Cell-Free Massive MIMO and Millimeter Wave Channel Modelling for 5G and Beyond

Manijeh Bashar

Doctor of Philosophy

University of York

Electronic Engineering

April 2019

Dedicated to Ali, who made my world a better place,
and who gives me joy in my times of trial,
and love in all I do.

Abstract

Huge demand for wireless throughput and number of users which are connected to the base station (BS) has been observed in the last decades. Massive multiple-input multiple-output (MIMO) is a promising technique for 5G for the following reasons; 1) high throughput; 2) serving large numbers of users at the same time; 3) energy efficiency. However, the low throughput of cell-edge users remains a limitation in realistic multi-cell massive MIMO systems. In cell-free massive MIMO, on the other hand, distributed access points (APs) are connected to a central processing unit (CPU) and jointly serve distributed users. This thesis investigates the performance of cell-free Massive MIMO with limited-capacity fronthaul links from the APs to the CPU which will be essential in practical 5G networks. To model the limited-capacity fronthaul links, we exploit the optimal uniform quantization. Next, closed-form expressions for spectral and energy efficiencies are presented. Numerical results investigate the performance gap between limited fronthaul and perfect fronthaul cases, and demonstrate that exploiting a relatively few quantization bits, the performance of limited-fronthaul cell-free Massive MIMO closely approaches the perfect-fronthaul performance. Next, the energy efficiency maximization problem and max-min fairness problems are considered with per-user power and fronthaul capacity constraints. We propose an iterative procedure which exploits a generalized eigen vector problem and geometric programming (GP) to solve the max-min optimization problem. Numerical results indicate the superiority of the proposed algorithms over the case of equal power allocation. On the other hand, the performance of communication systems depends on the propagation channel. To investigate the performance of MIMO systems, an accurate small scale fading channel model is necessary. Geometry-based stochastic channel models (GSCMs) are mathematically tractable models to investigate the performance of MIMO systems.

Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Alister G. Burr for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study. "*To our parents we owe our life, but to our teachers we owe the good life.*" Besides, the gratitude also goes to my co-supervisor, Dr. Kanapathippillai Cumanan, for his guidance. I have learned a lot from our day-to-day conversations and debates. Indeed, the accomplishment of this study would not have been possible without their generous lessons and support. I am very grateful to my thesis advisor, Dr. Yuriy Zakharov, whose insightful discussions and suggestions have benefited me.

I would like to thank Prof. Katsuyuki Haneda at Aalto University, Finland, for the cooperation, and for giving me a great opportunity to join his research group as a visiting scholar. I feel lucky to meet experts in this field. I am thankful to Dr. Hien Quoc Ngo at Queen's University Belfast, U.K., for the flourishing cooperation. I have learnt a lot from his knowledge and expertise. I would like to thank Dr. Nicolai Czink at Digital International GmbH, Austria, for the useful technical discussion and help. Kind thanks goes to Dr. Aki Karttunen at Aalto University, Finland, for the helpful guide and discussion. Many thanks to Dr. Amine Mezghani at University of Texas for technical issues and the useful comments. I am also very grateful to all of my co-authors for the great collaboration: Prof. Lajos Hanzo at University of Southampton, U.K., Prof. H. Vincent Poor at Princeton University, USA, Prof. Erik G. Larsson at Linköping University, Sweden, Prof. Mérouane Debbah at University of Paris-Saclay, Junior CentraleSupélec, France, and Prof. Pei Xiao, University of Surrey, U.K..

The kindest thank to my colleagues and friends at Communication Technologies

Research Group, University of York, for the helpful discussions, providing the nice and friendly environment. It was a profound experience during three years and half. Very warm thanks goes to Faezeh, Marjan, Kafi, Sunghyun, Yahya, Dick, Haitham and Abi. I would like to thank my husband, Ali, and to thank my parents and family, for their unbroken love and endless support throughout my life and from start to the end of this career.

York, May 2019
Manijeh Bashar

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. References and acknowledgements to other researchers have been given as appropriate.

Elements of the research presented in this thesis have resulted in some publications. A list of these publications can be found below.

Refereed Journal Publications

1. **M. Bashar**, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo "On the max-min SINR of cell-free massive MIMO," *IEEE Transactions on Wireless Communications*, Jan. 2019.
2. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, and H. V. Poor "Mixed quality of service in cell-free Massive MIMO," *IEEE Communications Letter*, Apr. 2018.
3. **M. Bashar**, A. G. Burr, K. Haneda, and K. Cumanan, "Robust user scheduling with COST 2100 channel model for Massive MIMO networks," *IET Microwaves, Antennas and Propagation*, Apr. 2018.
4. D. Maryopi, **M. Bashar**, and A. G. Burr "On The uplink throughput of zero-forcing in cell-free massive MIMO with coarse quantization," *IEEE Transactions on Vehicular Technology*, Accepted.

Conference Publications

1. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "On the energy efficiency of limited-backhaul cell-free massive MIMO," *IEEE ICC*,

May 2019.

2. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, L. Hanzo, and P. Xiao, "NOMA-based cell-free massive MIMO," *IEEE ICC*, May 2019.
3. **M. Bashar**, K. Haneda, A. G. Burr, and K. Cumanan, "A study of dynamic multipath clusters at 60 GHz in a large indoor environment," *IEEE Globecom Workshop*, Dec. 2018.
4. **M. Bashar**, H. Q. Ngo, K. Cumanan, A. G. Burr, and E. G. Larsson "On the performance of backhaul constrained cell-free Massive MIMO with linear receivers," *IEEE Asilomar*, Oct. 2018.
5. **M. Bashar**, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "Enhanced max-min SINR for uplink cell-free Massive MIMO systems," *IEEE ICC*, May 2018.
6. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, and M. Debbah, "Cell-free Massive MIMO with limited backhaul," *IEEE ICC*, May 2018.
7. **M. Bashar**, A. G. Burr, K. Haneda, and K. Cumanan, "Robust geometry-based user scheduling for large MIMO systems under realistic channel conditions," *IEEE EW*, May 2018.
8. **M. Bashar**, A. G. Burr, and K. Cumanan, "Low complexity Massive MIMO techniques for 5G communications," *International Union of Radio Science*, Aug. 2017.
9. A. G. Burr, **M. Bashar**, and D. Maryopi, "Ultra-dense radio access networks for smart cities: cloud-RAN, fog-RAN and cell-free massive MIMO," *IEEE PIMRC*, Sep. 2018.
10. A. G. Burr, **M. Bashar**, and D. Maryopi, "Cooperative access networks: optimum fronthaul quantization in distributed massive MIMO and cloud RAN," *IEEE VTC*, Jun. 2018.

Technical Meeting Papers

1. **M. Bashar**, K. Handea, and A. G. Burr, "Spatial consistency of clusters in mm-wave ray-tracing results for 5G communications," in the Meeting of European Cooperation in Science and Technology (eCOST), Cartagena, Spain, Jun. 2018.

2. **M. Bashar**, A. G. Burr, K. Handea, and K. Cumanan, "Correlation-based schemes in Massive MIMO systems with COST 2100 channel model," in the Meeting of eCOST, Lund, Sweden, Jun. 2017.
3. **M. Bashar**, A. G. Burr, and K. Cumanan, "Simplified user selection for large MIMO," *Festival of Radio Science*, York, U.K., Dec. 2016.
4. **M. Bashar**, A. G. Burr, and K. Cumanan, "On the scheduling for large MIMO systems with COST 2100 channel model," in the Meeting of eCOST, Durham, U.K., Oct. 2016.

Chapter 1

Introduction

Contents

1.1 Overview and Motivation

In recent years, urban population growth has attracted multi-disciplinary research attention. Not only are people using more mobile data per capita (2x per year), but the population density in urban areas has also grown. As a result, the data demand density per unit area has grown exponentially, and will continue to do so in most parts of the world. Sustainable growth of wireless infrastructure to match this demand is crucial to the digital economy and lifestyle. The main goal of wireless communications is to achieve high data capacity without delay. Over the past few years, an abundance of techniques have been proposed as a means to efficiently scale the wireless capacity. It remains unclear which technology or set of technologies can meet the demand.

Massive multiple input multiple output (MIMO) is one of the potential technologies for the 5th Generation (5G) cellular networks. The conventional way to support a large geographical area with wireless transmission is to exploit a cellular network methodology [3], where each base station (BS) serves a set of users. Note that it is not possible to use a single cell massive MIMO system which covers an effectively unlimited area, such as an entire country or continent. Hence, we have to divide the large geographical area into many cells resulting in a multi-cell massive MIMO system [3]. In the following, similar to the methodology in [4], we consider a given large geographical area and provide a basis for comparison between different scenarios.

A massive MIMO system can be defined by a 5-tuple $(N_a, N_u, R, \mathbf{G}, \mathbf{W})$, where N_a and N_u refer to the total number of service antennas and total number of users in the area with $N_a \gg N_u$, $\mathbf{G}$ is the $N_u \times N_a$ channel matrix between the N_u users and the N_a antennas, $\mathbf{W}$ refers to the $N_a \times N_u$ pre-coding/de-coding matrix between the N_a antennas and the N_u message symbols for/from the N_u users, and R km^2 is the size of the service area [4]. For a given area, there are different scenarios to implement massive MIMO technology. Based on [4], some of them are listed as follows: i) Single cell massive MIMO: a single BS equipped with N_a antennas communicates with N_u users distributed in the cell with size R km^2 ii) multi cell massive MIMO: the service area of size R km^2 is divided into N_c cells with one BS in each cell, where $N_c \times N_{a,\text{BS}} = N_a$, where $N_{a,\text{BS}}$ is the number of antennas at each BS. The BSs jointly serve distributed users in the cells, and iii) cell-free massive MIMO: large number of access points (APs) N_{AP}, with few antennas $N_{a,\text{AP}}$ at each, are randomly distributed through the coverage area with size R km^2 [4]. The bottleneck in single-cell and multi-cell networks is the poor performance of cell edge users [5]. To deal with this problem, cell-free massive MIMO is introduced. In cell-free massive MIMO, there are no cells, and hence, no boundaries. All users in the network are coherently served by several randomly distributed APs via a central processing unit (CPU) [2, 6]. In [7] a user-centric approach is proposed where each user is served by a small number of APs. Cell-free massive MIMO effectively implements a user-centric approach [8]. Cell-free massive MIMO is a scalable versions of the network-MIMO systems, or coordinated multipoint processing (CoMP) [9, 10]. Cell-free massive MIMO is likely a key element of next-generation networks [8]. In [11], Marzetta characterized the performance of massive MIMO systems in the context of time division duplexing (TDD), which has widely inspired the

community [12], which is capable of outperforming frequency division duplexing (FDD).

On the other hand, another promising set of technologies for 5G is reviewed in [13–15]: the combination of large antenna arrays and short wavelength carrier waves. To defeat the effect of path loss at millimetre-wave (mm-wave) bands necessitates highly directional antennas, which implies that mm-wave communication systems must deploy massive MIMO. Moreover, this combination allows for a greater bandwidth availability and extremely high spectral efficiency by utilizing a large number of antennas, whilst occupying a relatively small area. This technology is known as massive MIMO at the mm-wave spectrum. Therefore, we investigate an accurate channel model in mm-wave bands in Chapter 6 of this thesis.

The work in this thesis will consider several significant challenges with a view to contributing to 5G. Potential challenges that will be considered for this objective involve:

- One of the main issues of cell-free massive MIMO systems which requires more investigation is the limited-capacity fronthaul links from the APs to a CPU. The assumption of infinite fronthaul in [2, 16] is not realistic in practice. The fronthaul requirements for massive MIMO systems, including small-cell and macro-cell BSs have been investigated in [17]. The fronthaul load is the main challenge in any distributed antenna systems [17];
- A significant challenge of the 5G communication networks is the huge throughput difference among the users, which is addressed as fairness. Hence, one of the main tasks of massive MIMO is to provide good performance for all users in the network, which is referred to max-min fairness. This reveals the significance of max-min fairness power control problem which maximizes the smallest of all user rates;
- In [18], the authors show that the collocated massive MIMO is energy-efficient. On the other hand, the authors in [16] investigate the energy efficiency of cell-free massive MIMO downlink with error-free and unlimited-capacity fronthaul links. However, it is not yet obvious how energy efficient a cell-free massive MIMO with limited-fronthaul capacity is;
- Another important issue in massive MIMO systems is user scheduling to maximize

multiuser diversity gain with imperfect CSI. Recently, a range of user scheduling schemes have been proposed for large MIMO systems. Most of these, such as that described in [19], require accurate knowledge of the channel from all potential users to the BS -which in the massive MIMO case is completely infeasible to obtain. However, a simplified correlation-based user scheduling algorithm, by considering massive MIMO simplifications and the effect of the cell geometry is still an open problem.

- Spatially consistent channel modeling at mm-waves: Standardized MIMO channel models such as the 3GPP channel model for new radios [20] are geometry-based stochastic channel models (GSCMs) [21]. However, the available GSCMs at mm-waves do not necessarily retain spatial consistency in simulated channels, which is essential for small cells with ultra-dense users. In Chapter 6 of thesis, we work on parameterization of the COST 2100 channel model, which ensures the spatial consistency, using a ray-tracer which is adjusted to produce results consistent with measurements.

1.2 Aims

The aim of this work is to overcome challenges of implementing an ultra-dense radio access network using ”cell-free” massive MIMO techniques and COST 2100 channel model. The objectives are summarised as follows:

- Exploiting the knowledge channel statistics at the receiver (users' side in the downlink and CUP's end in the uplink) to derive a closed-form expression for spectral efficiency cell-free massive MIMO.
- To model the effect of quantization due to the limited-capacity fronthaul links from the APs to the CPU.
- Solving the non-convex max-min fairness problem to maximize the smallest rate of the users.
- Investigation of energy efficiency of cell-free massive MIMO assuming fronthaul limited-capacity fronthaul links.

- Finding a practical user assignment algorithm for cell-free massive MIMO.
- Presenting low complexity correlation-based user scheduling and beamforming design for collocated massive MIMO under a geometry-based stochastic channel model.
- Parameterizing COST2100 channel model at mm-wave band based on a ray-tracer optimized with measurements.

In this thesis, we investigate the practical requirements in cell-free massive MIMO, namely, the effect of quantization error and the problem of user assignment. Finally, in this thesis, we aim at bridging the gap between the two groups of research, i.e. this thesis develops a realistic and practical massive MIMO scheme with the realistic COST 2100 channel model.

1.3 Contributions

The contributions of this thesis are summarised as follows:

- We consider two cases in which: i) the quantized versions of the channel estimates and the received signals at the APs are available at the CPU and ii) the quantized versions of processed signals at the APs are available at the CPU. The corresponding achievable rates are derived by using the Use-and-then-Forget (UaF) bounding technique taking into account the effects of channel estimation error and quantization error. Next, to improve the performance of the cell-free massive MIMO system, we propose to use a novel receiver filter, operating at the CPU, which can be designed based only on the statistics of the channel.

- The uniform quantizer is investigated. Next, we make use of the Bussgang decomposition to model the effect of quantization and present the analytical solution to find the optimal step size of the quantizer.

- We propose a max-min fairness problem which maximizes the smallest of all user rates under the per-user power and fronthaul capacity constraints. To solve this problem, the original max-min fairness problem is decomposed into two sub-problems and an iterative algorithm is developed to solve the max-min optimization problem. The optimality of the proposed algorithm is proved through establishing the uplink-downlink duality for the cell-free massive MIMO system with limited-capacity fronthaul links. We provide the convergence and complexity analysis of the proposed scheme.

- A novel and efficient user assignment algorithm based on the capacity of fronthaul links is proposed which results in significant performance improvement.

- An expression for uplink energy efficiency is derived based on channel statistics and taking into account the effects of channel estimation errors, the effect of pilot contamination, and quantization error. A novel approach to solve the non-convex energy efficiency maximization problem is proposed, where we propose to decompose the original problem into two sub-problems and an iterative algorithm is developed to determine the optimal solution. Successive convex approximation (SCA) is used to efficiently solve the power allocation problem. Next, a heuristic sub-optimal energy efficiency maximization problem is proposed where the original optimization problem is transformed into a standard geometric programme (GP). The convergence and complexity analysis of the proposed schemes are presented.

- Exploiting the eigenvalue spectrum of the channel covariance matrix, we propose to use the angular bins to build up an approximate eigenchannel, which can be used for linear precoding design. Next, a new user scheduling scheme is proposed under the assumption of no instantaneous channel at the BS, other than the channel correlation. The complexity of the proposed scheme is presented.

- We work on channel parameterization for the COST 2100 channel model using mobile channel simulations at 61 GHz in Helsinki Airport. We study whether clusters exist or not. For the first time, we perform clustering of dynamic multipath

channels, where we identify and track clusters based on the spatial coordinates of the multi-path components (MPCs).

1.4 Thesis Outline

The thesis structure is as follows:

- Chapter 2: **Background**

  The chapter is dedicated the introduction of the basic concepts of multi-user MIMO, massive MIMO, and cell-free massive MIMO. Different ways to implement massive MIMO are described. Next, the fundamental theory of channel estimation is explained. The basic principle of uniform quantizer is reviewed. We next present the details of finding the optimal step size of uniform quantizer. The achievable rate is next explained for the case of unknown channel gain at the receiver. The details and parameters of COST 2100 channel model and the optimal uniform quantizer are next provided.

- Chapter 3: **Max-Min Rate of Cell-Free Massive MIMO Uplink**

  In this chapter, a cell-free massive MIMO with limited-capacity fronthaul links is considered. The optimal uniform quantizer is exploited to quantize the signals at the APs. The max-min rate optimization problem is investigated where the smallest rate of the users is maximized. An iterative algorithm is proposed to tackle the non-convexity of the problem. Next, we propose an efficient AP assignment to decreases the load of fronthaul links. The convergence and complexity of the proposed algorithm are provided. Finally, numerical results are provided to validate the proposed algorithms.

- Chapter 4: **Energy Efficiency of the Cell-Free Massive MIMO Uplink**

  In this chapter, the total energy consumption in a cell-free massive MIMO uplink is modeled, enabling us to define the energy efficiency optimization problem in cell-free massive MIMO uplink. The problem is not convex in its initial form and a

convex approximation is proposed to efficiently solve the original optimization problem. Next, to reduce the complexity of the system an sub-optimal algorithm is proposed. Numerical results confirm that the sub-optimal scheme has a very efficient performance while decreasing the complexity of the system. Next, we investigate the complexity and convergence of the proposed algorithms. Finally, numerical results are provided which show the effectiveness of the proposed algorithms compared to the scheme with equal power allocation.

- Chapter 5: **Low Complexity Massive MIMO Techniques Under Realistic Channel**

  In this chapter, a collocated massive MIMO system with realistic geometry-based COST 2100 channel model is considered. The problem of joint user scheduling and beamforming design is considered while the transmitter has access to the knowledge of the correlation of the channel. The proposed scheme exploits the Fourier transform of the correlation matrix at the transmitter. Numerical results demonstrate that the proposed algorithm achieve a performance very close to the performance of the scheme with the estimated channel available at the transmitter.

- Chapter 6: **Dynamic Multipath Clustering at 61 GHz**

  In this chapter, we consider ray-tracing results in a large scale environment in Helsinki airport. Using the $[x, y, z]$ information of the MPCs, we group the MPCs into clusters. A Kalman filter is exploited to track the cluster positions for different snapshots. Next, we extract the distribution of cluster life time, distribution of total number of clusters and distribution of number of MPCs per cluster.

- Chapter 7: **Conclusions and Future Work**

  In this chapter, we present the conclusions of the thesis and the possible future research directions.

1.5 Publication List

Refereed Journal Publications

1. **M. Bashar**, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo "On the max-min SINR of cell-free massive MIMO," *IEEE Transactions on Wireless Communications*, Jan. 2019.
2. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, and H. V. Poor "Mixed quality of service in cell-free massive MIMO," *IEEE Communications Letter*, Apr. 2018.
3. **M. Bashar**, A. G. Burr, K. Haneda, and K. Cumanan, "Robust user scheduling with COST 2100 channel model for massive MIMO networks," *IET Microwaves, Antennas and Propagation*, Apr. 2018.
4. D. Maryopi, **M. Bashar**, and A. G. Burr "On The uplink throughput of zero-forcing in cell-free massive MIMO with coarse quantization," *IEEE Transactions on Vehicular Technology*, May 2019.

Under Review Journal Papers

1. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, M. Debbah, and P. Xiao, "Max-min rate of cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Transactions on Communications (Passed the third review; two reviewers suggested to accept and one asked a few questions).*
2. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, L. Hanzo, and P. Xiao, "Analysis of the cell-free massive MIMO downlink relying on adaptive NOMA/OMA mode-switching," *IEEE Transactions on Communications (Passed the first review).*
3. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "On the energy efficiency of cell-free massive MIMO with optimal quantization," *IEEE Transactions on Green Communications and Networking (Passed the first review).*
4. **M. Bashar**, A. G. Burr, K. Haneda, K. Cumanan, and M. M. Molu, and P. Xiao, "Low complexity massive MIMO techniques under realistic channel conditions," *IEEE Transactions on Vehicular Technology (Passed the third review; two reviewers suggested to accept and one asked a few questions).*

To be submitted Journal Papers

1. **M. Bashar**, H. Q. Ngo, K. Cumanan, A. G. Burr, E. G. Larsson, and P. Xiao,"C-RAN and cell-free massive MIMO: Spectral and energy efficiency with optimal uniform quantization," *IEEE Transactions on Communications*, May. 2019.

Conference Publications

1. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "On the energy efficiency of limited-backhaul cell-free massive MIMO," *IEEE ICC*, May 2019.
2. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, L. Hanzo, and P. Xiao, "NOMA-based cell-free massive MIMO," *IEEE ICC*, May 2019.
3. **M. Bashar**, K. Haneda, A. G. Burr, and K. Cumanan, "A study of dynamic multipath clusters at 60 GHz in a large indoor environment," *IEEE Globecom Workshop*, Dec. 2018.
4. **M. Bashar**, H. Q. Ngo, K. Cumanan, A. G. Burr, and E. G. Larsson "On the performance of backhaul constrained cell-free massive MIMO with linear receivers," *IEEE Asilomar*, Oct. 2018.
5. **M. Bashar**, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "Enhanced max-min SINR for uplink cell-free massive MIMO systems," *IEEE ICC*, May 2018.
6. **M. Bashar**, K. Cumanan, A. G. Burr, H. Q. Ngo, and M. Debbah, "Cell-free massive MIMO with limited backhaul," *IEEE ICC*, May 2018.
7. **M. Bashar**, A. G. Burr, K. Haneda, and K. Cumanan, "Robust geometry-based user scheduling for large MIMO systems under realistic channel conditions," *IEEE EW*, May 2018.
8. **M. Bashar**, A. G. Burr, and K. Cumanan, "Low complexity massive MIMO techniques for 5G communications," *International Union of Radio Science*, Aug. 2017.
9. A. G. Burr, **M. Bashar**, and D. Maryopi, "Ultra-dense radio access networks for smart cities: cloud-RAN, fog-RAN and cell-free massive MIMO," *IEEE PIMRC*, Sep. 2018.

10. A. G. Burr, **M. Bashar**, and D. Maryopi, "Cooperative access networks: optimum fronthaul quantization in distributed massive MIMO and cloud RAN," *IEEE VTC*, Jun. 2018.

Under-Review Conference Papers

1. **M. Bashar**, A. Akbari, K. Cumanan, A. G. Burr, P. Xiao, H. Q. Ngo, and M. Debbah "On the gain of deep learning in limited-capacity fronthaul cell-free massive MIMO," *IEEE Globecom'19*.

Technical Meeting Papers

1. **M. Bashar**, K. Handea, and A. G. Burr, "Spatial consistency of clusters in mm-wave ray-tracing results for 5G communications," in the Meeting of European Cooperation in Science and Technology (eCOST), Cartagena, Spain, Jun. 2018.
2. **M. Bashar**, A. G. Burr, K. Handea, and K. Cumanan, "Correlation-based schemes in massive MIMO systems with COST 2100 channel model," in the Meeting of eCOST, Lund, Sweden, Jun. 2017.
3. **M. Bashar**, A. G. Burr, and K. Cumanan, "Simplified user selection for large MIMO," *Festival of Radio Science*, York, U.K., Dec. 2016.
4. **M. Bashar**, A. G. Burr, and K. Cumanan, "On the scheduling for large MIMO systems with COST 2100 channel model," in the Meeting of eCOST, Durham, U.K., Oct. 2016.

Chapter 2

Background

Contents

2.1 Introduction

Massive MIMO is a technology which is obtained by increasing the number of BS antennas and number of users in MU-MIMO systems. In massive MIMO, a very large

number of antennas (hundreds or even thousands of antennas) communicate with a large number of users, whereas the number of users is much smaller than the number of BS antennas. In this chapter we study the fundamentals of massive MIMO systems. Next, we present the basics of the COST 2100 channel model. The spectral and energy efficiency of the massive MIMO system depend on the geometry of area. Hence, to investigate the performance of massive MIMO in a real and practical propagation environment, an accurate geometry-based channel model is essential. Finally, we provide the details of the quantization model which we have used to model the quantization error in cell-free massive MIMO.

The chapter is organised as follows: First, an introduction to multi-user MIMO is presented in Section 2.3 which is followed by interesting practical aspects of massive MIMO in Section 2.4 and the details of channel estimation in Section 2.5, respectively. Next, the basics of massive MIMO and different types of massive MIMO are introduced in Sections 2.6 and 2.7, respectively. Then, we present the basic concepts of the realistic geometry-based COST 2100 channel model in Section 2.8 and detailed analysis of uniform quantization to find the optimal step size of the quantizer are given in Section 2.9. The details to calculate the achievable rate with unknown channel gain at the receiver and non-Gaussian noise is discussed in Section 2.10. Finally, a summary of the chapter is provided in Section 2.11.

2.2 Notation

The following notations are adopted in the rest of the chapter. Uppercase and lowercase boldface letters are used for matrices and vectors, respectively. The notation $\mathbb{E}\{\cdot\}$ denotes expectation. $|\cdot|$ stands for absolute value. The conjugate transpose of vector $\mathbf{x}$ is $\mathbf{x}^H$, and $\mathbf{X}^T$ denotes the transpose of matrix $\mathbf{X}$. In addition, it is assumed that $x \sim \mathcal{CN}(0, \sigma^2)$ represents a zero-mean circularly symmetric complex Gaussian RV with variance σ^2. The conjugate of the variable x is presented by $x^\star$. Moreover, $[\mathbf{x}]_n$, $\mathcal{R}(x)$ and $\mathcal{I}(x)$ represent the nth element of vector $\mathbf{x}$, the real part and imaginary part of the complex variable x, respectively. Next, diag$[\mathbf{x}]$ refers to a diagonal matrix whose diagonal elements are the elements of vector $\mathbf{x}$. Note that $\mathbf{I}_N$ refers to the $N \times N$ identity matrix. The notation $|x|$,

$|\mathbf{X}|_{\text{det}}$ and $|\mathbf{x}|_{\text{size}}$ stand for the absolute value of x, determinant of matrix $\mathbf{X}$, and the size of vector $\mathbf{x}$, respectively. Finally, the Kronecker product of $\mathbf{X}$ and $\mathbf{Y}$ is presented by $\mathbf{X} \otimes \mathbf{Y}$.

2.3 The Basics of Multi-User MIMO

In this section, the basic concepts of multi-user MIMO are presented. Note that massive MIMO can be regarded as multi-user MIMO with a large number of antennas at the BS. So, it is essential to understand the basics of multi-user MIMO.

2.3.1 System Model

Consider a single cell MU-MIMO system with M antennas at the BS and K single antenna users. For the sake of simplicity, we suppose all users and the BS work on the same time-frequency resource. Moreover, we assume a channel that is known at the BS and the users. The training schemes for frequency-division duplex (FDD) and time-division duplex (TDD) are discussed at the end of this chapter.

2.3.2 Uplink Transmission

We consider uplink transmission in a single cell massive MIMO system with M antennas at the BS and K single antenna users. The $M \times 1$ received signal at the BS is given by

$$\mathbf{y}_{up} = \sqrt{p_{up}}\mathbf{H}\mathbf{s} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{s} = [s_1, \cdots, s_K]^T$ represents the symbol vector of K users, p_{up} is the average power for each user in uplink mode. The elements of the noise $\mathbf{n}$ are independent and identically distributed (i.i.d.) Gaussian random variables (RVs) with zero mean and unit variance, and independent of the aggregate channel matrix $\mathbf{H} \in \mathcal{C}^{M \times K}$. Based on the analysis in [22], assuming knowledge of the channel at both transmitter and receiver, the

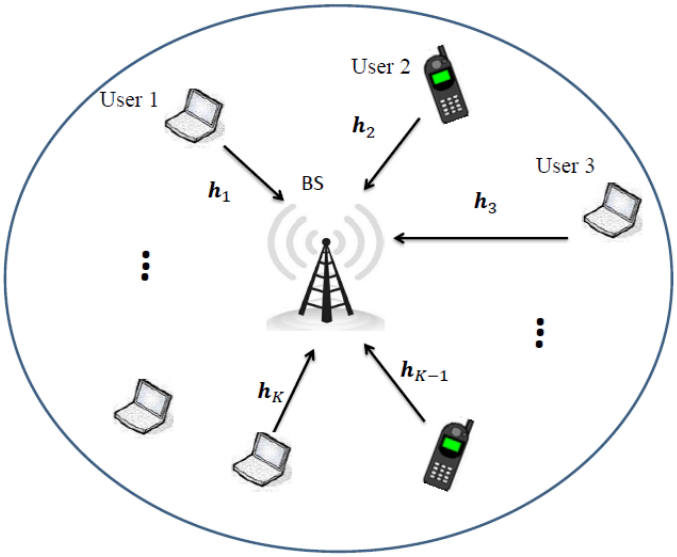

Figure 2.1: The uplink of a MU-MIMO system with K single-antenna users in a single cell.

Shannon sum-capacity is obtained by

$$C_{\text{sum cap}} = \log_2 \det \left(\mathbf{I}_K + p_{up} \mathbf{H}^H \mathbf{H} \right), \tag{2.2}$$

where $\mathbf{I}_K$ refers to $K \times K$ identity matrix. Note that the sum-capacity is achieved by successive interference cancellation (SIC), where users' signals are successively decoded and subtracted out of the received signal [22].

2.3.3 Downlink Transmission

Suppose $\mathbf{x} \in \mathcal{C}^{M \times 1}$ denotes the transmitted signal from the BS in downlink mode. Then, the received signal at the kth user is given by

$$y_{dl,k} = \sqrt{p_d} \mathbf{h}_k^T \mathbf{x} + z_k, \tag{2.3}$$

where $\sqrt{p_d}$ denotes the average signal-to-noise ratio (SNR). It is assumed that $\bar{p}_d$ denotes the data powers, where $p_d = \frac{\bar{p}_d}{p_z}$, where p_z is the power of noise. Moreover, z_k represents the Gaussian additive noise at the user k. Moreover, $\mathbf{h}_k \in \mathcal{C}^{M \times 1}$ is the channel from the

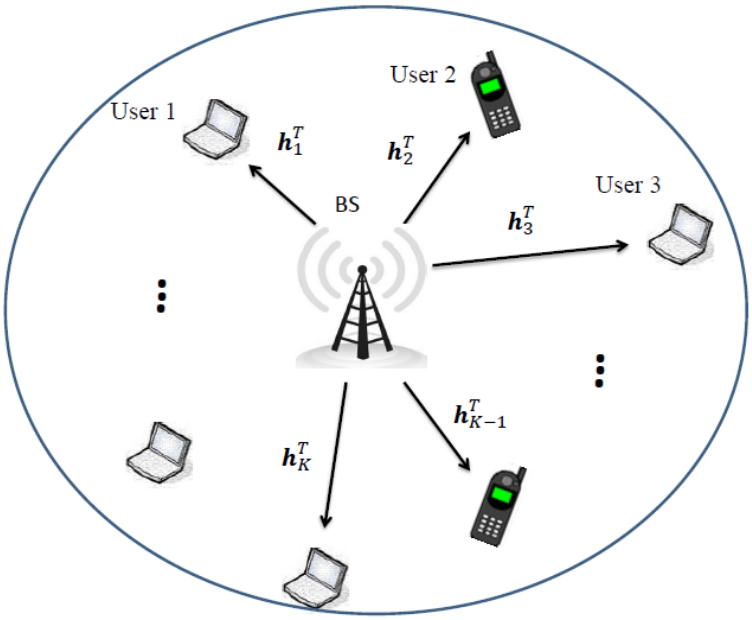

Figure 2.2: The downlink of a multi MU-MIMO system with K single-antenna users in a single cell.

kth user to the BS. The vector of received signals at the K users is given by

$$\mathbf{y}_{dl} = \sqrt{p_d}\mathbf{H}^T\mathbf{x} + \mathbf{z}, \tag{2.4}$$

where $\mathbf{y}_{dl} = [y_{dl,1}, y_{dl,2}, \cdots, y_{dl,K}]^T$, $\mathbf{z}_{dl} = [z_1, z_2, \cdots, z_K]^T$ and $\mathbf{H} = [\mathbf{h}_1 \cdots \mathbf{h}_K]$.

2.3.4 Linear Receivers in Uplink

The optimal performance in MU-MIMO is obtained by maximum-likelihood (ML) detection. However, as linear processing is nearly-optimal in massive MIMO [11], we consider linear processing in this thesis. The received signal $\mathbf{y}_{up}$ is multiplied by the $K \times M$ matrix $\mathbf{W}^H$ as follows:

$$\tilde{\mathbf{y}}_{up} = \sqrt{p_{up}}\mathbf{W}^H\mathbf{H}\mathbf{s} + \mathbf{W}^H\mathbf{n}. \tag{2.5}$$

Each stream of $\tilde{\mathbf{y}}_{up}$ is decoded. Then, the kth stream is given by

$$\tilde{y}_{up,k} = \sqrt{p_{up}}\mathbf{w}_k^H\mathbf{h}_k s_k + \sqrt{p_{up}}\sum_{k' \neq k}^{K}\mathbf{w}_k^H\mathbf{h}_{k'}s_{k'} + \mathbf{w}_k^H\mathbf{n}, \tag{2.6}$$

where $\mathbf{w}_k$ represents the kth column of $\mathbf{W}$. Hence, the SNR of the kth stream is given by

$$SINR_k = \frac{p_{up} \mid \mathbf{w}_k^H \mathbf{h}_k \mid^2}{p_{up} \sum_{k' \neq k} \mid \mathbf{w}_k^H \mathbf{h}_{k'} \mid^2 + \parallel \mathbf{w}_k \parallel^2} \tag{2.7}$$

We consider three linear multi user detectors: Maximum Ratio Combining (MRC), Zero-forcing (ZF), and Minimum mean-square error (MMSE) receiver. The details are described below.

1) Maximum Ratio Combining (MRC): where $\mathbf{w}_k = \mathbf{h}_k^H$. Hence, the received SINR of the kth symbol for MRC is given by

$$SINR_k = \frac{p_{up} \parallel \mathbf{h}_k \parallel^4}{p_{up} \sum_{k' \neq k}^{K} \mid \mathbf{h}_k^H \mathbf{h}_{k'} \mid^2 + \parallel \mathbf{h}_k \parallel^2}. \tag{2.8}$$

In this case, as the received signal at the BS is multiplied by $\mathbf{H}^H$, the signal processing of MRC is simple.

2) Zero-forcing (ZF) receiver: In ZF receiver, the multiuser interference is completely removed by projecting each stream onto the orthogonal complement of signal. For the case of ZF receiver, we can write:

$$\tilde{\mathbf{y}}_{up} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{y}_{up} = \sqrt{p_{up}} \mathbf{s} + (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{n}. \tag{2.9}$$

Note that to make the matrix $\mathbf{H}^H \mathbf{H}$ invertible, we need $M \geq K$. The kth stream of $\tilde{\mathbf{y}}_{up}$ is given by:

$$\tilde{y}_{up} = \sqrt{p_{up}} s_k + \tilde{n}_k, \tag{2.10}$$

where the term $\tilde{n}_k$ is the kth element of $(\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{n}$. The SINR of the kth user is given by

$$SINR_k = \frac{p_{up}}{\left[(\mathbf{H}^H \mathbf{H})^{-1} \right]_{kk}}. \tag{2.11}$$

3) Minimum mean-square error (MMSE) receiver: the MMSE receiver has the best performance among linear receiver, as it minimizes the mean-square error between the estimated signal and the transmitted signal. The MMSE receiver is obtained by solving the following equation

$$\mathbf{W}^{\text{MMSE}} = \min_{\mathbf{W}} \mathbb{E} \left\{ \left| \mathbf{W}^H \mathbf{H} - \mathbf{s} \right|^2 \right\}. \tag{2.12}$$

Based on the analysis in [23], the optimum MMSE receiver is obtained as follows:

$$\mathbf{w}_k^{\text{MMSE}} = \sqrt{p_{up}} \left(\mathbf{H}\mathbf{H}^H + \mathbf{I}_M \right)^{-1} \mathbf{h}_k, \tag{2.13}$$

where $\mathbf{w}_k^{\text{MMSE}}$ refers to the kth column of $\mathbf{W}^{\text{MMSE}}$.

