desired UE, excluding the unknown interference, by a fixed margin m to make up for the effect of the unknown interference from UEs in neighboring clusters. The SE calculation for the baseline scheme is detailed in [3, Eq. (20)].

It is clear that the baseline scheme cannot effectively determine the UL transmission rate to maintain a specific target outage probability. This is due to the fact that changing the number of unknown interferers K_u or the location of the desired UE results in a significant change in the outage probability. However, our proposed analytical model is able to determine the correct SE that results in any target predefined outage probability; thus, providing an effective tool for UL rate adaptation with guaranteed target outage performance.

4.5 Paper V - Cell-Free Beamforming Design for Physical-Layer Multigroup Multicasting

4.5.1 Introduction

With the growing appetite for multimedia services and the surge in the number of UEs, multimedia multicasting provides an efficient solution to transmit data to a group of UEs. It can support various applications such as television channels' broadcasting over mobile networks, videoconferencing, and emergency broadcast systems. Due to the nature of multicasting, where the UEs are served using the same transmission and time-frequency resources, this technique mitigates the problem of co-channel interference within the group of UEs requesting the same data and provides a remarkable improvement in radio resource utilization [29].

Cell-free massive MIMO has the ability to provide uniform coverage within the network area and to suppress inter-cell interference [2, 17]. Such network infrastructure focuses on boosting the poor cell-edge UEs' performance. This is particularly desirable for multicast transmissions as the minimum SINR determines the common information rate of a UE group [48].

4.5.2 Related Work

QoS (guaranteeing a minimum received SNR to every UE) and MMF (maximizing the smallest received SNR) designs have been first introduced for a single-cell single-multicast group in [42]. The idea was later extended in [48] to consider multiple co-channel multicast groups. The core problem was shown to be NP-hard and approximate solutions were obtained using relaxation techniques on the basis of semidefinite programming (SDP). The works in [42, 48] suggested different randomization procedures to extract the sub-optimal rank-1 beamforming solutions from the higher-rank solutions attained through SDP. However, as the system size grows large, this method does not scale well in terms of the computational complexity required to test enough random vectors to obtain a satisfactory spectral efficiency (SE) performance. Particularly, in case of having more than one multicast group, a multigroup multicast power control (MMPC) optimization problem needs to be solved for each candidate set of beamforming vectors obtained through this method. Moreover, the approximation quality degrades considerably as the total number of transmit antennas increases [42, 65].

A recent line of research has suggested other iterative optimization procedures to solve the QoS and MMF multicast problems. For instance, the authors in [29] develop three algorithms to solve the MMF problem with per-cell power constraints while performing joint beamforming among the cooperating BSs. Their numerical evaluation shows that the difference-of-convex approximation (DCA) algorithm yields the highest SE performance with improved computational complexity compared to the other SDR with randomization-based algorithms. The DCA algorithm relies

on the successive convex approximation (SCA) technique, which represents the state-of-the-art in multicast beamforming optimization.

Moving towards the massive MIMO multicasting, several works have considered single- and multiple-group multicasting in cooperative and non-cooperative transmission topologies considering large massive arrays equipped at the BSs. As an example, [66, 67] utilize MR and RZF based precoding for the composite multicast channel (sum of the channels in a UE group) with a pilot power allocation scheme that compensates for the pathloss differences between each of the UEs in the group and the serving BS; while [66] considers a single massive MIMO BS and a single multicast channel, whereas [67] presents similar results for a relatively small cell-free massive MIMO system with a few APs in close proximity to each other and a small number of UEs. The main idea of these papers is to rely on asymptotic expressions to design beamforming vectors that approach the optimum solution when the number of antennas at the BSs grows exceedingly large. Although such simple design criterion works well in the unicast scheme, the convergence of this approach for the multicast case is slower as the number of antennas increases, especially when the number of UEs per group grows large, i.e., as the multicast problem structure becomes more distant to that of the unicast problem.

4.5.3 Contribution

In this paper, we tackle the problem of multigroup multicast beamforming optimization. The paper offers an algorithmic contribution and derives a novel iterative optimization procedure that is applicable to a wide range of SDP problems. In particular, we propose a novel optimization procedure to find near-global optimum MMF beamforming vectors in a cell-free network providing service to multiple multicast UE groups. Further, we develop a low-complexity heuristic algorithm that provides an efficient solution to the multigroup multicast MMF problem. We highlight that the proposed procedures are straightforwardly applicable to any multigroup multicast problem and can provide a near-optimal solution, regardless of the network architecture and size. Throughout this paper, we assume perfect CSI acquisition at the transmitting APs to keep the focus on the multicast beamforming optimization problem. The main contributions can be summarized as follows:

- We propose a new formulation for the QoS problem that directly attempts to minimize the per-AP powers rather than the total power available at all APs, which is common in the previous literature. In this way, we alleviate the need for the redundant total power constraint commonly used in the previous literature, which reduces the number of constraints and the required computational time for the algorithm.
- We propose a novel iterative optimization procedure to solve the multigroup multicast MMF problem. The procedure utilizes the SDR technique along with successive elimination of the higher-rank solutions produced by SDR to achieve the near-global optimum rank-1 beamforming solution to the MMF problem

as opposed to the state-of-the-art SCA based methods that can only guarantee convergence to a stationary point of the original non-convex problem.