2.3.5 Linear Precoding in Downlink

In the downlink, the signal transmitted from M antennas, $\mathbf{x}$, is a linear combination of the symbols of the all K users. Suppose s_k with $\mathbb{E}\{|s_k|^2\} = 1$ is the symbol intended for the kth user. The precoded signal is given by

$$\mathbf{x} = \sqrt{c}\mathbf{W}\mathbf{s}, \tag{2.14}$$

where $\mathbf{x} = [s_1 \cdots s_K]^T$, $\mathbf{W} \in \mathbb{C}^{M \times K}$ refers to the precoding matrix, and $c = \dfrac{1}{\mathbb{E}\{\text{tr}(\mathbf{W}\mathbf{W}^H)\}}$ is a normalization constant chosen to satisfy the power constraint $\mathbb{E}\{||s||^2\} = 1$. Hence, the received signal at the kth user is given by

$$\tilde{y}_{\text{dl},k} = \sqrt{cp_{dl}}\mathbf{h}_k^T\mathbf{w}_k s_k + \sqrt{cp_{dl}} \sum_{k' \neq k}^{K} \mathbf{h}_k^T \mathbf{w}_{k'} s_{k'} + n_k. \tag{2.15}$$

Hence, the SINR of the kth symbol for Conjugate Beamforming is given by

$$SINR_k = \frac{p_{up}|\mathbf{h}_k^T\mathbf{w}_k|^2}{p_{up}\sum_{k' \neq k}^{K} \mid \mathbf{h}_k^H\mathbf{w}_{k'} \mid^2 + 1}. \tag{2.16}$$

Three conventional linear beamformers are conjugate beamforming, ZF, and MMSE precoders. The basic concept and properties of these beamformers are the same as linear receivers. Hence, the we provide their definitions as follows:

$$\mathbf{W} = \begin{cases} \mathbf{H}^*, & \text{Conjugate beamforming}, \\ \mathbf{H}^* \left(\mathbf{H}^T\mathbf{H}^* \right)^{-1}, & \text{ZF}, \\ \mathbf{H}^* \left(\mathbf{H}^T\mathbf{H}^* + \frac{K}{p_{dl}}\mathbf{I}_K \right)^{-1}, & \text{MMSE}. \end{cases} \tag{2.17}$$

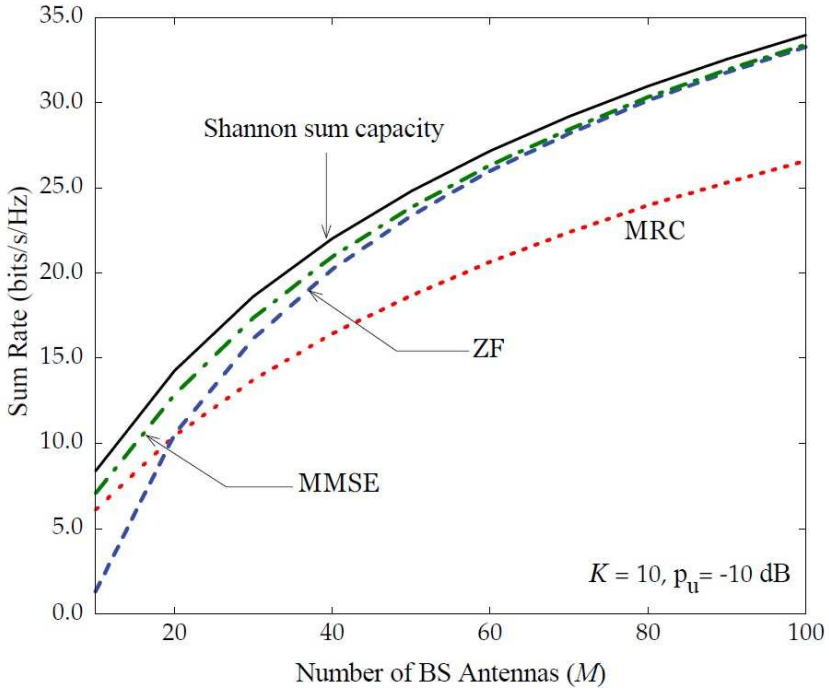

Figure 2.3: The performance of uplink sum rate for the linear receivers and the optimal reciever. Taken from [1].

2.4 Why Is Massive MIMO Practically Interesting?

In massive MIMO, when the number of BS antennas is large, due to the law of large numbers, the channels of users become orthogonal (please see Section 2.6.2 for the details). As a result, linear processing is almost optimal. Fig. 2.3 demonstrates the sum rate performance of the system versus the number of BS antennas with linear receivers. Let us assume perfect channel knowledge at both receiver and transmitter and assuming successive decoding, in which users' signals are successively decoded and subtracted out of the received signal. Hence the Shannon sum capacity is obtained by [24]

$$C_{\text{sum cap}} = \log_2 \det \left(\mathbf{I}_K + p_{up} M \mathbf{I}_K \right). \tag{2.18}$$

As the figure shows, by increasing number of antennas M, the sum rate performance with linear processing becomes very close to the Shannon sum capacity.

2.5 Channel Estimation

The channel needs to be estimated at the BS and users. The estimation scheme depends on the operating modes; FDD and TTD. In this section, we review the general approach to estimate the channel at both TDD and FDD.

2.5.1 Channel Estimation in TDD Mode

In TDD, the uplink and downlink channel are the same. Hence, the CSI can be obtained by the following approach [1]:

- Uplink mode: The BS requires CSI for detecting the transmitted signal by the users. At the first step, the K users transmit K orthogonal pilot sequences to the BS and the BS estimates CSI based on the received data.

- Downlink mode: The BS requires CSI to design the precoding matrix and each user needs the effective channel gain to detect the signal of interest. At the first step, similar to the uplink mode, the K users transmit K orthogonal pilot sequences to the BS and the BS estimates CSI. As the next step, the BS beamforms pilot sequences and sends them to the K users. The users estimate the effective channel gains via the received signal.

As a result, the training scheme in the TDD mode requires $2K$ symbols (K symbols for uplink and K symbols for downlink). Let us suppose channel is constant during T symbols. Hence, the following condition should be satisfied:

$$2K < T. \tag{2.19}$$

2.5.2 Channel Estimation in FDD Mode

Different frequencies are used in the uplink and downlink of FDD mode. Hence, the uplink and downlink channels in FDD systems are not the same. In this subsection, we describe a scheme to obtain the channel in the uplink and downlink transmissions of FDD systems.

- Downlink mode: CSI is required at the BS to design the beamforming matrix. The BS sends M orthogonal pilots to the K users. Each user estimates the channel based on the received signal. At the next step, the K users feed back the estimated channel to the BS. So, M symbols for the downlink and K symbols for the uplink are required by this scheme [1].

- Uplink mode: The K users send K orthogonal pilots to the BS. Next, the BS estimates the channels based on the received signal. This scheme needs K channels uses [1].

As a result, one could conclude that the coherence time of the channel (in symbol) should be large enough to transmit $M + K$ symbols in downlink and M symbols in uplink mode. Hence, the channel estimation scheme for FDD mode requires the following condition:

$$T > \begin{cases} M + K, & \text{Downlink} \\ M, & \text{Uplink .} \end{cases} \tag{2.20}$$

2.6 The Basics of Massive MIMO

First, we discuss the basic concept that TDD provides better performance than FDD in massive MIMO systems. Next, favourable propagation and channel hardening in massive MIMO is provided.

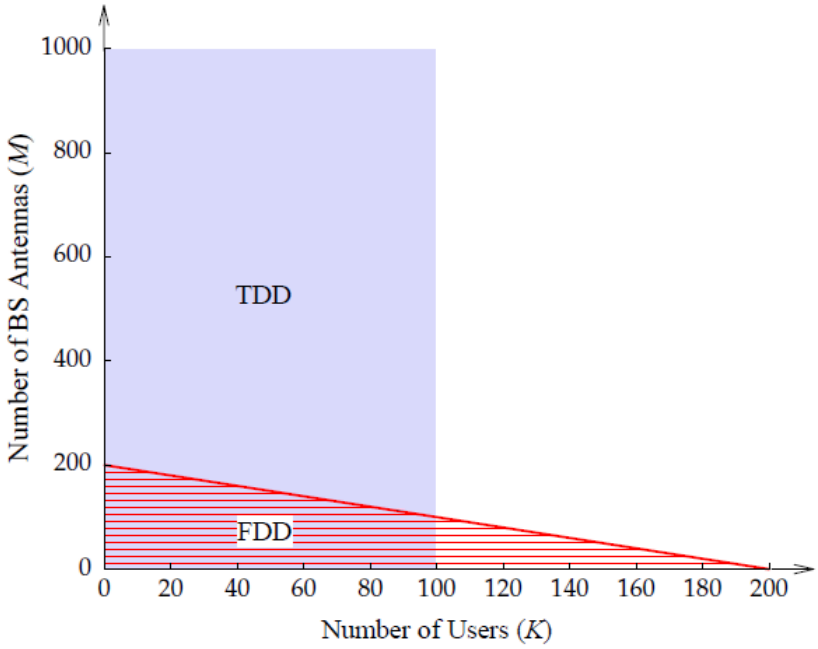

Figure 2.4: The possible regions for TDD and FDD modes of massive MIMO for the coherence time $T = 200$. Taken from [1].

2.6.1 Is TDD or FDD Best Suited to Massive MIMO Systems?

Channel estimation load in TDD mode does not depend on M, while channel estimation load in FDD systems is a function of M. As a large number of antennas are used in massive systems, TDD mode is more suitable for massive MIMO systems. Let us consider an example from [1]: let us assume $T = 200$ symbols denote the coherence interval, which corresponds to a coherence bandwidth of 200 KHz and coherence time of 1ms. Then (2.20) results in $M + K < 200$. However, exploiting (2.19), we have $K < 100$. To compare TDD and FDD mode, we use Fig. 2.4 which is extracted from [1]. The figure shows the regions for possible values of M and K for $T = 200$. Fig. 2.4 investigates the operating feasible regions for different values of (M, K) in the FDD and TDD modes of massive MIMO systems. As the figure shows, the FDD region is very small compared to the TDD region. By having a large number of antennas in massive MIMO, we have to use the whole channel coherence time for channel estimation, which does not leave any time to transmit data.

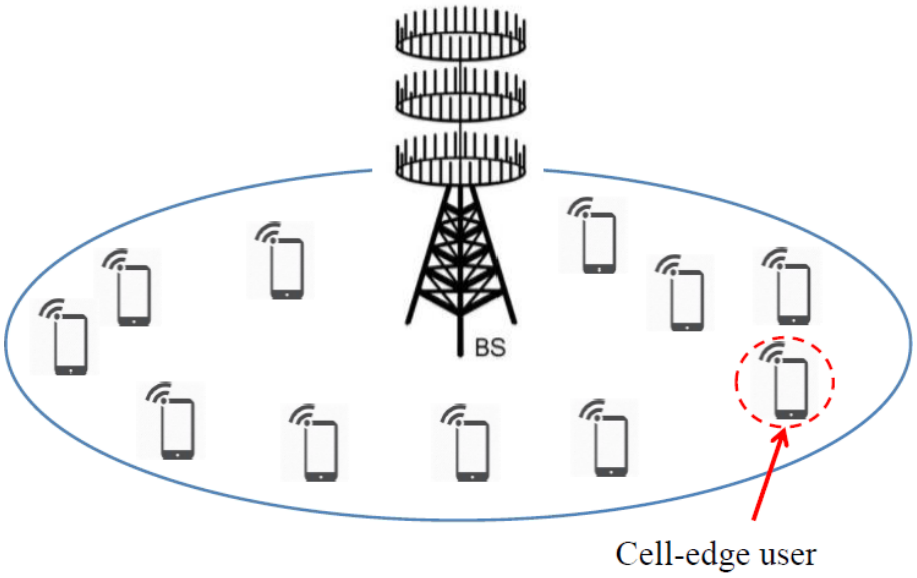

Figure 2.5: A single-cell massive MIMO system with K single-antenna users and a BS equipped with M antennas.

2.6.2 Favourable Propagation

Massive MIMO systems benefit from *favourable propagation* which introduces the desirable condition to increase sum-capacity [25]. The mutual orthogonality among the channel vectors of the users is called favourable propagation [25]. Let us assume the vectors $\mathbf{h}_k$, $k = 1, \cdots, K$ denote the size $M \times 1$ channels. Based on [25], the channels introduce favourable propagation if the following condition hold

$$\mathbf{h}_i^H \mathbf{h}_j = \begin{cases} 0, & i \neq j \\ \|\mathbf{h}_i\|^2 \neq 0, & i = j, \end{cases} \tag{2.21}$$

where in (2.21), we use the law of large numbers. Note that based on the law of large numbers, the average of the results obtained from a large number of trials should be close to the expected value.

2.7 Different Scenarios to Implement Massive MIMO

As mentioned earlier, a massive MIMO system is determined by the 5-tuple $(M, K, R, \mathbf{H}, \mathbf{W})$, where M and K refer to the total number of service antennas and

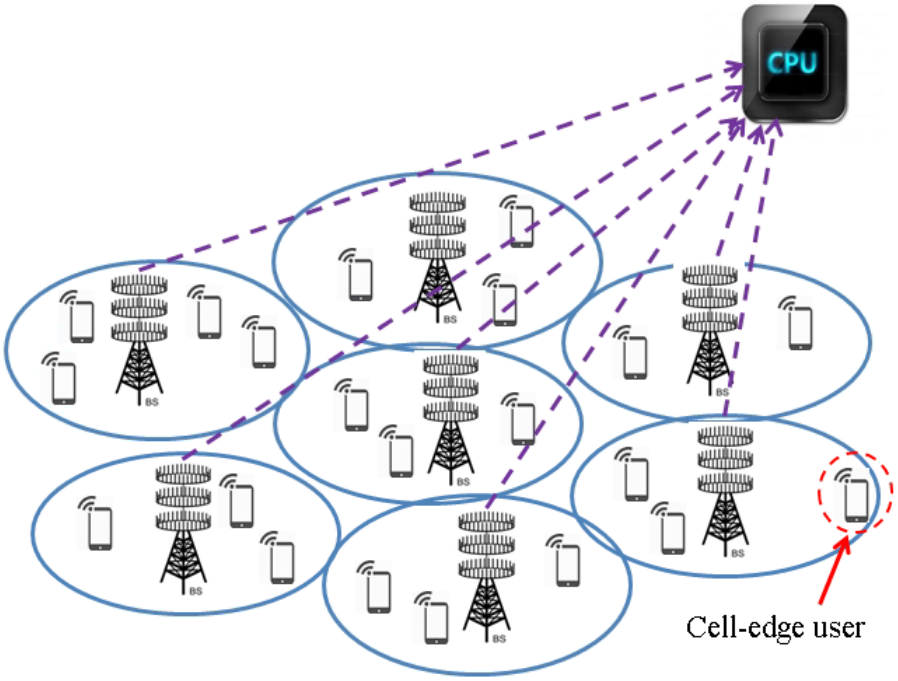

Figure 2.6: A multi-cell massive MIMO system with K single-antenna users and L cells, with one BS at the centre of each cell. The dashed lines denote the fronthaul links from the APs to the CPU.

total number of users in the area with $M \gg K$, and $\mathbf{H}$ and $\mathbf{W}$ are the channel and pre-coding/de-coding matrix in the downlink/uplink modes. Note that R km^2 refers to the size of the area. In this section we present a performance comparison between three scenarios as follows: i) Single cell massive MIMO: a single BS equipped with M antennas communicates with K users distributed across a single cell with size R km^2, which constitutes the entire service area; ii) multi cell massive MIMO with joint processing: the service area of size R km^2 is divided into N_c cells with one BS in each cell, with total number of service antennas at all BSs is M. The BSs jointly serve distributed users in the cells, which is combination of massive MIMO and network MIMO or CoMP [26–28], and iii) cell-free massive MIMO: a large number M of APs, with a small number N of antennas at each, are randomly distributed through the coverage area with size R km^2 [4]. Hence the total number of service antennas in cell-free massive MIMO is NM. Note that a genuine single-cell massive MIMO system is not realistic in practice- an entire network of users in a country (or even in a city) is never going to be served by just one cell. In the following, we explain the above-mentioned system models.

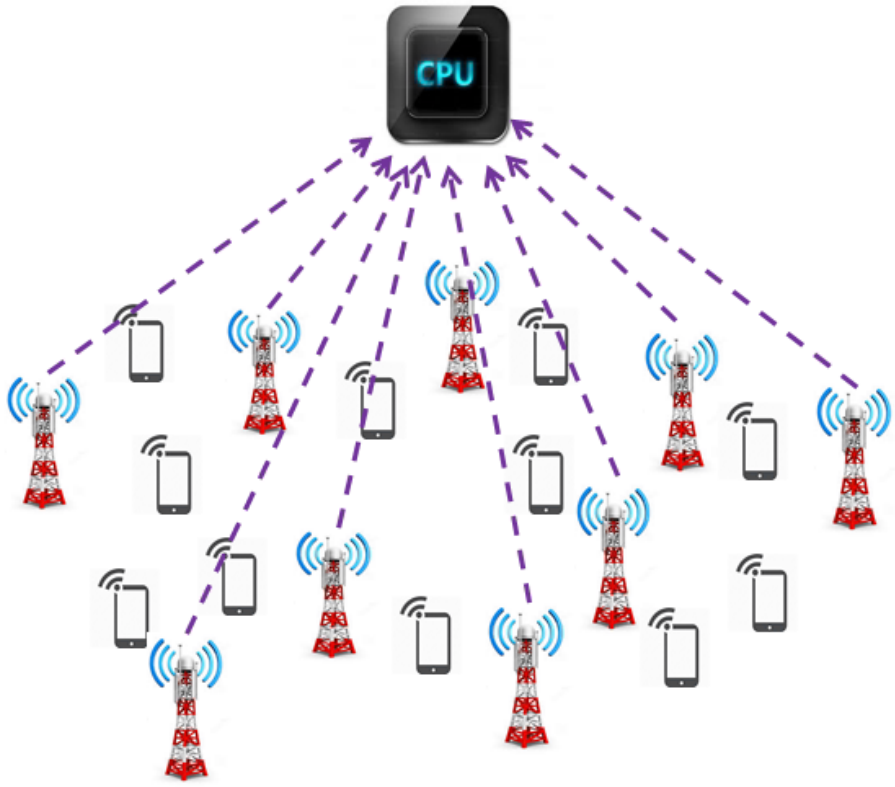

Figure 2.7: A cell-free massive MIMO system with K single-antenna users and M APs. The dashed lines denote the fronthaul links from the APs to the CPU.

2.7.1 Single-Cell Massive MIMO

A single BS equipped with large number of antennas communicates with a much smaller number of users distributed in the cell. Fig. 2.5 represents a single-cell massive MIMO with a BS with M antennas and K single antenna users. As the figure shows, the "cell edge" users are far away from the BS and suffer from very high pathloss, which results in poor service for the cell edge users.

2.7.2 Multi-Cell Massive MIMO

In multi-cell massive MIMO, the area is divided into several cells and there is one BS at the centre of each cell. The BSs could be connected to the CPU via fronthaul links. However, the low throughput of cell edge users remains a limitation factor in realistic multi-cell massive MIMO systems [5]. The system model of multi-cell massive MIMO is demonstrated in Fig. 2.6.

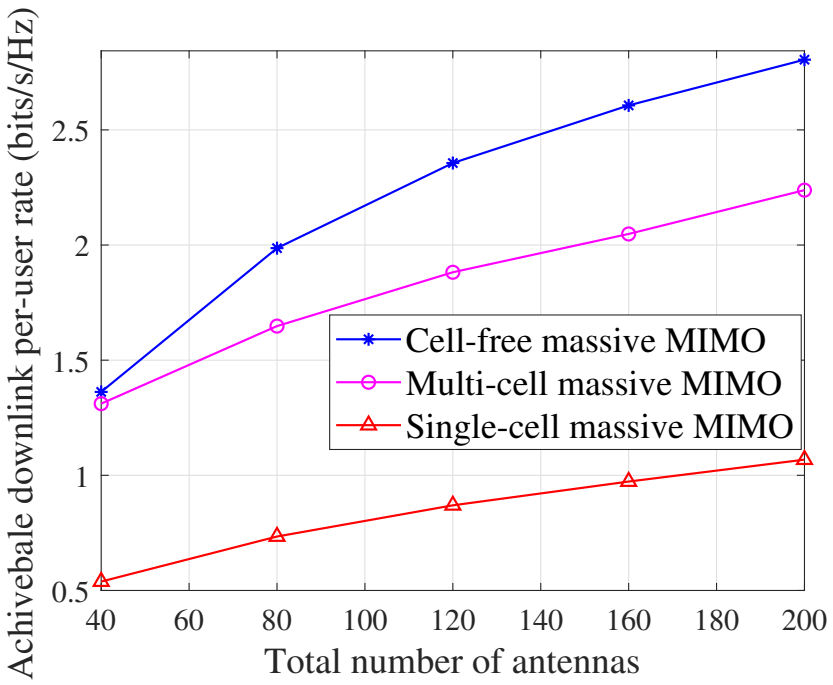

Figure 2.8: The achievable rate versus total number of antennas with conjugate beamforming, orthogonal pilot sequences, pilot power $p_p = 100$ mW, downlink data power $p_d = 200$ mW, and size of area 2×2 km^2. For the case of cell-free massive MIMO, there are M single-antenna APs which are uniformly distributed through the area. In case of single-cell massive MIMO, there is one BS with M antennas at the centre of cell. In multi-cell massive MIMO, we divide the area into 4 cells and assume one BS at the centre of each cell. Moreover, it is assumed that each BS has $\frac{M}{4}$ antennas. Note that this result is a contribution of this thesis and is included here to the sake of completeness.

2.7.3 Cell-Free Massive MIMO

In this subsection, we explain the concept of cell-free massive MIMO which has gained a lot of attention recently because of its potential to ensure uniformly good service rates for all users [2]. Cell-free massive MIMO system is a combination of distributed MIMO and massive MIMO systems [2]. The distributed APs are connected to a CPU via fronthaul links [2]. The system model of cell-free massive MIMO is presented in Fig. 2.7.

In cell-free massive MIMO, many distributed APs are connected to one or several CPUs and coherently serve many distributed users [2]. In other words, here the service antennas can be assumed as the elements of a very large distributed antenna array in a multiuser MIMO system. This approach provides many of the benefits of cloud radio access network (C-RAN), enjoying lower path loss as well as distributed signal processing. Note that in C-RAN the radio access network (RAN) over an area of

possible tens of km^2 is treated as a distributed antenna system (DAS) in which AP antennas are connected to a large central baseband processing unit (BBU). Cell-free massive MIMO is an effectively scalable version of network MIMO, or CoMP [9, 10], and can effectively eliminate the concept of cells. One of the main issues of cell-free massive MIMO systems which requires more investigation is the limited-capacity backhaul links from the APs to the CPU. Due to the combination of massive MIMO and distributed MIMO, cell-free massive MIMO benefits from high system throughput and energy efficiency [2, 29]. Fig. 2.8 presents a performance comparision between cell-free massive MIMO and single-cell massive MIMO versus total number of antennas M. Note that we deploy M single-antenna APs in cell-free massive MIMO. As the figure shows cell-free massive MIMO provides a better performance thanks to the fact that APs have a smaller geographical distance to the users.

2.7.4 Cell-Free Massive MIMO System Model

We consider uplink transmission in a cell-free massive MIMO system with M APs and K single-antenna users randomly distributed in a large area. Moreover, we assume each AP has N antennas, as shown in Fig. 2.9. The channel coefficient vector between the kth user and the mth AP, $\mathbf{g}_{mk} \in \mathbb{C}^{N\times 1}$, is modeled as

$$\mathbf{g}_{mk} = \sqrt{\beta_{mk}}\mathbf{h}_{mk}, \tag{2.22}$$

where β_{mk} denotes the large-scale fading, the elements of $\mathbf{h}_{mk}$ are i.i.d. $\mathcal{CN}(0, 1)$ RVs, and represent the small-scale fading [2]. Note that throughout this thesis we refer to β_{mk} by the statistics of the channel. The large-scale fading β_{mk} changes very slowly with time. Compared to the small-scale fading, the large-scale fading changes much more slowly, some 40 times slower according to [12, 30]. Therefore, β_{mk} can be estimated in advance. One simple way is that the AP takes the average of the power level of the received signal over a long time period. A similar technique for collocated massive MIMO is discussed in Section III-D of [12].

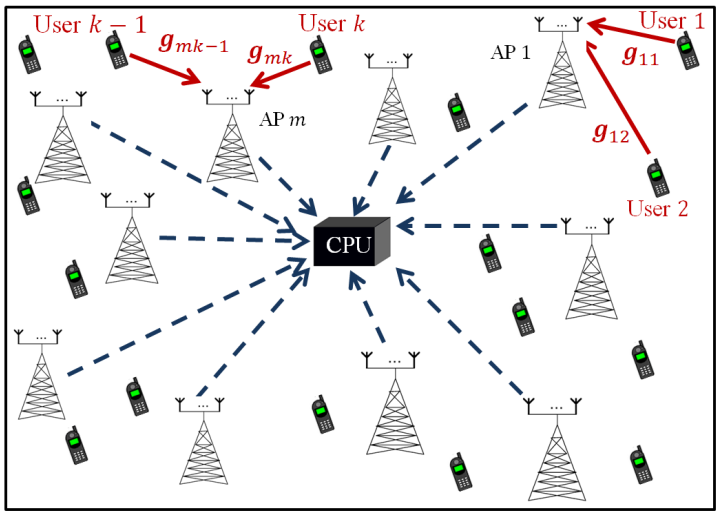

Figure 2.9: The uplink of a cell-free massive MIMO system with K single-antenna users and M APs. Each AP is equipped with N antennas. The solid lines denote the uplink channels and the dashed lines present the fronthaul links from the APs to the CPU. Similar to [2], we assume that the simulation area is wrapped around at the edges which can simulate an area without boundaries. Hence, the square simulation area has eight neighbours.

2.7.5 Uplink Channel Estimation in Cell-Free Massive MIMO

In order to estimate channel coefficients in the uplink, the APs employ a minimum mean-square error (MMSE) estimator. During the training phase, all K users simultaneously transmit their pilot sequences of length τ symbols to the APs. Let $\sqrt{\tau}\boldsymbol{\phi}_k \in C^{\tau \times 1}$, where $\|\boldsymbol{\phi}_k\|^2 = 1$, be the pilot sequence assigned to the kth user. Then, the received signal at the mth AP is given by

$$\mathbf{Y}_m^p = \sqrt{\tau p_p} \sum_{k=1}^{K} \mathbf{g}_{mk} \boldsymbol{\phi}_k^H + \mathbf{W}_m^p, \tag{2.23}$$

where vector $\mathbf{W}_m^p \in \mathbb{C}^{N \times \tau}$ is the noise whose elements are i.i.d. $\mathcal{CN}(0,1)$. Moreover, p_p is the normalized transmit SNR of each pilot symbol, where $p_p = \frac{\bar{p}_p}{p_n}$, where $\bar{p}_p$ and p_n refer to the transmit pilot power and the noise power, respectively. Next, the APs exploit the pilot sequence $\boldsymbol{\phi}_k$ to correlate the received signal with the pilot sequence as follows [2]:

$$\check{\mathbf{y}}_{m,k}^p = \mathbf{Y}_m^p \boldsymbol{\phi}_k = \sqrt{\tau p_p} \mathbf{g}_{mk} + \sqrt{\tau p_p} \sum_{k' \neq k}^{K} \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \dot{\mathbf{w}}_{mk}^p, \tag{2.24}$$

where $\dot{\mathbf{w}}_{mk}^{p} \triangleq \mathbf{W}_{m}^{p}\boldsymbol{\phi}_{k}^{H}$. The linear MMSE estimate of g_{mk} is

$$\hat{\mathbf{g}}_{mk} = \frac{\mathbb{E}\left\{\mathbf{g}_{mk}\check{\mathbf{y}}_{m,k}^{p}\right\}}{\mathbb{E}\left\{\left|\check{\mathbf{y}}_{m,k}^{p}\right|^{2}\right\}}\check{\mathbf{y}}_{m,k}^{p} = c_{mk}\left(\sqrt{\tau p_p}g_{mk} + \sqrt{\tau p_p}\sum_{k'\neq k}^{K}\mathbf{g}_{mk'}\boldsymbol{\phi}_{k'}^{H}\boldsymbol{\phi}_{k} + \dot{\mathbf{w}}_{mk}^{p}\right), \quad (2.25)$$

where that the expectations are taken over small-scale fading and noise. Next, c_{mk} is obtained as [2]

$$c_{mk} = \frac{\sqrt{\tau p_p}\beta_{mk}}{\tau p_p \sum_{k'=1}^{K}\beta_{mk'}\left|\boldsymbol{\phi}_{k'}^{H}\boldsymbol{\phi}_{k}\right|^{2} + 1}. \quad (2.26)$$

Note that, as in [2], we assume that the large-scale fading, β_{mk}, is known. The estimated channels in (2.25) are used by the APs to design the receiver filter coefficients and determine power allocations at users to maximize the minimum rate of the users. Note that the channel $\mathbf{g}_{mk}$ has N i.i.d. Gaussian elements. Therefore, the power on the nth component of $\mathbf{g}_{mk}$ is defined as follows:

$$\gamma_{mk} = E\left\{\left|[\hat{\mathbf{g}}_{mk}]_{n}\right|^{2}\right\} = \frac{\tau p_p \beta_{mk}^{2}}{\tau p_p \sum_{k'=1}^{K}\beta_{mk'}\left|\boldsymbol{\phi}_{k'}^{H}\boldsymbol{\phi}_{k}\right|^{2} + 1} = \sqrt{\tau p_p}\beta_{mk}c_{mk}. \quad (2.27)$$

In this thesis, we investigate the cases of both random pilot assignment and orthogonal pilots in cell-free massive MIMO. Here the term "orthogonal pilots" refers to the case where unique orthogonal pilots are assigned to all users. In the case of orthogonal pilots, the length of pilots is $\tau \geq K$. Hence, the pilots assigned to the users are orthogonal. Moreover, for case of "random pilot assignment", we have $\tau < K$. Hence, similar to the scheme in [2,31], each user is randomly assigned a pilot sequence from a set of orthogonal sequences of length τ ($< K$).

2.8 The Basics of COST 2100 Channel Model

The propagation channel model is the basis of wireless communications. An accurate and realistic channel model is a requirement for 5G systems. This section investigates a geometry-based channel model in detail.

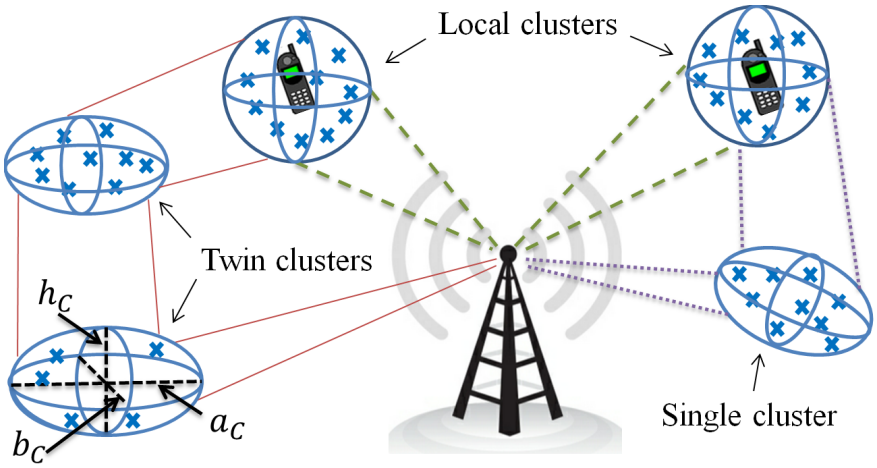

Figure 2.10: The general description of the cluster model. The spatial spreads for cth cluster are given.

COST 2100 Channel Model

In geometry-based stochastic channel models (GSCMs), the double directional channel impulse response is a superposition of MPCs. The channel is given by [32]

$$h(\tau, \phi, \theta) = \sum_{j=1}^{N_C} \sum_{i=1}^{N_p} a_{i,j} \delta(\phi - \phi_{i,j}) \delta(\theta - \theta_{i,j}) \delta(\tau - \tau_{i,j}), \tag{2.28}$$

where N_p denotes the number of multipath components, τ denotes the delay, δ denotes the Dirac delta function, and ϕ and θ represent the direction of arrival (DoA) and direction of departure (DoD) respectively. Similar to [32], we group the multipath components with similar delay and directions (depending on cluster delay spread) into clusters. Three kinds of clusters are defined; local clusters, single clusters and twin clusters. Local clusters are located around users and the BS while single clusters represented by one cluster and twin clusters are characterized by two clusters related respectively to the user and the BS side as shown in Fig. 2.10. A local cluster is a single cluster that surrounds a user: single clusters can also occur in a different position. Twin clusters consist of a linked pair of clusters, one of which defines the angles of departure of multipaths from the transmitter, while the other defines the angles of arrival at the receiver [32]. There is a large number of clusters in the area, however only some of them can contribute to the channel. The circular visibility region (VR) determines whether the cluster is active or not for a given user. The MPC's gain scales by a transition function that is given by

$$A_{VR}(\bar{r}_{MS}) = \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{2\sqrt{2}\left(L_C + d_{MS,VR} - R_C\right)}{\sqrt{\lambda L_C}}\right), \tag{2.29}$$

where $\bar{r}_{MS}$ is the centre of the VR, R_C denotes the VR radius, L_C represents the size of the transition region and $d_{MS,VR}$ refers to the distance between the mobile stations (MS)s and the VR centre. Moreover, λ denotes the wavelength (in m). For a constant expected number of clusters N_C, the area density of VRs is given by

$$\rho_C = \frac{N_C - 1}{\pi \left(R_C - L_C\right)^2}. \tag{2.30}$$

All clusters are ellipsoids in the environment and can be characterized by the cluster spatial delay spread, elevation spread and azimuth spread. Once the position of the BS and users are fixed, we need to determine the positions of the clusters in the area by geometrical calculations. For the local clusters, we consider a circle around the users and the BS, so that the size of the local cluster can be characterized by the cluster delay spread (a_C), elevation spread (h_C) and the position of MPCs [32]. For local clusters the cluster delay, azimuth and elevation spreads can be given by

$$a_C = \frac{\Delta\tau c_0}{2}, \tag{2.31a}$$

$$b_C = a_C, \tag{2.31b}$$

$$h_C = d_{C,BS} \tan\theta_{BS}, \tag{2.31c}$$

where c_0 denotes the speed of light, $d_{C,BS}$ is the distance between the cluster and the BS, $\Delta\tau$ refers to the delay spread and θ_{BS} is the elevation spread seen by the BS. The delay spread, angular spreads and shadow fading are correlated RVs and for all kinds of clusters are given by [33]

$$\Delta\tau_c = \mu_\tau (\tfrac{d}{1000})^{\frac{1}{2}} 10^{\sigma_\tau \frac{Z_c}{10}}, \tag{2.32a}$$

$$\beta_c = \tau_\beta 10^{\sigma_\beta \frac{Y_c}{10}}, \tag{2.32b}$$

$$S_m = 10^{\sigma_s \frac{X_c}{10}}, \tag{2.32c}$$

where $\Delta\tau_c$ refers to the delay spread, β_c denotes angular spread, and S_m is the shadow fading of cluster c. Moreover, X_c, Y_c and Z_c denote correlated RVs with zero mean and unit variance. Correlated random process can be simulated using Cholesky factorization

[33]. Cholesky factorization can be used to generate a random vector with a desired covariance matrix [34]. The MPCs' positions can be drawn from the truncated Gaussian distribution given by [32]

$$f(r)=\begin{cases} \frac{1}{\sqrt{2\pi\sigma_{r,o}^2}}\exp\left(-\left[\frac{r-\mu_{r,o}}{\sqrt{2}\sigma_{r,o}}\right]^2\right) & |r|\leq r_T, \\ 0 & \text{otherwise}, \end{cases} \tag{2.33}$$

where r_T denotes the truncation value. For single clusters, the cluster delay, azimuth and elevation spreads can be given by

$$a_C = \Delta\tau c_0/2, \tag{2.34a}$$

$$b_C = d_{C,BS}\tan\phi_{BS}, \tag{2.34b}$$

$$h_C = d_{C,BS}\tan\theta_{BS}. \tag{2.34c}$$

To get the fixed positions of the single clusters, the radial distance of the cluster from the BS is drawn from the exponential distribution [32]

$$f(r) = \begin{cases} 0 & r < r_{min}, \\ \frac{1}{\sigma_r}\exp\left(-(r-\frac{r_{min}}{\sigma_r})\right) & \text{otherwise}. \end{cases} \tag{2.35}$$

To determine the fixed position of the cluster, the angle of the cluster can be drawn from the Gaussian distribution with a standard deviation $\sigma_{\phi,C}$. For the twin clusters, for both the BS and user side clusters we have

$$a_C = \frac{\Delta\tau c_0}{2}, \tag{2.36a}$$

$$b_C = d_{C,BS}\tan\phi_{BS}. \tag{2.36b}$$

For the BS side cluster, the elevation spread can be given by

$$h_C = d_{C,BS}\tan\theta_{BS}, \tag{2.37}$$

while for the MS side cluster, we have

$$h_C = d_{C,MS}\tan\theta_{MS}. \tag{2.38}$$

Fig. 2.10 gives an example of the geometry of the Cth cluster. For twin clusters, the distance between the cluster and the BS and the distance from the VR centre and the MS is given by [32]

$$d_{C,BS} \tan \Phi_{C,BS} = d_{C,MS} \tan \Phi_{C,MS}. \tag{2.39}$$

The delay of a cluster is represented by [32]

$$\tau_C = (d_{C,BS} + d_{C,MS} + d_C)/c_0 + \tau_{C,link}, \tag{2.40}$$

where the geometrical distance between twin clusters is represented by d_C, $d_{C,MS}$ denotes the geometrical distance between the user and the centre of the visibility region, $d_{C,BS}$ refers to the distance between the BS and the cluster, and finally $\tau_{C,link}$ is the cluster link delay between the twin clusters. Hence, the cluster power attenuation is given by [32]

$$A_C = \max\left(\exp\left[-k_\tau(\tau_C - \tau_0)\right], \exp\left[-k_\tau(\tau_B - \tau_0)\right]\right), \tag{2.41}$$

where k_τ denotes the decay parameter, and τ_B is the cut-off delay. We assume Rayleigh fading for the MPCs within each cluster. Hence, the complex amplitude of the ith MPC in the jth cluster in (2.28) is given by

$$a_{i,j} = \sqrt{L_p} A_{VR} \sqrt{A_C A_{MPC}}\, e^{-j2\pi f_c \tau_{i,j}}, \tag{2.42}$$

where L_p is the channel path loss, A_{MPC} is the power of each MPC which is characterized by the Rayleigh fading distribution and $\tau_{i,j}$ is the delay of the ith MPC in cluster j given by [32]

$$\tau_{i,j} = \frac{\left(d_{MPC_{i,j},BS} + d_{MPC_{i,j},MS}\right)}{c_0} + \tau_{i,C,link}. \tag{2.43}$$

By assuming a fixed OFDM subcarrier, we can drop the variable $\tau_{i,j}$ from (5.6). For the non-line-of-sight (NLoS) case of the micro-cell scenario, the path loss expression can be given by [35]

$$L_p = 26 \log_{10} d + 20 \log_{10}(\frac{4\pi}{\lambda}), \tag{2.44}$$

where d and again λ denote the distance (in m) and the wavelength (in m), respectively. Note that the basics of COST 2100 channel model presented in this section are exploited in Chapters 5 and 6.

2.9 The Basics of Optimal Uniform Quantization

Based on Bussgang's theorem [36], the output of a non-linear function such as a quantizer with Gaussian input can be represented as a linear function as follows:

$$\mathcal{Q}(z) = h(z) = az + n_d, \ \forall k, \tag{2.45}$$

where a is a constant value and n_d refers to the distortion noise which is uncorrelated with the input of the quantizer, z. The term a is given by

$$a = \frac{\mathbb{E}\left\{zh(z)\right\}}{\mathbb{E}\{z^2\}} = \frac{1}{p_z}\int_{\mathcal{Z}} \tilde{z}h(\tilde{z})f_z(\tilde{z})d\tilde{z}, \tag{2.46}$$

where $p_z = \mathbb{E}\{|z|^2\} = \mathbb{E}\{z^2\}$ is the power of z and we drop absolute value as z is a real number, and $f_z(z)$ is the probability distribution function of z. Define

$$b = \frac{\mathbb{E}\left\{h^2(z)\right\}}{\mathbb{E}\{z^2\}} = \frac{1}{p_z}\int_{\mathcal{Z}} h^2(\tilde{z})f_z(\tilde{z})d\tilde{z}. \tag{2.47}$$

Then, the signal-to-distortion noise ratio (SDNR) is

$$\text{SDNR} = \frac{\mathbb{E}\left\{(az)^2\right\}}{\mathbb{E}\{n_d^2\}} = \frac{p_z a^2}{p_z\left(b - a^2\right)} = \frac{a^2}{b - a^2}, \tag{2.48}$$

According to [36], the midrise uniform quantizer function $h(z)$ is given by

$$h(z) = \begin{cases} -\frac{L-1}{2}\Delta & z \leq -\left(\frac{L}{2}+1\right)\Delta, \\ \left(l+\frac{1}{2}\right)\Delta & l\Delta \leq z \leq (l+1)\Delta, l = -\frac{L}{2}+1, \cdots, \frac{L}{2}-2, \\ \frac{L-1}{2}\Delta & z \geq \left(\frac{L}{2}-1\right)\Delta, \end{cases} \tag{2.49}$$

where Δ is the step size of the quantizer and $L = 2^{\alpha}$, where α is number of quantization bits. Note that the lemma and the analysis provided below are novelty of this thesis, and we include them in this chapter for the sake of completeness.

Lemma 1. *The terms a and b are obtained as follows:*

$$a = \Delta\sqrt{\frac{2}{\pi p_z}}\left(\sum_{l=1}^{\frac{L}{2}-1} e^{-\frac{l^2\Delta^2}{2p_z}} + 1\right), \tag{2.50a}$$

$$b = \frac{\Delta^2}{p_z}\left(\tfrac{1}{4} + 4\sum_{l=1}^{\frac{L}{2}-1} l\, Q\left(\tfrac{l\Delta}{\sqrt{p_z}}\right)\right), \tag{2.50b}$$

where Q is the Q-function and is given by $Q = \frac{1}{2}erfc\left(\frac{z}{\sqrt{2}}\right)$, where erfc refers to the complementary error function [37].

Proof: We exploit (2.46) and (2.49) to find a and b for uniform quantizer as follows:

$$\begin{aligned}
a &= \frac{1}{p_z}\int_{-\infty}^{\infty} z h(z) f_z(z) dz \\
&= \frac{1}{p_z}\left(\int_{-\infty}^{-\frac{L}{2}+1} -z\frac{L-1}{2}\Delta f_z(z)dz + \sum_{l=-\frac{L}{2}+1}^{l=-\frac{L}{2}+1}\int_{l\Delta}^{(l+1)\Delta} x\left(l+\frac{1}{2}\right)\Delta f_z(z)dz\right. \\
&\left. + \int_{\frac{L}{2}-1}^{\infty} z\frac{L-1}{2}\Delta f_z(z)dz\right) \overset{a_1}{=} \frac{(L-1)\Delta}{2\sqrt{2\pi p_z}}\exp\left(-\frac{\left(\frac{L}{2}-1\right)^2}{2}\right) \\
&+ \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{\left(l+\frac{1}{2}\right)\Delta}{\sqrt{2\pi p_z}} \\
&\quad \left(\exp\left(-\frac{l^2\Delta^2}{2}\right) - \exp\left(\frac{(l+1)^2\Delta^2}{2}\right)\right) \\
&+ \frac{(L-1)\Delta}{2\sqrt{2\pi p_z}}\exp\left(-\frac{\left(\frac{L}{2}-1\right)^2}{2}\right) \\
&\overset{a_2}{=} \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{\left(l+\frac{1}{2}\right)\Delta}{\sqrt{2\pi p_z}}\exp\left(-\frac{l^2\Delta^2}{2}\right) - \sum_{l'=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{\left(l'+\frac{1}{2}\right)\Delta}{\sqrt{2\pi p_z}}\exp\left(-\frac{l'^2\Delta^2}{2}\right) \\
&= \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{\Delta}{\sqrt{2\pi p_z}}\exp\left(\frac{-l^2\Delta^2}{2}\right) \\
&= \sum_{l=1}^{\frac{L}{2}-1} \frac{2\Delta}{\sqrt{2\pi}}\exp\left(\frac{-l^2\Delta^2}{2}\right) + \frac{\Delta}{\sqrt{2\pi p_z}} \\
&= \Delta\sqrt{\frac{2}{\pi p_z}}\left(\sum_{l=1}^{\frac{L}{2}-1}\exp\left(\frac{-l^2\Delta^2}{2}\right) + 1\right),
\end{aligned} \tag{2.51}$$

$$
\begin{aligned}
b &= \frac{\mathbb{E}\left\{h^2\left(z\right)\right\}}{\mathbb{E}\left\{z^2\right\}} = \frac{1}{p_z}\int_{-\infty}^{\infty} h^2(z) f_z(z) dz \\
&= \frac{2}{p_z}\left(\sum_{l=1}^{\frac{L}{2}-1}\int_{(l-1)\Delta}^{l\Delta}\left(l-\frac{1}{2}\right)^2\Delta^2 f_z(z)dz\right. \\
&+ \left.\int_{\left(\frac{L}{2}-1\right)\Delta}^{l\Delta}\left(\frac{L-1}{2}\right)^2\Delta^2 f_z(z)dz\right) \\
&\overset{a_1}{=} \frac{2\Delta^2}{p_z}\left(\sum_{l=1}^{\frac{L}{2}-1}\left(l-\frac{1}{2}\right)^2\left(Q\left(\frac{(l-1)\,\Delta}{\sqrt{p_z}}\right)-Q\left(\frac{l\Delta}{\sqrt{p_z}}\right)\right)\right. \\
&+ \left.\left(\frac{L-1}{2}\right)^2 Q\left(\frac{(L-2)\,\Delta}{2p_z}\right)\right) \\
&\overset{a_2}{=} \frac{2\Delta^2}{p_z}\left(\sum_{l'=0}^{\frac{L}{2}-1}\left(rl'+\frac{1}{2}\right)^2 Q\left(\frac{l'\Delta}{p_z}\right)-\sum_{l=1}^{\frac{L}{2}-1}\left(l-\frac{1}{2}\right)^2 Q\left(\frac{l\Delta}{p_z}\right)\right) \\
&= \frac{2\Delta^2}{p_z}\left(\left(\frac{1}{2}\right)^2 Q(0)+\sum_{l=1}^{\frac{L}{2}-1}\left(\left(l+\frac{1}{2}\right)^2-\left(l-\frac{1}{2}\right)^2\right)Q\left(\frac{l\Delta}{p_z}\right)\right) \\
&= \frac{\Delta^2}{p_z}\left(\frac{1}{4}+4\sum_{l=1}^{\frac{L}{2}-1} lQ\left(\frac{l\Delta}{p_z}\right)\right)
\end{aligned}
\tag{2.52}
$$

where the steps a_1 and a_2 come from the property that the input of the quantizer has the Gaussian distribution, and $l' = l + 1$, respectively.

In general, terms a and b are functions of the power of the quantizer input, p_z. To remove this dependency, we normalize the input signal by dividing the input signal, z, by the square root of its power, $\sqrt{p_z}$, and then multiply the quantizer output by its square root, $\sqrt{p_z}$. Hence, by introducing a new variable $\tilde{z} = \frac{z}{\sqrt{p_z}}$, we have

$$
\begin{aligned}
\mathcal{Q}(z) &= \sqrt{p_z}\mathcal{Q}(\tilde{z}) \\
&= \tilde{a}\sqrt{p_z}\tilde{z} + \sqrt{p_z}\tilde{n}_d \\
&= \tilde{a}z + \sqrt{p_z}\tilde{n}_d.
\end{aligned}
\tag{2.53}
$$

Note that (2.53) enables us to find the optimum step size of the quantizer and the corresponding $\tilde{a}$. The optimal step size of the quantizer is obtained by solving the

following maximization problem:

$$\begin{aligned}
\Delta_{\text{opt}} &= \arg\max_{\Delta} \text{SDNR} \\
&= \arg\max_{\Delta} \frac{a^2}{b - a^2} \\
&\overset{I_1}{=} \arg\max_{\Delta} \frac{\tilde{a}^2}{\tilde{b} - \tilde{a}^2} \\
&= \arg\max_{\Delta} \frac{\tilde{a}^2}{\tilde{b}} \\
&\overset{I_2}{\underset{\Delta}{=}} \arg\max_{\Delta} \left(\frac{\frac{2\Delta^2}{\pi} \left(\sum_{l=1}^{\frac{L}{2}-1} \exp\left(\frac{-l^2\Delta^2}{2} \right) + 1 \right)^2}{\Delta^2 \left(\frac{1}{4} + 4 \sum_{l=1}^{\frac{L}{2}-1} lQ\left(\frac{l\Delta}{p_z} \right) \right)} \right) \\
&= \arg\max_{\Delta} \left(\frac{\left(\sum_{l=1}^{\frac{L}{2}-1} 2\exp\left(-\frac{l^2\Delta^2}{2} \right) + 1 \right)^2}{\frac{1}{4} + 4 \sum_{l=1}^{\frac{L}{2}-1} l\, Q\left(l\Delta \right)} \right),
\end{aligned} \tag{2.54}$$

where in step I_1, we have used (2.53) and step I_2 comes from results in Lemma 1. The maximization problem in (2.54) can be solved through a one-dimensional search over Δ for a given L in a symbolic mathematics tool such as Mathematica. For the input $\tilde{z}$ with $p_{\tilde{z}} = 1$, the optimal step size of the quantizer, the resulting distortion noise power, $p_{\tilde{n}_d} = \mathbb{E}\{|\tilde{n}_d|^2\} = \tilde{b} - \tilde{a}^2$, and the resulting $\tilde{a}$ are summarized in Table 2.1.

Remark 1. *Interestingly, the optimal values for quantization step size, Δ_{opt}, given in Table 2.1, are exactly the same as the optimal values of quantization step size without using Bussgang theorem in [38]. In [38], J. Max did not provide any analytical solution to obtain the optimal quantization step size. Moreover, J. Max only calculates the optimal step size and the resulting distortion power for $\alpha = 1, \cdots, 5$ whereas Lemma 1 enables us to calculate the optimal step size and the resulting distortion power for any quantization resolution values for α up to 18 listed in Table 2.1.*

Table 2.1: The optimal step size and distortion power of a uniform quantizer with Bussgang decomposition.

α	Δ_{opt}	$p_{\tilde{n}_d} = \tilde{b} - \tilde{a}^2$	$\tilde{a}$
1	1.596	0.2313	0.6366
2	0.9957	0.10472	0.88115
3	0.586	0.036037	0.96256
4	0.3352	0.011409	0.98845
5	0.1881	0.003482	0.996505
6	0.1041	0.0010389	0.99896
7	0.0568	0.0003042	0.99969
8	0.0307	0.0000876	0.999912
9	0.0165	0.0000249	0.999975
10	0.0088	6.99696×10^{-6}	0.999993
11	0.004649	1.94441×10^{-6}	0.999998
12	0.0024484	5.35536×10^{-7}	0.999999
13	0.001283	1.46369×10^{-7}	1
14	0.001283	3.97394×10^{-8}	1
15	0.000349	1.0727×10^{-8}	1
16	0.0001812	2.88095×10^{-9}	1
17	0.0000938	7.70251×10^{-10}	1
18	0.0000485	2.05146×10^{-10}	1

2.10 Achievable Rate with Unknown Channel Gain at the Receiver and Non-Gaussian Noise

It is a common assumption that in massive MIMO, the BS or users do not have knowledge of the channel gain [4, 24]. So, in this section, we investigate the achievable rate without assuming knowledge of channel at the receiver. Let us assume g is a scalar RV representing the point-to-point fading channel. Hence, the received signal is given by

$$y = \sqrt{p}gx + w, \tag{2.55}$$

where x is the transmitted symbol ($\mathbb{E}\{|x|^2\}] \leq 1$) and p is the power and w is noise. Similar to [24, Section 2.3.4], we assume g and w are independent, and there is no assumption considered on the statistical relationship between g and w. Based on [24, Section 2.3.4], to calculate the achievable rate, we need to re-write the received

signal as follows:

$$y = \sqrt{p}\mathbb{E}\{g\}\, x + \underbrace{\sqrt{p}\,(g - \mathbb{E}\{g\})\, x + w}_{\text{effective noise}}. \tag{2.56}$$

Note that the term $\sqrt{p}\,(g - \mathbb{E}\{g\})\, x$ is due to the lack of the information about g. Following the methodology in [24, Section 2.3.4], the second and third terms in (2.56) are mutually uncorrelated, and moreover, they are uncorrelated with x. The term w is the effective noise. Finally, the achievable rate is given by

$$R = \log_2\left(1 + \frac{p|\mathbb{E}\{g\}|^2}{p\text{Var}\{g\} + 1}\right). \tag{2.57}$$

Assuming C refers to the capacity of this channel, we have [24]

$$C \geq R. \tag{2.58}$$

By using (2.57) and (2.58) and the analysis in [24], one could conclude that the bound in (2.58) is helpful if g oscillates closely around its expected value $\mathbb{E}\{g\}$ resulting small variance Var$\{g\}$ [24]. Finally, exploiting the effect of channel hardening, although g is random, the value of $g - \mathbb{E}\{g\}$ is small, resulting in small Var$\{g\}$.

2.11 Summary

The chapter is summarised as follows:

- A general overview of the basics related to the current thesis has been provided. Some background knowledge which is essential for the upcoming chapters of the thesis has been presented.
- In this chapter we have given an overview of basics of MU-MIMO systems. The linear transmission schemes including ZFBF and MRT and the SINR of performance of the system with linear precoding have been presented.
- Different ways of massive MIMO implementations have been discussed: single-cell massive MIMO; multi-cell massive MIMO; cell-free massive MIMO.

- The principles of channel estimation in cell-free massive MIMO uplink have been investigated.
- The detailed parameter set of realistic geometry-based COST 2100 channel model has been presented.
- The general concept of the uniform quantizer has been introduced. Next a detailed analysis to find the optimal step size of the uniform quantizer has been provided.
- The details of calculating the achievable rate with unknown channel gain at the receiver and non-Gaussian noise have been presented.

Chapter 3

Max-Min Rate of Cell-Free Massive MIMO Uplink

Contents

3.1 Introduction

One of the main issues of cell-free massive MIMO systems which requires more investigation is the limited-capacity fronthaul links from the APs to a CPU. The

assumption of infinite fronthaul in [2, 16] is not realistic in practice. First, we consider the case where all APs send back the quantized version of the minimum mean-square error (MMSE) estimate of the channel from each user and the quantized version of the received signal to the CPU. We next study the case when each AP multiplies the received signal by the conjugate of the estimated channel from each user, and sends back a quantized version of this weighted signal to the CPU. We derive the total number of bits for both cases and show that given the same fronthaul capacity for both cases, the relative performance of the aforementioned cases depends on the number of antennas at each AP, the total number of APs and the channel coherence time. A new approach is provided to the analysis of the effect of fronthaul quantization on the uplink of cell-free massive MIMO. While there has been significant work in the context of network MIMO on compression techniques such as Wyner-Ziv coding for interconnection of distributed BSs, here for simplicity (and hence improved scalability) we assume simple uniform quantization. We exploit the Bussgang decomposition [36] to model the effect of quantization.

In [2, 6, 16] the authors propose that the APs design the linear receivers based on the estimated channels, and that this is carried out locally at the APs. Hence, the CPU exploits only the statistics of the channel for data detection. However, in this thesis, we propose to exploit a new receiver filter at the CPU to improve the performance of cell-free massive MIMO systems. The coefficients of the proposed receiver filter are designed based on only the statistics of the channel, which is different from the linear receiver at the APs. The proposed receiver filter significantly improves the performance of the uplink of cell-free massive MIMO.

We next investigate an uplink max-min rate problem with limited fronthaul links. In particular, the receiver filter coefficients and power allocation are optimized in the proposed scheme whereas the work in [2] only considered user power allocations. In particular, we propose a new approach to solve this max-min problem. A similar max–min rate problem based on signal-to-interference-plus-noise ratio (SINR) known as *SINR balancing* in the literature has been considered for cognitive radio networks in [39, 40]. In [41, 42], the authors consider MIMO systems and study the problem of max-min user rate to maximize the smallest user SINR. The problem of uplink-downlink duality has been investigated in [43, 44]. Note that none of the previous works on

uplink-downlink duality consider massive MIMO and the SINR formula in single-cell does not include any pilot contamination, channel estimation and quantization distortions.

To tackle the non-convexity of the original max-min rate problem, we propose to decouple the original problem into two sub-problems, namely, receiver filter coefficient design, and power allocation. We next show that the receiver filter coefficient design problem may be solved through a generalized eigenvalue problem [45]. Moreover, the user power allocation problem is solved through standard geometric programming (GP) [46, 47]. We present an iterative algorithm to alternately solve each sub-problem while one of the design parameters is fixed. Next an uplink-downlink duality for cell-free massive MIMO system with limited fronthaul links is established to validate the optimality of the proposed scheme. We show that there exists an equivalent problem related to virtual downlink SINR to realize the same user rate in the uplink with an equivalent total power constraint and the same receiver filter coefficients. By solving this equivalent virtual max-min SINR problem, the optimality of the proposed scheme in the uplink is validated. We finally propose an efficient user assignment algorithm and show that further improvement is achieved by the proposed user assignment algorithm. The contributions of the chapter are summarized as follows:

1. We consider two cases: i) the quantized versions of the channel estimates and the received signals at the APs are available at the CPU and ii) the quantized versions of processed signals at the APs are available at the CPU. The corresponding achievable rates are derived taking into account the effects of channel estimation error and quantization distortion.

2. We make use of the Bussgang decomposition to model the effect of quantization and present the analytical solution to find the optimal step size of the quantizer.

3. We propose a max-min fairness power control problem which maximizes the smallest of all user rates under the per-user power and fronthaul capacity constraints. To solve this problem, the original problem is decomposed into two sub-problems and an iterative algorithm is developed. The optimality of the proposed algorithm is proved through establishing the uplink-downlink duality for the cell-free massive MIMO system with limited fronthaul link capacities.

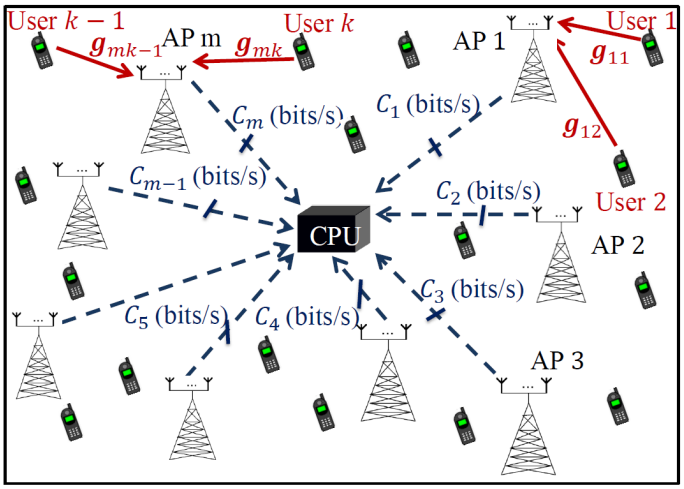

Figure 3.1: The uplink of a cell-free massive MIMO system with K single-antenna users and M APs. Each AP is equipped with N antennas. The solid lines denote the uplink channels and the dashed lines present the limited-capacity fronthaul links from the APs to the CPU.

4. A novel and efficient user assignment algorithm based on the capacity of fronthaul links is proposed which results in significant performance improvement.

The chapter is outlined as follows; Section II describes the system model and Section III provides performance analysis. The proposed max-min rate scheme is presented in Section IV and the convergence is provided in Section V. The optimality of the proposed scheme is proved in Section VI. Section VII investigates the proposed user assignment algorithm. Numerical results are presented in Section VIII, and finally Section IX concludes the chapter.

3.2 System Model

We consider uplink transmission in a cell-free massive MIMO system with M APs and K single-antenna users randomly distributed in a large area. Moreover, we assume each AP has N antennas. The channel coefficient vector between the kth user and the mth AP, $\mathbf{g}_{mk} \in \mathbb{C}^{N\times 1}$, is modeled as $\mathbf{g}_{mk} = \sqrt{\beta_{mk}}\mathbf{h}_{mk}$, where β_{mk} denotes the large-scale fading, the elemnts of $\mathbf{h}_{mk}$ are i.i.d. $\mathcal{CN}(0,1)$ RVs, and represent the small-scale fading [2]. Note that the uplink channel estimation is presented in Section 2.7.5.

Uplink Transmission

In this subsection, we consider the uplink data transmission, where all users send their signals to the APs. The transmitted signal from the kth user is represented by $x_k = \sqrt{\rho q_k} s_k$, where s_k ($\mathbb{E}\{|s_k|^2\} = 1$) and q_k denotes the transmitted symbol and the transmit power from the kth user, respectively. The $N \times 1$ received signal at the mth AP from all users is given by

$$\mathbf{y}_m = \sqrt{\rho} \sum_{k=1}^{K} \mathbf{g}_{mk} \sqrt{q_k} s_k + \mathbf{n}_m, \tag{3.1}$$

where ρ represents the normalized uplink signal-to-noise ratio (SNR) (see Section VIII for more details). Moreover, each element of $\mathbf{n}_m \in \mathbb{C}^{N \times 1}$, $n_{n,m} \sim \mathcal{CN}(0, 1)$ is the noise at the mth AP.

3.3 Performance Analysis

In this section, the performance analysis for two cases is presented. First we consider the case when the quantized versions of the channel estimates and the received signals are available at the CPU. Next, it is assumed that only the quantized versions of the weighted signals are available at the CPU.

Case 1. *Quantized Estimate of the Channel and Quantized Signal Available at the CPU*: The mth AP quantizes the terms $\hat{\mathbf{g}}_{mk}$, $\forall k$, and $\mathbf{y}_m$, and forwards the quantized CSI and the quantized signals in each symbol duration to the CPU. The quantized signal can be obtained as:

$$[\tilde{\mathbf{y}}_m]_n = \tilde{a}[\mathbf{y}_m]_n + [\mathbf{e}_m^y]_n = [\boldsymbol{\zeta}_m]_n + j[\boldsymbol{\nu}_m]_n, \ \forall m, n, \tag{3.2}$$

where $[\mathbf{e}_m^y]_n$ refers to the quantization distortion, and $[\boldsymbol{\zeta}_m]_n$ and $[\boldsymbol{\nu}_m]_n$ are the real and imaginary parts of $[\tilde{\mathbf{y}}_m]_n$, respectively. The analog-to-digital converter (ADC) quantizes the real and imaginary parts of $[\mathbf{y}_m]_n$ with α bits each, which introduces quantization distortions $[\mathbf{e}_m^y]_n$ to the received signals [48]. In addition, the ADC quantizes the MMSE

estimate of CSI as:

$$[\tilde{\mathbf{g}}_{mk}]_n = \tilde{a}[\hat{\mathbf{g}}_{mk}]_n + [\mathbf{e}^g_{mk}]_n = [\boldsymbol{\varrho}_{mk}]_n + j[\boldsymbol{\kappa}_{mk}]_n, \forall k, n, \tag{3.3}$$

where $[\boldsymbol{\varrho}_{mk}]_n$ and $[\boldsymbol{\kappa}_{mk}]_n$ denote the real and imaginary parts of $[\tilde{\mathbf{g}}_{mk}]_n$, respectively. For simplicity, we assume all APs use the same number of bits to quantize the received signal, $\mathbf{y}_m$, and the estimated channel, $\hat{\mathbf{g}}_{mk}$. Therefore, $\mathbb{E}\left\{|[\tilde{\mathbf{e}}^y_m]_n|^2\right\} = \mathbb{E}\left\{|[\tilde{\mathbf{e}}^g_{mk}]_n|^2\right\} = \sigma^2_{\tilde{e}}$, where $\mathbb{E}\left\{|[\tilde{\mathbf{e}}^y_m]_n|^2\right\}$ and $\mathbb{E}\left\{|[\tilde{\mathbf{e}}^g_{mk}]_n|^2\right\}$ are quantization distortions of a quantizer with normalized input $\frac{[\mathbf{y}_m]_n}{\sqrt{\mathbb{E}\{|[\mathbf{y}_m]_n|^2\}}}$ and $\frac{[\hat{\mathbf{g}}_{mk}]_n}{\sqrt{\mathbb{E}\{|[\hat{\mathbf{g}}_{mk}]_n|^2\}}}$, respectively, and $\sigma^2_{\tilde{e}} = p_{\tilde{n}_d}$. Note that due to power normalization, $p_{\tilde{n}_d}$, $\tilde{a}$, $\tilde{b}$, and optimal Δ for (3.2) and (3.3) are the same and provided in Table 2.1. The received signal for the kth user after using the MRC detector

at the CPU is given by

$$
\begin{aligned}
r_k &= \sum_{m=1}^{M} u_{mk} \tilde{\mathbf{g}}_{mk}^H \tilde{\mathbf{y}}_m = \sum_{m=1}^{M} u_{mk} \left(\tilde{a} \hat{\mathbf{g}}_{mk} + \mathbf{e}_{mk}^{\hat{g}} \right)^H \left(\tilde{a} \mathbf{y}_m + \mathbf{e}_m^y \right) \\
&= \sum_{m=1}^{M} u_{mk} \left(\tilde{a} \hat{\mathbf{g}}_{mk} + \mathbf{e}_{mk}^{\hat{g}} \right)^H \left(\tilde{a} \sqrt{\rho} \sum_{k=1}^{K} \mathbf{g}_{mk} \sqrt{q_k} s_k + \tilde{a} \mathbf{n}_m + \mathbf{e}_m^y \right) \\
&= \tilde{a}^2 \underbrace{\sqrt{\rho} \mathbb{E} \left\{ \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\}}_{\mathrm{DS}_k} s_k \\
&\quad + \tilde{a}^2 \underbrace{\sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m}_{\mathrm{TN}_k} \\
&\quad + \tilde{a}^2 \underbrace{\sqrt{\rho} \left(\sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} - \mathbb{E} \left\{ \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\} \right)}_{\mathrm{BU}_k} s_k \\
&\quad + \tilde{a}^2 \sum_{k' \neq k}^{K} \underbrace{\sqrt{\rho} \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'}}}_{\mathrm{IUI}_{kk'}} s_{k'} \\
&\quad + \sum_{k'=1}^{K} \tilde{a} \underbrace{\sqrt{\rho} \sum_{m=1}^{M} u_{mk} (\mathbf{e}_{mk}^{\hat{g}})^H \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'}}_{\mathrm{TQD}_{kk'}} \\
&\quad + \tilde{a} \underbrace{\sum_{m=1}^{M} u_{mk} (\mathbf{e}_{mk}^{g})^H \mathbf{n}_m}_{\mathrm{TQD}_k^g} \\
&\quad + \tilde{a} \underbrace{\sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{e}_m^y}_{\mathrm{TQD}_k^y} \\
&\quad + \underbrace{\sum_{m=1}^{M} u_{mk} \left(\mathbf{e}_{mk}^{\hat{g}} \right)^H \mathbf{e}_m^y}_{\mathrm{TQD}_k^{gy}},
\end{aligned}
\tag{3.4}
$$

where DS_k and BU_k denote the desired signal (DS) and beamforming uncertainty (BU) for the kth user, respectively, and IUI_k represents the inter-user-interference (IUI) caused by the k'th user. In addition, TN_k accounts for the total noise (TN) following the MRC detection, and finally the terms $\mathrm{TQD}_k^{\mathrm{y}}$, $\mathrm{TQD}_k^{\mathrm{g}}$, $\mathrm{TQD}_k^{\mathrm{gy}}$ and $\mathrm{TQD}_{kk'}$ refer to the total quantization distortion (TQD) at the kth user due to the quantization distortions at the

channel and signal. Moreover, by collecting all the coefficients $u_{mk},\ \forall m$, corresponding to the *k*th user, we define $\mathbf{u}_k = [u_{1k}, u_{2k}, \cdots, u_{Mk}]^T$ and without loss of generality, it is assumed that $||\mathbf{u}_k|| = 1$. The optimal values of u_{mk} are investigated in Section IV.

Proposition 1. *Terms DS_k, BU_k, $IUI_{kk'}$, $TQN_{kk'}$, TQN_k^g, TQN_k^y, TQN_k^{gy} are mutually uncorrelated.*

Proof: Please refer to Appendix 3.A. ■

To obtain an achievable rate, we use the analysis in [2]. This techniques is commonly used in massive MIMO [23, 49] since it yields a simple and tight achievable rate which enables us to further design the systems. The tightness of this bound for cell-free massive MIMO is presented in [2]. Using Proposition 1 and the lower bound technique in [2], we can obtain an achievable rate as $R_k^{\text{Case 1}} = \log_2(1 + \text{SINR}_k^{\text{Case 1}})$, where $\text{SINR}_k^{\text{Case 1}}$ is given by

$$\text{SINR}_k^{\text{Case 1}} = \frac{\tilde{a}^4 \left|\text{DS}_k\right|^2}{\tilde{a}^4\mathbb{E}\left\{\left|\text{BU}_k\right|^2\right\} + \tilde{a}^4\mathbb{E}\left\{\left|\text{TN}_k\right|^2\right\} + \tilde{a}^4\sum\limits_{k'\neq k}^{K}\mathbb{E}\left\{\left|\text{IUI}_{kk'}\right|^2\right\} + \tilde{a}^2\mathbb{E}\left\{\left|\text{TQD}_k^{\text{y}}\right|^2\right\} + \tilde{a}^2\mathbb{E}\left\{\left|\text{TQD}_k^{\text{g}}\right|^2\right\} + \tilde{a}^2\sum\limits_{k'=1}^{K}\mathbb{E}\left\{\left|\text{TQD}_{kk'}\right|^2\right\} + \mathbb{E}\left\{\left|\text{TQD}_k^{\text{gy}}\right|^2\right\}}. \tag{3.5}$$

The closed-form expression for the achievable uplink rate of the *k*th user is given in the following theorem.

Theorem 1. *Having the quantized CSI and the quantized signal at the CPU and employing MRC detection at the CPU, the closed-form expression for the achievable rate of the kth user is given by $R_k^{Case\ 1} = \log_2(1 + SINR_k^{Case\ 1})$, where the $SINR_k^{Case\ 1}$ is given by*

$$SINR_k^{Case\ 1} = \frac{N^2 q_k \left(\sum_{m=1}^{M} u_{mk}\gamma_{mk}\right)^2}{N^2\sum\limits_{k'\neq k}^{K} q_{k'}\left(\sum\limits_{m=1}^{M} u_{mk}\gamma_{mk}\frac{\beta_{mk'}}{\beta_{mk}}\right)^2\left|\boldsymbol{\phi}_k^H\boldsymbol{\phi}_{k'}\right|^2 + N\left(\frac{C_{tot}}{\tilde{a}^4}+1\right)\sum\limits_{m=1}^{M} u_{mk}\gamma_{mk}\sum\limits_{k'=1}^{K} q_{k'}\beta_{mk'} + \frac{N}{\rho}\left(\frac{C_{tot}}{\tilde{a}^4}+1\right)\sum\limits_{m=1}^{M} u_{mk}\gamma_{mk}}, \tag{3.6}$$

where $\gamma_{mk} = \sqrt{\tau_p p_p}\beta_{mk}c_{mk}$ and $C_{tot} = 2\tilde{a}^2\sigma_{\tilde{e}}^2 + \sigma_{\tilde{e}}^4$.

Proof: The power of quantization distortions can be obtained as

$$\mathbb{E}\left\{\left|[\mathbf{e}_m^y]_n\right|^2\right\} = \mathbb{E}\left\{\left|[\tilde{\mathbf{e}}_m^y]_n\right|^2\right\}\left(\rho\sum_{k'=1}^{K} q_{k'}\beta_{mk'} + 1\right), \tag{3.7a}$$

$$\mathbb{E}\left\{\left|[\mathbf{e}_{mk}^g]_n\right|^2\right\} = \mathbb{E}\left\{\left|[\tilde{\mathbf{e}}_{mk}^g]_n\right|^2\right\}\gamma_{mk}. \tag{3.7b}$$

Since $\mathbb{E}\left\{|[\tilde{\mathbf{e}}_m^y]_n|^2\right\} = \mathbb{E}\left\{|[\tilde{\mathbf{e}}_{mk}^g]_n|^2\right\} = \sigma_{\tilde{e}}^2$, we have:

$$\mathbb{E}\left\{|[\mathbf{e}_m^y]_n|^2\right\} = \sigma_{\tilde{e}}^2\left(\rho\sum_{k'=1}^{K} q_{k'}\beta_{mk'} + 1\right), \tag{3.8a}$$

$$\mathbb{E}\left\{|[\mathbf{e}_{mk}^g]_n|^2\right\} = \sigma_{\tilde{e}}^2\gamma_{mk}. \tag{3.8b}$$

Using (3.7) and the fact that quantization distortion is indepnedent with the input of the quantizer, after some mathematical manipulations, we have:

$$\begin{aligned}\tilde{a}^2\mathbb{E}\left\{|\mathrm{TQD}_k^{\mathrm{y}}|^2\right\} + \tilde{a}^2\mathbb{E}\left\{|\mathrm{TQD}_k^{\mathrm{g}}|^2\right\} + \tilde{a}^2\sum_{k'=1}^{K}\mathbb{E}\left\{|\mathrm{TQD}_{kk'}|^2\right\} \\ + \mathbb{E}\left\{|\mathrm{TQD}_k^{\mathrm{gy}}|^2\right\} = NC_{\mathrm{tot}}\sum_{m=1}^{M} u_{mk}\gamma_{mk}\left(\rho\sum_{k'=1}^{K} q_{k'}\beta_{mk'} + 1\right).\end{aligned} \tag{3.9}$$

Note that the terms $|\mathrm{DS}_k|^2$, $\mathbb{E}\{|\mathrm{BU}_k|^2\}$, and $\mathbb{E}\{|\mathrm{IUI}_{kk'}|^2\}$ are derived in (3.45), (3.46) and (3.51), respectively. Finally substituting (3.9), (3.45), (3.46) and (3.51) into (3.5) results in (3.6), which completes the proof of Theorem 1. ■

Case 2. *Quantized Weighted Signal Available at the CPU:* The mth AP quantizes the terms $z_{m,k} = \hat{\mathbf{g}}_{mk}^H\mathbf{y}_m$, $\forall k$, and forwards the quantized signals in each symbol duration to the CPU as

$$z_{mk} = \hat{\mathbf{g}}_{mk}^H\mathbf{y}_m = r_{mk} + js_{mk},\ \forall k, \tag{3.10}$$

where r_{mk} and s_{mk} represent the real and imaginary parts of z_{mk}. An ADC quantizes the real and imaginary parts of $z_{m,k}$ with α bits each, which introduces quantization distortions to the received signals [48]. Let us consider the term e_{mk}^z as the quantization distortion of the mth AP. Hence, using the Bussgang decomposition, the relation between z_{mk} and its quantized version, $\tilde{z}_{mk}$, can be written as

$$\tilde{z}_{mk} = \tilde{a}z_{mk} + e_{mk}^z. \tag{3.11}$$

Remark 2. *Note that we assume that only the statistics of the channel are known, and the variance at the quantization input is defined on average over the channels. Although the input of the quantizer is not precisely Gaussian, since it is the sum of many random*

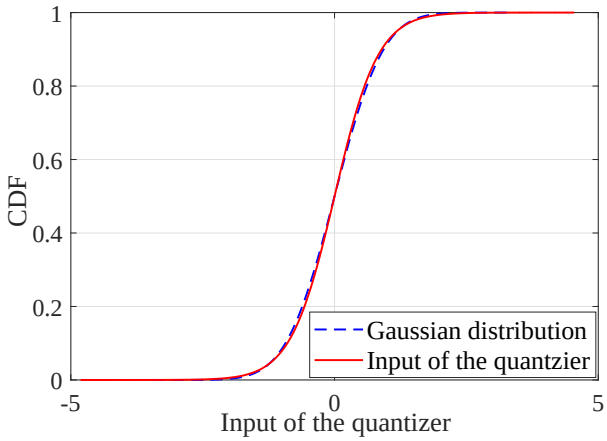

Figure 3.2: Cumulative distribution of the input of the quantizer with $K = 40$, $N = 10$ and $\tau_p = K$.