- We propose a novel low-complexity heuristic algorithm that provides an effective solution to the multicast beamforming optimization problem at a vastly reduced computational time. The algorithm utilizes a newly proposed phase alignment and UE emphasis procedure that aims at iteratively maximizing the minimum SEs across the multicast UE groups. The proposed heuristic marks the first low-complexity algorithm that is specifically tailored for the multicast problem with computational requirements that are in the milliseconds range.
- We provide an extensive performance evaluation of the proposed algorithms and compare them to the state-of-the-art in multigroup multicast beamforming optimization. Numerical results show significant improvements in the minimum achievable SE and the computational time required by the algorithms.

4.5.4 QoS and MMF Multicast Problems

Building on the system model described in [5], we define the concatenated channel vector between all APs and UE k in multicast UE group $g \in \{1, \dots, G\}$ as $\mathbf{h}_{kg} = [\mathbf{h}_{kg,1}^T, \mathbf{h}_{kg,2}^T, \dots, \mathbf{h}_{kg,L}^T]^T \in \mathbb{C}^{LN \times 1}$ and the precoding vector for group g as $\mathbf{w}_g = [\mathbf{w}_{g,1}^T, \mathbf{w}_{g,2}^T, \dots, \mathbf{w}_{g,L}^T]^T \in \mathbb{C}^{LN \times 1}$. Further, we define $\mathbf{D}_l = \text{blkdiag}(\mathbf{Z}_{l1}, \dots, \mathbf{Z}_{lL})$, with $\mathbf{Z}_{ll} = \mathbf{I}_N$ and $\mathbf{Z}_{li} = \mathbf{0}_{N \times N}, \forall i \neq l$. Hence, the multigroup multicast MMF problem is given by

$$\underset{\{\mathbf{W}_g\}_{g=1}^G}{\text{maximize}} \quad \underset{\substack{\forall k \in \{1, \dots, K_g\} \\ \forall g \in \{1, \dots, G\}}}{\min} \quad \frac{1}{\eta_g} \frac{\text{tr}(\mathbf{H}_{kg} \mathbf{W}_g)}{\sum_{j=1, j \neq g}^G \text{tr}(\mathbf{H}_{kj} \mathbf{W}_j) + \sigma_{kg}^2} \quad (4.12a)$$

$$\text{s.t.} \quad \mathbf{W}_g \succeq \mathbf{0} \quad \forall g \in \{1, \dots, G\}, \quad (4.12b)$$

$$\text{rank}(\mathbf{W}_g) = 1 \quad \forall g \in \{1, \dots, G\}, \quad (4.12c)$$

$$\sum_{g=1}^G \text{tr}(\mathbf{D}_l \mathbf{W}_g) \leq P_{l,\max} \quad \forall l \in \{1, \dots, L\}, \quad (4.12d)$$

where $\mathbf{H}_{kg} = \mathbf{h}_{kg} \mathbf{h}_{kg}^H$, $\mathbf{W}_g = \mathbf{w}_g \mathbf{w}_g^H$, $\forall g$ and η_g represents the per-group normalized weight of the SINR target, such that $0 \leq \eta_g \leq 1$, $\forall g$, which allows for having different SINR targets among the different multicast groups. σ_{kg}^2 represents the noise power and $P_{l,\max}$ is the power budget at AP l . The above formulation makes it suitable to construct the SDP of the multigroup multicast MMF problem.

The corresponding QoS problem to the MMF problem can be formulated directly in terms of the normalized per-AP powers as

$$\text{minimize}_{\{\mathbf{W}_g\}_{g=1}^G} \max_{\forall l \in \{1, \dots, L\}} \frac{1}{P_{l, \max}} \sum_{g=1}^G \text{tr}(\mathbf{D}_l \mathbf{W}_g) \quad (4.13a)$$

$$\text{s.t.} \quad \frac{1}{\eta_g} \frac{\text{tr}(\mathbf{H}_{kg} \mathbf{W}_g)}{\sum_{\substack{j=1 \\ j \neq g}}^G \text{tr}(\mathbf{H}_{kg} \mathbf{W}_j) + \sigma_{kg}^2} \geq \gamma, \quad (4.13b)$$

$$\forall k \in \{1, \dots, K_g\}, \forall g \in \{1, \dots, G\}, \quad (4.13c)$$

$$\mathbf{W}_g \succeq \mathbf{0} \quad \forall g \in \{1, \dots, G\}, \quad (4.13c)$$

$$\text{rank}(\mathbf{W}_g) = 1 \quad \forall g \in \{1, \dots, G\}, \quad (4.13d)$$

where γ represents the SINR target. The relation between the QoS and MMF problems can be briefly described as follows. If the minimum power that can be achieved for a given QoS problem is equal to the power budget at the APs, then the target SINR of the QoS problem is the MMF SINR that can be achieved for that power budget.

4.5.5 Selected Results

In this section, we use Monte Carlo simulations to evaluate the performance of the proposed optimization procedure and heuristic algorithm to solve the multigroup multicast MMF problem. Unless otherwise indicated, we assume $G = 3$ multicast groups each comprising the same number of $K_g = K_G$ UEs, $\forall g$, that are randomly and uniformly distributed within the area of interest. The simulation parameters are summarized in Table 4.6. In the following, the proposed successive elimination algorithm will be referred to as “SEA”, while SDR followed by choosing the dominant eigenvector or Gaussian randomization as “SDR-D” and “SDR-G”, respectively.