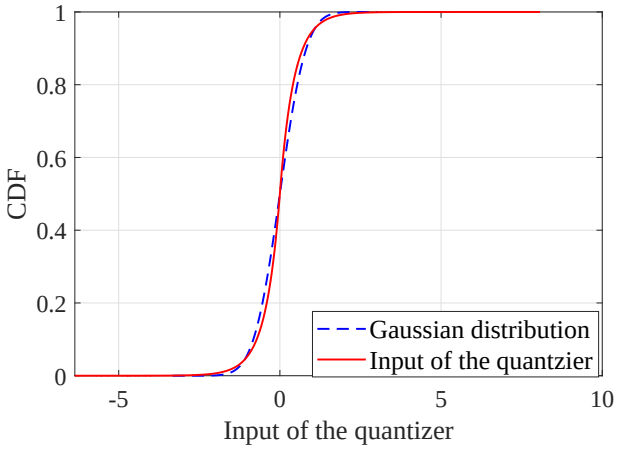

Figure 3.3: Cumulative distribution of the input of the quantizer with $K = 40$, $N = 1$ and $\tau_p = K$.

variates, from the central limit theorem, it has a near Gaussian distribution. Therefore, we use the Bussgang decomposition, making the approximation that the input of the quantizer is Gaussian distributed. The Gaussian approximation can be verified numerically, for typical parameter values, as shown in Figs. 3.2-3.4, we can see that the cumulative distribution of the empirical distribution matches very well with that of the Gaussian distribution.

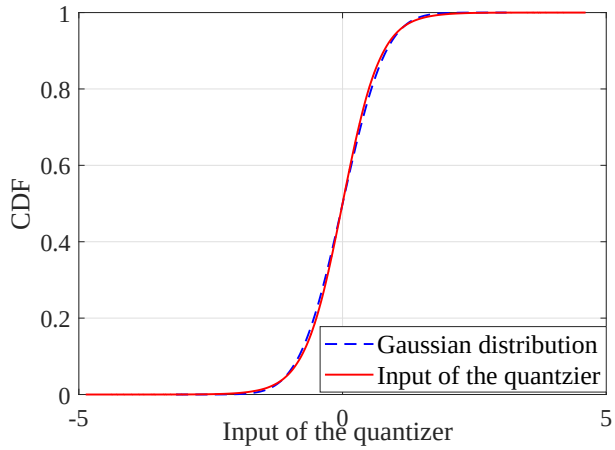

Figure 3.4: Cumulative distribution of the input of the quantizer with $K = 40$, $N = 1$ and $\tau_p = 30$.

The aggregated received signal at the CPU can be written as

$$
\begin{aligned}
r_k &= \sum_{m=1}^{M} u_{mk} \left(\tilde{a} \underbrace{\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m}_{z_{mk}} + e_{mk}^z \right) \\
&= \tilde{a} \sqrt{\rho} \sum_{k'=1}^{K} \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'} + \tilde{a} \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m + \sum_{m=1}^{M} u_{mk} e_{mk}^z \\
&= \underbrace{\tilde{a} \sqrt{\rho} \mathbb{E} \left\{ \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\} s_k}_{\mathrm{DS}_k} \\
&+ \underbrace{\tilde{a} \sqrt{\rho} \left(\sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} - \mathbb{E} \left\{ \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\} \right) s_k}_{\mathrm{BU}_k} \\
&+ \sum_{k' \neq k}^{K} \underbrace{\tilde{a} \sqrt{\rho} \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'}}}_{\mathrm{IUI}_{kk'}} s_{k'} \\
&+ \underbrace{\tilde{a} \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m}_{\mathrm{TN}_k} \\
&+ \underbrace{\sum_{m=1}^{M} u_{mk} e_{mk}^z}_{\mathrm{TQD}_k},
\end{aligned}
\tag{3.12}
$$

where TQD_k refers to the total quantization distortion (TQD) at the kth user. Note that

in cell-free massive MIMO with $M \to \infty$, due to the channel hardening property (please refer to Section 2.10 for more details), detection using only the channel statistics is nearly optimal. This is shown in [2] (see Fig. 2 of reference [2] and its discussion). Moreover, in [2] the authors show that in cell-free massive MIMO with $M \to \infty$, the received signal includes only the desired signal plus interference from the pilot sequence non-orthogonality. Finally, using the analysis in [2], the corresponding SINR of the received signal in (3.12) can be defined by considering the worst-case of the uncorrelated Gaussian noise is given by

$$\text{SINR}_k^{\text{Case 2}} = \frac{|\text{DS}_k|^2}{\mathbb{E}\{|\text{BU}_k|^2\} + \sum\limits_{k' \neq k}^{K} \mathbb{E}\{|\text{IUI}_{kk'}|^2\} + \mathbb{E}\{|\text{TN}_k|^2\} + \frac{1}{\tilde{a}^2}\mathbb{E}\{|\text{TQD}_k|^2\}}. \quad (3.13)$$

Based on the SINR definition in (3.13), the achievable uplink rate of the *k*th user is given in the following theorem.

Theorem 2. *Having the quantized weighted signal at the CPU and employing MRC detection at the CPU, the achievable uplink rate of the kth user in the cell-free massive MIMO system is* $R = \log_2(1 + \textit{SINR}^{\textit{Case 2}})$*, where* $\textit{SINR}^{\textit{Case 2}}$ *is given by*

$$\textit{SINR}_k^{\textit{Case 2}} = \frac{N^2 q_k \left(\sum_{m=1}^{M} u_{mk}\gamma_{mk}\right)^2}{N^2 \sum\limits_{k' \neq k}^{K} q_{k'} \left(\sum\limits_{m=1}^{M} u_{mk}\gamma_{mk}\frac{\beta_{mk'}}{\beta_{mk}}\right)^2 \left|\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}\right|^2 + N\sum\limits_{m=1}^{M} u_{mk}\left(\frac{\sigma_{\tilde{e}}^2(2\beta_{mk} - \gamma_{mk})}{\tilde{a}^2} + \gamma_{mk}\right)\sum\limits_{k'=1}^{K} q_{k'}\beta_{mk'} + \frac{N}{\rho}\left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2} + 1\right)\sum\limits_{m=1}^{M} u_{mk}\gamma_{mk}}. \quad (3.14)$$

Proof: Please refer to Appendix 3.B. ■

Required Fronthaul Capacity

Let τ_f be the length of the uplink payload data transmission for each coherence interval, i.e., $\tau_f = \tau_c - \tau_p$, where τ_c denotes the number of samples for each coherence interval and τ_p represents the length of pilot sequence. Defining the number of quantization bits as $\alpha_{m,i}$, for $i = 1, 2$, corresponding to Cases 1 and 2, and m refers to the mth AP. For Case 1, the required number of bits for each AP during each coherence interval is $2\alpha_{m,1} \times (NK + N\tau_f)$ whereas Case 2 requires $2\alpha_{m,2} \times (K\tau_f)$ bits for each AP during each coherence interval (note that the factor 2 indicates that we separately quantize the real

and imaginary parts of the signals). Hence, the total fronthaul capacity required between the mth AP and the CPU for all schemes is defined as

$$C_m = \begin{cases} \frac{2\left(NK + N\tau_f\right)\alpha_{m,1}}{T_c}, & \text{Case 1}, \\ \frac{2\left(K\tau_f\right)\alpha_{m,2}}{T_c}, & \text{Case 2}, \end{cases} \tag{3.15}$$

where T_c (in sec.) refers to coherence time. In the following, we present a comparison between two cases of uplink transmission. To make a fair comparison between Case 1 and Case 2, we use the same total number of fronthaul bits for both cases, that is $2(NK + N\tau_f)\alpha_{m,1} = 2(K\tau_f)\alpha_{m,2}$.

3.4 Proposed Max-Min Rate Scheme

In this section, we formulate the max-min rate problem for Case 2 of uplink transmission in cell-free massive MIMO system, where the minimum uplink rate of all users is maximized while satisfying the transmit power constraint at each user and the fronthaul capacity constraint. Note that the same approach can be used to investigate the max-min rate problem for Case 1. The achievable user SINR for the system model considered in the previous section is obtained by following a similar approach to that in [2]. Note that the main difference between the proposed approach and the scheme in [2] is the new set of receiver coefficients which are introduced at the CPU to improve the achievable user rates. The benefits of the proposed approach in terms of the achieved user uplink rate is demonstrated through numerical simulation results in Section V. In deriving the achievable rates of each user, it is assumed that the CPU exploits only the knowledge of channel statistics between the users and APs to detect data from the received signal in (3.12). Using the SINR given in (3.14), the achievable rate is obtained

$R_k^{\text{UP}} = \log_2(1 + \text{SINR}_k^{\text{Case 2}})$. Defining

$$\mathbf{u}_k = [u_{1k}, u_{2k}, \cdots, u_{Mk}]^T, \tag{3.16a}$$

$$\mathbf{\Gamma}_k = [\gamma_{1k}, \gamma_{2k}, \cdots, \gamma_{Mk}]^T, \tag{3.16b}$$

$$\mathbf{\Upsilon}_{kk'} = \text{diag}\left[\beta_{1k'}\left(\frac{\sigma_{\tilde{e}}^2\left(2\beta_{1k}-\gamma_{1k}\right)}{\tilde{a}^2}+\gamma_{1k}\right), \cdots, \beta_{Mk'}\left(\frac{\sigma_{\tilde{e}}^2\left(2\beta_{Mk}-\gamma_{Mk}\right)}{\tilde{a}^2}+\gamma_{Mk}\right)\right], \tag{3.16c}$$

$$\mathbf{\Lambda}_{kk'} = \left[\frac{\gamma_{1k}\beta_{1k'}}{\beta_{1k}}, \frac{\gamma_{2k}\beta_{2k'}}{\beta_{2k}}, \cdots, \frac{\gamma_{Mk}\beta_{Mk'}}{\beta_{Mk}}\right]^T, \tag{3.16d}$$

$$\mathbf{R}_k = \text{diag}\left[\left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2}+1\right)\gamma_{1k}, \cdots, \left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2}+1\right)\gamma_{Mk}\right]. \tag{3.16e}$$

The achievable uplink rate of the *k*th user is given by

$$R_k^{\text{UP}} = \log_2\left(1+\frac{\mathbf{u}_k^H\left(N^2 q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H\right)\mathbf{u}_k}{\mathbf{u}_k^H\left(N^2\sum\limits_{k'\neq k}^{K} q_{k'}|\boldsymbol{\phi}_k^H\boldsymbol{\phi}_{k'}|^2\mathbf{\Lambda}_{kk'}\mathbf{\Lambda}_{kk'}^H + N\sum\limits_{k'=1}^{K} q_{k'}\mathbf{\Upsilon}_{kk'} + \frac{N}{\rho}\mathbf{R}_k\right)\mathbf{u}_k}\right). \tag{3.17}$$

Next, the max-min rate problem can be formulated as follows:

$$P_1: \max_{q_k, \mathbf{u}_k, \alpha_2} \quad \min_{k=1,\cdots,K} \quad R_k^{\text{UP}} \tag{3.18a}$$

$$\text{subject to } ||\mathbf{u}_k|| = 1, \ \forall k, \quad 0 \leq q_k \leq p_{\text{max}}^{(k)}, \ \forall k, \tag{3.18b}$$

$$C_m \leq C_{\text{fh}}, \ \forall m, \tag{3.18c}$$

where $p_{\text{max}}^{(k)}$ and C_{fh} refer to the maximum transmit power available at user k and the capacity of fronthaul link between the mth AP and the CPU, respectively. Note that using (3.15), C_m is given as $C_m = \frac{2(K\tau_f)\alpha_{m,2}}{T_c}, \forall m$. Throughout the rest of the chapter, the index m is dropped from $\alpha_{m,i}, \ i = 1, 2$, as we consider the same number of bits to quantize the signal at all APs. Problem P_1 is a discrete optimization with integer decision variables (i.e., quantization bits) and it is obvious that the achievable user rates monotonically increase with the capacity of the fronthaul link between the mth AP and the CPU. Hence, the optimal solution is achieved when $C_m = C_{\text{fh}}, \ \forall m$, which leads to fixed values for the number of quantization bits. As a result, the max-min rate problem

can be re-formulated as follows:

$$P_2: \quad \max_{q_k, \mathbf{u}_k} \quad \min_{k=1,\cdots,K} \quad R_k^{\text{UP}} \tag{3.19a}$$

$$\text{subject to} \quad ||\mathbf{u}_k|| = 1, \ \forall k, \tag{3.19b}$$

$$0 \leq q_k \leq p_{\text{max}}^{(k)}, \ \ \forall k. \tag{3.19c}$$

Problem P_2 is not jointly convex in terms of $\mathbf{u}_k$ and power allocation $q_k, \ \forall k$. Therefore, it cannot be directly solved through existing convex optimization software. To tackle this non-convexity issue, we decouple Problem P_2 into two sub-problems: receiver coefficient design (i.e. $\mathbf{u}_k$) and the power allocation problem. The optimal solution for Problem P_2, is obtained through alternately solving these sub-problems, as explained in the following subsections.

3.4.1 Receiver Filter Coefficient Design

In this subsection, the problem of designing the receiver coefficients is considered. We solve the max-min rate problem for a given set of allocated powers at all users, $q_k, \forall k$, and fixed values for the number of quantization bits. These coefficients (i.e., $\mathbf{u}_k$, $\forall k$) are obtained by independently maximizing the uplink SINR of each user. Therefore, the optimal receiver filter coefficients can be determined by solving the following optimization problem:

$$P_3: \max_{\mathbf{u}_k} \frac{N^2 \mathbf{u}_k^H \left(q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H z\right) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 \sum\limits_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{kk'} \mathbf{\Lambda}_{kk'}^H + N \sum\limits_{k'=1}^{K} q_{k'} \mathbf{\Upsilon}_{kk'} + \frac{N}{\rho} \mathbf{R}_k\right) \mathbf{u}_k} \tag{3.20a}$$

$$\text{subject to} \quad ||\mathbf{u}_k|| = 1, \quad \forall k. \tag{3.20b}$$

Problem P_3 is a generalized eigenvalue problem [45], where the optimal solutions can be obtained by determining the generalized eigenvector of the matrix pair $\mathbf{A}_k = N^2 q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H$ and $\mathbf{B}_k = N^2 \sum_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{kk'} \mathbf{\Lambda}_{kk'}^H + N \sum_{k'=1}^{K} q_{k'} \mathbf{\Upsilon}_{kk'} + \frac{N}{\rho} \mathbf{R}_k$ corresponding to the maximum generalized eigenvalue.

Algorithm 1 Proposed algorithm to solve Problem P_2

1. Initialize $\mathbf{q}^{(0)} = [q_1^{(0)}, q_2^{(0)}, \cdots, q_K^{(0)}]$, $i = 1$
2. Repeat steps 3-5 until $\dfrac{\text{SINR}_k^{\text{UP},(i)} - \text{SINR}_k^{\text{UP},(i-1)}}{\text{SINR}_k^{\text{UP},(i-1)}} \leq \epsilon, \forall k$
3. Determine the optimal receiver coefficients $\mathbf{U}^{(i)} = [\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \cdots, \mathbf{u}_K^{(i)}]$ through solving the generalized eigenvalue Problem P_3 in (3.20) for a given $\mathbf{q}^{(i-1)}$,
4. Compute $\mathbf{q}^{(i)}$ through solving Problem P_5 in (3.22) for a given $\mathbf{U}^{(i)}$
5. $i = i + 1$

3.4.2 Power Allocation

In this subsection, we solve the power allocation problem for a given set of fixed receiver filter coefficients, $\mathbf{u}_k$, $\forall k$, and fixed values of quantization bits. The optimal transmit power can be determined by solving the following max-min problem:

$$P_4 : \max_{q_k} \min_{k=1,\cdots,K} \quad \text{SINR}_k^{\text{UP}} \tag{3.21a}$$

$$\text{subject to} \quad 0 \leq q_k \leq p_{\max}^{(k)}. \tag{3.21b}$$

Without loss of generality, Problem P_4 can be rewritten by introducing a new slack variable as

$$P_5 : \max_{t, q_k} \quad t \tag{3.22a}$$

$$\text{subject to} \quad 0 \leq q_k \leq p_{\max}^{(k)}, \ \ \forall k, \text{SINR}_k^{\text{UP}} \geq t, \ \ \forall k. \tag{3.22b}$$

Proposition 2. *Problem P_5 can be formulated into a standard GP.*

Proof: Please refer to Appendix 3.C. ■

Therefore, Problem P_5 is efficiently solved through existing convex optimization software. Based on these two sub-problems, an iterative algorithm has been developed by alternately solving both sub-problems at each iteration. The proposed algorithm is summarized in Algorithm 1. Note that ϵ in Step 2 of Algorithm 1 refers to a small predetermined value.

3.5 Convergence

In this section, we present the convergence of the proposed Algorithm 1. We propose to alternatively solve two sub-problems to find the solution of the original Problem P_2, where at each iteration, one of the design parameters is determined by solving the corresponding sub-problem while other design variable is fixed. We showed that each sub-problem provides an optimal solution for the other given design variable. Let us assume at the $i-1$th iteration, that the receiver filter coefficients $\mathbf{u}_k^{(i-1)},\ \forall k$ are obtained for a given power allocation $\mathbf{q}^{(i-1)}$ and similarly, the power allocation $\mathbf{q}^{(i)}$ is determined for a fixed set of receiver filter coefficients $\mathbf{u}_k^{(i-1)},\ \forall k$. Note that, the optimal power allocation $\mathbf{q}^{(i)}$ determined for a given $\mathbf{u}_k^{(i-1)}$ achieves an uplink rate greater than or equal to that of the previous iteration. In addition, the power allocation $\mathbf{q}^{(i-1)}$ is a feasible solution to find $\mathbf{q}^{(i)}$ as the receiver filter coefficients $\mathbf{u}_k^{(i)},\ \forall k$ are determined for a given $\mathbf{q}^{(i-1)}$. Note that the uplink rate of the system monotonically increases with the power. As a result, the achievable uplink rate of the system monotonically increases at each iteration. Note that the achievable uplink max-min rate is bounded from above for a given set of per-user power constraints and fronthaul link capacity constraint. Hence the proposed algorithm converges to a specific solution. Note that to the best of our knowledge and referring to [16, 40] this is a common way to show the convergence. In the next section, we prove the optimality of the proposed Algorithm 1 through the principle of uplink-downlink duality.

3.6 Optimality of the Proposed Max-Min Rate Algorithm

In this section, we present a method to prove the optimality of the proposed Algorithm 1. The proof is based on two main observations: we first demonstrate that the original max-min Problem P_2 with per-user power constraint is equivalent to an uplink problem with an equivalent total power constraint. We next prove that the same SINRs can be achieved in both the uplink and the downlink with an equivalent total power constraint, which enables us to establish an uplink-downlink duality. The concept of the virtual downlink max-min

SINR problem is explained in Section 3.6.2. Finally, we show that the virtual downlink max-min SINR problem is quasi-convex and can be optimally solved through a bisection search [50]. Note that the uplink Problem P_1 and the equivalent virtual downlink max-min SINR problem achieve the same SINRs and the solution of the virtual downlink max-min SINR problem is optimal. As a result, the optimality of the proposed Algorithm 1 is guaranteed. The details of the proof are provided in the following subsections. The key steps of the optimality of proposed scheme to solve Problem P_2 are explained in Fig. 3.5.

3.6.1 Equivalent Max-Min Uplink Problem

We aim to show the equivalence of Problem P_2 with a per-user power constraint and the uplink max-min rate problem with a total power constraint. Note that in the total power constraint, the maximum available transmit power is defined as the sum of all users' transmit powers from the solution of Problem P_2, which is formulated as:

$$P_6: \quad \max_{q_k, \mathbf{u}_k} \quad \min_{k=1,\cdots,K} \quad R_k^{\text{UP}} \tag{3.23a}$$

$$\text{subject to} \quad ||\mathbf{u}_k|| = 1, \ \forall k, \tag{3.23b}$$

$$\sum_{k=1}^{K} q_k \leq P_{\text{tot}}^c. \tag{3.23c}$$

Problem P_6 is not convex in terms of receiver filter coefficients $\mathbf{u}_k$ and power allocation $q_k, \forall k$. To deal with this non-convexity, similar to the proposed method to solve problem P_2, we propose to modify Algorithm 1 to incorporate the total power constraint in Problem P_6. Hence, we decompose Problem P_6 into receiver filter coefficient design and power allocation sub-problems. The same generalized eigenvalue problem in Problem P_3 is solved to determine the receiver filter coefficients whereas the GP formulation in P_5 is modified to incorporate the total power constraint (3.23c). Note that, the total power constraint is a convex constraint (posynomial function in terms of power allocation) and GP with the equivalent total power constraint can be used to find the optimum solution.

Lemma 2. *The original Problem P_2 (with per-user power constraint) and the equivalent Problem P_6 (with the equivalent total power constraint) have the same optimal solution.*

$$\mathrm{SINR}_k^{\mathrm{DL}}(\mathbf{U},\mathbf{p}) = \frac{\mathbf{u}_k^H \left(N^2 p_k \boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_k^H\right) \mathbf{u}_k}{N^2 \sum_{k' \neq k}^{K} \mathbf{u}_{k'}^H p_{k'} |\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k|^2 \boldsymbol{\Delta}_{k'k} \boldsymbol{\Delta}_{k'k}^H \mathbf{u}_{k'} + N \sum_{k'=1}^{K} \mathbf{u}_{k'}^H p_{k'} \mathbf{F}_{k'k} \mathbf{u}_{k'} + \frac{N}{\rho}}. \tag{3.24}$$

$$\mathrm{SINR}_k^{\mathrm{UP}}(\mathbf{U},\mathbf{q}) = \frac{\mathbf{u}_k^H \left(N^2 q_k \boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_k^H\right) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 \sum_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \boldsymbol{\Lambda}_{kk'} \boldsymbol{\Lambda}_{kk'}^H + N \sum_{k'=1}^{K} q_{k'} \boldsymbol{\Upsilon}_{kk'} + \frac{N}{\rho} \mathbf{R}_k\right) \mathbf{u}_k}. \tag{3.25}$$

Proof: Please refer to Appendix 3.D. ■

3.6.2 Uplink-Downlink Duality for Cell-Free Massive MIMO

This subsection demonstrates an uplink-downlink duality for cell-free massive MIMO systems. In particular, it is shown that the same SINRs (or rate regions) can be realized for all users in the uplink and the downlink with the equivalent total power constraints [44, 51], respectively. In other words, based on the principle of uplink-downlink duality, the same set of filter coefficients can be utilized in the uplink and the downlink to achieve the same SINRs for all users with different user power allocations. The following theorem defines the achievable virtual downlink rate for cell-free massive MIMO systems:

Theorem 3. *By employing conjugate beamforming at the APs, the achievable virtual downlink rate of the kth user in the cell-free massive MIMO system with K randomly distributed single-antenna users, M APs where each AP is equipped with N antennas and limited-capacity fronthaul links is given by (3.24) (defined at the top of this page).*

Proof: This can be derived by following the same approach as for uplink transmission in Theorem 2. ■

Note that in (3.24), $p_k,\ \forall k$ denotes the downlink power allocation for the kth user and

the following equalities hold:

$$\mathbf{\Gamma}_k = [\gamma_{1k}, \gamma_{2k}, \cdots, \gamma_{Mk}]^T, \tag{3.26a}$$

$$\mathbf{F}_{k'k} = \text{diag}\Big[\beta_{1k}(\frac{w^2(2\beta_{1k'} - \gamma_{1k'})}{3Q_1^2} + \gamma_{1k'}), \cdots, \beta_{Mk}(\frac{w^2(2\beta_{Mk'} - \gamma_{Mk'})}{3Q_M^2} + \gamma_{Mk'})\Big], \tag{3.26b}$$

$$\mathbf{\Delta}_{k'k} = [\frac{\gamma_{1k'}\beta_{1k}}{\beta_{1k'}}, \frac{\gamma_{2k'}\beta_{2k}}{\beta_{2k'}}, \cdots, \frac{\gamma_{Mk'}\beta_{Mk}}{\beta_{Mk'}}]^T, \tag{3.26c}$$

The following Theorem provides the required condition to establish the uplink-downlink duality for cell-free massive MIMO systems with limited-capacity fronthaul links:

Theorem 4. *By employing MRC detection in the uplink and conjugate beamforming in the downlink, to realize the same SINR tuples in both the uplink and the downlink of a cell-free massive MIMO system, with the same fronthaul loads, the same filter coefficients and different transmit power allocations, the following condition should be satisfied:*

$$N \sum_{m=1}^{M} \sum_{k=1}^{K} \left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2} + 1 \right) \gamma_{mk} \left| w_{mk} \right|^2 = \sum_{k=1}^{K} q_k^* = P_{tot}^c \,, \tag{3.27}$$

where q_k^*, $\forall k$ refer to the optimal solution of Algorithm 1, and w_{mk} denotes the (m,k)-th entry of matrix $\mathbf{W}$ which is defined as follows:

$$\mathbf{W} = [\sqrt{p_1}\mathbf{u}_1, \sqrt{p_2}\mathbf{u}_2, \cdots, \sqrt{p_K}\mathbf{u}_K]. \tag{3.28}$$

Proof: Please refer to Appendix 3.E. ■

Let us again consider the uplink and downlink SINR formulas in (3.24) and (3.25), respectively. Considering the downlink SINR in (3.24), it is possible to define the precoding matrix $\mathbf{W} = [\sqrt{p_1}\mathbf{u}_1, \sqrt{p_2}\mathbf{u}_2, \cdots, \sqrt{p_K}\mathbf{u}_K]$, which let us solve the downlink max-min SINR problem jointly in terms of $\sqrt{p_k}$s and $\mathbf{u}_k$s. However, it is not possible to define such a $\mathbf{W}$ matrix for the uplink SINR in (3.25). In other words, in the downlink, it is not required to define the separate variables $\sqrt{p_k}$s and $\mathbf{u}_k$s as both the power and transmitter filter are designed in a centralized manner at the CPU whereas in the uplink, the power elements q_ks are used in a distributed manner at the users' ends to transmit the uplink data, while the receiver filters $\mathbf{u}_k$s are designed at the CPU. Hence, the equivalent downlink SINR defined with $\sqrt{p_k}$s and $\mathbf{u}_k$s is called virtual downlink SINR.

3.6.3 Equivalent Max-Min Downlink Problem

Note that the uplink Problems P_2 and P_6 are not convex in terms of receiver filter coefficients $\mathbf{U}$ and power allocation $q_k, \forall k$, and hence they are iteratively solved. However, it can be shown that the downlink problem it can be solved using bisection method as both receiver filter coefficients and power allocation can be combined into a singe variable. In particular, it is shown that the same SINRs (or rate regions) can be realized for all users in the uplink and the downlink with the equivalent total power constraints [44, 51], respectively. Next, let us consider the following *virtual* downlink max-min SINR problem:

$$P_7: \quad \max_{p_k, \mathbf{u}_k} \min_{k=1,\cdots,K} \quad R_k^{\text{DL}} \tag{3.29a}$$

$$\text{subject to} \quad ||\mathbf{u}_k|| = 1, \forall k, \tag{3.29b}$$

$$\sum_{k=1}^{K} c_k^{\text{vir}} p_k \leq P_{\text{tot}}^c, \tag{3.29c}$$

where R_k^{DL} is the downlink SINR of user k, where p_k is the downlink power allocated for user k, and using the analysis in Theorem 4, we have $c_k^{\text{vir}} = \sum_{m=1}^{M} \left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2} + 1 \right) \gamma_{mk} \left| u_{mk} \right|^2$. It is easy to show that Problem P_7 can be re-formulated as a quasi-concave and a bisection search can be used to obtain the optimal solution. This is why the virtual downlink max-min SINR problem is interesting to us (because the optimal solution can be determined through a bisection search as explained in Problem P_8 as explained below). Therefore, Problem P_7 is difficult to jointly solve in terms of transmit filter coefficients $\mathbf{u}_k$'s and power allocations p_k's. However, it can be represented by introducing a new variable $\mathbf{W}$ to decouple the variables $\mathbf{U}$ and $\mathbf{q}$ as follows:

$$P_8: \max_{\mathbf{W}} \min_{k=1,\cdots,K} \quad R_k^{\text{DL}} \tag{3.30a}$$

$$\text{subject to} \quad N \sum_{m=1}^{M} \sum_{k=1}^{K} \left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2} + 1 \right) \gamma_{mk} |w_{mk}|^2 \leq P_{\text{tot}}^c. \tag{3.30b}$$

It is easy to show that Problem P_8 is quasi-convex. Hence, a bisection [50] approach can be used to obtain the optimal solution for the original Problem P_8 by sequentially solving

the following power minimization problem for a given target SINR t at all users:

$$P_9: \min_{\mathbf{W}} \quad \sum_{m=1}^{M}\sum_{k=1}^{K}\gamma_{mk}|w_{mk}|^2 \tag{3.31a}$$

$$\text{subject to } \frac{\mathbf{w}_k^H\left(N^2\mathbf{\Gamma}_k\mathbf{\Gamma}_k^H\right)\mathbf{w}_k}{N^2\sum\limits_{k'\neq k}^{K}\mathbf{w}_{k'}^H|\boldsymbol{\phi}_{k'}^H\boldsymbol{\phi}_k|^2\mathbf{\Delta}_{k'k}\mathbf{\Delta}_{k'k}^H\mathbf{w}_{k'}+N\sum\limits_{k'=1}^{K}\mathbf{w}_{k'}^H\mathbf{F}_{k'k}\mathbf{w}_{k'}+\frac{N}{\rho}}\geq t, \tag{3.31b}$$

$$N\sum_{m=1}^{M}\sum_{k=1}^{K}\left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}}+1\right)\gamma_{mk}|w_{mk}|^2\leq P_{\text{tot}}^c, \tag{3.31c}$$

where $\mathbf{w}_k$ represents the kth column of the matrix $\mathbf{W}$ defined in (3.28). Problem P_9 can be reformulated by exploiting a second order cone programming (SOCP). Note that the objective function in (3.31) refers to the total transmit power. As a result, the optimal solution for Problem P_7 can be obtained by extracting the normalized transmit filter coefficients $\mathbf{u}_k$'s and power allocations p_k's as

$$p_k^*=||\mathbf{w}_k^*||^2,\ \forall k,\ \ \&\ \ \mathbf{u}_k^*=\frac{\mathbf{w}_k^*}{||\mathbf{w}_k^*||},\ \forall k, \tag{3.32}$$

where the $\mathbf{w}_k^*$'s refer to the optimal solution of Problem P_8. Note that constraint (3.31c) is an equivalent total power constraint to the per-user power constraint in the original Problem P_2, which is a more relaxed constraint than (3.19c).

3.6.4 Proof of Optimality of the Proposed Algorithm to Solve Problem P_2:

In Lemma 2, we prove that Problems P_2 and P_6 are equivalent, and have the same solution. Next, in Proposition 1, using uplink-downlink duality, we prove that Problem P_6 and the virtual downlink max-min SINR Problem P_7 are equivalent. Note that the SINR achieved by solving Problem P_7 are optimal (the optimal solution is obtained by a bisection search approach). This confirms that the proposed algorithm to solve Problem P_2 is optimal. The steps of this proof are presented in Fig. 3.5.

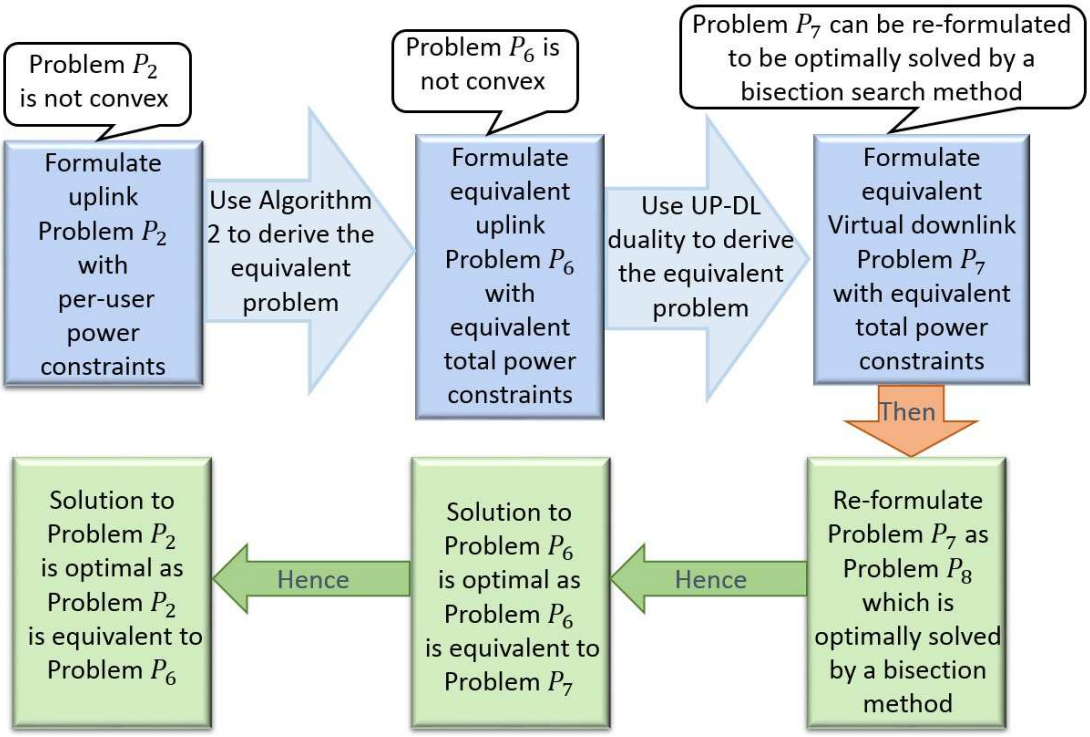

Figure 3.5: The key steps of the optimality of proposed scheme to solve Problem P_2.

3.7 User Assignment

Exploiting (3.15), it is obvious that the total fronthaul capacity required between the mth AP and the CPU increases linearly with the total number of users served by the mth AP. This motivates the need to pick a proper set of active users for each AP. Using (3.15), we have

$$\alpha_2 \times K_m \leq \frac{C_{\text{fh}} T_c}{2\tau_f}, \tag{3.33}$$

where K_m denotes the size of the set of active users for the mth AP. From (3.33), it can be seen that decreasing the size of the set of active users allows for a larger number of quantization bits. Motivated by this fact, and to exploit the capacity of fronthaul links more efficiently, we investigate all possible combinations of α_2 and K_m. First, for a fixed value of α_2, we find an upper bound on the size of the set of active users for each AP. In the next step, we propose for all APs that the users are sorted according to β_{mk}, $\forall k$, and find the K_m users which have the highest values of β_{mk} among all users. If a user is not selected by any AP, we propose to find the AP which has the best link to this user (in Algorithm 6, $\pi(j) = \underset{m}{\text{argmax}}\, \beta_{mj}$ determines best link to the jth user, i.e., the index of the AP which is closest to the jth user). Note that to only consider the users that have

Algorithm 2 User Assignment

1. Using (3.33), find the maximum possible integer value for $K_m, \forall m$
2. Sort users according to the ascending channel gain: $\beta_{m1} \geq \beta_{m2} \geq \cdots \geq \beta_{mK}, \forall m$
3. Assign K_m users with the highest values of $\beta_{mk}, \forall m$ to each AP, i.e., $\mathcal{T}_m \leftarrow \{k^{(1)}, k^{(2)}, \cdots, k^{(K_m)}\}, \forall m$
4. Find set of active APs for each user; $\mathcal{S}_k \leftarrow \{m^{(1)}, m^{(2)}, \cdots, m^{(M_k)}\}, \forall k$
5. for $j = 1 : K$
if size $\{\mathcal{S}_j\} = 0$
$\pi(j) = \underset{m}{\operatorname{argmax}}\, \beta_{mj}$, $\delta(j) = \underset{k}{\operatorname{argmin}}\, \beta_{\pi(j)k}, k|\mathcal{S}_k\pi_j \neq \varnothing$, $\mathcal{T}_{\pi(j)} \leftarrow \mathcal{T}_{\pi(j)} \backslash \delta(j)$, $\mathcal{T}_{\pi(j)} \leftarrow \mathcal{T}_{\pi(j)} \cup j$
end
end
6. If $m \in \mathcal{S}_k$, then $\tilde{\gamma}_{mk} \leftarrow \gamma_{mk}$, otherwise $\tilde{\gamma}_{mk} = 0$ and solve the max-min rate problem P_2

links to other APs, we use $k|\mathcal{S}_k\pi_j \neq \varnothing$, where $\varnothing$ refers to the empty set. Then we drop the user which has the lowest $\beta_{mk},\ \forall k$, among the set of active users for that AP, which has links to other APs as well. Finally, we add the user which is not selected by any AP to the set of active users for this AP. We next solve the virtual downlink max-min SINR problem to maximize the minimum uplink rate of the users as follows

$$P_{10} : \max_{\mathbf{W}} \min_{k=1,\cdots,K} \quad R_k^{\text{DL}}\left(\tilde{\gamma}_{mk}\right) \tag{3.34a}$$

$$\text{subject to} \quad N \sum_{m=1}^{M} \sum_{k=1}^{K} \left(\frac{\sigma_e^2}{\tilde{a}^2} + 1\right) \tilde{\gamma}_{mk} |w_{mk}|^2 \leq P_{\text{tot}}^c, \tag{3.34b}$$

where

$$\tilde{\gamma}_{mk} = \begin{cases} \gamma_{mk}, & m \in \mathcal{S}_k \\ 0, & \text{otherwise} \end{cases} \tag{3.35}$$

where $\mathcal{S}_k$ refers to the set of active APs for the kth user. The proposed algorithm is summarized in Algorithm 6.