Figure 4.11 plots the average minimum SE versus the number of UEs per group K_G . It can be seen that the optimum unicast max-min solution presents a lower bound on the achievable average minimum SE via multicasting, which shows the possible improvements of multicast transmissions over unicasting when delivering the same data content to multiple UEs. The figure shows that the proposed SEA and the DCA algorithm maintain an average minimum SE that is close to the upper bound up to $K_G = 15$ UEs per group. The SDR-D and SDR-G provide relatively good performance for small K_G . The reason is that when the number of UEs is small, the solution matrices produced by SDR have ranks that are close to 1. As a result, the approximations made by picking the dominant eigenvector or performing randomization do not sacrifice much to achieve an approximate rank-1 solution. As the number of UEs increase, a faster degradation in performance is seen for these approximation techniques compared to other algorithms. The proposed heuristic algorithm shows 5%, 28%, and 60% improvement in the average minimum SE over SDR-G, SDR-D, and unicast max-min for $K_G = 10$ UEs per group, respectively. Further, the heuristic algorithm is able to attain a constant gap to the optimum

Table 4.6: Network simulation parameters.

Area of interest	750 m \times 750 m
Bandwidth	20 MHz
Number of APs	$L = 9$
Number of antennas per AP	$N = 4$
Number of multicast groups	$G = 3$
Maximum AP transmit power	$P_{l,\max} = 1$ W
Pathloss exponent	$\alpha = 3.67$
DL noise power	-94 dBm

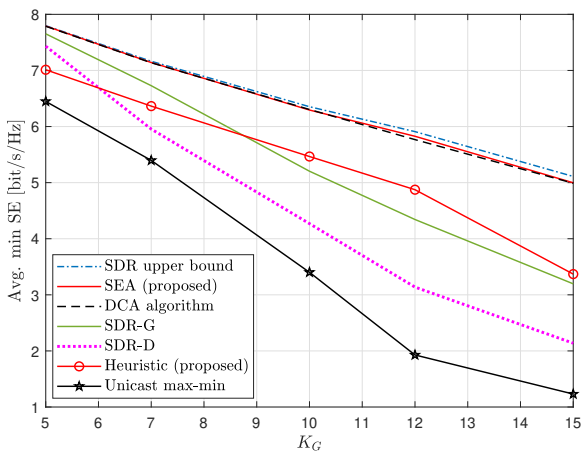


Figure 4.11: Average minimum SE for different numbers of UEs.

rank-1 solutions up to $K_G = 12$ UEs per group, which translates to about 87% of the achievable average minimum SE of the optimum rank-1 solution for $K_G = 10$ UEs per group.

Since order expressions provide limited information about an algorithm's complexity in normal-sized scenarios due to the lack of scaling factors and lower-dimensional terms as well as rely on worst-case bounds which could lead to ill-informed conclusions, we supplement the analysis with numerical measurements for the average runtimes required by each algorithm. We use the same platform for performing all simulations for a fair comparison: a 4 core Intel(R) Core i5-10310U CPU with 1.7 GHz base frequency and 4.4 GHz turbo frequency. All programs are written in Matlab and utilize CVX [68] to solve the optimization problems. Figure 4.12 plots the average computational time against K_G . The figure shows that the proposed SEA achieves about 35% reduction in computational time compared to the state-of-the-art DCA algorithm for $K_G = 10$ UEs per group, and the reduction is

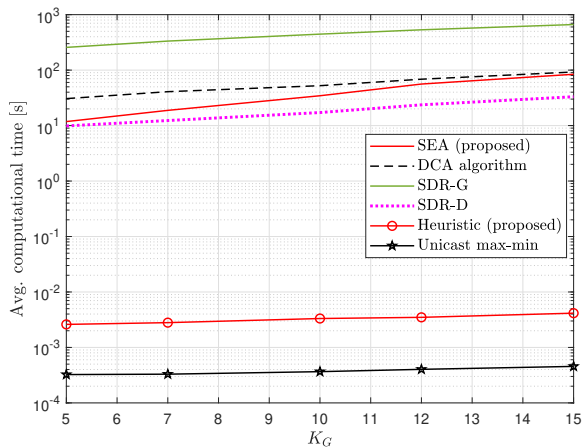


Figure 4.12: Average computational time in seconds.

particularly large when K_G is small. Moreover, the proposed heuristic algorithm is able to provide orders-of-magnitude reduction in computational time compared to the other multigroup optimization procedures, setting a record-low runtime for multicast beamforming design of 3.3 ms for $K_G = 10$ UEs per group. Hence, this algorithm could likely be directly used in real-time applications. In comparison to the unicast max-min scheme, the proposed heuristic strikes a balance in the inevitable trade-off between SE performance and computational complexity.

4.6 Paper VI - Low-Complexity SDP-ADMM for Physical-Layer Multicasting in Massive MIMO Systems

4.6.1 Introduction

There is a demand for the same data content from several UEs in many wireless communication applications. Physical-layer multicasting combines the beamforming capability of massive MIMO and the broadcast nature of the wireless channel to efficiently deliver the same data to a group of UEs using a single transmission. This paper tackles the MMF multicast beamforming optimization, which is an NP-hard problem.