3.8 Numerical Results and Discussion

In this section, we provide numerical simulation results to validate the performance of the proposed max-min rate scheme with different parameters. A cell-free massive

MIMO system with M APs and K single-antenna users is considered in a $D \times D$ simulation area, where both APs and users are uniformly distributed at random. In the following subsections, we define the simulation parameters and then present the corresponding simulation results. The channel coefficients between users and APs are modeled in Section 2.7.4 where the coefficient β_{mk} is given by $\beta_{mk} = \text{PL}_{mk} 10^{\frac{\sigma_{sh}\, z_{mk}}{10}}$, where PL_{mk} is the path loss from the kth user to the mth AP and the second term$10^{\frac{\sigma_{sh} z_{mk}}{10}}$, denotes the shadow fading with standard deviation $\sigma_{sh} = 8$ dB, and $z_{mk} \sim \mathcal{N}(0,1)$ [2]. In the simulation, an uncorrelated shadowing model is considered and a three-slope model for the path loss similar to [2]. The noise power is given by $p_n = \text{BW} \times k_B \times T_0 \times W$, where $\text{BW} = 20$ MHz denotes the bandwidth, $k_B = 1.381 \times 10^{-23}$ represents the Boltzmann constant, and $T_0 = 290$ (Kelvin) denotes the noise temperature. Moreover, $W = 9$dB, and denotes the noise figure. It is assumed that $\bar{p}_p$ and $\bar{\rho}$ denote the power of pilot sequence and the uplink data powers, respectively, where $p_p = \frac{\bar{p}_p}{p_n}$ and $\rho = \frac{\bar{\rho}}{p_n}$. In simulations, we set $\bar{p}_p = 200$ mW and $\bar{\rho} = 200$ mW. Similar to [2], we assume that the simulation area is wrapped around at the edges which can simulate an area without boundaries. Hence, the square simulation area has eight neighbours. We evaluate the average rate of the system over 300 random realizations of the locations of APs, users and shadow fading. Similar to the model in [52], the fronthaul links establish communications through wireless microwave links with limited capacity. Hence, we use $C_{\text{fh}} = 100$ Mbits/s [52], unless otherwise is indicated. In this chapter, the term "orthogonal pilots" refers to the case where unique orthogonal pilots are assigned to all users, while in "random pilot assignment" each user is randomly assigned a pilot sequence from a set of τ_p orthogonal sequences of length τ_p ($< K$), following the approach of [2] (for more details refer to Section 2.7.5).

Performance of Different Cases of Uplink Transmission

Fig. 3.6 presents the average per-user uplink rate is obtained by solving Problem P_4, given by (3.21) for Cases 1 and 2. The values of $\alpha_1 = 9$ and $\alpha_2 = 2$ correspond to a total number of 14,400 bits for each AP during each coherence time (or frame). In addition, similar to [48] we use a uniform quantizer with fixed step size. As Fig 3.6 shows the performance of Case 1 is slightly better than Case 2 for $K = 20$. Next, the performance of the cell-free massive MIMO system is evaluated for a system with $K = 40$ in which each AP is

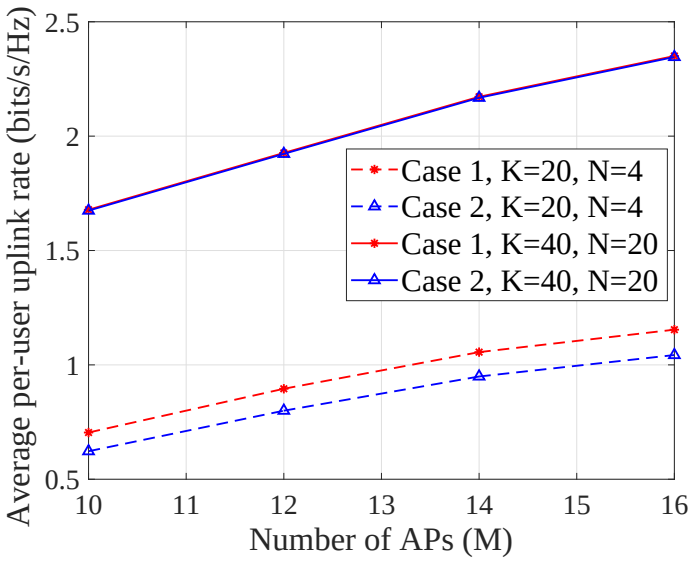

Figure 3.6: Average per-user uplink rate for cases 1 and 2, with ($N = 4$, $K = 20$, $\tau_p = 20$, $\alpha_1 = 9$, $\alpha_2 = 2$), and ($N = 20$, $K = 40$, $\tau_p = 40$, $\alpha_1 = 8$, $\alpha_2 = 5$) with $D = 1$ km and $\tau_c = 200$. Note that here $\tau_f = \tau_c - \tau_p = 160$.

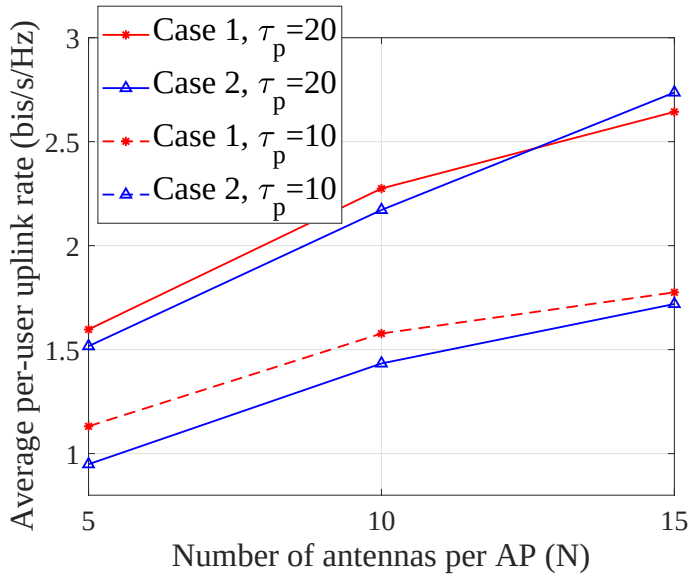

Figure 3.7: Average per-user uplink rate for cases 1 and 2, for $M = 20$, $K = 20$, $\tau_p = 20$, $\tau_p = 10$, $D = 1$ km and $\tau_c = 200$ versus number of antennas per AP. Note that we consider $(\alpha_1 = 18, \alpha_2 = 5)$, $(\alpha_1 = 18, \alpha_2 = 10)$, $(\alpha_1 = 18, \alpha_2 = 15)$ for the cases of $N = 5$, $N = 10$, $N = 15$, respectively. This results in total number of $18,000$ bits for all values of N.

equipped with $N = 20$ antennas. Fig. 3.6 shows the average rate of the cell-free massive MIMO system, where for Case 1 and Case 2, we set $\alpha_1 = 8$ and $\alpha_2 = 5$, respectively which leads to a total number of 64,000 fronthaul bits per AP per frame. Fig. 3.6 shows that the performances of Case 1 and Case 2 depend on the values of N, K and τ_f. Next,

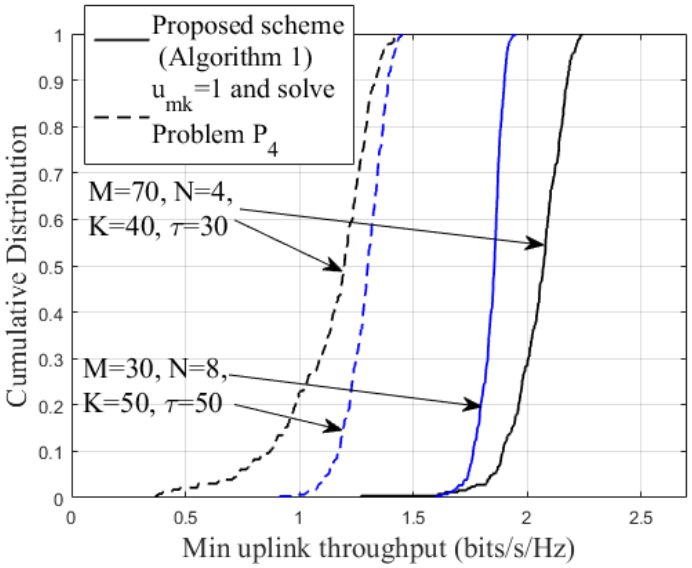

Figure 3.8: The cumulative distribution of the per-user uplink rate, for $\{M = 70, N = 4, K = 40\}$, $\{M = 50, N = 8, K = 50\}$, and $\tau_p = 30$, $\alpha_1 = 1$ and $D = 1$ km.

we investigate the effect of number of antennas per AP and τ_f for $K = 20$. Fig. 3.7 shows the average per-user uplink rate of cell-free massive MIMO versus number of antennas per AP and two cases of $\tau_p = 20$ ($\tau_f = 180$) and $\tau_p = 10$ ($\tau_f = 190$). Moreover, we consider $(\alpha_1 = 18, \alpha_2 = 5)$, $(\alpha_1 = 18, \alpha_2 = 10)$, $(\alpha_1 = 18, \alpha_2 = 15)$ for the cases of $N = 5$, $N = 10$, $N = 15$, respectively, resulting $18,000$ bits for all values of N. As the figure shows the difference between Case 1 and Case 2 decreases as N increases. Moreover, for the case of orthogonal pilots and $N = 15$, the performance of Case 2 is better than the performance of Case 1. Since in case 1, the CPU knows the quantized channel estimates, other signal processing techniques (e.g., ZF processing) can be implemented to improve the system performance and can be considered in future work.

Performance of the Proposed User Max-Min Rate Algorithm

In this subsection, we evaluate the performance of the proposed uplink max-min rate scheme. To assess the performance, a cell-free massive MIMO system is considered with 70 APs ($M = 70$) where each AP is equipped with $N = 4$ antennas and 40 users ($K = 40$) which are randomly distributed over the simulation area of size 1×1 km meters. Moreover, we consider the case $\{M = 50, N = 8, K = 50\}$. Fig. 3.8 presents the cumulative distribution of the achievable uplink rates for the proposed Algorithm 1

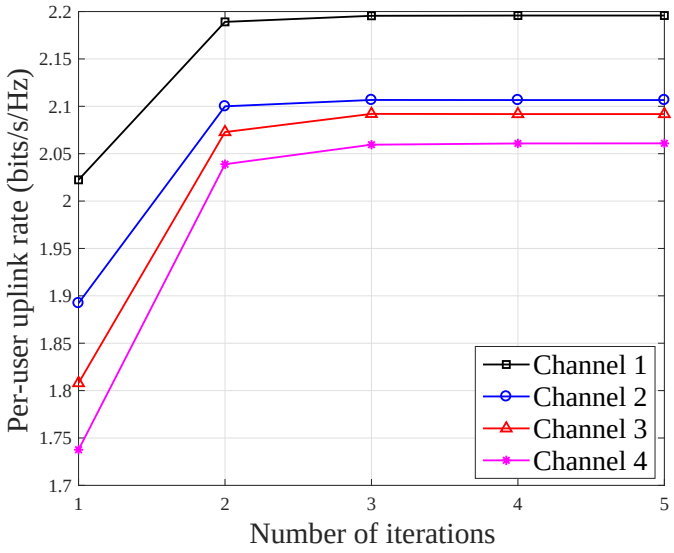

Figure 3.9: The convergence of the proposed max-min SINR approach (Algorithm 1) for $M = 70$, $N = 4$, $K = 40$, $\tau_p = 30$, $\alpha_1 = 1$ and $D = 1$ km.

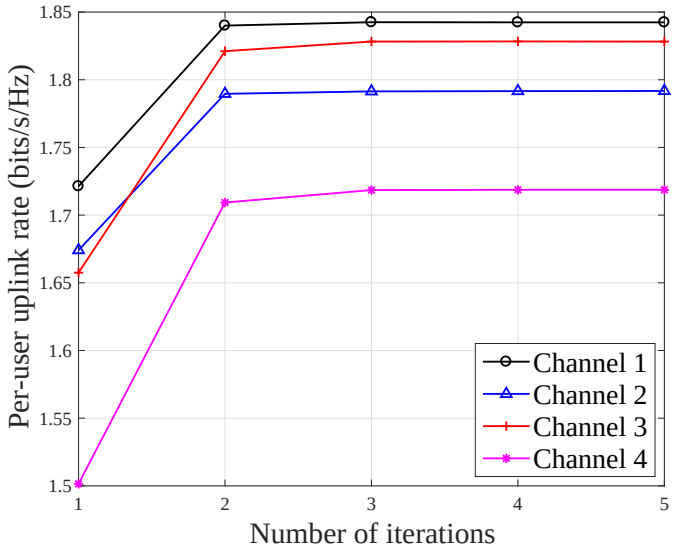

Figure 3.10: The convergence of the proposed max-min SINR approach (Algorithm 1) for $M = 30$, $N = 8$, $K = 50$, $\tau_p = 50$, $\alpha_1 = 1$ and $D = 1$ km.

in the case similar to [2], without defining the coefficients $\mathbf{u}_k$, (i.e., $u_{mk} = 1 \; \forall m, k$) and solving Problem P_4, with random pilot sequences with length $\tau_p = 30$. As seen in Fig. 3.8, the performance (i.e. the 10%-outage rate, R_{out}, refers to the case when $P_{\text{out}} = \Pr(R_k < R_{\text{out}}) = 0.1$, where Pr refers to the probability function) of the proposed scheme is almost three times than that of the case with $u_{mk} = 1 \; \forall m, k$. Note that the authors in [2] do not consider any receiver filter $\mathbf{U}$, which is equivalent to $u_{mk} = 1 \; \forall m, k$. Moreover, the authors in [2] consider the error-free and unlimited-capacity fronthaul links.

Convergence

Next, we provide simulation results to validate the convergence of the proposed algorithm for a set of different random realizations of the locations of APs, users and shadow fading. These results are generated over the simulation area of size 1×1 km^2 with random and orthogonal pilot sequences. Fig. 3.9 investigates the convergence of the proposed Algorithm 1 with 70 APs (M = 70) and 40 users (K=40) and random pilot sequences with length $\tau_p = 30$, whereas Fig. 3.10 demonstrates the convergence of the proposed Algorithm 1 for the case of $M = 30$ APs and $K = 50$ with orthogonal pilot sequences. The figures confirm that the proposed algorithm converges after a few iterations, while the minimum rate of the users increases with the iteration number.

Uplink-Downlink Duality in Cell-Free Massive MIMO System

Here, the simulation results are provided to support the theoretical derivations of the uplink-downlink duality and the optimality of Algorithm 1. It is assumed that users are randomly distributed through the simulation area of size 1×1 km. Figs. 3.11 compares the cumulative distribution of the achievable uplink rates between the original uplink max-min problem (Problem P_1), the equivalent uplink problem (Problem P_6) and the equivalent downlink problem (Problem P_7). In Fig. 3.11, the minimum uplink rate is obtained for a system with 30 APs ($M = 30$) where each is equipped with $N = 8$ antennas and has 50 users ($K = 50$) for two cases of orthogonal pilot sequences and random pilot sequences with length $\tau_p = 30$. Moreover, Fig. 3.11 demonstrates the same results for 70 APs ($M = 70$), $N = 4$, 40 users ($K = 40$), and $\tau_p = 30$. The simulation results provided in Fig. 3.11 validate our result that the problem formulations P_1, P_6 and P_7 are equivalent and achieve the same minimum user rate. In addition, these results support our result on the uplink-downlink duality for cell-free massive MIMO in Section VI and the proof of optimality of Algorithm 1.

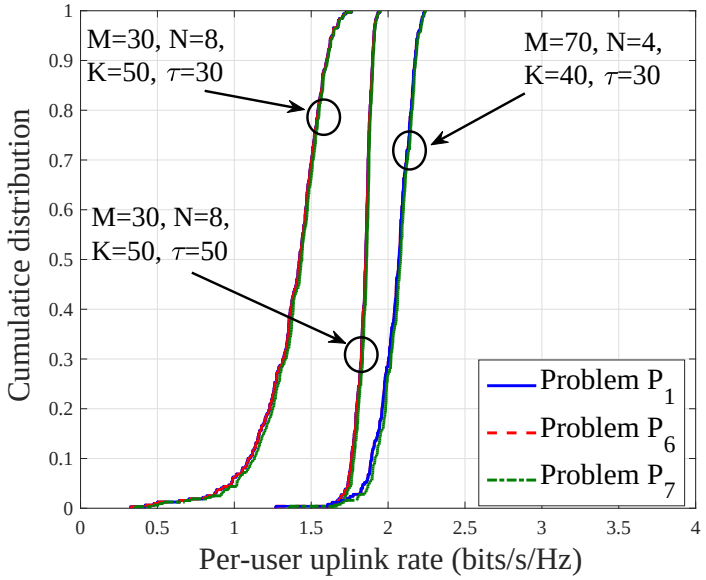

Figure 3.11: The cumulative distribution of the per-user uplink rate for the original problem with per-user power constraint (Problem P_1), the equivalent uplink problem with total power constraint (Problem P_5), and the equivalent downlink problem (Problem P_6), with $\alpha_1 = 1$ and $D = 1$ km.

Performance of the Proposed User Assignment Algorithm 6

This subsection investigates the performance of the proposed user assignment Algorithm 6. In Fig. 3.12, the average per-user uplink rate is presented with $M = 120$, $N = 2$, $K = 50$, orthogonal pilot sequences and random pilot assignment with $D = 1$ km, versus the total number of active users per AP. Here, we used inequality (3.33) and set $\alpha_2 \times K_m = 100$ for all curves in Fig. 3.12. The optimum value of K_m, (K_m^{opt}), depends on the system parameters and as Fig. 3.12 shows for both cases of $\tau_p = 50$ and $\tau_p = 30$, the optimum value is achieved by $K_m^{\text{opt}} = 20$. As a result, the proposed user assignment scheme can efficiently improve the performance of cell-free massive MIMO systems with limited fronthaul capacity. For instance, using the proposed user assignment scheme for the case of $\tau_p = 50$ in Fig. 3.12, one can achieve per-user uplink rate of 2.442 bits/s/Hz by setting $K_m^{\text{opt}} = 20$, instead of quantizing the signals of all $K = 40$ users and achieving per-user uplink rate of 2.3 bits/s/Hz, which indicates more than 5.2% in the performance of cell-free massive MIMO systems with limited fronthaul capacity.

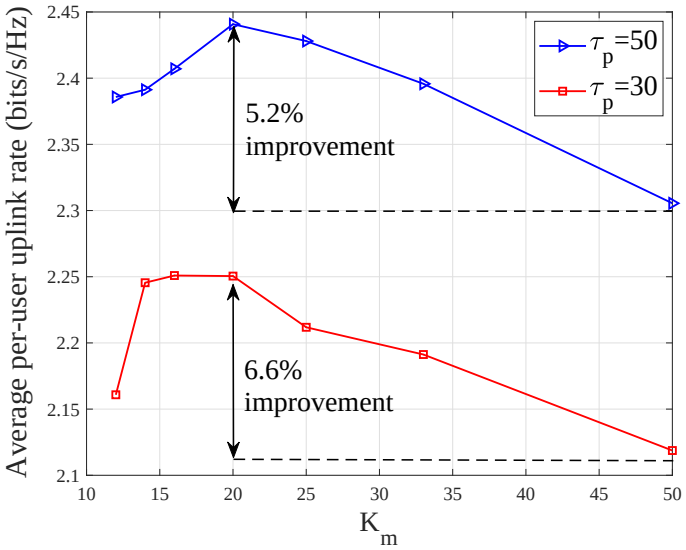

Figure 3.12: Average per-user uplink rate versus K_m (the total number of active users for each AP) with $M = 120$, $N = 2$, $K = 50$ and $\alpha_2 \times K_m = 100$.

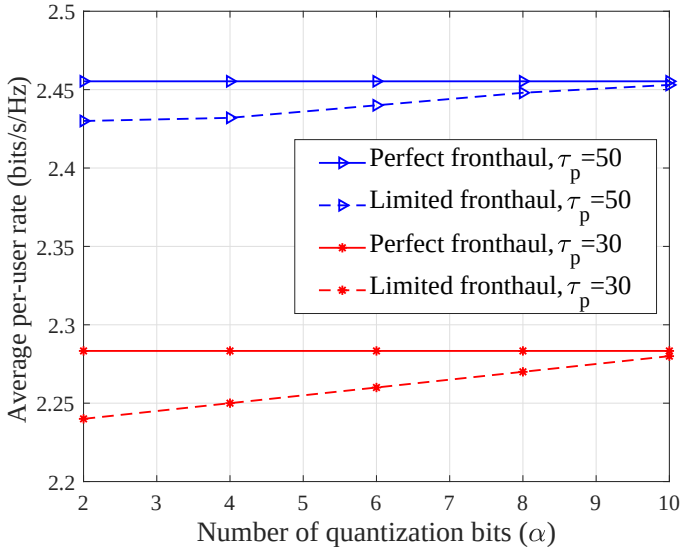

Figure 3.13: Average per-user uplink rate versus the number of quantization bits, α_2, with limited and perfect fronthaul links and $M = 120$, $K = 50$, $N = 2$, $D = 1$ km, $\tau_p = 30$ and $\tau_p = 50$.

Effect of the Capacity of Fronthaul Links

What is the optimal capacity of fronthaul links in cell-free massive MIMO systems to approach the performance of the system with perfect and error-free fronthaul links? The aim of this subsection is to answer this fundamental question. In this subsection, we evaluate the performance of the cell-free massive MIMO system with two cases of perfect and limited fronthaul links. To assess the performance, a cell-free massive MIMO system is considered with $M = 120$, $K = 50$, $N = 2$, $D = 1$ km, $\tau_p = 30$ and $\tau_p = 50$. To improve the performance of the system, we exploit the proposed user assignment algorithm. Fig. 3.13 presents average per-user uplink rate with the proposed max-min rate algorithm versus number of quantization bits, α_1 with the use of proposed user assignment algorithm. As Fig. 3.13 shows, for both cases of random and orthogonal pilots to closely approach the performance of perfect fronthaul links, we need to set $\alpha_1 \geq 8$.

3.9 Summary

We have studied the uplink max-min rate problem in cell-free massive MIMO with the realistic assumption of limited-capacity fronthaul links, and have proposed an optimal solution to maximize the minimum user rate. The original max-min problem was divided into two sub-problems which were iteratively solved by formulating them into generalized eigenvalue problem and GP. The optimality of the proposed solution has been validated through establishing an uplink-downlink duality. Numerical results have been provided to demonstrate the optimality of the proposed scheme in comparison with the existing schemes. In addition, these results confirmed that the proposed max-min rate algorithm can increase the median of the CDF of the minimum uplink rate of the users by more than two times, compared to existing algorithms. We finally showed that further improvement in average rate of the users can be achieved by the proposed user assignment algorithm.

3.10 Appendix

Appendix 3.A: Proof of Proposition 1

Terms $\mathbf{e}_m^y$ and $\mathbf{e}_{mk}^{\hat{g}}$ have i.i.d. RVs with zero mean [48]. The value of the quantization distortion is uncorrelated with the input of the quantizer. This can be achieved by exploiting the Bussgang decomposition [36]. In this chapter, we do not address the details of Bussgang decomposition and it can be considered an an interesting future direction. As a result, we have

$$\mathbb{E}\left\{[\mathbf{e}_m^y]_n\right\} = 0, \tag{3.36a}$$

$$\mathbb{E}\left\{\left[\mathbf{e}_{mk}^{\hat{g}}\right]_n\right\} = 0, \tag{3.36b}$$

$$\mathbb{E}\left\{(\mathbf{e}_m^y)^H \mathbf{e}_{mk}^{\hat{g}}\right\} = 0, \tag{3.36c}$$

$$\mathbb{E}\left\{\hat{\mathbf{g}}_{mk}^H \mathbf{e}_{mk}^{\hat{g}}\right\} = 0, \tag{3.36d}$$

$$\mathbb{E}\left\{\mathbf{y}_m^H \mathbf{e}_{mk}^{\hat{g}}\right\} = 0, \tag{3.36e}$$

$$\mathbb{E}\left\{\mathbf{y}_m^H \mathbf{e}_m^y\right\} = 0, \tag{3.36f}$$

$$\mathbb{E}\left\{\hat{\mathbf{g}}_{mk}^H \mathbf{e}_m^y\right\} = 0. \tag{3.36g}$$

In addition, based on [16], we have

$$\mathbf{g}_{mk} = \hat{\mathbf{g}}_{mk} + \bar{\mathbf{g}}_{mk}, \tag{3.37}$$

where $\bar{\mathbf{g}}_{mk}$ has i.i.d. $\mathcal{CN}(0,1)$ elements. Hence,

$$\mathbb{E}\left\{\mathbf{g}_{mk}^H \mathbf{e}_m^y\right\} = 0, \tag{3.38a}$$

$$\mathbb{E}\left\{\mathbf{g}_{mk}^H \mathbf{e}_{mk}^{\hat{g}}\right\} = 0. \tag{3.38b}$$

These result in

$$\mathbb{E}\left\{\mathrm{TQN}_{kk'}\right\} = 0, \tag{3.39a}$$

$$\mathbb{E}\left\{\mathrm{TQN}_{k}^{g}\right\} = 0, \tag{3.39b}$$

$$\mathbb{E}\left\{\mathrm{TQN}_{k}^{y}\right\} = 0, \tag{3.39c}$$

$$\mathbb{E}\left\{\mathrm{TQN}_{k}^{gy}\right\} = 0. \tag{3.39d}$$

Moreover, note that as the term DS_k is a constant, we have

$$\mathbb{E}\left\{\mathrm{DS}_k^H \mathrm{TQN}_k^y\right\} = \mathrm{DS}_k^H \mathbb{E}\left\{\mathrm{TQN}_k^y\right\} = 0, \tag{3.40}$$

and similarly

$$\mathbb{E}\left\{\mathrm{DS}_k^H \mathrm{TQN}_k^g\right\} = 0, \tag{3.41a}$$

$$\mathbb{E}\left\{\mathrm{DS}_k^H \mathrm{TQN}_k^{gy}\right\} = 0, \tag{3.41b}$$

$$\mathbb{E}\left\{\mathrm{DS}_k^H \mathrm{TQN}_{kk'}\right\} = 0. \tag{3.41c}$$

In addition, we have

$$\begin{aligned}
&\mathbb{E}\left\{\mathrm{BU}_k^H \mathrm{TQN}_{kk'}\right\} \\
&= \mathbb{E}\left\{\left(\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k} - \underbrace{\mathbb{E}\left\{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right\}}_{A_1}\right)\sum_{m=1}^{H_M} u_{mk}(\mathbf{e}_{mk}^{\hat{g}})^H \mathbf{g}_{mk'}\sqrt{q_{k'}}\right\} \\
&= \mathbb{E}\left\{\left(\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right)^H \left(\sum_{m=1}^{M} u_{mk}(\mathbf{e}_{mk}^{\hat{g}})^H \mathbf{g}_{mk'}\sqrt{q_{k'}}\right)\right\} \\
&- \mathbb{E}\left\{{A_1}^H\left(\sum_{m=1}^{M} u_{mk}(\mathbf{e}_{mk}^{\hat{g}})^H \mathbf{g}_{mk'}\sqrt{q_{k'}}\right)\right\}.
\end{aligned} \tag{3.42}$$

For the first term of (3.42), we have

$$\begin{aligned}
&\mathbb{E}\left\{\left(\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right)^H \left(\sum_{m=1}^{M} u_{mk}(\mathbf{e}_{mk}^{\hat{g}})^H \mathbf{g}_{mk'}\sqrt{q_{k'}}\right)\right\} \\
&= \sqrt{q_{k'}}\sqrt{q_k}\mathbb{E}\left\{\sum_{m=1}^{M}\sum_{n=1}^{M} u_{mk}u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}(\mathbf{e}_{nk}^{\hat{g}})^H \mathbf{g}_{nk'}\right\} = 0,
\end{aligned} \tag{3.43}$$

where the last equality is due to $\mathbb{E}\left\{\mathbf{g}_{mk}^H \mathbf{e}_m^y\right\} = 0$, $\mathbb{E}\left\{\mathbf{g}_{mk}^H \mathbf{e}_{mk}^{\hat{g}}\right\} = 0$, and $\mathbb{E}\left\{\hat{\mathbf{g}}_{mk}^H \mathbf{e}_{mk}^{\hat{g}}\right\} = 0$. For the second term of (3.42), as A_1 is a constant, and using $\mathbb{E}\left\{\mathbf{g}_{mk}^H \mathbf{e}_{mk}^{\hat{g}}\right\} = 0$, we

have

$$\mathbb{E}\left\{A_1^H\left(\sum_{m=1}^{M} u_{mk}(\mathbf{e}_{mk}^{\hat{g}})^H \mathbf{g}_{mk'}\sqrt{q_{k'}}\right)\right\} = 0. \tag{3.44}$$

Finally, using (3.43) and (3.44), we have $\mathbb{E}\left\{\text{BU}_k^H \text{TQN}_{kk'}\right\} = 0$. Using the same approach, it is easy to show that the terms DS_k, BU_k, $\text{IUI}_{kk'}$, $\text{TQN}_{kk'}$, TQN_k^g, TQN_k^y, and TQN_k^{gy} are mutually uncorrelated, which completes the proof of Proposition 1. ■

Appendix 3.B: Proof of Theorem 2

The desired signal for the user k is given by

$$\begin{aligned}\text{DS}_k &= \sqrt{\rho}\mathbb{E}\left\{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right\} \\ &= N\sqrt{pq_k}\sum_{m=1}^{M} u_{mk}\gamma_{mk}.\end{aligned} \tag{3.45}$$

Hence, $|\text{DS}_k|^2 = \rho q_k\left(N\sum_{m=1}^{M} u_{mk}\gamma_{mk}\right)^2$. Moreover, the term $\mathbb{E}\{|\text{BU}_k|^2\}$ can be obtained as

$$\begin{aligned}\mathbb{E}\left\{|\text{BU}_k|^2\right\} &= \rho\mathbb{E}\left\{\left|\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k} - \rho\mathbb{E}\left\{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right\}\right|^2\right\} \\ &= \rho\sum_{m=1}^{M} q_k u_{mk}^2\left(\mathbb{E}\left\{\left|\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} - \mathbb{E}\left\{\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\right\}\right|^2\right\}\right) \\ &= \rho N\sum_{m=1}^{M} q_k u_{mk}^2\gamma_{mk}\beta_{mk},\end{aligned} \tag{3.46}$$

where the last equality comes from the analysis in [2, Appendix A], and using $\gamma_{mk} = \sqrt{\tau_p p_p}\beta_{mk}c_{mk}$. The term $\mathbb{E}\{|\mathrm{IUI}_{kk'}|^2\}$ is obtained as

$$\begin{aligned}\mathbb{E}\{|\mathrm{IUI}_{kk'}|^2\} &= \rho\mathbb{E}\left\{\left|\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'}\sqrt{q_{k'}}\right|^2\right\} \\ &= \underbrace{\rho\, q_{k'}\mathbb{E}\left\{\left|\sum_{m=1}^{M} c_{mk}u_{mk}\mathbf{g}_{mk'}^H\tilde{\mathbf{w}}_{mk}\right|^2\right\}}_{A} \\ &\quad + \underbrace{\rho\,\tau_p p_p\mathbb{E}\left\{q_{k'}\left|\sum_{m=1}^{M} c_{mk}u_{mk}\left(\sum_{i=1}^{K}\mathbf{g}_{mi}\boldsymbol{\phi}_k^H\boldsymbol{\phi}_i\right)^H\mathbf{g}_{mk'}\right|^2\right\}}_{B}, \end{aligned} \tag{3.47}$$

where the third equality in (3.47) is due to the fact that for two independent RVs X and Y and $\mathbb{E}\{X\} = 0$, we have $\mathbb{E}\{|X+Y|^2\} = \mathbb{E}\{|X|^2\} + \mathbb{E}\{|Y|^2\}$ [2]. Since $\tilde{\mathbf{w}}_{mk} = \boldsymbol{\phi}_k^H\mathbf{W}_{\mathbf{p},\mathbf{m}}$ is independent from the term $g_{mk'}$ similar to [2], *Appendix A*, the term A in (3.47) immediately is given by $A = Nq_{k'}\sum_{m=1}^{M} c_{mk}^2 u_{mk}^2\beta_{mk'}$. The term B in (3.47) can be obtained as

$$\begin{aligned} B &= \underbrace{\tau_p p_p q_{k'}\mathbb{E}\left\{\left|\sum_{m=1}^{M} c_{mk}u_{mk}||\mathbf{g}_{mk'}||^2\boldsymbol{\phi}_k^H\boldsymbol{\phi}_{k'}\right|^2\right\}}_{C} \\ &\quad + \underbrace{\tau_p p_p q_{k'}\mathbb{E}\left\{\left|\sum_{m=1}^{M} c_{mk}u_{mk}\left(\sum_{i\neq k'}^{K}\mathbf{g}_{mi}\boldsymbol{\phi}_k^H\boldsymbol{\phi}_i\right)^H\mathbf{g}_{mk'}\right|^2\right\}}_{D}. \end{aligned} \tag{3.48}$$

The first term in (3.48) is given by

$$
\begin{aligned}
C &= \tau_p p_p q_{k'} \mathbb{E}\left\{ \left| \sum_{m=1}^{M} c_{mk} u_{mk} ||\mathbf{g}_{mk'}||^2 \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \right\} \\
&= \tau_p p_p q_{k'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \mathbb{E}\left\{ \sum_{m=1}^{M} c_{mk} u_{mk} ||\mathbf{g}_{mk'}||^4 \right\} + \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \\
&\quad \tau_p p_p q_{k'} \mathbb{E}\left\{ \sum_{m=1}^{M} \sum_{n \neq m}^{M} c_{mk} c_{nk} u_{mk} u_{nk} ||\mathbf{g}_{mk'}||^2 ||\mathbf{g}_{nk'}||^2 \right\} \\
&= N \tau_p p_p q_{k'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \sum_{m=1}^{M} c_{mk}^2 u_{mk}^2 \beta_{mk'}^2 \\
&+ N^2 q_{k'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \left(\sum_{m=1}^{M} u_{mk} \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2 ,
\end{aligned} \tag{3.49}
$$

where the last equality is derived based on the fact: $\gamma_{mk} = \sqrt{\tau_p p_p} \beta_{mk} c_{mk}$. The second term in (3.48) can be obtained as

$$
\begin{aligned}
D &= \tau_p p_p q_{k'} \mathbb{E}\left\{ \left| \sum_{m=1}^{M} c_{mk} u_{mk} \Big(\sum_{i \neq k'}^{K} \mathbf{g}_{mi} \boldsymbol{\phi}_k^H \boldsymbol{\phi}_i \Big)^H \mathbf{g}_{mk'} \right|^2 \right\} \\
&= N \tau_p p_p q_{k'} \left[\sum_{m=1}^{M} c_{mk}^2 u_{mk}^2 \beta_{mk'} \left(\sum_{i \neq k'}^{K} \beta_{mi} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_i \right|^2 \right) \right] \\
&= N \tau_p p_p q_{k'} \left[\sum_{m=1}^{M} c_{mk}^2 u_{mk}^2 \beta_{mk'} \left(\sum_{i=1}^{K} \beta_{mi} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_i \right|^2 - \beta_{mk'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \right) \right] \\
&= N \sqrt{\tau_p p_p} q_{k'} \sum_{m=1}^{M} u_{mk}^2 c_{mk} \beta_{mk'} \beta_{mk} \\
&- N q_{k'} \sum_{m=1}^{M} u_{mk}^2 c_{mk}^2 \beta_{mk'} \\
&- N \tau_p p_p q_{k'} \sum_{m=1}^{M} u_{mk}^2 c_{mk}^2 \beta_{mk'}^2 \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 .
\end{aligned} \tag{3.50}
$$

Finally by substituting (3.49) and (3.50) into (3.48), and substituting (3.48) into (3.47),

we obtain

$$\mathbb{E}\{|\mathrm{IUI}_{kk'}|^2\} = N\rho q_{k'} \left(\sum_{m=1}^{M} u_{mk}^2 \beta_{mk'} \gamma_{mk} \right) + N^2 \rho q_{k'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \left(\sum_{m=1}^{M} u_{mk} \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2 . \tag{3.51}$$

The total noise for the user k is given by

$$\mathbb{E}\left\{|\mathrm{TN}_k|^2\right\} = \mathbb{E}\left\{ \left| \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m \right|^2 \right\} = N \sum_{m=1}^{M} u_{mk}^2 \gamma_{mk}, \tag{3.52}$$

where the last equality is due to the fact that the terms $\hat{\mathbf{g}}_{mk}$ and $\mathbf{n}_m$ are uncorrelated. The power of the quantization distortion for user k is given by

$$\mathbb{E}\left\{|\mathrm{TQD}_k|^2\right\} = \mathbb{E}\left\{ \left| \sum_{m=1}^{M} u_{mk} e_{mk}^z \right|^2 \right\} = \sum_{m=1}^{M} u_{mk}^2 \mathbb{E}\left\{|e_{mk}^z|^2\right\}, \tag{3.53}$$

where in the last equality, we used the fact that using the Bussgang decomposition, with the correlated inputs, the covariance matrix of the quantization distortion is approximated with a diagonal matrix [53–57]. Finally, the power of the quantization distortion is obtained as the following:

$$\mathbb{E}\left\{|e_{mk}^z|^2\right\} = \mathbb{E}\left\{|\tilde{e}_{mk}^z|^2\right\} \sigma_{z_{mk}}^2 = \sigma_{\tilde{e}}^2 \sigma_{z_{mk}}^2, \tag{3.54}$$

where we used the fact that all APs use the same number of bits to quantize the weighted

signal z_{mk} in (3.12). Next, the term $\sigma^2_{z_{mk}}$ is obtained as

$$\begin{aligned}
\sigma^2_{z_{mk}} &= \mathbb{E}\left\{ \left(\hat{\mathbf{g}}^H_{mk}\mathbf{y}_m\right)^H \left(\hat{\mathbf{g}}^H_{mk}\mathbf{y}_m\right) \right\} \\
&= \mathbb{E}\Bigg\{ \left(\sqrt{\rho}\sum_{k'=1}^{K} u_{mk}\hat{\mathbf{g}}^H_{mk}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} + u_{mk}\hat{\mathbf{g}}^H_{mk}\mathbf{n}_m \right)^H \\
&\quad \left(\sqrt{\rho}\sum_{k'=1}^{K} u_{mk}\hat{\mathbf{g}}^H_{mk}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} + u_{mk}\hat{\mathbf{g}}^H_{mk}\mathbf{n}_m \right) \Bigg\} \\
&= \rho\mathbb{E}\left\{ \left| \sum_{k'=1}^{K} u_{mk}\hat{\mathbf{g}}^H_{mk}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} \right|^2 \right\} \\
&\quad + \mathbb{E}\left\{ \left| u_{mk}\hat{\mathbf{g}}^H_{mk}\mathbf{n}_m \right|^2 \right\},
\end{aligned}$$

where the last inequality in (3.47) is due to the fact that for two independent RVs X and Y and $\mathbb{E}\{X\} = 0$, we have $\mathbb{E}\{|X+Y|^2\} = \mathbb{E}\{|X|^2\} + \mathbb{E}\{|Y|^2\}$. For the second term of (4.69), we have $\mathbb{E}\left\{ \left| \hat{\mathbf{g}}^H_{mk}\mathbf{n}_m \right|^2 \right\} = N\gamma_{mk}$. The first term in (4.69) can be obtained as

$$\begin{aligned}
\mathbb{E}\left\{ \left| \sum_{k'=1}^{K} u_{mk}\hat{\mathbf{g}}^H_{mk}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} \right|^2 \right\} &= \mathbb{E}\left\{ \left| \sum_{k'=1}^{K} u_{mk}(\mathbf{g}_{mk} - \boldsymbol{\epsilon}_{mk})^H\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} \right|^2 \right\} \\
&= \underbrace{\mathbb{E}\left\{ \left| \sum_{k'=1}^{K} u_{mk}\mathbf{g}^H_{mk}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} \right|^2 \right\}}_{\text{I}} \\
&\quad + \underbrace{\mathbb{E}\left\{ \left| \sum_{k'=1}^{K} u_{mk}\boldsymbol{\epsilon}^H_{mk}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} \right|^2 \right\}}_{\text{II}}, \qquad (3.55)
\end{aligned}$$

where each element of $\boldsymbol{\epsilon}$ is given by $\epsilon_{mk} = \mathcal{CN}(0, \beta_{mk} - \gamma_{mk})$. The terms I and II in (3.55) are given as following:

$$\text{I} = N\beta_{mk}\sum_{k'=1}^{K} q_{k'}\beta_{mk'}, \qquad (3.56)$$

and

$$\text{II} = N(\beta_{mk} - \gamma_{mk})\sum_{k'=1}^{K} q_{k'}\beta_{mk'}. \qquad (3.57)$$