4.6.2 Related Work

Physical-layer multicasting beamforming optimization has been considered for the QoS and MMF objectives utilizing SDR followed by Gaussian randomization [42, 48] and SCA [29, 47]. The SCA based methods have shown superior performance since randomization requires excessive computations and suffers from significant performance degradation as the system size grows large. However, SCA can only guarantee convergence to a local optimum of the non-convex beamforming problems and requires initialization with a feasible solution which might be hard to obtain particularly for large systems. In [9], we have proposed a successive elimination algorithm (SEA) that relies on SDR followed by iterative elimination of higher-rank solutions to extract a near-global optimal rank-1 solution to the MMF multicast problem. Other works, such as [69–71], have utilized the SCA technique together with ADMM to achieve a low-complexity local optimum solution to the QoS and MMF multicast problems.

4.6.3 Contribution

In this paper, we propose a novel SDP-ADMM algorithm to solve the MMF multicast problem in massive MIMO systems. Notably, this represents the first application of SDP-ADMM in multicast beamforming optimization, demonstrating its strong potential for improving efficiency in this field. We develop a new problem reformulation and ADMM-based implementation, extending our previously proposed SEA framework in [9], to obtain a near-global optimal rank-1 beamforming solution. A key advantage of our approach is its adaptability to various multicast optimization objectives and network configurations without requiring ADMM parameter tuning. Throughout the paper, we focus on multicast beamforming design, assuming perfect CSI at the transmitting BS. Our results confirm that the proposed SDP-ADMM algorithm achieves comparable SE performance to the SEA that relies on conventional SDP solvers, while significantly reducing computational complexity.

4.6.4 SDP-ADMM: Efficient Multicast Beamforming

Building on the system model in [6], we define the notation $\mathbf{H}_k = \mathbf{h}_k \mathbf{h}_k^H / \sigma_k^2$ and $\mathbf{W} = \mathbf{w} \mathbf{w}^H$. Utilizing the fact that $|\mathbf{h}_k^H \mathbf{w}|^2 = \text{tr}(\mathbf{h}_k \mathbf{h}_k^H \mathbf{w} \mathbf{w}^H)$, the SNR of UE k in (3.2) can be written as

$$\text{SNR}_k^{\text{dl}} = \frac{|\mathbf{h}_k^H \mathbf{w}|^2}{\sigma_k^2} = \text{tr}(\mathbf{H}_k \mathbf{W}). \quad (4.14)$$

This SNR formulation is suitable to construct the SDP of the MMF multicast problem.

Let \mathbb{H}^N denote the set of $N \times N$ Hermitian matrices. The linear map $\mathcal{H}(\cdot) : \mathbb{H}^N \rightarrow \mathbb{R}^K$ is defined as

$$\mathcal{H}(\mathbf{W}) = \left(\langle \mathbf{H}_1^H, \mathbf{W} \rangle, \dots, \langle \mathbf{H}_K^H, \mathbf{W} \rangle \right), \quad (4.15)$$

where the inner product between two matrices is given by $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{A}^H \mathbf{B})$. The QoS problem can then be written compactly as

$$\min_{\mathbf{W}} \text{tr}(\mathbf{W}) \quad (4.16a)$$

$$\text{s.t. } \mathcal{H}(\mathbf{W}) \geq \boldsymbol{\gamma}, \quad (4.16b)$$

$$\mathbf{W} \succeq \mathbf{0}, \quad \text{rank}(\mathbf{W}) = 1, \quad (4.16c)$$

where $\boldsymbol{\gamma} \in \mathbb{R}^K$ represents the SNR targets of the UEs. The above problem is non-convex due to the non-convex rank-1 constraint in (4.16c). A relaxation of the problem is devised by dropping the rank-1 constraint to obtain the relaxed convex QoS problem. In general, such relaxation results in the solution matrix \mathbf{W} having a high rank that does not satisfy the rank constraint of the original problem, requiring post-processing of the output matrix to extract a feasible rank-1 beamforming solution. Finding the optimal post-processing is challenging.

In this paper, we utilize the state-of-the-art SEA proposed in our primal work [9]. We propose a new formulation that is suitable for developing a computationally fast and effective ADMM algorithm to solve the QoS problem. The penalized relaxed QoS problem is formulated as

$$\min_{\mathbf{W}} \langle \boldsymbol{\Lambda}, \mathbf{W} \rangle \quad (4.17a)$$

$$\text{s.t. } \mathcal{H}(\mathbf{W}) \geq \boldsymbol{\gamma}, \quad (4.17b)$$

$$\mathbf{W} \succeq \mathbf{0}, \quad (4.17c)$$

where $\boldsymbol{\Lambda} \in \mathbb{C}^{N \times N}$ is a design parameter that is responsible for the elimination of the higher-rank solutions and improves the convergence speed of the proposed algorithm.

Table 4.7: Network simulation parameters.

Area of interest	$750 \text{ m} \times 750 \text{ m}$
Bandwidth	20 MHz
Number of BS antennas	$N = 36$
Number of UEs	$K = 15$
BS transmit power	$P_T = 40 \text{ W}$
DL noise power	-94 dBm

4.6.5 Selected Results

In this section, we use Monte Carlo simulations to verify the effectiveness of the proposed SDP-ADMM optimization procedure to solve the MMF multicast problem. We consider a massive MIMO system with a half-wavelength-spaced uniform linear array of $N = 36$ antennas deployed at the BS, serving $K = 15$ UEs in an area of $750 \text{ m} \times 750 \text{ m}$ using a single multicast transmission. The simulation parameters are summarized in Table 4.7.