Finally, we have

$$\mathbb{E}\left\{|\mathrm{TQD}_k|^2\right\} = N\sigma_{\tilde{e}}^2 \sum_{m=1}^{M} u_{mk}^2 \left[\sqrt{\rho}\left(2\beta_{mk} - \gamma_{mk}\right) \sum_{k'=1}^{K} q_{k'}\beta_{mk'} + \gamma_{mk}\right], \quad (3.58)$$

By substituting (3.45), (3.46), (3.51) and (3.52) into (3.13), the corersponding SINR of the kth user is obtained by (3.17), which completes the proof of Theorem 2. ■

Appendix 3.C: Proof of Proposition 2

The standard form of GP is defined as follows [50]:

$$P_{12}: \quad \min \quad f_0(\mathbf{x}), \quad (3.59a)$$

$$\text{subject to} \quad f_i(\mathbf{x}) \leq 1, \;\; i = 1, \cdots, m, g_i(\mathbf{x}) = 1, \;\; i = 1, \cdots, p, \quad (3.59b)$$

where f_0 and f_i are posynomial and g_i are monomial functions. Moreover, $\mathbf{x} = \{x_1, \cdots, x_n\}$ represent the optimization variables. The SINR constraint in (3.59) is not a posynomial function in this form, however it can be rewritten as the following posynomial function:

$$\frac{\mathbf{u}_k^H \left(N^2 \sum\limits_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \boldsymbol{\Lambda}_{kk'} \boldsymbol{\Lambda}_{kk'}^H + N \sum_{k'=1}^{K} q_{k'} \boldsymbol{\Upsilon}_{kk'} + \frac{N}{\rho}\mathbf{R}_k\right) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 q_k \boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_k^H\right) \mathbf{u}_k} < \frac{1}{t}, \forall k. \quad (3.60)$$

By applying a simple transformation, (3.60) is equivalent to the following inequality:

$$q_k^{-1} \left(\sum_{k' \neq k}^{K} a_{kk'} q_{k'} + \sum_{k'=1}^{K} b_{kk'} q_{k'} + c_k\right) < \frac{1}{t}, \quad (3.61)$$

where

$$a_{kk'} = \frac{\mathbf{u}_k^H \left(\left|\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}\right|^2 \boldsymbol{\Lambda}_{kk'} \boldsymbol{\Lambda}_{kk'}^H\right) \mathbf{u}_k}{\mathbf{u}_k^H \left(\boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_k^H\right) \mathbf{u}_k}, \quad (3.62a)$$

$$b_{kk'} = \frac{\mathbf{u}_k^H \boldsymbol{\Upsilon}_{kk'} \mathbf{u}_k}{\mathbf{u}_k^H \left(\boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_k^H\right) \mathbf{u}_k}, \quad (3.62b)$$

$$c_k = \frac{\mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k}{\rho \mathbf{u}_k^H \left(\boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_k^H\right) \mathbf{u}_k}. \quad (3.62c)$$

The transformation in (3.61) shows that the left-hand side of (3.60) is a posynomial function. Hence, the power allocation Problem P_4 is a GP (convex problem), where the objective function and constraints are monomial and posynomial, respectively, which completes the proof of Proposition 2. ■

Appendix 3.D: Proof of Lemma 2

This lemma is proven by exploiting the unique optimal solution of the uplink max-min SINR problem with total power limitation through an eigensystem [43]. This problem is iteratively solved and the optimal receiver filter coefficients $\check{\mathbf{U}}$ are determined by solving Problem P_3 of submitted manuscript. Next, we scale the power allocation at each user such that the per-user power constraints are satisfied. Let us consider the following optimization problem for a given receiver filter coefficients $\check{\mathbf{U}}$:

$$P_{11}: C^{\text{UP}}\left(\check{\mathbf{U}}, P_{\text{tot}}\right) = \max_{q_k} \quad \min_{k=1,\cdots,K} \quad \text{SINR}_k^{\text{UP}}\left(\check{\mathbf{U}}, \mathbf{q}\right) \tag{3.63a}$$

$$\text{subject to} \quad \sum_{k=1}^{K} q_k \leq P_{\text{tot}}. \tag{3.63b}$$

3.D.1 Optimal Solution of Problem P_{11} Exploiting an Eigensystem

The optimal solution of Problem P_{11} can be determined by finding the unique eigenvector associated with unique positive eigenvalue of an eigensystem and the power allocation $\check{\mathbf{q}}$ that satisfies the following condition [43]:

$$\sum_{k=1}^{K} \check{q}_k = P_{\text{tot}}. \tag{3.64}$$

The SINRs of all users can be collectively written as

$$\check{\mathbf{q}} \frac{1}{C_k^{\text{UP}}\left(\check{\mathbf{U}}, P_{\text{tot}}\right)} = \mathbf{D}\boldsymbol{\Psi}\left(\check{\mathbf{U}}\right)\check{\mathbf{q}} + \mathbf{D}\boldsymbol{\sigma}\left(\check{\mathbf{U}}\right), \tag{3.65}$$

where $\boldsymbol{\sigma}\left(\check{\mathbf{U}}\right) \in \mathbb{C}^{K\times 1}$, $\sigma_k\left(\mathbf{u}_k\right) = \dfrac{N}{\rho}\left(\dfrac{\sigma_{\tilde{e}}^2}{\tilde{a}^2}+1\right)\sum\limits_{m=1}^{M}\check{u}_{mk}\gamma_{mk}$ and $\mathbf{D}$ and $\boldsymbol{\Psi}\left(\check{\mathbf{U}}\right)$ are defined as

$$\mathbf{D} = \text{diag}\left[\frac{1}{\check{\mathbf{u}}_1^H\check{\mathbf{D}}_1\check{\mathbf{u}}_1},\cdots,\frac{1}{\check{\mathbf{u}}_K^H\check{\mathbf{D}}_K\check{\mathbf{u}}_K}\right], \left[\boldsymbol{\Psi}\left(\check{\mathbf{U}}\right)\right]_{kk'} = \begin{cases}\check{\mathbf{u}}_k^H\check{\check{\mathbf{R}}}_{kk}\check{\mathbf{u}}_k, & k=k',\\ \check{\mathbf{u}}_k^H\check{\mathbf{R}}_{kk'}\check{\mathbf{u}}_k+\check{\mathbf{u}}_k^H\check{\check{\mathbf{R}}}_{kk'}\check{\mathbf{u}}_k, & k\neq k',\end{cases} \tag{3.66}$$

where $\check{\mathbf{D}}_k$, $\check{\mathbf{R}}_{kk'}$ and $\check{\check{\mathbf{R}}}_{kk'}$ are defined as

$$\text{SINR}_k^{\text{UP}} = \frac{q_k\mathbf{u}_k^H\Big(\overbrace{N^2\boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^H}^{\check{\mathbf{D}}_k}\Big)\mathbf{u}_k}{\mathbf{u}_k^H\Big(\sum_{k'\neq k}^{K}q_{k'}\underbrace{N^2|\boldsymbol{\phi}_k^H\boldsymbol{\phi}_{k'}|^2\boldsymbol{\Lambda}_{kk'}\boldsymbol{\Lambda}_{kk'}^H}_{\check{\mathbf{R}}_{kk'}}+\sum_{k'=1}^{K}q_{k'}\underbrace{N\boldsymbol{\Upsilon}_{kk'}}_{\check{\check{\mathbf{R}}}_{kk'}}+\frac{N}{\rho}\mathbf{R}_k\Big)\mathbf{u}_k}. \tag{3.67}$$

Having both sides of (3.65) multiplied by $\mathbf{1}^T = [1,\cdots,1]$, we obtain

$$\frac{1}{C_k^{\text{UP}}\left(\check{\mathbf{U}},P_{\text{tot}}\right)} = \frac{1}{P_{\text{tot}}}\mathbf{1}^T\check{\mathbf{D}}\boldsymbol{\Psi}\left(\check{\mathbf{U}}\right)\check{\mathbf{q}}+\frac{1}{P_{\text{tot}}}\mathbf{1}^T\mathbf{D}\boldsymbol{\sigma}\left(\check{\mathbf{U}}\right), \tag{3.68}$$

which can be combined with (3.65) to define the following eigensystem:

$$\boldsymbol{\Lambda}\left(\check{\mathbf{U}},P_{\text{tot}}\right)\check{\mathbf{q}}_{\text{ext}} = \frac{1}{C_k^{\text{UP}}\left(\check{\mathbf{U}},P_{\text{tot}}\right)}\check{\mathbf{q}}_{\text{ext}},\ \boldsymbol{\Lambda}\left(\check{\mathbf{U}},P_{\text{tot}}\right) = \begin{bmatrix}\mathbf{D}\boldsymbol{\Psi}^T\left(\check{\mathbf{U}}\right) & \mathbf{D}\boldsymbol{\sigma}\left(\check{\mathbf{U}}\right)\\ \dfrac{1}{P_{\text{tot}}}\mathbf{1}^T\mathbf{D}\boldsymbol{\Psi}^T\left(\check{\mathbf{U}}\right) & \dfrac{1}{P_{\text{tot}}}\mathbf{1}^T\mathbf{D}\boldsymbol{\sigma}\left(\check{\mathbf{U}}\right)\end{bmatrix}. \tag{3.69}$$

The optimal power allocation $\check{\mathbf{q}}$ is obtained by determining the eigenvector corresponding to the maximum eigenvalue of $\boldsymbol{\Lambda}\left(\check{\mathbf{U}},P_{\text{tot}}\right)$ and scaling the last element to one as follows:

$$\check{\mathbf{q}}_{\text{ext}} = \begin{bmatrix}\check{\mathbf{q}}\\ 1\end{bmatrix}, \tag{3.70a}$$

$$\boldsymbol{\Lambda}\left(\check{\mathbf{U}},P_{\text{tot}}\right)\check{\mathbf{q}}_{\text{ext}} = \lambda_{\max}\left(\boldsymbol{\Lambda}\left(\check{\mathbf{U}},P_{\text{tot}}\right)\right)\check{\mathbf{q}}_{\text{ext}}. \tag{3.70b}$$

Note that to find the optimal power allocation $\check{\mathbf{q}}$, the elements of eigenvector of $\boldsymbol{\Lambda}\left(\check{\mathbf{U}},P_{\text{tot}}\right)$ should be scaled such that the last element is one to satisfy the total power constraint. In particular, the element of the eigenvector that needs to be scaled depends on the type of power constraint in the problem. For example, to meet the total power constraint, the last element is scaled to one. Similarly, to meet the other types of power

constraints (for example, per-user power constraint), the components of this eigenvector can be scaled by any positive value to satisfy a given condition as follows:

$$\mathbf{\Lambda}\left(\check{\mathbf{U}}, P_{\text{tot}}\right) \delta_{\text{cons}} \check{\mathbf{q}}_{\text{ext}} = \lambda_{\max}\left(\mathbf{\Lambda}\left(\check{\mathbf{U}}, P_{\text{tot}}\right)\right) \delta_{\text{cons}} \check{\mathbf{q}}_{\text{ext}}, \tag{3.71}$$

where δ_{cons} is a positive constant. This is the key fact that exploited to show that both Problems P_2 and P_6 provide the same optimal solution.

In the next subsection, we propose to scale the elements of $\check{\mathbf{q}}$ such that the per-user power constraints are satisfied at each iteration.

3.D.2 Modified Eigensystem to Satisfy the Per-User Power Constraints

vspace.5cm We further scale the power allocation $\check{\mathbf{q}}$ to satisfy the per-user power constraints which is performed through carrying out the following two steps:

$$\bar{\mathbf{q}} = \begin{bmatrix} \dfrac{\check{q}_1}{p_{\max}^{(1)}} \\ \vdots \\ \dfrac{\check{q}_K}{p_{\max}^{(K)}}. \end{bmatrix}. \tag{3.72}$$

Next we find the maximum value among the elements of $\bar{\mathbf{q}}$, i.e., $\max(\bar{\mathbf{q}})$, and divide all elements of $\bar{\mathbf{q}}$ by it. Hence the power allocation $\check{\bar{\mathbf{q}}}$ is defined as follows:

$$\check{\bar{\mathbf{q}}} = \begin{bmatrix} \dfrac{\check{q}_1}{\max(\bar{\mathbf{q}})} \\ \vdots \\ \dfrac{\check{q}_K}{\max(\bar{\mathbf{q}})}, \end{bmatrix}. \tag{3.73}$$

In the next iteration, the same max-min problem is solved with a new total power constraint obtained by summing up the allocated power to all users in the previous

iteration:

$$L_1: \quad C^{\text{UP}}\left(\check{\mathbf{U}}, P_{\text{tot}}^{\text{new}}\right) = \max_{q_k} \quad \min_{k=1,\cdots,K} \quad \text{SINR}_k^{\text{UP}}\left(\check{\mathbf{U}}, \mathbf{q}\right), \tag{3.74a}$$

$$\text{subject to} \quad \sum_{k=1}^{K} q_k \leq P_{\text{tot}}^{\text{new}}, \tag{3.74b}$$

$$\text{where} \quad P_{\text{tot}}^{\text{new}} = \sum_{k=1}^{K} \check{\check{q}}_k. \tag{3.74c}$$

At the convergence of the algorithm, the per-user power constraints are satisfied with achieving the same uplink SINR for each user. Interestingly, if this max-min problem is solved with the corresponding total power constraint, then it will converge to the same optimal solution of max-min problem with per-user power constraints. This is due to the property that the eigensystem exploited to obtain the power allocation in (3.69) has a unique positive eigenvalue and a corresponding unique eigenvector. Furthermore, in both Problems P_2 and P_6, different elements of the same eigenvector are scaled to meet the corresponding constraints on the power allocation. In other words, the last element is scaled to meet the total power constraint in P_6 whereas the element with the highest ratio as in (3.71) is scaled to meet the per-user power constraint. As the equivalent total power P_{tot}^{c} for Problem P_6 chosen from the solution of the original P_2, both of them will converge to the same solution whose optimality is proven later by considering an equivalent problem related to the virtual downlink SINR. Therefore, Problems P_2 and P_6 of the revised manuscript are equivalent and have the same optimal solution. The steps of proof involved in solving both Problems P_6 and P_2 are provided in Algorithm 3. Step 5 in Algorithm 3 is the key step in which the different elements of the eigen vector (i.e., power allocations) are scaled to meet the constraints on transmit power. This completes the proof of Lemma 2. ■

Appendix 3.E: Proof of Theorem 4

To achieve the same SINR tuples in both the uplink and the downlink, we need:

$$\text{SINR}_k^{\text{DL}}\left(\mathbf{U}, \mathbf{p}\right) = \text{SINR}_k^{\text{UP}}\left(\mathbf{U}, \mathbf{q}\right), \forall k. \tag{3.75}$$

Algorithm 3 Proposed algorithm to solve Problems P_2 and P_6

1. Initialize $\check{P}_{\text{tot}}^{\text{new},(0)} = P_{\text{tot}}^{c}$ for P_2, and $\check{\check{P}}_{\text{tot}}^{\text{new},(0)} = P_{\text{tot}}^{c}$ for P_6, $\check{\check{\mathbf{q}}}^{(0)} = \check{\mathbf{q}}^{(0)} = \left[\frac{P_{\text{tot}}}{K}^{c}, \cdots, \frac{P_{\text{tot}}}{K}^{c}\right], i = 1$

2. Repeat steps 3-5 until $\dfrac{\text{SINR}_k^{\text{UP},(i+1)} - \text{SINR}_k^{\text{UP},(i)}}{\text{SINR}_k^{\text{UP},(i)}} \leq \epsilon, \forall k$

3. Determine the optimal receiver coefficients $\check{\mathbf{U}}^{(i)}$ for P_2, and $\check{\check{\mathbf{U}}}^{(i)}$ for P_6, by solving the generalized eigenvalue Problem P_3, for a given $\check{\mathbf{q}}^{(i-1)}$ for P_2 and $\check{\check{\mathbf{q}}}^{(i-1)}$ for P_6

4. Compute $\check{\mathbf{q}}^{(i)}$ for P_2, and $\check{\check{\mathbf{q}}}^{(i)}$ for P_6, through solving Problem P_{11} in (3.63) for a given $\check{\mathbf{U}}^{(i)}$ for P_2 and $\check{\check{\mathbf{U}}}^{(i)}$ for P_6

5. Scale the last element to 1, i.e., $[\check{\mathbf{q}}_{\text{ext}}]_{K+1} = 1$ as given in (3.70) for P_2, and Scale the elements of $\check{\check{\mathbf{q}}}$ based on (3.72) and (3.73) for P_6

6. $\check{P}_{\text{tot}}^{\text{new},(i)} = \sum_{k=1}^{K} \check{q}_k^{(i)}$ for P_2, and $\check{\check{P}}_{\text{tot}}^{\text{new},(i)} = \sum_{k=1}^{K} \check{\check{q}}_k^{(i)}$ for P_6, and $i = i + 1$

By substituting uplink and downlink SINRs, in (3.25) and (3.24), respectively, in equation (3.75) and summing all equations by both sides, we have

$$p_1 N \sum_{m=1}^{M} \left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2} + 1 \right) u_{m1}^2 \gamma_{m1} + \cdots + p_K N \sum_{m=1}^{M} \left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2} + 1 \right) u_{mK}^2 \gamma_{mK} = \sum_{k=1}^{K} q_k. \quad (3.76)$$

Therefore, this condition between the total transmit power on the uplink and the equivalent total transmit power on the downlink should be satisfied to realize the same SINRs for all set of users, which completes the proof of Theorem 4. ■

Chapter 4

Energy Efficiency of the Cell-Free Massive MIMO Uplink

Contents

4.1 Introduction

For a given fronthaul capacity, we show that the total power consumption in the cell-free massive MIMO system depends on the length of uplink pilot vectors, channel coherence time and the total number of quantization bits. The uplink energy efficiency of the

cell-free massive MIMO system is investigated in this chapter. In particular, optimal power allocation strategies which maximize the uplink energy efficiency are investigated for a system in which the quantized version of the weighted signals obtained from MRC weighting at APs are available at the CPU. The contributions of this chapter are summarized as follows:

1. An expression for uplink energy efficiency is derived based on channel statistics and taking into account the effects of channel estimation errors, the effect of pilot contamination, and quantization distortion.

2. A novel approach to solve the non-convex energy efficiency maximization problem is proposed, where we propose to decompose the original problem into two sub-problems and an iterative algorithm is developed to determine the optimal solution. A successive convex approximation (SCA) is used to efficiently solve the power allocation problem. Next, a heuristic sub-optimal energy efficiency maximization problem is proposed where the original optimization problem is transformed into a standard GP.

3. The convergence and complexity analysis of the proposed schemes are presented.

4. Numerical results demonstrate that the proposed scheme substantially outperforms the case with equal power allocation. Finally, numerical results demonstrate that although the proposed sub-optimal scheme has a lower complexity, it provides a performance fairly close to that of the SCA scheme.

4.2 System Model

In this chapter, the system model is the same as the one used in Chapter 3. However, as a reminder, we briefly review the system model. We consider uplink transmission in a cell-free massive MIMO system with M APs and K single-antenna users randomly distributed in a large area, as shown in Fig. 4.1. Moreover, we assume each AP has N antennas. The channel coefficient vector between the kth user and the mth AP, $\mathbf{g}_{mk} \in \mathbb{C}^{N\times 1}$, is modeled as $\mathbf{g}_{mk} = \sqrt{\beta_{mk}}\mathbf{h}_{mk}$, where β_{mk} denotes the large-scale fading and $\mathbf{h}_{mk} \sim \mathcal{CN}(0, \mathbf{I_N})$

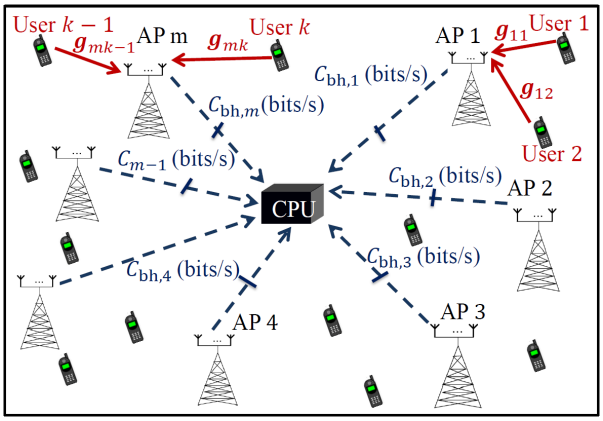

Figure 4.1: The uplink of a cell-free massive MIMO system with K single-antenna users and M APs. Each AP is equipped with N antennas. The solid lines denote the uplink channels and the dashed lines present the limited capacity fronthaul links between the APs and the CPU.

represents the small-scale fading [2]. Note that the uplink channel estimation is presented in Section 2.7.5.

4.2.1 Uplink Transmission

In the uplink data transmission, where all users send their signals to the AP, the transmitted signal from the kth user is represented by $x_k = \sqrt{q_k}s_k$, where s_k ($\mathbb{E}\{|s_k|^2\} = 1$) and q_k denotes the transmitted symbol and the transmit power coefficient from the kth user, respectively. The $N \times 1$ received signal at the mth AP from all users is given by

$$\mathbf{y}_m = \sqrt{\rho}\sum_{k=1}^{K}\mathbf{g}_{mk}\sqrt{q_k}s_k + \mathbf{n}_m, \tag{4.1}$$

where $\mathbf{n}_m \in \mathbb{C}^{N\times 1}$ is the noise at the mth AP and ρ is the normalized uplink SNR. We assume that elements of $\mathbf{n}_m$ are i.i.d. $\mathcal{CN}(0,1)$ RVs.

4.2.2 Received Signal

The mth AP quantizes the terms $z_{m,k} = \hat{\mathbf{g}}_{mk}^H\mathbf{y}_m$, $\forall k$, and forwards the quantized signals in each symbol duration to the CPU as $z_{mk} = \hat{\mathbf{g}}_{mk}^H\mathbf{y}_m = r_{mk} + js_{mk}$, $\forall k$, where r_{mk} and

s_{mk} represent the real and imaginary parts of z_{mk}. An ADC quantizes the real and imaginary parts of $z_{m,k}$ with α bits each, which introduces quantization distortions to the received signals [48]. In order to improve the performance, the forwarded signal is further multiplied by the receiver filter coefficients at the CPU. Finally, using the Bussgang decomposition and the receiver filter coefficients $u_{mk}, \forall m, k$ at the CPU, the aggregate received signal at the CPU can be written as

$$\begin{aligned} r_k &= \sum_{m=1}^{M} u_{mk} \mathcal{Q}\left(\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m\right) \\ &= \sum_{m=1}^{M} u_{mk} \left(a\, \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m + n_{d,mk}\right) \\ &= \sum_{m=1}^{M} u_{mk} \left(\tilde{a}\, \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m + \underbrace{\sigma_{\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m} \tilde{n}_{d,mk}}_{n_{d,mk}} \right), \end{aligned} \tag{4.2}$$

where the terms a and $\tilde{a}$ are defined in Section 2.9. Collecting all the receiver filter coefficients $u_{mk}, \forall m$, corresponding to the *k*th user, we define $\mathbf{u}_k = [u_{1k}, u_{2k}, \cdots, u_{Mk}]^T$, without loss of generality, it is assumed that $||\mathbf{u}_k|| = 1$.

4.3 Performance Analysis

In this section, we derive the spectral efficiency for the considered system model by following a similar approach in [2]. Note that the main difference between the proposed approach and the scheme in [2] is the new set of receiver coefficients which are introduced at the CPU to improve the spectral efficiency. The benefits of the proposed approach in terms of the spectral efficiency is demonstrated through numerical results. In deriving the spectral efficiency of each user, it is assumed that the CPU exploits only the knowledge of channel statistics between the users and APs in detecting data from the

received signal in (4.2). The aggregated received signal in (4.2) can be written as

$$
\begin{aligned}
r_k = \tilde{a}\, \underbrace{\sqrt{\rho}\mathbb{E}\left\{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right\}}_{\mathrm{DS}_k} s_k \\
+ \tilde{a}\, \underbrace{\sqrt{\rho}\left(\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k} - \mathbb{E}\left\{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right\}\right)}_{\mathrm{BU}_k} s_k \\
+ \tilde{a} \sum_{k'\neq k}^{K} \underbrace{\sqrt{\rho}\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'}\sqrt{q_{k'}}}_{\mathrm{IUI}_{kk'}} s_{k'} \\
+ \tilde{a}\underbrace{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{n}_m}_{\mathrm{TN}_k} + \underbrace{\sum_{m=1}^{M} u_{mk} n_{d,mk}}_{\mathrm{TQD}_k},
\end{aligned}
\tag{4.3}
$$

where DS_k, BU_k and IUI_k denote the desired signal (DS), beamforming uncertainty (BU) for the kth user, and the inter-user-interference (IUI) caused by the k'th user, respectively. In addition, TN_k accounts for the total noise (TN) following the MRC detection, and finally TQD_k refers to the total quantization distortion (TQD) at the kth user. The elements of quantization distortion are i.i.d. RVs [48]. Moreover, if the probability density function of the input of the quantizer is even and we use a symmetrical quantizer, the quantization noise has zero mean [56, 58, 59]. In addition, note that using Bussgang decomposition the elements of the quantization distortion are uncorrelated with the input of the quantizer [36], i.e.,

$$
\mathbb{E}\left\{\left(\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m\right)^H n_{d,mk}\right\} = 0. \tag{4.4}
$$

Exploiting (4.4), we have

$$
\mathbb{E}\left\{\left(\mathrm{DS}_k.s_k + \mathrm{BU}_k.s_k\right) \times \mathrm{TQD}_k\right\} = 0. \tag{4.5}
$$

Hence, exploiting the analysis in [2], it can be shown that terms $\mathrm{DS}_k.s_k$, $\mathrm{BU}_k.s_k$, $\mathrm{IUI}_{kk'}.s_{k'}$, TN_k and TQD_k are mutually uncorrelated. Using the analysis in Section 2.10,

the corresponding spectral efficiency (in bit/s/Hz) is obtained as follows:

$$S_k = \left(1 - \frac{\tau_p}{\tau_c}\right) \log_2\left(1 + \text{SINR}_k\right) = \left(1 - \frac{\tau_p}{\tau_c}\right) \log_2\left(1 + \frac{|\text{DS}_k|^2}{\mathbb{E}\left\{|\text{BU}_k|^2\right\} + \sum_{k' \neq k}^{K} \mathbb{E}\{|\text{IUI}_{kk'}|^2\} + \mathbb{E}\{|\text{TN}_k|^2\} + \frac{1}{\tilde{a}^2}\mathbb{E}\{|\text{TQD}_k|^2\}}\right). \quad (4.6)$$

where SINR_k refers to the signal-to-interference-plus-noise ratio (SINR) of the *k*th user and its closed-form expression is provided in the following theorem.

Theorem 5. *By employing MRC detection at the APs, the achievable uplink SINR of the kth user in the cell-free massive MIMO system with K randomly distributed single-antenna users and M APs, each is equipped with N antennas, is given by*

$$\textit{SINR}_k = \frac{N^2 \mathbf{u}_k^H \left(q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H\right) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 \sum\limits_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Delta}_{kk'} \mathbf{\Delta}_{kk'}^H + N^2 \sum\limits_{k'=1}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{k'} + N \sum\limits_{k'=1}^{K} q_{k'} \mathbf{D}_{kk'} + \frac{N}{\rho} \mathbf{R}_k\right) \mathbf{u}_k}, \quad (4.7)$$

where

$$\mathbf{\Gamma}_k = [\gamma_{1k}, \gamma_{2k}, \cdots, \gamma_{Mk}]^T, \quad (4.8a)$$

$$\mathbf{\Delta}_{kk'} = \left[\frac{\gamma_{1k}\beta_{1k'}}{\beta_{1k}}, \frac{\gamma_{2k}\beta_{2k'}}{\beta_{2k}}, \cdots, \frac{\gamma_{Mk}\beta_{Mk'}}{\beta_{Mk}}\right]^T, \quad (4.8b)$$

$$\mathbf{\Lambda}_{k'} = \frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2}\, \textit{diag}\left[\gamma_{1k'}^2, \cdots, \gamma_{Mk'}^2\right], \quad (4.8c)$$

$$\mathbf{D}_{kk'} = \textit{diag}\left(\frac{\sigma_{\tilde{e}}^2}{\tilde{a}^2} + 1\right)\left[\beta_{1k'}\gamma_{1k}, \cdots, \beta_{Mk'}\gamma_{Mk}\right], \quad (4.8d)$$

and where $\gamma_{mk} = \sqrt{\tau_p p_p}\beta_{mk} c_{mk}$.

Proof: Please refer to Appendix 4.A. ■

Finally, the sum spectral efficiency is given by

$$S\left(q_k, \mathbf{u}_k, \alpha\right) = \sum_{k=1}^{K} S_k\left(q_k, \mathbf{u}_k, \alpha\right). \quad (4.9)$$

4.4 Total Energy Efficiency Model

4.4.1 Power Consumption Model

The total power consumption can be defined as follows [60]:

$$P_{\text{total}} = P_{\text{TX}} + P_{\text{CP}}, \tag{4.10}$$

where P_{TX} is the uplink power amplifiers (PAs) due to transmit power at the users and PA dissipation [60], and P_{CP} refers to the circuit power (CP) consumption. The power consumption P_{TX} is given by [2]

$$P_{\text{TX}} = \frac{1}{\zeta}\rho p_n \sum_{k=1}^{K} q_k, \tag{4.11}$$

where ζ is the PA efficiency at each user, and p_n is the noise power. The power consumption P_{CP} is obtained as [2]

$$P_{\text{CP}} = MP_{\text{fix}} + KP_{\text{U}} + \sum_{m=1}^{M} P_{\text{fh},m}, \tag{4.12}$$

where P_{fix} is a fixed power consumption (including control signals and fronthaul) at each AP, P_U denotes the required power to run circuit components at each user and finally, fronthaul power consumption from the mth AP to the CPU is obtained as follows [52,61–63]:

$$P_{\text{fh},m} = P_{\text{FT}} \ \frac{R_{\text{fh},m}}{C_{\text{fh},m}}, \tag{4.13}$$

where P_{FT} is the total power required for fronthaul traffic at full capacity, $C_{\text{fh},m}$ is the capacity of the fronthaul link between the mth AP and the CPU, and finally $R_{\text{fh},m}$ is the actual fronthaul rate between the mth AP and the CPU and is given by [52,61–63]

$$R_{\text{fh},m} = \frac{2\ K\ \tau_f\ \alpha_m}{T_c}, \tag{4.14}$$

where α_m denotes the number of quantization bits at each AP and for simplicity we consider the same number of bits at all APs, drop the index m and use α as the number

of quantization bits. Moreover, τ_f introduces the length of the uplink data (in symbols) and is given by $\tau_f = \tau_c - \tau_p$, where τ_c denotes the number of samples for each coherence interval, τ_p represents the length of pilot sequence, and finally T_c refers to coherence time in seconds. Note that in (4.14) α is related to the total uplink spectral efficiency, since it will affect the TQD term and hence the total spectral efficiency in (4.6).

4.4.2 Total Energy Efficiency

In this section, we formulate the total energy efficiency of cell-free massive MIMO uplink. The total energy efficiency is obtained by dividing the sum throughput (bit/s) by the total consumed power (Watt) which is given by

$$E_e\left(q_k, \mathbf{u}_k, \alpha\right) = \frac{B\, S\left(q_k, \mathbf{u}_k, \alpha\right)}{P_{\text{total}}} \left(\frac{\text{bit}}{\text{Joule}}\right), \tag{4.15}$$

where B is the frequency bandwidth.

4.5 Total Energy Efficiency Maximization

In this section, we propose a total energy efficiency maximization problem in cell-free massive MIMO, where we design the number of quantization bits α, the receiver filter coefficients $\mathbf{u}_k$ and the power coefficients q_k to maximize the total energy efficiency under per-user power and per-user spectral efficiency constraints. Hence, the total energy efficiency maximization can be modeled as follows:

$$P_1: \max_{q_k, \mathbf{u}_k, \alpha} \quad E_e\left(q_k, \mathbf{u}_k, \alpha\right) \tag{4.16a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k\right) \geq S_k^{(\text{r})}, \ \ \forall k, \tag{4.16b}$$

$$||\mathbf{u}_k|| = 1, \ \forall k, \tag{4.16c}$$

$$0 \leq q_k \leq p_{\text{max}}^{(k)}, \ \ \forall k, \tag{4.16d}$$

$$R_{bh,m} \leq C_{\text{fh},m}, \ \ \forall m, \tag{4.16e}$$

where $S_k^{(\mathrm{r})}$ is the required spectral efficiency of the kth user, $p_{\max}^{(k)}$ and $C_{\mathrm{fh},m}$ refer to the maximum transmit power available at user k and the capacity of fronthaul link between the mth AP and the CPU, respectively. Assuming the same amount of fronthaul capacity between all APs and the CPU, we drop the index m, and use C_{fh} for simplicity. Problem P_1 can be written as

$$P_2: \max_{q_k,\mathbf{u}_k,\alpha} \quad \frac{B\,.\,S\left(q_k,\mathbf{u}_k,\alpha\right)}{\frac{1}{\zeta}\rho p_n \sum_{k=1}^{K} q_k + MP_{\mathrm{fix}} + KP_{\mathrm{U}} + P_{\mathrm{FT}}\frac{2\,K\,\tau_f\,\alpha}{T_c}\frac{P_{\mathrm{FT}}}{C_{\mathrm{fh}}}} \tag{4.17a}$$

$$\text{s.t.} \quad S_k\left(q_k,\mathbf{u}_k,\alpha\right) \geq S_k^{(\mathrm{r})},\ \ \forall k, \tag{4.17b}$$

$$||\mathbf{u}_k|| = 1,\ \forall k, \tag{4.17c}$$

$$0 \leq q_k \leq p_{\max}^{(k)},\ \ \forall k, \tag{4.17d}$$

$$R_{fh} \leq C_{\mathrm{fh}},\ \ \forall m. \tag{4.17e}$$

Problem P_2 contains one discrete variable (the number of quantization bits). Hence, we can formulate the problem for fixed values of the number of quantization bits α, and we investigate the optimal values of α numerically. As a result, for a given α, the total energy efficiency maximization problem can be re-formulated as follows:

$$P_3: \max_{q_k,\mathbf{u}_k} \quad \frac{B\,.\,S\left(q_k,\mathbf{u}_k,\alpha\right)}{\frac{1}{\zeta}\rho p_n \sum_{k=1}^{K} q_k + MP_{\mathrm{fix}} + KP_{\mathrm{U}} + P_{\mathrm{FT}}\frac{2\,K\,\tau_f\,\alpha}{T_c}\frac{P_{\mathrm{FT}}}{C_{\mathrm{fh}}}} \tag{4.18a}$$

$$\text{s.t.} \quad S_k\left(q_k,\mathbf{u}_k,\alpha\right) \geq S_k^{(\mathrm{r})},\ \ \forall k, \tag{4.18b}$$

$$||\mathbf{u}_k|| = 1,\ \forall k, \tag{4.18c}$$

$$0 \leq q_k \leq p_{\max}^{(k)},\ \ \forall k. \tag{4.18d}$$

We then reformulate Problem P_3 into the following problem:

$$P_4: \max_{q_k,\mathbf{u}_k,\nu} \quad \frac{B\,.\,S\left(q_k,\mathbf{u}_k,\alpha\right)}{\frac{1}{\zeta}\rho p_n \nu \sum_{k=1}^{K} p_{\max}^{(k)} + MP_{\mathrm{fix}} + KP_{\mathrm{U}} + P_{\mathrm{FT}}\frac{2\,K\,\tau_f\,\alpha}{T_c}\frac{P_{\mathrm{FT}}}{C_{\mathrm{fh}}}} \tag{4.19a}$$

$$\text{s.t.} \quad S_k\left(q_k,\mathbf{u}_k,\alpha\right) \geq S_k^{(\mathrm{r})},\ \ \forall k, \tag{4.19b}$$

$$||\mathbf{u}_k|| = 1,\ \forall k, \tag{4.19c}$$

$$0 \leq q_k \leq p_{\max}^{(k)}, \forall k, \tag{4.19d}$$

$$\sum_{k=1}^{K} q_k \leq \nu \sum_{k=1}^{K} p_{\max}^{(k)}, \tag{4.19e}$$

$$\nu^{\star} \leq \nu \leq 1, \tag{4.19f}$$

where ν is an auxiliary variable and $\nu^\star$ and is obtained through the following remark.

Remark 3. *Based on the analysis in [64, 65], the slack variable $\nu^\star$ is obtained by solving a power minimization problem subject to the same per-user power constraints in (4.19d) and throughput requirement constraints in (4.19b). For details, please refer to Appendix 4.B.* ■

Theorem 6. *The optimal solution of Problem P_3 and problem P_4 are equal.*

Proof: The proof of Theorem 6 follows the same approach in the proof of [64, Theorem 1]. Let us assume $\{\mathbf{U}^{\text{opt}}, \mathbf{q}^{\text{opt}}\}$ and $\{\dot{\mathbf{U}}^{\text{opt}}, \dot{\mathbf{q}}^{\text{opt}}, \dot{\nu}\}$ are the optimal solution of Problems P_3 and P_4, respectively. It is easy to show that $\sum_{k=1}^{K} \dot{q}_k = \dot{\nu} \sum_{k=1}^{K} p_{\max}^{(k)}$. Moreover, based on [64], it is clear that $\dot{\mathbf{U}}^{\text{opt}}$ and $\dot{\mathbf{q}}^{\text{opt}}$ provide a feasible solution to Problem P_3. Exploiting the per-user power constraints, using $\nu = \frac{1}{\sum_{k=1}^{K} p_{\max}^{(k)}} \sum_{k=1}^{K} q_k$ and $0 \leq \nu \leq 1$, and by considering the throughput requirement constraints, one can conclude that $\{\mathbf{U}^{\text{opt}}, \mathbf{q}^{\text{opt}}\}$ provide a feasible solution to Problem P_4. Through these two facts, it is not difficult to show that the optimal solutions of Problems P_3 and P_4 are equal, which completes the proof of Theorem 6. ■

Hence, we can convert the original total energy efficiency maximization problem into a total energy efficiency maximization problem with per-user power constraints, throughput requirement constraints and the new total power constraint. Next, Problem P_4 is iteratively solved by performing a one-dimensional search over the variable $\nu^\star \leq \nu \leq 1$ [64]. Therefore, for a given ν, the denominator of the objective function of Problem P_4 is a constant, which enables us to define the following equivalent optimization problem:

$$P_5 : \max_{q_k, \mathbf{u}_k} \quad S\left(q_k, \mathbf{u}_k, \alpha\right) \tag{4.20a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\text{r})}, \ \forall k, \tag{4.20b}$$

$$||\mathbf{u}_k|| = 1, \ \forall k, \tag{4.20c}$$

$$0 \leq q_k \leq p_{\max}^{(k)}, \forall k, \tag{4.20d}$$

$$\sum_{k=1}^{K} q_k \leq \nu \sum_{k=1}^{K} p_{\max}^{(k)}. \tag{4.20e}$$

Problem P_5 is not convex in terms of $\mathbf{u}_k$ and power allocation $q_k, \ \forall k$. Therefore, it

cannot be directly solved through existing convex optimization software. To tackle this non-convexity issue, we decouple Problem P_5 into two sub-problems: receiver coefficient design (i.e. $\mathbf{u}_k$) and the power allocation problem. The optimal solution for Problem P_5, is obtained through alternately solving these sub-problems, as explained in the following subsections.

4.5.1 Receiver Filter Coefficient Design

In this subsection, the problem of designing the receiver filter coefficient vector is considered. We solve the total energy efficiency maximization problem for a given set of power allocations at all users, $q_k, \forall k$, and fixed values for the number of quantization bits, $\alpha_m,\ \forall m$. These coefficients (i.e., $\mathbf{u}_k,\ \forall\ k$) are obtained by independently maximizing the total uplink energy efficiency of the system. Note that the spectral efficiency of the kth user, i.e., $S_k\left(q_k, \mathbf{u}_k, \alpha\right)$, is a function of only $\mathbf{u}_k$ (it does not depend on $\mathbf{u}_{k'}$, where $k' \neq k$), and hence, the optimal receiver filter coefficients can be determined by solving the following optimization problem:

$$P_6 : \max_{\mathbf{u}_k} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \tag{4.21a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\mathrm{r})},\ \ \forall k, \tag{4.21b}$$

$$||\mathbf{u}_k|| = 1,\ \forall k. \tag{4.21c}$$

Note that the satisfaction of constraints in (4.21b) will be ensured in the power allocation problem. Hence, we drop constraint (4.21b) and Problem P_6 can be reformulated as:

$$P_7 : \max_{\mathbf{u}_k} \tag{4.22a}$$

$$\frac{N^2 \mathbf{u}_k^H \left(q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H\right) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 \sum\limits_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Delta}_{kk'} \mathbf{\Delta}_{kk'}^H + N^2 \sum\limits_{k'=1}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{k'} + N \sum\limits_{k'=1}^{K} q_{k'} \mathbf{D}_{kk'} + \frac{N \mathbf{R}_k}{\rho}\right) \mathbf{u}_k}$$

$$\text{s.t.} \quad ||\mathbf{u}_k|| = 1,\ \forall k. \tag{4.22b}$$

Problem P_7 is a generalized eigenvalue problem [45], where the optimal solutions can be obtained by determining the generalized eigen vector of the matrix pair $\mathbf{A}_k = N^2 q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H$ and $\mathbf{B}_k = N^2 \sum_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Delta}_{kk'} \mathbf{\Delta}_{kk'}^H + N^2 \sum_{k'=1}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{k'} + N \sum_{k'=1}^{K} q_{k'} \mathbf{D}_{kk'} + \frac{N}{\rho} \mathbf{R}_k$

corresponding to the maximum generalized eigenvalue.

4.5.2 Power Allocation

In this subsection, we solve the power allocation problem for a given set of fixed receiver filter coefficients, $\mathbf{u}_k$, $\forall\ k$, and fixed values of quantization bits. The optimal transmit power can be determined by solving the following total spectral efficiency maximization problem:

$$P_8 : \max_{q_k} \quad S\left(q_k, \mathbf{u}_k, \alpha\right) \tag{4.23a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\mathrm{r})}, \ \ \forall k, \tag{4.23b}$$

$$0 \leq q_k \leq p_{\max}^{(k)}, \ \ \forall k, \tag{4.23c}$$

$$\sum_{k=1}^{K} q_k \leq \nu \sum_{k=1}^{K} p_{\max}^{(k)}. \tag{4.23d}$$

Problem P_8 can be reformulated as follows:

$$P_9 : \min_{q_k} \quad \prod_{k=1}^{K} \left(1 + \text{SINR}_k\left(q_k, \mathbf{u}_k, \alpha\right)\right)^{-1} \tag{4.24a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\mathrm{r})}, \ \ \forall k, \tag{4.24b}$$

$$0 \leq q_k \leq p_{\max}^{(k)}, \ \ \forall k, \tag{4.24c}$$

$$\sum_{k=1}^{K} q_k \leq \nu \sum_{k=1}^{K} p_{\max}^{(k)}. \tag{4.24d}$$

Problem P_9 is generally a non-convex problem, however, it can be reformulated as a standard GP problem [46]. We first rewrite Problem P_9 as follows:

$$P_{10} : \min_{q_k, t_k} \quad \prod_{k=1}^{K} \left(1 + t_k\right)^{-1} \tag{4.25a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\mathrm{r})}, \ \ \forall k, \tag{4.25b}$$

$$0 \leq q_k \leq p_{\max}^{(k)}, \forall k, \tag{4.25c}$$

$$\text{SINR}_k \geq t_k, \forall k, \tag{4.25d}$$

$$\sum_{k=1}^{K} q_k \leq \nu \sum_{k=1}^{K} p_{\max}^{(k)}, \tag{4.25e}$$

where $t_k, \forall k$ refers to the slack variables. Problem (4.25) is a non-convex signomial problem. However, in Appendix 4.B, we will show that all constraints in (4.25) can be reformulated into posynomial functions. Hence, if the objective function in (4.25) can be reformulated into a posynomial function, problem (4.25) is a standard GP and has an optimal solution [46]. This motivates us to propose two schemes to transform Problem (4.25) into a standard GP.

Efficient Power Allocation Scheme

We use the SCA scheme proposed in [66] to convert Problem (4.25) into a standard GP. This scheme is referred to as the "inner approximation algorithm for non-convex problems" in [66], and introduces an efficient solution for the original problem [64, 66]. Based on the analysis in [66], it is possible to search for a local optimum through solving a sequence of GPs which locally approximate the original optimization problem. This scheme is called the "inner approximation algorithm for non-convex problems" in [66]. This scheme provides an efficient solution for the original problem [64, 66]. Next, the following lemma using SCA is required [64, Lemma 1]:

Lemma 3. *Function $\Theta(x) = \kappa t^{\xi}$ can be used to approximate function $\Pi(x) = 1 + t$, near the point $\hat{t}$. The best monomial local approximation is obtained by the following parameters:*

$$\xi = \frac{\hat{t}}{1+\hat{t}}, \tag{4.26a}$$

$$\kappa = \frac{1+\hat{t}}{\hat{t}^{\xi}}, \tag{4.26b}$$

where $\Theta(t) \leq \Pi(t)$, $\forall t > 0$.

Using the local approximation in Lemma 3, we can tackle the non-convexity of Prob-

lem P_{10}, which enables us to reformulate Problem P_{10} as follows:

$$P_{11}: \min_{q_k, t_k} \quad \left(\prod_{k=1}^{K} t_k^{-\frac{\hat{t}_k}{1+\hat{t}_k}} \right) \tag{4.27a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\mathrm{r})}, \ \ \forall k, \tag{4.27b}$$

$$0 \leq q_k \leq p_{\max}^{(k)}, \forall k, \tag{4.27c}$$

$$\mathrm{SINR}_k \geq t_k, \forall k, \tag{4.27d}$$

$$\sum_{k=1}^{K} q_k \leq \nu \sum_{k=1}^{K} p_{\max}^{(k)}, \tag{4.27e}$$

$$\left((1-\delta)\hat{t}_k\right) \leq t_k \leq \left((1-\delta)\hat{t}_k\right), \forall k, \tag{4.27f}$$

where δ is a constant value to control the approximation accuracy [64].

Proposition 3. *Problem P_{11} can be formulated into a standard GP.*

Proof: Please refer to Appendix 4.C. ■

Therefore, Problem P_{11} is efficiently solved through existing convex optimization software. Based on these two sub-problems (P_7 and P_{11}), an iterative algorithm has been developed by alternately solving both sub-problems at each iteration. The proposed algorithm is summarized in Algorithm 4, where ϵ_1 and ϵ_2 are small values, and we set $\epsilon_1 = \epsilon_2 = 0.01$.

Sub-Optimal Power Allocation Scheme

In this section, we present a heuristic solution to tackle the non-convexity issue of Problem P_{10}. Exploiting the analysis in [67], we propose to reformulate the energy efficiency

Algorithm 4 Proposed algorithm to solve Problem P_5

1. Initialize $\mathbf{q}^{(0)}$, $\mathbf{U}^{(0)}$. Calculate the uplink $\text{SINR}_k^{(0)}$, $t_0^{(0)}$ and $S_k^{(r)}$ using $\mathbf{q}^{(0)}$ and $\mathbf{U}^{(0)}$, and set the initial SINR guess and initial auxiliary variables as $\hat{t}_k = \text{SINR}_k^{(0)}, \forall k$, and $t_k^{(0)} = \text{SINR}_k^{(0)}, \forall k$, respectively.
2. Set $\mathbf{q}^{(\star)} = 0$, $t_k^{(\star)} = t_k^{(0)}$, $\mathbf{U}^{(\star)} = \mathbf{U}^{(0)}$, and $\tilde{E}_{e,k}^{(\star)} = 0, \forall k$.
3. Calculate the constants ξ and κ using (4.26), and solve problem P_{11} with $t_k^{(\star)}$ and $\mathbf{U}^{(\star)}$, and find $\mathbf{q}^{(\star\star)}$ and calculate $t_0^{(\star\star)}$ and $t_k^{(\star\star)}$.
4. If $\left|t_k^{(\star\star)} - t_k^{(\star)}\right| \leq \epsilon_1$, then set $t_k^{(\star\star)} = t_k^{(\star)}$ and $\mathbf{q}^{(\star\star)} = \mathbf{q}^{(\star)}$ and go to step 8, otherwise, $t_k^{(\star)} = t_k^{(\star\star)}$ and go to step 3.
5. Solve the generalized eigenvalue Problem P_7 using $\mathbf{q}^{(\star)}$ and calculate $\mathbf{U}$. Next, let $\mathbf{U}^{(\star\star)} = \mathbf{U}$.
6. Compute the objective value of Problem P_{11} with $\mathbf{U}^{(\star\star)}$ and $\mathbf{q}^{(\star)}$ and call it $\tilde{E}_{e,k}^{(\star\star)}, \forall k$.
7. If $\left|\tilde{E}_{e,k}^{(\star\star)} - \tilde{E}_{,k}e^{(\star)}\right| \leq \epsilon_2, \forall k$, then $\mathbf{U}^{(\star)} = \mathbf{U}^{(\star\star)}$ and go to step 8, otherwise, go to step 3.
8. If the stop criteria is satisfied stop, otherwise, go to step 3.

Algorithm 5 Proposed sub-optimal algorithm to solve Problem P_5

1. Initialize $\mathbf{q}^{(0)}$, $i = 1$.
2. Repeat steps 3-5 until $\left|\tilde{\tilde{E}}_{e,k}^{(i+1)} - \tilde{\tilde{E}}_{e,k}^{(i)}\right| \leq \epsilon_3, \forall k$, where $\tilde{\tilde{E}}_{e,k}$ is the objective value of Problem P_{10}.
3. $i = i + 1$.
4. Set $\mathbf{q}^{(i)} = \mathbf{q}^{(i-1)}$ and determine the optimal receiver coefficients $\mathbf{U}^{(i)}$ through solving the generalized eigenvalue Problem P_7.
5. Compute $\mathbf{q}^{(i+1)}$ through solving Problem P_{12}.

maximization Problem P_8 as follows:

$$P_{12}: \min_{q_k, t_k} \quad \prod_{k=1}^{K} t_k^{-1} \tag{4.28a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\text{r})}, \ \ \forall k, \tag{4.28b}$$

$$0 \leq q_k \leq p_{\text{max}}^{(k)}, \forall k, \tag{4.28c}$$

$$\text{SINR}_k \geq t_k, \forall k, \tag{4.28d}$$

$$\sum_{k=1}^{K} q_k \leq \nu \sum_{k=1}^{K} p_{\text{max}}^{(k)}. \tag{4.28e}$$

Proposition 4. *Problem P_{12} can be formulated into a standard GP.*

Proof: The objective function in (4.25a) and the power constraint in (4.25e) are posynomial functions. The spectral efficiency constraint in (4.25b) and the SINR constraint in (4.25d) can be rewritten into the posynomial functions similar to (4.85) and (4.88), which completes the proof. ■

Hence, existing convex optimization software can be used to solve problem P_{12}. As in the previous section, here we propose an iterative algorithm to iteratively solve sub-problems P_7 and P_{12}. Finally, Algorithm 5 summarizes the proposed scheme.

4.5.3 Convergence

In this section, the convergence analysis of the proposed Algorithms 4 and 5 are provided. Two sub-problems are alternately solved to determine the solution to Problem P_2. At each iteration, one of the design parameters is determined by solving the corresponding sub-problem while other design variables are kept fixed. Note that each sub-problem provides an optimal solution for the other given design variables. At the nth iteration, the receiver filter coefficients $\mathbf{u}_k^{(n)},\ \forall k$ are determined for a given power allocation $\mathbf{q}^{(n)}$ and similarly, the power allocation $\mathbf{q}^{(n+1)}$ is updated for a given set of receiver filter coefficients $\mathbf{u}_k^{(n)},\ \forall k$. The optimal power allocation $\mathbf{q}^{(n+1)}$ obtained for a given $\mathbf{u}_k^{(n)}$ achieves an uplink spectral efficiency greater than or equal to that of the previous iteration. In addition, the power allocation $\mathbf{q}^{(n)}$ is also a feasible solution in determining $\mathbf{q}^{(n+1)}$ as the receiver filter coefficients $\mathbf{u}_k^{(n+1)},\ \forall k$ are determined for a given $\mathbf{q}^{(n)}$. This reveals that the achieved uplink spectral efficiency monotonically increases with each iteration, which can also be observed from the numerical results presented in Figs. 4.2, 4.3 and 4.4. As the achievable uplink energy efficiency is upper bounded by a certain value for a given set of per-user power and spectral efficiency constraints, the proposed algorithms converges to a particular solution. Note that to the best of our knowledge and referring to [16, 40] this is a common way to show the convergence.

Table 4.1: Computational Complexity of Different Problems

Problem	**Required arithmetic operations**
Problem P_7	$\frac{14}{3}KM^3$
Problem P_{11}	$n_{\text{iter}} \times \mathcal{O}\Big((4K-1)^{\frac{1}{2}}(24K^3-20K^2+8K-1)\Big)$
Problem P_{12}	$\mathcal{O}\Big((4K-1)^{\frac{1}{2}}(24K^3-20K^2+8K-1)\Big)$

4.5.4 Complexity analysis

Here, we provide the computational complexity analysis for the proposed Algorithms, which solve a generalized eigenvalue problem P_7 and a GP (convex optimization problem) given by P_{11} and P_{12}, respectively, at each iteration. For the receiver filter coefficient design in P_7, an eigenvalue solver requires $\frac{14}{3}KM^3$ flops for K users using the QR algorithm [68].

Proposition 5. *Problem P_{11}, can be solved with complexity equivalent to $n_{iter} \times \mathcal{O}\Big((4K-1)^{\frac{1}{2}}(24K^3-20K^2+8K-1)\Big)$, where n_{iter} refers to the number of iterations in P_{11} which depends on δ in (4.27f). Note that the term $\mathcal{O}$ means there is an unknown factor. Moreover, it can be shown that Problem P_{12} can be solved with a complexity of $\mathcal{O}\Big((4K-1)^{\frac{1}{2}}(24K^3-20K^2+8K-1)\Big)$.*

The number of arithmetic operations required for Algorithms are provided in Table 4.1.

Proof: Let us consider the following GP problem:

$$P^{\text{GP}} : \min f_0(\mathbf{x}) = \sum_{i \in \mathcal{I}_0} c_{i0} \exp\{a_i^T \mathbf{x}\} \tag{4.29a}$$

$$\text{s.t. } f_j(\mathbf{x}) = \sum_{i \in \mathcal{I}_j} c_{ij} \exp\{a_i^T \mathbf{x}\} \leq d_j, j = 1, \cdots, n_3, \tag{4.29b}$$

where $\mathbf{x} = \{x_1, \cdots, x_{n_1}\}$ represents the optimization variables, $\mathcal{I}_j$ are subset of the index set $\mathcal{I} = 1, \cdots, n_2$, and all coefficients c_{ij} are positive, $j = 1, \cdots, n_3$ [69, Chapter 10]. Based on the analysis in [69, Chapter 10], the complexity of solving the GP problem given in (4.29) is given by $\mathcal{C} = \mathcal{O}\left((n_2 + n_3)^{\frac{1}{2}} \left(n_3 n_2^2 + n_2^3 + n_1^3\right)\right)$. Therefore, exploiting P_{11} defined in (4.27) and the transformation in (4.84)-(4.88), we have $n_1 = K$, $n_2 = 2K - 1$ and $n_3 = 2K$. Note that $n_2 = 2K - 1$ is obtained using the transformation in (4.88) for the constraint in (4.27b), and also the transformation in (4.85) for constraint (4.27d). Hence, Problem P_{11}, can be solved with a complexity equivalent to $n_{\text{iter}} \times \mathcal{O}\left((2K - 1 + 2K)^{\frac{1}{2}} \left((2K)(2K-1)^2 + (2K-1)^3 + (2K)^3\right)\right)$, where n_{iter} refers to the number of iterations to solve P_{11} which depends on δ in (4.27f). Moreover, it can be shown that Problem P_{12} can be solved with a complexity of $\mathcal{O}\left((2K-1+2K)^{\frac{1}{2}} \left((2K)(2K-1)^2 + (2K-1)^3 + (2K)^3\right)\right)$. After some manipulations, we end up with the values given in Table 4.1, which completes the proof of Proposition 5. ■

4.6 User Assignment

Let τ_f be the length of the uplink payload data transmission for each coherence interval, i.e., $\tau_f = \tau_c - \tau_p$, where τ_c denotes the number of samples for each coherence interval and τ_p represents the length of pilot sequence. Note that we need $2\alpha_m \times (K\tau_f)$ bits for each AP during each coherence interval. Hence, the total fronthaul capacity required between the mth AP and the CPU for all schemes is defined as

$$C_m = \frac{2\,(K\tau_f)\,\alpha_m}{T_c}, \tag{4.30}$$

where T_c (in sec.) refers to coherence time. Exploiting (4.30), it is obvious that the total fronthaul capacity required between the mth AP and the CPU increases linearly with the

Algorithm 6 User Assignment

1. Using (4.31), find the maximum possible integer value for $K_m, \forall m$
2. Sort users according to the ascending channel gain: $\beta_{m1} \geq \beta_{m2} \geq \cdots \geq \beta_{mK}, \forall m$
3. Assign K_m users with the highest values of $\beta_{mk}, \forall m$ to each AP, i.e., $\mathcal{T}_m \leftarrow \{k^{(1)}, k^{(2)}, \cdots, k^{(K_m)}\}, \forall m$
4. Find set of active APs for each user; $\mathcal{S}_k \leftarrow \{m^{(1)}, m^{(2)}, \cdots, m^{(M_k)}\}, \forall k$
5. for $j = 1 : K$
if size $\{\mathcal{S}_j\} = 0$
$\pi(j) = \underset{m}{\operatorname{argmax}}\, \beta_{mj}$, $\delta(j) = \underset{k}{\operatorname{argmin}}\, \beta_{\pi(j)k}, k|\mathcal{S}_k\pi_j \neq \varnothing$, $\mathcal{T}_{\pi(j)} \leftarrow \mathcal{T}_{\pi(j)} \backslash \delta(j)$, $\mathcal{T}_{\pi(j)} \leftarrow \mathcal{T}_{\pi(j)} \cup j$
end
end
6. If $m \in \mathcal{S}_k$, then $\vec{\gamma}_{mk} \leftarrow \gamma_{mk}$, otherwise $\vec{\gamma}_{mk} = 0$ and solve the max-min rate problem P_2

total number of users served by the mth AP. This motivates the need to pick a proper set of active users for each AP. Using (4.30), we have

$$\alpha_m \times K_m \leq \frac{C_{\text{fh}} T_c}{2\tau_f}, \tag{4.31}$$

where K_m denotes the size of the set of active users for the mth AP. From (4.31), it can be seen that decreasing the size of the set of active users allows for a larger number of quantization bits. Motivated by this fact, and to exploit the capacity of fronthaul links more efficiently, we investigate all possible combinations of α_m and K_m. First, for a fixed value of α_m, we find an upper bound on the size of the set of active users for each AP. In the next step, we propose for all APs that the users are sorted according to $\beta_{mk}, \ \forall k$, and find the K_m users which have the highest values of β_{mk} among all users. If a user is not selected by any AP, we propose to find the AP which has the best link to this user ($\pi(j) = \underset{m}{\operatorname{argmax}}\, \beta_{mj}$ determines best link to the jth user, i.e., the index of the AP which is closest to the jth user). Note that to consider only the users that have links to other APs, we use $k|\mathcal{S}_k\pi_j \neq \varnothing$, where $\varnothing$ refers to empty set. Then we drop the user which has the lowest $\beta_{mk}, \ \forall k$, among the set of active users for that AP, which has links to other APs as well. Finally, we add the user which is not selected by any AP to the set of active users for this AP. The proposed algorithm is summarized in Algorithm 6. We next solve the uplink energy efficiency maximization problem as follows

$$P^{\text{user assignment}}: \max_{q_k, \mathbf{u}_k, \alpha} \quad E_e\left(q_k, \mathbf{u}_k, \alpha, \tilde{\gamma}_{mk}\right) \tag{4.32a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \tilde{\gamma}_{mk}\right) \geq S_k^{(\text{r})}, \ \forall k, \tag{4.32b}$$

$$||\mathbf{u}_k|| = 1, \ \forall k, \ 0 \leq q_k \leq p_{\text{max}}^{(k)}, \ \forall k, \tag{4.32c}$$

$$R_{bh,m} \leq C_{\text{fh},m}, \ \forall m, \tag{4.32d}$$

where

$$\tilde{\gamma}_{mk} = \begin{cases} \gamma_{mk}, & m \in \mathcal{S}_k \\ 0, & \text{otherwise} \end{cases} \tag{4.33}$$

where $\mathcal{S}_k$ refers to the set of active APs for the kth user. Finally, note that this reduces the complexity of the optimization problem, as some entries of $\tilde{\gamma}_{mk}$ are zero. Finally, note that we turn off the mth AP, if the set of active users for the mth AP is empty, after performing the user assignment scheme. Hence, we put the number of active APs instead of M.

4.7 Numerical Results and Discussion

In this section, we provide numerical results to validate the performance of the proposed scheme. A cell-free massive MIMO system with M APs and K single-antenna users is considered in a $D \times D$ coverage area, where both APs and users are randomly distributed. In the following subsections, we define the numerical parameters and then present the corresponding numerical results.

4.7.1 Simulation Parameters

The channel coefficients between users and APs are modeled in Section 2.7.4, where the coefficient β_{mk} is given by [2]

$$\beta_{mk} = \text{PL}_{mk} 10^{\frac{\sigma_{sh} \ z_{mk}}{10}}, \tag{4.34}$$

where PL_{mk} is the path loss from the kth user to the mth AP and the second term in (4.34), $10^{\frac{\sigma_{sh}z_{mk}}{10}}$, denotes the shadow fading with standard deviation $\sigma_{sh} = 8$ dB, and $z_{mk} \sim \mathcal{N}(0,1)$. In the simulation, an uncorrelated shadowing model is considered and a three-slope model for the path loss is given by [2]

$$\text{PL}_{mk} = \begin{cases} -L - 35\log_{10}(d_{mk}), & d_{mk} > d_1, \\ -L - 15\log_{10}(d_1) - 20\log_{10}(d_{mk}), & d_0 < d_{mk} \leq d_1, \\ -L - 15\log_{10}(d_1) - 20\log_{10}(d_0), & d_{mk} \leq d_0, \end{cases} \tag{4.35}$$

and $L = 46.3 + 33.9\log_{10}(f) - 13.82\log_{10}(h_{AP}) - (1.1\log_{10}(f) - 0.7)\,h_k + (1.56\log_{10}(f) - 0.8)$, where f denotes the carrier frequency (in MHz), h_{AP} and h_k represent the AP antenna height (in m) and user height (in m), respectively. The noise power is given by $p_n = \text{BW} \times k_B \times T_0 \times W$, where $\text{BW} = 20$ MHz denotes the bandwidth, $k_B = 1.381 \times 10^{-23}$ represents the Boltzmann constant, and $T_0 = 290$ (Kelvin) denotes the noise temperature. Moreover, $W = 9$ dB, and denotes the noise figure. It is assumed that that $\bar{p}_p$ and $\bar{\rho}$ denote the power of the pilot sequence and the uplink data powers, respectively, where $p_p = \frac{\bar{p}_p}{p_n}$ and $\rho = \frac{\bar{\rho}}{p_n}$. In simulations, we set $\bar{p}_p = 200$ mW and $\bar{\rho} = 1$ Watt. Similar to [2], we assume that the simulation area is wrapped around at the edges which can simulate an area without boundaries. Hence, the square simulation area has eight neighbours. Moreover, we set $\zeta = 0.3$, $P_U = 0.1$ Watt, $P_{\text{fix}} = .825$ Watt [52, 60–63]. Moreover, hereafter the term "orthogonal pilots" refers to the case where unique orthogonal pilots are assigned to all users, while in "random pilot assignment" each user is randomly assigned a pilot sequence from a set of orthogonal sequences of length τ_p ($< K$), following the approach of [2, 31].

4.7.2 Numerical Results

Convergence of the Proposed Schemes

In this section, the convergence of the proposed Algorithms 4 and 5 is investigated. Figs. 4.2, 4.3 and 4.4 present the convergence of the proposed Algorithms 4 and 5 with $M = 100$ and $M = 200$ APs, and $K = 20$ and $K = 40$ users with the length of pilot $\tau_p = 20$. Note that in the figures, the term "Channel" means a random realization of positions of

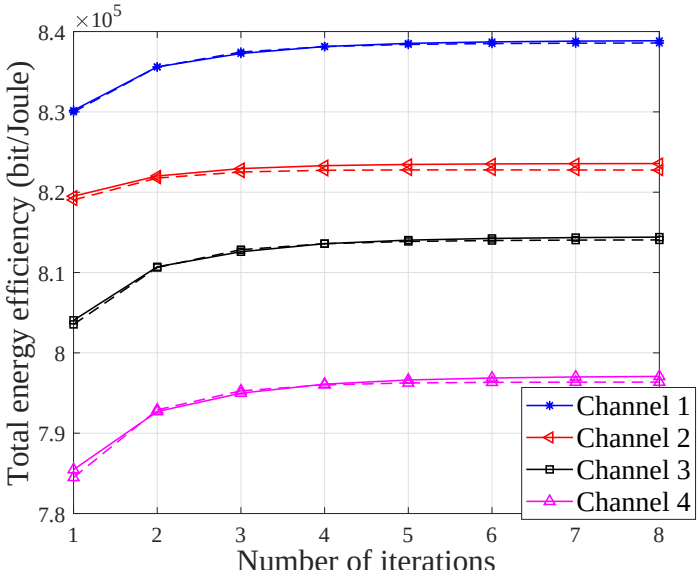

Figure 4.2: The total energy efficiency of proposed Algorithm 4 (solid curves) and proposed Algorithm 5 (dashed curves) versus number of iterations with $K = 20$, $M = 100$, $N = 1$, $\alpha = 2$, $\tau_p = 20$, and $D = 1$ km.

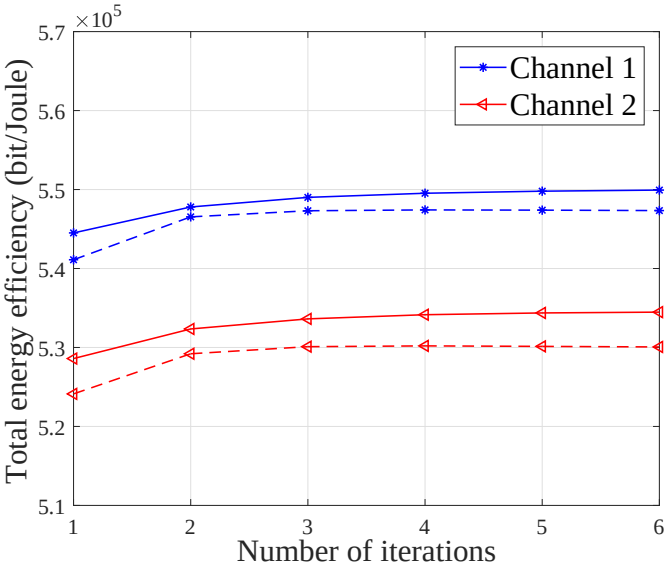

Figure 4.3: The total energy efficiency of proposed Algorithm 4 (solid curves) and proposed Algorithm 5 (dashed curves) versus number of iterations with $K = 40$, $M = 100$, $N = 1$, $\alpha = 2$, $\tau_p = 20$, and $D = 1$ km.

users and APs and small-scale fading. Note that we choose different values of K, M and α to show that the proposed algorithm converges very fast for different parameters.

Note that in Figs. 4.2, 4.3 and 4.4, the solid and dashed curves represent the performance of proposed Algorithm 4 and Algorithm 5, respectively. The figures

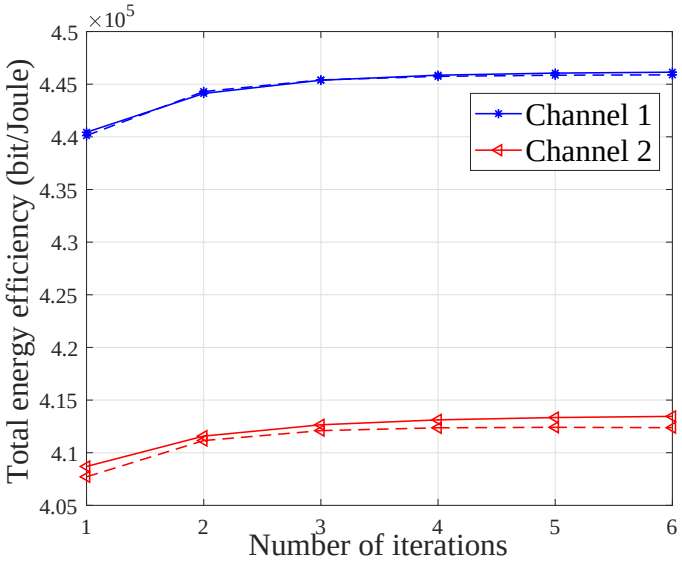

Figure 4.4: The total energy efficiency of proposed Algorithm 4 (solid curves) and proposed Algorithm 5 (dashed curves) versus number of iterations with $K = 40$, $M = 200$, $N = 1$, $\alpha = 2$, $\tau_p = 20$, and $D = 1$ km.

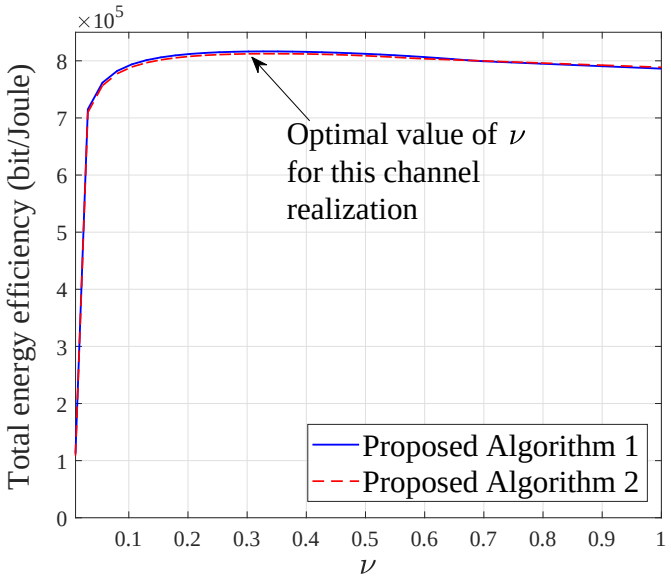

Figure 4.5: The total energy efficiency of proposed Algorithm 4 and proposed Algorithm 5 versus ν for one channel realization with $K = 20$, $M = 100$, $N = 1$, $\alpha = 2$, $\tau_p = 20$, and $D = 1$ km.

confirm that the proposed Algorithms 4 and 5 converge in a few iterations.

Figs. 4.2, 4.3 and 4.4 demonstrate that the proposed sub-optimal scheme has a performance fairly close to the performance of the proposed Algorithm 4. As Algorithm 5 has a lower complexity and good performance, in the rest of numerical results, we investigate

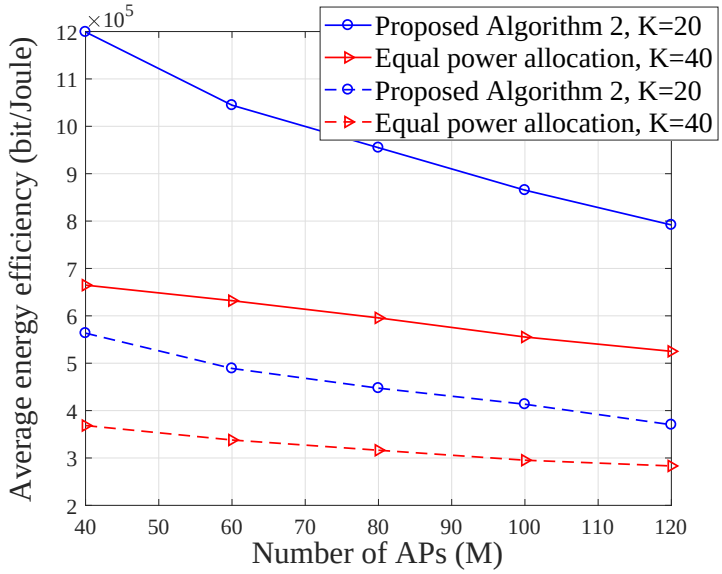

Figure 4.6: The average total energy efficiency versus number of APs with proposed Algorithm 4 and equal power allocation with $N = 1$, $\alpha = 2$, $\tau_p = 20$, and $D = 1$ km.

the performance using only the proposed Algorithm 5.

The Optimal Value of ν

To study the effect of ν in Problem P_5, we solve Problem P_5 with different values of ν and plot the total energy efficiency versus ν in Fig. 4.5. For this channel realization, for both proposed Algorithms 4 and 5, the optimal value of ν has a range from $0.25 - 0.35$, and we set the optimal value to $\nu^{\text{opt}} = 0.3$.

Performance Comparison

Fig. 4.6 presents the total energy efficiency of the proposed Algorithm 5 and the scheme with the equal power allocation with $M = 100$, $N = 1$, $\alpha = 2$, $\tau_p = 20$, and $D = 1$ km. As seen in Fig. 4.6, the proposed scheme significantly improves the total energy efficiency of cell-free massive MIMO compared to equal power allocation scheme (i.e., $q_k = 1, \forall k, \mathbf{u}_k = [1, \cdots, 1], \forall k$).