Figure 4.13 plots the cumulative distribution function (CDF) of the max-min SE for the proposed SDP-ADMM algorithm for $K = 15$ UEs. The state-of-the-art SEA algorithm [9], which utilizes CVX [68], and the SDR upper bound are shown for comparison. It is clear that the proposed low-complexity SDP-ADMM algorithm exhibits similar SE performance to that of the SEA. Later, we will show the tremendous saving in computational time for the proposed SDP-ADMM algorithm. Both the SDP-ADMM and SEA algorithms almost achieve the SDR upper bound for the this scenario.

We highlight that we have selected our previously proposed SEA as a benchmark as it outperforms state-of-the-art SCA-based methods when standard solvers were used to solve the optimization problems, as demonstrated in [9]. The SE improvement over SCA-based methods stems from their use of a gradient descent approach, which ensures convergence only to a stationary point of the non-convex MMF problem. In contrast, the proposed algorithm begins with the optimal higher-rank solution obtained via SDR and iteratively reduces its rank by penalizing eigenvectors associated with the second-largest eigenvalues. This process continues until a near-optimal rank-1 solution is reached in the orthogonal subspace of these eigenvectors. At each iteration, the solution matrices are Hermitian positive semidefinite with distinct non-zero eigenvalues, ensuring that the penalty minimally impacts the optimality of subsequent iterations. As a result, this approach efficiently converges to a near-global optimal rank-1 solution to the NP-hard MMF multicast problem.

Table 4.8 presents the average computational time of the proposed SDP-ADMM algorithm as compared to the benchmark. The proposed algorithm is able to achieve more than 10 times reduction in computational time as compared to the state-of-the-art SEA utilizing CVX and attains the same SE performance. We use the same platform for performing the simulations, a 4 core Intel(R) Core i5-10310U CPU with

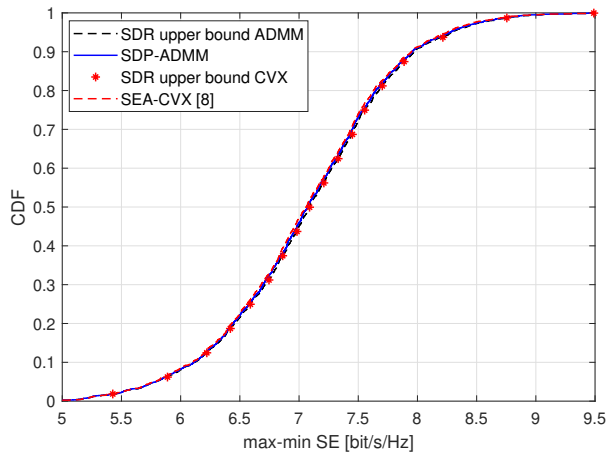
Figure 4.13: CDF of the max-min SE for $K = 15$ UEs.

Table 4.8: Average computational time in seconds.

K	SDP-ADMM	SEA [9]
15	0.6	8.6

1.7 GHz base frequency and 4.4 GHz turbo frequency. All programs are written in Matlab. The CDF curves are generated using 2000 simulation samples.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

Cell-Free massive MIMO is a new promising mobile network architecture that enables ubiquitous connectivity by exploiting spatial macro-diversity. However, there are inherent challenges to deploying this superior architecture. This thesis has contributed to addressing key issues in power allocation, mobility management, unknown interference modeling, and physical-layer multicasting by developing novel algorithms and optimization techniques.

The work presented in this thesis has demonstrated innovative approaches to power allocation and mobility management in large-scale cell-free networks. First, we formulated the network-wide sum-SE and PF maximization problems and utilize the WMMSE-ADMM algorithm to generate a large amount of data in a reasonable time. We then propose different distributed and clustered DNNs to approximate the power coefficients using only the LSF coefficients that are locally available at an AP/EP, and so allowing for reduced fronthaul/midhaul requirements. We draw attention to that our DNN models find an unknown mapping from a smaller domain input, than the conventional WMMSE-ADMM algorithm, to the desired output power coefficients. Moreover, the run-time for the learning-based algorithms has been shown to be much less than that of the ADMM algorithm, which is the most efficient conventional approach for solving the power allocation optimization problem, to the best of our knowledge. The reduced fronthaul and computational time indicate that our proposed models are particularly useful for implementation in large-scale systems. Numerical results show that our distributed and clustered DNN implementations outperform low-complexity practical heuristic allocations and provide a reasonable approximation to the centralized WMMSE-ADMM benchmark. Additionally, we proposed a distributed algorithm designed to dynamically update AP clusters and pilot assignments in response to user mobility and changing channel conditions. The main goal is to extend traditional handover concepts to the more complicated AP-UE

associations of cell-free networks, with focus on mmWave scenarios with APs along streets. To numerically evaluate the algorithm, we developed a novel mobility model to simulate realistic UE mobility in a site map including buildings/obstructions. The results indicate that our proposed distributed algorithms outperform the cellular UDN architecture, providing a remarkable improvement in the 95%-likely SE thereby delivering more stable SE under mobility, while minimizing unnecessary handovers and pilot reassignments. We conclude that the combination of the cell-free massive MIMO architecture and mmWave bands (known for high bandwidth but poor channel reliability) lead to unprecedented data rates with high reliability and network coverage.

Afterwards, we shift the focus to the prevalent problem of unknown UL interference in an MU-MIMO wireless network. We derive an analytical model for the unknown interference power distribution and show the accuracy of the analysis numerically. The model is useful for robust rate allocation to handle outages caused by rapidly changing inter-cell interference. The proposed method demonstrates clear advantages over traditional fixed-margin schemes, which suffers from significant change in the resulting outage performance due to unpredictable interference variations. In contrast, our framework is capable of always choosing the correct SE that offers a specific predetermined target outage probability. The work is then extended to model the total unknown UL interference at the CPU experienced by the serving AP cluster in a cell-free massive MIMO network.

In the context of physical-layer multicasting, we revisited the multigroup multicast QoS and MMF optimization problems with particular emphasis on cell-free massive MIMO networks. We have proposed a novel iterative optimization procedure that is able to provide a near-global optimum rank-1 beamforming solution to this NP-hard problem. The proposed algorithm is compared against state-of-the-art SDR-based and SCA-based algorithms in multicast beamforming optimization. The numerical results show significant improvements in terms of SE performance and computational complexity while utilizing the same platform and software to solve optimization problems. The rationale behind the improvements is that the proposed procedure relies on eliminating high-rank solutions while imposing a negligible effect on the optimal achievable SE, whereas previous methods rely on approximations that degrade the performance, particularly when the number of UEs grows large. The significance of the proposed algorithm is that it can be adopted for any type of problem where SDR is utilized and a low-rank solution is desirable, which is a wide class of optimization problems. A fast ADMM algorithm is then designed to achieve the near-global optimum rank-1 beamforming solution to the multigroup multicast QoS and MMF problems at reduced computational complexity. A novel problem reformulation and penalization scheme is proposed that is specifically tailored for our SDP-ADMM implementation. Compared to standard SDP solvers, SDP-ADMM attains similar SE performance with about 10 times reduction in computational time. Numerical results show the robustness of the proposed SDP-ADMM algorithm where it maintains its performance for different simulation settings without the need for ADMM parameter tuning. Moreover, we have developed a low-complexity

phase alignment and UE emphasis scaling heuristic algorithm that is able to achieve about 87% of the minimum SE of the optimum rank-1 solution with orders-of-magnitude lower computational time. The heuristic algorithm enables a real-time implementation of multigroup multicast beamforming.

Overall, the proposed methodologies in this thesis pave the way for scalable and high-performance implementations, contributing to the advancement of next-generation wireless communication systems.

5.2 Future Work

In the following, we list possible future research directions.

- **Extending the work on ADMM algorithms for physical-layer multicasting.**

Physical-layer multicasting presents an efficient transmission technique that is well-suited for many current and emerging wireless communication applications. In a cell-free network, the problem would require more involved solutions to deal with per-AP power constraints. Further, the effect of imperfect CSI needs to be evaluated for the proposed algorithms.

- **Federated learning for resource allocation in cell-free networks.**

The architecture of cell-free networks calls for distributed solutions as previously explained in the thesis. Nevertheless, to fully exploit the potential of cell-free networks, some less-complex tasks that require limited information may be performed at the CPU. This idea fits well with federated learning concepts.

References

- [1] M. Zaher, Ö. T. Demir, E. Björnson, and M. Petrova, “Learning-based downlink power allocation in cell-free massive MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 174–188, 2023.
- [2] M. Zaher, E. Björnson, and M. Petrova, “Soft handover procedures in mmWave cell-free massive MIMO networks,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 6, pp. 6124–6138, 2024.
- [3] M. Zaher, E. Björnson, and M. Petrova, “A bayesian approach to characterize unknown interference power in wireless networks,” *IEEE Wireless Communications Letters*, vol. 12, no. 8, pp. 1374–1378, 2023.
- [4] M. Zaher, E. Björnson, and M. Petrova, “Unknown interference modeling for rate adaptation in cell-free massive MIMO networks,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2024.
- [5] M. Zaher, E. Björnson, and M. Petrova, “Cell-free beamforming design for physical layer multigroup multicasting,” *under review in IEEE Transactions on Wireless Communications*, 2024.
- [6] M. Zaher and E. Björnson, “Low-complexity SDP-ADMM for physical-layer multicasting in massive MIMO systems,” *Accepted in 23rd International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt)*, 2025.
- [7] M. Zaher, Ö. T. Demir, E. Björnson, and M. Petrova, “Distributed DNN power allocation in cell-free massive MIMO,” in *Proc. 55th Asilomar Conference on Signals, Systems, and Computers*, pp. 722–726, IEEE, 2021.
- [8] M. Zaher, E. Björnson, and M. Petrova, “Mobility management in mmWave cell-free massive MIMO networks,” in *Proc. IEEE International Conference on Communications*, pp. 6492–6497, 2023.
- [9] M. Zaher, E. Björnson, and M. Petrova, “Near-optimal cell-free beamforming for physical layer multigroup multicasting,” in *Proc. IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2024.

- [10] M. Zaher and E. Björnson, “Low-complexity optimal beamforming for physical-layer multicasting in cell-free massive MIMO,” *to be submitted to IEEE Transactions on Wireless Communications*, 2025.
- [11] S. Imam, M. Zaher, and A. El-Mahdy, “Interference cancellation in full duplex uplink multi-user MIMO DF relaying system with imperfect CSI,” in *Proc. International Wireless Communications and Mobile Computing (IWCMC)*, pp. 740–745, 2021.
- [12] A. Mahmoudi, M. Zaher, and E. Björnson, “Joint energy and latency optimization in federated learning over cell-free massive mimo networks,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2024.
- [13] A. Mahmoudi, M. Zaher, and E. Björnson, “Low-latency and energy-efficient federated learning over cell-free networks: A trade-off analysis,” *minor revision, in IEEE Open Journal of the Communications Society*, 2025.
- [14] P. Foti, B. Mouris, T. Voigt, M. Zaher, and M. Afshang, “On the performance of harmonic backscattering for zero-energy devices,” in *Proc. IEEE Wireless Power Technology Conference and Expo (WPTCE)*, 2025.
- [15] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge University Press, 2016.
- [16] E. Björnson, J. Hoydis, and L. Sanguinetti, “Massive MIMO networks: Spectral, energy, and hardware efficiency,” *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.
- [17] Ö. T. Demir, E. Björnson, and L. Sanguinetti, “Foundations of user-centric cell-free massive MIMO,” *Foundations and Trends® in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021.
- [18] E. Jay, J. P. Ovarlez, D. Declercq, and P. Duvaut, “BORD: Bayesian optimum radar detector,” *Signal Processing*, vol. 83, no. 6, pp. 1151–1162, 2003.
- [19] J. Wang, A. Dogandzic, and A. Nehorai, “Maximum likelihood estimation of compound-gaussian clutter and target parameters,” *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3884–3898, 2006.
- [20] E. Björnson and L. Sanguinetti, “Making cell-free massive MIMO competitive with MMSE processing and centralized implementation,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2019.
- [21] S. Chakraborty, Ö. T. Demir, E. Björnson, and P. Giselsson, “Efficient downlink power allocation algorithms for cell-free massive MIMO systems,” *IEEE Open Journal of the Communications Society*, vol. 2, pp. 168–186, 2021.

- [22] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, “Ubiquitous cell-free massive MIMO communications,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–13, 2019.
- [23] A. Burr, M. Bashar, and D. Maryopi, “Cooperative access networks: Optimum fronthaul quantization in distributed massive MIMO and cloud RAN-invited paper,” in *IEEE 87th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, 2018.
- [24] H. Masoumi and M. J. Emadi, “Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1038–1053, 2019.
- [25] E. Björnson and L. Sanguinetti, “Scalable cell-free massive MIMO systems,” *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.
- [26] M. Bashar, A. Akbari, K. Cumanan, H. Q. Ngo, A. G. Burr, P. Xiao, M. Debbah, and J. Kittler, “Exploiting deep learning in limited-fronthaul cell-free massive MIMO uplink,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1678–1697, 2020.
- [27] S. Chakraborty, E. Björnson, and L. Sanguinetti, “Centralized and distributed power allocation for max-min fairness in cell-free massive MIMO,” in *53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 576–580, 2019.
- [28] G. Interdonato, P. Frenger, and E. G. Larsson, “Scalability aspects of cell-free massive MIMO,” in *IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2019.
- [29] G.-W. Hsu, B. Liu, H.-H. Wang, and H.-J. Su, “Joint beamforming for multicell multigroup multicast with per-cell power constraints,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4044–4058, 2016.
- [30] L. Simic, J. Riihijärvi, A. Venkatesh, and P. Mahonen, “Demo abstract: An open source toolchain for planning and visualizing highly directional mm-wave cellular networks in the 5G era,” in *IEEE International Conference on Computer Communications (INFOCOM) WKSHPs*, pp. 966–967, IEEE, 2017.
- [31] T. S. Rappaport, “The wireless revolution,” *IEEE Communications Magazine*, vol. 29, no. 11, pp. 52–71, 1991.
- [32] S. Asif, *5G mobile communications: Concepts and technologies*. CRC Press, 2018.
- [33] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology*. Academic Press, 2020.

- [34] E. Björnson, E. Jorswieck, *et al.*, “Optimal resource allocation in coordinated multi-cell systems,” *Foundations and Trends® in Communications and Information Theory*, vol. 9, no. 2–3, pp. 113–381, 2013.
- [35] R. Apelfröjd and M. Sternad, “Design and measurement-based evaluations of coherent JT CoMP: a study of precoding, user grouping and resource allocation using predicted CSI,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, pp. 1–20, 2014.
- [36] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfröjd, and T. Svensson, “The role of small cells, coordinated multipoint, and massive MIMO in 5G,” *IEEE Communications Magazine*, vol. 52, no. 5, pp. 44–51, 2014.
- [37] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, “Precoding and power optimization in cell-free massive MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, 2017.
- [38] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, “Learning to optimize: Training deep neural networks for interference management,” *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5438–5453, 2018.
- [39] W. Lee, M. Kim, and D.-H. Cho, “Deep power control: Transmit power control scheme based on convolutional neural network,” *IEEE Communications Letters*, vol. 22, no. 6, pp. 1276–1279, 2018.
- [40] C. Zhang, P. Patras, and H. Haddadi, “Deep learning in mobile and wireless networking: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.
- [41] E. Björnson and Ö. T. Demir, “Introduction to multiple antenna communications and reconfigurable surfaces,” 2024.
- [42] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, “Transmit beamforming for physical-layer multicasting,” *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, 2006.
- [43] S. Y. Park and D. J. Love, “Capacity limits of multiple antenna multicasting using antenna subset selection,” *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2524–2534, 2008.
- [44] M. Dong and Q. Wang, “Multi-group multicast beamforming: Optimal structure and efficient algorithms,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 3738–3753, 2020.
- [45] A. d’Aspremont and S. Boyd, “Relaxations and randomized methods for non-convex QCQPs,” *EE392o Class Notes, Stanford University*, vol. 1, pp. 1–16, 2003.

- [46] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, and N. D. Sidiropoulos, “Feasible point pursuit and successive approximation of non-convex QCQPs,” *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 804–808, 2015.
- [47] M. Sadeghi, L. Sanguinetti, R. Couillet, and C. Yuen, “Reducing the computational complexity of multicasting in large-scale antenna systems,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2963–2975, 2017.
- [48] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, “Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups,” *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1268–1279, 2008.
- [49] Y. Zhao, I. G. Niemegeers, and S. H. De Groot, “Power allocation in cell-free massive MIMO: A deep learning method,” *IEEE Access*, vol. 8, pp. 87185–87200, 2020.
- [50] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1. MIT press Cambridge, 2016.
- [51] T. Van Chien, T. N. Canh, E. Björnson, and E. G. Larsson, “Power control in cellular massive MIMO with varying user activity: A deep learning solution,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5732–5748, 2020.
- [52] S. Buzzi, C. D’Andrea, M. Fresia, Y.-P. Zhang, and S. Feng, “Pilot assignment in cell-free massive MIMO based on the hungarian algorithm,” *IEEE Wireless Communications Letters*, vol. 10, no. 1, pp. 34–37, 2020.
- [53] R. Sabbagh, C. Pan, and J. Wang, “Pilot allocation and sum-rate analysis in cell-free massive MIMO systems,” in *Proc. International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2018.
- [54] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, “Structured massive access for scalable cell-free massive MIMO systems,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1086–1100, 2020.
- [55] C. D’Andrea, G. Interdonato, and S. Buzzi, “User-centric handover in mmWave cell-free massive MIMO with user mobility,” in *Proc. European Signal Processing Conference (EUSIPCO)*, pp. 1–5, IEEE, 2021.
- [56] H. Gao, P. J. Smith, and M. V. Clark, “Theoretical reliability of MMSE linear diversity combining in rayleigh-fading additive interference channels,” *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 666–672, 1998.
- [57] O. B. S. Ali, C. Cardinal, and F. Gagnon, “Performance of optimum combining in a poisson field of interferers and rayleigh fading channels,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2461–2467, 2010.

- [58] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, “On the distribution of SINR for the MMSE MIMO receiver and performance analysis,” *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 271–286, 2005.
- [59] J. Ma, Y. J. Zhang, X. Su, and Y. Yao, “On capacity of wireless ad hoc networks with MIMO MMSE receivers,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5493–5503, 2008.
- [60] Z. Zhang, Y. Li, R. Wang, Y. Chen, and K. Huang, “Learning-based rate adaptation for uplink massive MIMO with a cooperative data-assisted detector,” in *Proc. IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2019.
- [61] Z. Zhang, Y. Li, R. Wang, and K. Huang, “Rate adaptation for downlink massive MIMO networks and underlaid D2D links: A learning approach,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1819–1833, 2019.
- [62] H. Lim and D. Yoon, “On the distribution of SINR for MMSE MIMO systems,” *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4035–4046, 2019.
- [63] L. Tan, Z. Zhang, and D. Wang, “Performance of multiuser downlink cell-free massive MIMO systems with hard deadlines,” *IEEE Access*, vol. 10, pp. 62910–62919, 2022.
- [64] S. Kurma, K. Singh, P. K. Sharma, and C.-P. Li, “Outage probability analysis of uplink cell-free massive MIMO with user mobility,” in *Proc. IEEE MILCOM*, pp. 37–42, 2022.
- [65] L. Zhou, L. Zheng, X. Wang, W. Jiang, and W. Luo, “Coordinated multicell multicast beamforming based on manifold optimization,” *IEEE Communications Letters*, vol. 21, no. 7, pp. 1673–1676, 2017.
- [66] A. de la Fuente, G. Interdonato, and G. Araniti, “User subgrouping and power control for multicast massive MIMO over spatially correlated channels,” *IEEE Transactions on Broadcasting*, vol. 68, no. 4, pp. 834–847, 2022.
- [67] J. Li, Q. Pan, Z. Wu, P. Zhu, D. Wang, and X. You, “Spectral efficiency of unicast and multigroup multicast transmission in cell-free distributed massive MIMO systems,” *IEEE Transactions on Vehicular Technology*, vol. 71, no. 12, pp. 12826–12839, 2022.
- [68] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1.” <https://cvxr.com/cvx>, Mar. 2014.
- [69] E. Chen and M. Tao, “ADMM-based fast algorithm for multi-group multicast beamforming in large-scale wireless systems,” *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2685–2698, 2017.

- [70] A. Konar and N. D. Sidiropoulos, “Fast approximation algorithms for a class of non-convex QCQP problems using first-order methods,” *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3494–3509, 2017.
- [71] C. Zhang, M. Dong, and B. Liang, “Ultra-low-complexity algorithms with structurally optimal multi-group multicast beamforming in large-scale systems,” *IEEE Transactions on Signal Processing*, 2023.