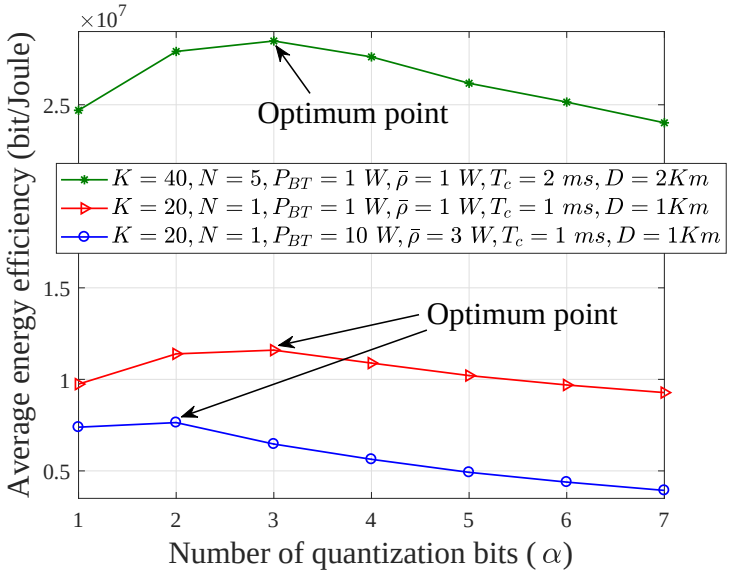

Figure 4.7: The average total energy efficiency of proposed Algorithm 5 versus number of quantization bits with $K = 20$, $N = 1$, $\tau_p = 20$, and $D = 1$ km.

Effect of the Number of Quantization Bits

This section investigates the optimum values of number of quantization bits to maximize the energy efficiency of cell-free massive MIMO. Increasing the number of quantization bits introduces spectral efficiency improvement whereas it increases the fronthaul power consumption from the APs to the CPU. Therefore, there is an optimum value in terms of number of quantization bits to maximize the total energy efficiency of the cell-free massive MIMO system. The average energy efficiency versus the number of quantization bits is shown in Fig. 4.7 for the system with $\{K = 40, N = 5, P_{\text{FT}} = 1$ Watt, $\rho = 3$ Watt, $T_c = 2$ ms, $D = 2$ Km$\}$, $\{K = 20, N = 1, P_{\text{FT}} = 1$ Watt, $\rho = 1$ Watt, $T_c = 1$ ms, $D = 1$ Km$\}$, $\{K = 40, N = 5, P_{\text{FT}} = 10$ Watt, $\rho = 3$ Watt, $T_c = 1$ ms, $D = 1$ Km$\}$ with orthogonal pilots. Optimally, we need only 2-4 bits to quantize the data. Again, note that ρ is defined in Section 4.7.1.

Effect of the Number of Antennas Per AP

In this section, the performance of cell-free massive MIMO is studied with different numbers of antennas per AP. Similar to the methodolgy in [16], we set $MN = 256$ as the total number of service antennas. The average energy efficiency of the system is shown

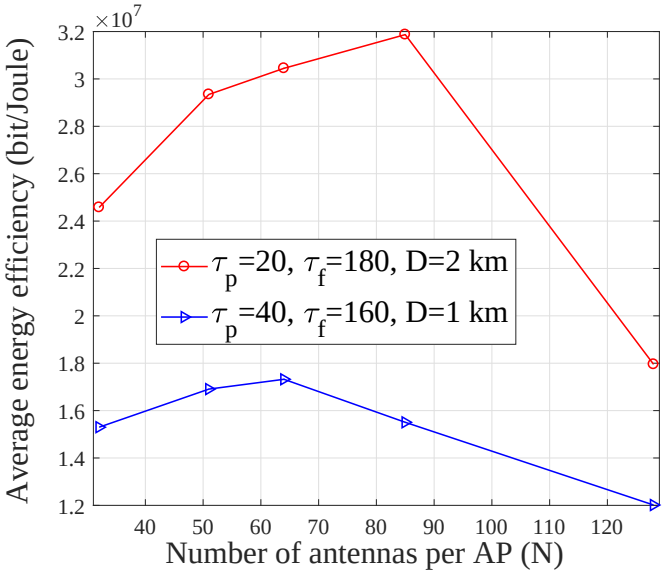

Figure 4.8: The average total energy efficiency of proposed Algorithm 5 versus the number of antennas per AP with $K = 40$, $MN = 256$, $P_{FT} = 10$ Watt, $C_{fh} = 100$ Mbps, and $\alpha = 4$ bits.

in Fig. 4.8, for $K = 40$, $\alpha = 4$ bits, and $P_{\text{FT}} = 10$ Watt. Moreover, we provide numerical results for two cases of orthogonal and random pilot assignment. It can be seen for a fixed total number of service antennas, by reducing the total number of APs, M (which is equivalent to increasing number of antennas per APs, N), the total power consumption will decrease. On the other hand, reducing M results in throughput reduction. As a result, one can find a trade off between M and N. Fig. 4.8 reveals the optimum values of M and N to have the largest total energy efficiency.

Effect of Power of Fronthaul Links

Fig. 4.9 shows the average energy efficiency of the cell-free massive MIMO system versus the total fronthaul traffic power, P_{FT}, for $K = 20$, $N = 1$, $\tau_p = 20$, $D = 1$ km, $C_{fh} = 102.4$ Mbps, and two cases of $M = 60$ and $M = 120$. As the figure demonstrates, the average energy efficiency decreases as the total power for fronthaul traffic increases.

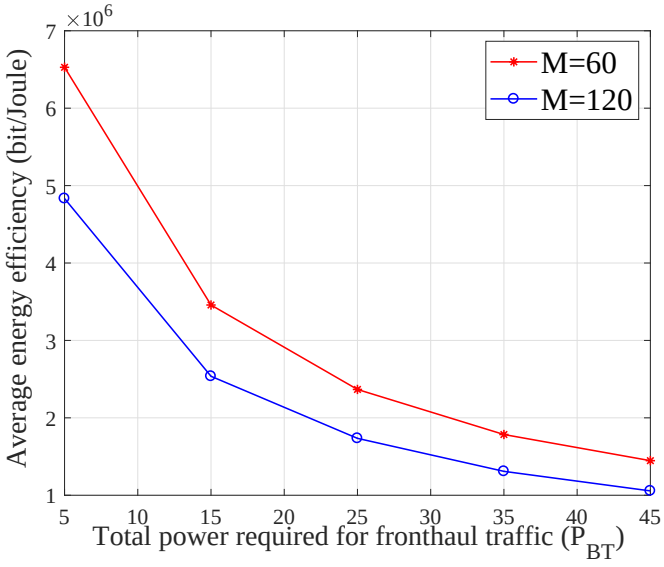

Figure 4.9: The average total energy efficiency of proposed Algorithm 5 versus number of quantization bits with $K = 20$, $N = 1$, $\tau_p = 20$, $D = 1$ km, $C_{fh} = 102.4$ Mbps, and two cases of $M = 60$ and $M = 120$.

Energy Efficiency vs Relative Loss in Max-Min Spectral Efficiency

It is interesting to evaluate how much we can gain with the proposed energy efficiency power control by sacrificing the required spectral efficiency. To investigate this, we consider the max-min spectral efficiency problem with a given fronthaul rate, which is defined as follows:

$$P^{\text{max-min}}: \max_{q_k, \mathbf{u}_k} \quad \min_{k=1,\cdots,K} \quad R_k, \tag{4.36a}$$

$$\text{s.t.} \quad ||\mathbf{u}_k|| = 1, \ \forall k, \tag{4.36b}$$

$$0 \leq q_k \leq p_{\text{max}}^{(k)}, \ \ \forall k. \tag{4.36c}$$

where R_k refers to the rate of the kth user. Next, we define the following optimization problem:

$$P^{\text{sac}}: \max_{q_k, \mathbf{u}_k} \quad E_e\left(q_k, \mathbf{u}_k\right), \tag{4.37a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k\right) \geq \left(\text{th}_{\text{sac}} \times S_k^{(\text{max-min})}\right), \ \ \forall k, \tag{4.37b}$$

$$||\mathbf{u}_k|| = 1, \ \forall k, \tag{4.37c}$$

$$0 \leq q_k \leq p_{\text{max}}^{(k)}, \ \ \forall k, \tag{4.37d}$$

where $S_k^{(\text{max-min})} = (1 - \frac{\tau_p}{\tau_c}) R_k^{\text{max-min}}$, where $R_k^{\text{max-min}}$ is the optimal solution of Problem $P^{\text{max-min}}$. Figure 4.10 presents the average energy efficiency performance of the cell-free

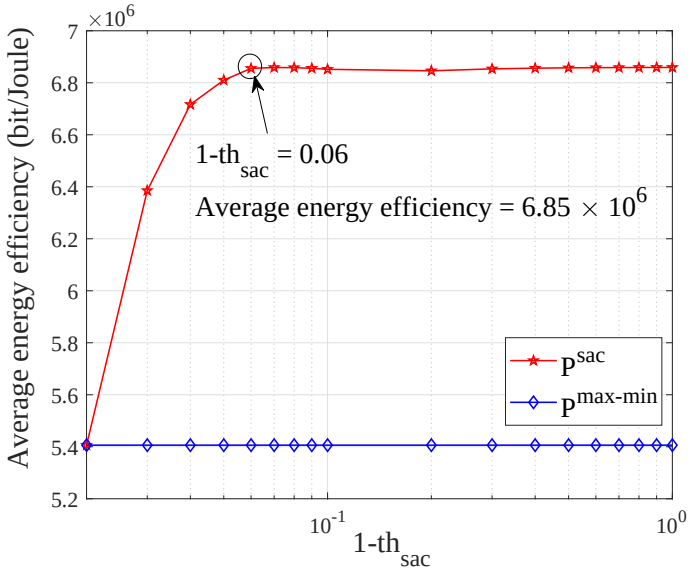

Figure 4.10: The average energy efficiency of proposed Algorithm 5 versus the sacrifice in max-min spectral efficiency for $K = 15$, $M = 80$, $N = 1$, $\tau_p = 15$, $D = 1$ km, $\alpha = 2$, $P_{\text{FT}} = 1$ Watt and $C_{fh} = 100$ Mbps.

massive MIMO with $M = 80$, $K = 15$, $N = 1$, $\alpha = 2$ and orthogonal pilots, obtained by solving Problems $P^{\text{max-min}}$ and P^{sac}. Note that we use the sub-optimal power allocation scheme presented in Subsection 4.5.2 to solve Problem P^{sac}. The figure shows that by sacrificing 6% of the max-min spectral efficiency (i.e., $1 - \text{th}_{\text{sac}} = 0.06$), one could gain $\frac{6.85\times10^6 - 5.4\times10^6}{5.4\times10^6} = 26.8\%$ improvement in the average energy efficiency of the system.

Performance of the Proposed User Assignment Scheme

This subsection investigates the performance of the proposed user assignment scheme. In Fig. 4.11, the average energy efficiency proposed using Algorithm 5 is presented with $M = 40$, $N = 4$, $K = 50$, and $\tau_p = 30$ versus the total number of active users per AP. Here, we used inequality (4.31) and set $\alpha_m \times K_m = 100$. As Fig. 4.11 shows, the optimum value of K_m, (K_m^{opt}) is achieved by $K_m^{\text{opt}} = 33$. As a result, the proposed user assignment scheme can effectively improve the energy efficiency performance of cell-free massive MIMO systems with limited fronthaul capacity.

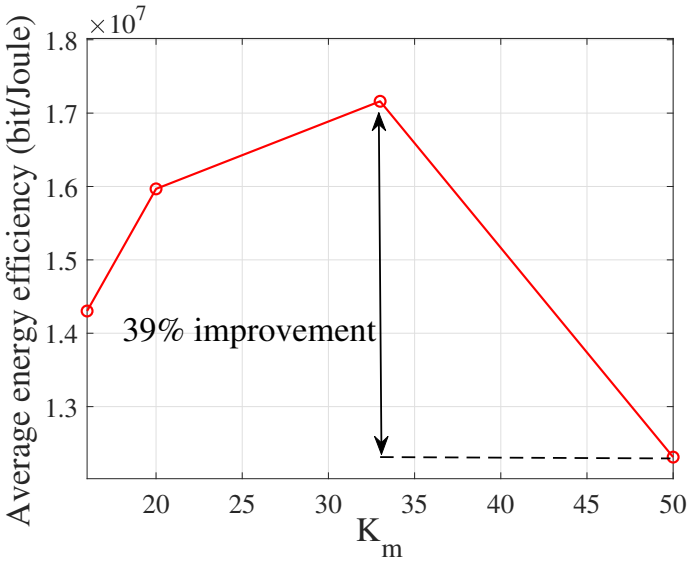

Figure 4.11: The average energy efficiency of proposed Algorithm 5 versus the total number of active users for each AP with $M = 40$, $N = 4$, $K = 50$, $\tau_p = 30$ and $\alpha_m \times K_m = 100$.

4.8 Summary

We have considered cell-free massive MIMO when the quantized version of the weighted signals are available at the CPU. Bussgang decomposition has been used to model the quantization effects. A closed-form expression for spectral efficiency has been derived. We have then studied the problem of the energy efficiency maximization with per-user power constraints, fronthaul capacity constraints and throughput requirements. We have developed an SCA to efficiently solve this non-convex problem. Next a low-complexity sub-optimal scheme is proposed. In addition, complexity and convergence of the proposed schemes have been investigated. Numerical results confirmed that the limited-fronthaul cell-free massive MIMO system with the proposed algorithm can reach almost twice the uplink total energy efficiency compared to the case of equal power allocation. In addition, a trade-off between the total number of APs and the number of antennas at the APs has been shown. Moreover, we investigated the optimal number of AP antennas along with the optimal number of quantization bits to maximize the uplink total energy efficiency of cell-free massive MIMO. Finally, we have presented the energy efficiency performance as a function of relative loss in the max-min spectral efficiency and evaluated the energy efficiency improvement achieved by sacrificing some of the max-min spectral efficiency.

4.9 Appendix

Appendix 4.A: Proof of Theorem 5

The desired signal for the user k is given by

$$\begin{aligned}\mathrm{DS}_k &= \sqrt{\rho}\mathbb{E}\left\{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\right\} \\ &= N\sqrt{pq_k}\sum_{m=1}^{M} u_{mk}\gamma_{mk}. \end{aligned} \tag{4.38}$$

Hence, $|\mathrm{DS}_k|^2 = \rho q_k \left(N\sum_{m=1}^{M} u_{mk}\gamma_{mk}\right)^2$. Moreover, the term $\mathbb{E}\{|\mathrm{BU}_k|^2\}$ can be obtained as

$$\begin{aligned}\mathbb{E}\left\{|\mathrm{BU}_k|^2\right\} &= \rho\mathbb{E}\left\{\left|\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k} - \rho\mathbb{E}\Big\{\sum_{m=1}^{M} u_{mk}\hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk}\sqrt{q_k}\Big\}\right|^2\right\} \\ &= \rho N\sum_{m=1}^{M} q_k u_{mk}^2 \gamma_{mk}\beta_{mk},\end{aligned}$$

where the last equality comes from the analysis in [2, Appendix A], and using $\gamma_{mk} = \sqrt{\tau_p p_p}\beta_{mk}c_{mk}$. The term $\mathbb{E}\{|\mathrm{IUI}_{kk'}|^2\}$ is obtained as

$$\begin{aligned}\mathbb{E}\{|\mathrm{IUI}_{kk'}|^2\} = \underbrace{\rho\, q_{k'}\mathbb{E}\left\{\left|\sum_{m=1}^{M} c_{mk}u_{mk}\mathbf{g}_{mk'}^H \tilde{\mathbf{w}}_{mk}\right|^2\right\}}_{A} \\ + \underbrace{\rho\,\tau_p p_p \mathbb{E}\left\{q_{k'}\left|\sum_{m=1}^{M} c_{mk}u_{mk}\Big(\sum_{i=1}^{K}\mathbf{g}_{mi}\boldsymbol{\phi}_k^H\boldsymbol{\phi}_i\Big)^H \mathbf{g}_{mk'}\right|^2\right\}}_{B},\end{aligned} \tag{4.39}$$

where the third equality in (4.39) is due to the fact that for two independent RVs X and Y and $\mathbb{E}\{X\} = 0$, we have $\mathbb{E}\{|X+Y|^2\} = \mathbb{E}\{|X|^2\} + \mathbb{E}\{|Y|^2\}$ [2]. Since $\tilde{\mathbf{w}}_{mk} = \boldsymbol{\phi}_k^H\mathbf{W}_{\mathbf{p},\mathbf{m}}$ is independent of the term $g_{mk'}$ similar to [2, Appendix A], the term A in (4.39) immediately is given by $A = Nq_{k'}\sum_{m=1}^{M} c_{mk}^2 u_{mk}^2\beta_{mk'}$. The term B in (4.39) can be

obtained as

$$B = \underbrace{\tau_p p_p q_{k'} \mathbb{E}\left\{ \left| \sum_{m=1}^{M} c_{mk} u_{mk} ||\mathbf{g}_{mk'}||^2 \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \right\}}_{C} + \underbrace{\tau_p p_p q_{k'} \mathbb{E}\left\{ \left| \sum_{m=1}^{M} c_{mk} u_{mk} \Big(\sum_{i \neq k'}^{K} \mathbf{g}_{mi} \boldsymbol{\phi}_k^H \boldsymbol{\phi}_i \Big)^H \mathbf{g}_{mk'} \right|^2 \right\}}_{D}. \tag{4.40}$$

The first term in (4.40) is given by

$$C = N \tau_p p_p q_{k'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \sum_{m=1}^{M} c_{mk}^2 u_{mk}^2 \beta_{mk'} + N^2 q_{k'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \left(\sum_{m=1}^{M} u_{mk} \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2, \tag{4.41}$$

where the last equality is derived based on the fact that $\gamma_{mk} = \sqrt{\tau_p p_p} \beta_{mk} c_{mk}$. The second term in (4.40) can be obtained as

$$D = N \sqrt{\tau p_p} q_{k'} \sum_{m=1}^{M} u_{mk}^2 c_{mk} \beta_{mk'} \beta_{mk} - N q_{k'} \sum_{m=1}^{M} u_{mk}^2 c_{mk}^2 \beta_{mk'} - N \tau p_p q_{k'} \sum_{m=1}^{M} u_{mk}^2 c_{mk}^2 \beta_{mk'}^2 \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2. \tag{4.42}$$

Finally by substituting (4.41) and (4.42) into (4.40), and substituting (4.40) into (4.39), we obtain

$$\mathbb{E}\{|\mathrm{IUI}_{kk'}|^2\} = N \rho q_{k'} \left(\sum_{m=1}^{M} u_{mk}^2 \beta_{mk'} \gamma_{mk} \right) + N^2 \rho q_{k'} \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \left(\sum_{m=1}^{M} u_{mk} \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2. \tag{4.43}$$

The total noise for the user k is given by

$$\mathbb{E}\left\{ |\mathrm{TN}_k|^2 \right\} = \mathbb{E}\left\{ \left| \sum_{m=1}^{M} u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m \right|^2 \right\} = N \sum_{m=1}^{M} u_{mk}^2 \gamma_{mk}, \tag{4.44}$$

where the last equality is due to the fact that the terms $\hat{\mathbf{g}}_{mk}$ and $\mathbf{n}_m$ are uncorrelated. Next, we calculate the power of the quantization distortion. Let us assume z_{mk} is the input of the quantizer at the mth AP, where $z_{mk} = \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m$. Using the Bussgang decomposition and the receiver filter coefficients $u_{mk}, \forall m, k$, at the CPU, the aggregate received signal at the CPU can be written as

$$r_k = \sum_{m=1}^{M} \mathcal{Q}\left(z_{mk}\right) = \sum_{m=1}^{M} \mathcal{Q}\left(\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m\right) = \sum_{m=1}^{M} a\, \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m + \underbrace{\sum_{m=1}^{M} d_{mk}^z}_{\text{TQD}_k}, \tag{4.45}$$

where TQD_k refers to the total quantization distortion (TQD) at the kth user. The power of the quantization distortion for user k is given by

$$\begin{aligned}
\mathbb{E}\left\{\left|\text{TQD}_k\right|^2\right\} &= \mathbb{E}\left\{\left|\sum_{m=1}^{M} d_{mk}^z\right|^2\right\} \\
&= \mathbb{E}\left\{\left(\sum_{m=1}^{M} d_{mk}^z\right)\left(\sum_{m=1}^{M} d_{mk}^z\right)^*\right\} \\
&= \sum_{m=1}^{M} \mathbb{E}\left\{|d_{mk}^z|^2\right\} + \sum_{m=1}^{M}\sum_{n\neq 1}^{M} \mathbb{E}\left\{d_{mk}^z \left(d_{nk}^z\right)^*\right\} \\
&= \sum_{m=1}^{M} \left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mm} + \sum_{m=1}^{M}\sum_{n\neq m}^{M} \left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{nm},
\end{aligned} \tag{4.46}$$

where $\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z} = \mathbb{E}\left\{\mathbf{d}_k^z \left(\mathbf{d}_k^z\right)^H\right\}$ and it is the covariance matrix of the quantization distortion, where $\mathbf{d}_k^z = \left[d_{1k}^z \cdots d_{Mk}^z\right]^T$ is the quantization distortion vector. Note that $\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mn}$ is the mnth element of $\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}$. To calculate $\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}$, we first re-write the aggregate received signal at the CPU as follows:

$$\mathbf{r}_k = Q(\mathbf{z}_k) = \mathbf{A}\mathbf{z}_k + \mathbf{d}_k^z, \tag{4.47}$$

where $\mathbf{r}_k = \left[r_{1k} \cdots r_{Mk}\right]^T$ and $\mathbf{z}_k = \left[z_{1k} \cdots z_{Mk}\right]^T$. Moreover, using the analysis in Appendix 4.A, it can be shown that the matrix $\mathbf{A}$ is diagonal. $\mathbf{A}$ is determined by the linear minimum mean square (MMSE) estimation of $\mathbf{r}_k$ from $\mathbf{x}_k$ as follows [55]:

$$\mathbf{A} = \mathbb{E}\left\{\mathbf{r}_k \mathbf{z}_k^H\right\} \mathbb{E}\left\{\mathbf{z}_k \mathbf{z}_k^H\right\}^{-1} = \mathbf{C}_{\mathbf{r}_k \mathbf{z}_k} \mathbf{C}_{\mathbf{x}_k \mathbf{x}_k}^{-1}, \tag{4.48}$$

and the error has the following covariance matrix [55]

$$\begin{aligned}\mathbf{C}_{\mathbf{d}_k^z\mathbf{d}_k^z} &= \mathbb{E}\left\{\left(\mathbf{r}_k-\mathbf{A}\mathbf{z}_k\right)\left(\mathbf{r}_k-\mathbf{A}\mathbf{z}_k\right)^H\right\}\mathbb{E}\left\{\mathbf{z}_k\mathbf{z}_k^H\right\}^{-1}\\ &= \mathbf{C}_{\mathbf{r}_k\mathbf{r}_k}-\mathbf{C}_{\mathbf{r}_k\mathbf{z}_k}\mathbf{A}^H-\mathbf{A}\mathbf{C}_{\mathbf{z}_k\mathbf{r}_k}+\mathbf{A}\mathbf{C}_{\mathbf{z}_k\mathbf{z}_k}\mathbf{A}^H\\ &= \mathbf{C}_{\mathbf{r}_k\mathbf{r}_k}-\mathbf{C}_{\mathbf{r}_k\mathbf{z}_k}\mathbf{C}_{\mathbf{z}_k\mathbf{z}_k}^{-1}\mathbf{C}_{\mathbf{z}_k\mathbf{r}_k}.\end{aligned} \tag{4.49}$$

Remark 4. *The covariance matrices* $\mathbf{C}_{\mathbf{r}_k\mathbf{r}_k}$, $\mathbf{C}_{\mathbf{r}_k\mathbf{z}_k}$, *and* $\mathbf{C}_{\mathbf{z}_k\mathbf{r}_k}$ *are obtained using the Price Theorem.*

Proof: To characterize the cross-correlation and auto-correlation properties of Gaussian input signals, we exploit "the Price Theorem" [70]. Based on the Price theorem, the correlation coefficient at the output of nonlinear functions $f_1(x_1)$ and $f_2(x_2)$ with correlated inputs x_1 and x_2 having zero-mean and the variances σ_1 and σ_2, respectively, and the correlation coefficient $\rho_{x_1x_2} = \frac{\mathbb{E}\{x_1x_2^*\}}{\sigma_{x_1}\sigma_{x_2}}$, has the following derivatives [70]

$$\frac{\partial^k\mathbb{E}\left\{f_1(\tilde{x}_1)f_2(\tilde{x}_2)\right\}}{\partial\rho_{\tilde{x}_1\tilde{x}_2}^k} \tag{4.50}$$

$$= \sigma_1^k\sigma_2^k\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f_1^{(k)}(\tilde{x}_1)f_2^{(k)}(\tilde{x}_2)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{x_1x_2}^2}}\exp\left(-\frac{1}{2\left(1-\rho_{x_1x_2}^2\right)}\left[\frac{\tilde{x}_1^2}{\sigma_1^2}+\frac{\tilde{x}_2^2}{\sigma_2^2}-\frac{2\rho_{x_1x_2}\tilde{x}_1\tilde{x}_2}{\sigma_1\sigma_2}\right]\right)d\tilde{x}_1d\tilde{x}_2. \tag{4.51}$$

Next, for the special case $f_1(x_1)=x_1$, then we have [55]

$$\frac{\partial\mathbb{E}\left\{x_1f_2(x_2)\right\}}{\partial\rho_{x_1x_2}} = \sigma_1\sigma_2\int_{-\infty}^{\infty}\frac{1}{\sigma_2\sqrt{2\pi}}f_2'(\tilde{x}_2)\exp\left(\frac{\tilde{x}_2^2}{\sigma_2^2}\right)d\tilde{x}_2, \tag{4.52}$$

resulting in

$$\mathbb{E}\left\{x_1f_2(x_2)\right\} = \sigma_1\sigma_2\rho_{x_1x_2}\int_{-\infty}^{\infty}\frac{1}{\sigma_2\sqrt{2\pi}}f_2'(\tilde{x}_2)\exp\left(\frac{\tilde{x}_2^2}{\sigma_2^2}\right)d\tilde{x}_2. \tag{4.53}$$

Next, we use

$$f_2(x)=Q(x)=\sum_{i=1}^{2^\alpha}l^i\left(u\left(x-l^{\mathrm{lo},i}\right)-u\left(x-l^{\mathrm{up},i}\right)\right)=l^1+\sum_{i=2}^{2^\alpha}\left(l^i-l^{i-1}\right)u\left(x-l^{\mathrm{lo},i}\right). \tag{4.54}$$

Using the fact $l^{\text{up},i} = l^{\text{lo},i+1}$, we have

$$\frac{\partial Q(x)}{\partial x} = \sum_{i=1}^{2^\alpha} l^i \left(\delta\left(x - l^{\text{lo},i}\right) - \delta\left(x - l^{\text{up},i}\right)\right) = l^1 + \sum_{i=2}^{2^\alpha} \left(l^i - l^{i-1}\right) \delta\left(x - l^{\text{lo},i}\right), \quad (4.55)$$

where δ is the Dirac Delta function. Therefore

$$\mathbb{E}\left\{x_1 Q(x_2)\right\} = \sigma_1 \rho_{x_1 x_2} \sum_{i=1}^{2^\alpha} \frac{1}{\sqrt{2\pi}} l^i \left(\exp\left(-\frac{\left(l^{\text{lo},i}\right)^2}{2\sigma_2^2}\right) - \exp\left(-\frac{\left(l^{\text{up},i}\right)^2}{2\sigma_2^2}\right)\right). \quad (4.56)$$

Next, we find the covariance at the output of the quantizer as follows [55]

$$\mathbb{E}\left\{Q(x_1)Q(x_2)\right\} = \sum_{i=2}^{K}\sum_{k=2}^{K} \Delta^2 \int_0^{\rho_{x_1 x_2}} \frac{\exp\left(-\frac{1}{2\left(1-\rho'^2\right)}\left[\frac{\left(l^{\text{lo},i}\right)^2}{\sigma_1^2} + \frac{\left(l^{\text{lo},k}\right)^2}{\sigma_2^2} - \frac{2\rho' l^{\text{lo},i} l^{\text{lo},k}}{\sigma_1 \sigma_2}\right]\right)}{2\pi\sqrt{1-\rho'^2}} d\rho'. \quad (4.57)$$

Note that in cell-free massive MIMO, we have $\rho_{x_1 x_2} = \rho_{mnk}$, where ρ_{mnk} is defined in Appendix 4.B. Finally, $\mathbf{C}_{\mathbf{r}_k \mathbf{r}_k}$, $\mathbf{C}_{\mathbf{r}_k \mathbf{z}_k}$ and $\mathbf{C}_{\mathbf{z}_k \mathbf{r}_k}$ are determined using (4.56) and (4.57) and the following equalities:

$$\mathbf{C}_{\mathbf{r}_k \mathbf{r}_k} = \mathbb{E}\left\{\mathbf{r}_k \mathbf{r}_k\right\} = \mathbb{E}\left\{Q(\mathbf{z}_k) Q(\mathbf{z}_k)\right\}, \quad (4.58a)$$

$$\mathbf{C}_{\mathbf{z}_k \mathbf{r}_k} = \mathbb{E}\left\{\mathbf{z}_k \mathbf{r}_k\right\} = \mathbb{E}\left\{\mathbf{z}_k Q(\mathbf{z}_k)\right\}, \quad (4.58b)$$

$$\mathbf{C}_{\mathbf{r}_k \mathbf{z}_k} = \mathbb{E}\left\{\mathbf{r}_k \mathbf{z}_k\right\} = \mathbb{E}\left\{Q(\mathbf{z}_k) \mathbf{z}_k\right\}. \quad (4.58c)$$

Based on the above derivations, $\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mn}$ is a function of the number of quantization bits α, the step size of the quantizer Δ, and the correlation coefficient between the inputs of the quantizer at the mth and nth APs, i.e., ρ_{mnk} (which is given in the following proposition).

Proposition 6. *The correlation coefficient between the inputs of the quantizers at the mth and the nth APs is obtained as follows:*

$$\rho_{mnk} = \frac{N^2 c_{mk} c_{nk} \tau_p p_p \sum_{k'=1}^{K} \rho_d q_{k'} \beta_{mk'} \beta_{nk'} \left|\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k\right|^2}{\sqrt{\left(N^2 \sum_{k'=1}^{K} \gamma_{mk'}^2 \left|\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k\right|^2 \rho_d q_{k'} + N\gamma_{mk} \sum_{k'=1}^{K} \beta_{mk'} \rho_d q_{k'} + N\gamma_{mk}\right)\left(N^2 \sum_{k'=1}^{K} \gamma_{nk'}^2 \left|\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k\right|^2 \rho_d q_{k'} + N\gamma_{nk} \sum_{k'=1}^{K} \beta_{nk'} \rho_d q_{k'} + N\gamma_{nk}\right)}}. \quad (4.59)$$

Proof: To remind, note that the input of the quantizer at the mth AP, z_m, is given by

$$z_{mk} = \hat{\mathbf{g}}_{mk}^{*} \mathbf{y}_m, \tag{4.60}$$

where the MMSE estimate of the channel coefficient between the kth user and the mth AP, i.e., $\hat{\mathbf{g}}_{mk}$, is given by

$$\hat{\mathbf{g}}_{mk} = c_{mk} \left(\sqrt{\tau_p p_p} \mathbf{g}_{mk} + \sqrt{\tau_p p_p} \sum_{k' \neq k}^{K} \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^{H} \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \right), \tag{4.61}$$

and the received signal at the mth AP from all users is given by

$$\mathbf{y}_m = \sqrt{\rho} \sum_{k=1}^{K} \mathbf{g}_{mk} \sqrt{q_k} s_k + \mathbf{n}_m. \tag{4.62}$$

We aim to calculate the correlation coefficient between the mth and nth APs, which is given by

$$\rho_{mnk} = \frac{\left| \mathbb{E}\left\{ z_{mk} z_{nk}^{*} \right\} \right|}{\sigma_{z_{mk}} \sigma_{z_{nk}}}, \tag{4.63}$$

where $\sigma_{z_{mk}}^2$ and $\mathbb{E}\left\{ z_{mk} z_{nk}^{*} \right\}$ and we have $\sigma_{z_{mk}}^2 = \mathbb{E}\left\{ \left| z_{mk} \right|^2 \right\}$. Next, using (4.60)-(4.62), we have

$$\begin{aligned}
&\mathbb{E}\left\{ \left(z_{mk} \right)^{*} \times z_{nk} \right\} = \mathbb{E}\left\{ \left(\hat{\mathbf{g}}_{mk}^{H} \mathbf{y}_m \right)^{*} \left(\hat{\mathbf{g}}_{nk}^{H} \mathbf{y}_n \right) \right\} \\
&= \mathbb{E}\left\{ \left(\hat{\mathbf{g}}_{mk}^{H} \left(\sqrt{\rho} \sum_{k'=1}^{K} \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'} + \mathbf{n}_m \right) \right)^{H} \left(\hat{\mathbf{g}}_{nk}^{H} \left(\sqrt{\rho} \sum_{k'=1}^{K} \mathbf{g}_{nk'} \sqrt{q_{k'}} s_{k'} + \mathbf{n}_n \right) \right) \right\} \\
&= \mathbb{E}\left\{ \left(\sqrt{\rho} \sum_{k'=1}^{K} \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'} \right)^{H} \hat{\mathbf{g}}_{mk} \hat{\mathbf{g}}_{nk}^{H} \left(\sqrt{\rho} \sum_{k'=1}^{K} \mathbf{g}_{nk'} \sqrt{q_{k'}} s_{k'} \right) \right\} \\
&= \underbrace{\mathbb{E}\left\{ \left(\sqrt{\rho} \mathbf{g}_{mk} \sqrt{q_k} s_k \right)^{H} \hat{\mathbf{g}}_{mk} \hat{\mathbf{g}}_{nk}^{H} \left(\sqrt{\rho} \mathbf{g}_{nk} \sqrt{q_k} s_k \right) \right\}}_{A} \\
&+ \sum_{k' \neq k}^{K} \underbrace{\mathbb{E}\left\{ \left(\sqrt{\rho} \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'} \right)^{H} \hat{\mathbf{g}}_{mk} \hat{\mathbf{g}}_{nk}^{H} \left(\sqrt{\rho} \mathbf{g}_{nk'} \sqrt{q_{k'}} s_{k'} \right) \right\}}_{B},
\end{aligned} \tag{4.64}$$

where in the last equality we used the fact that $\mathbb{E}\{s_k^* s_{k'}\} = 0$. Next, we have

$$\begin{aligned} A &= \rho q_k c_{mk} c_{nk} \mathbb{E}\left\{ (\mathbf{g}_{mk} s_k)^H \left(\sqrt{\tau_p p_p}\mathbf{g}_{mk} + \sqrt{\tau_p p_p} \sum_{k'' \neq k}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \right) \right. \\ &\quad \left. \left(\sqrt{\tau_p p_p}\mathbf{g}_{nk} + \sqrt{\tau_p p_p} \sum_{k'' \neq k}^{K} \mathbf{g}_{nk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,n} \boldsymbol{\phi}_k \right)^H (\mathbf{g}_{nk} s_k) \right\} \\ &= \rho q_k c_{mk} c_{nk} \tau_p p_p \mathbb{E}\left\{ \mathbf{g}_{mk}^H \mathbf{g}_{mk} \mathbf{g}_{nk}^H \mathbf{g}_{nk} \right\} \underbrace{\mathbb{E}\left\{ |s_k|^2 \right\}}_{1} \\ &= N^2 \rho q_k c_{mk} c_{nk} \tau_p p_p \beta_{mk} \beta_{nk}, \end{aligned} \tag{4.65}$$

and

$$B = \rho q_{k'} \underbrace{\mathbb{E}\left\{ \mathbf{g}_{mk'}^H \hat{\mathbf{g}}_{mk} \hat{\mathbf{g}}_{nk}^H \mathbf{g}_{nk'} \right\}}_{I} \underbrace{\mathbb{E}\left\{ |s_{k'}|^2 \right\}}_{1}, \forall k' \neq k, \tag{4.66}$$

where for $k' \neq k$, we have

$$\begin{aligned} I &= \mathbb{E}\left\{ c_{mk} \mathbf{g}_{mk'}^H \left(\sqrt{\tau_p p_p} \sum_{l=1}^{K} \mathbf{g}_{ml} \boldsymbol{\phi}_l^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \right) \right. \\ &\quad \left. c_{nk} \left(\sqrt{\tau_p p_p} \sum_{l=1}^{K} \mathbf{g}_{nl} \boldsymbol{\phi}_l^H \boldsymbol{\phi}_k + \mathbf{W}_{p,n} \boldsymbol{\phi}_k \right)^H \mathbf{g}_{nk'} \right\} \\ &= c_{mk} c_{nk} \tau_p p_p \mathbb{E}\left\{ \mathbf{g}_{mk'}^H \mathbf{g}_{mk'} \mathbf{g}_{nk'}^H \mathbf{g}_{nk'} \right\} \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 \\ &= N^2 c_{mk} c_{nk} \tau_p p_p \beta_{mk'} \beta_{nk'} \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2. \end{aligned} \tag{4.67}$$

Finally, we have

$$\begin{aligned} \mathbb{E}\left\{ (z_{mk})^* \times z_{nk} \right\} &= N^2 \rho q_k c_{mk} c_{nk} \tau_p p_p \beta_{mk} \beta_{nk} \\ &\quad + N^2 c_{mk} c_{nk} \tau_p p_p \rho \sum_{k' \neq k}^{K} \beta_{mk'} \beta_{nk'} q_{k'} \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 \\ &= N^2 c_{mk} c_{nk} \tau_p p_p \rho \sum_{k'=1}^{K} q_{k'} \beta_{mk'} \beta_{nk'} \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2. \end{aligned} \tag{4.68}$$

Next, we calculate the power of the input signal of the quantizer, i.e., $\mathbb{E}\left\{ |z_{mk}|^2 \right\}$ as

follows

$$\begin{aligned}
&\mathbb{E}\left\{|z_{mk}|^2\right\} = \mathbb{E}\left\{\left(\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m\right)^H \left(\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m\right)\right\} \\
&= \mathbb{E}\left\{\left(\hat{\mathbf{g}}_{mk}^H \left(\sqrt{\rho}\sum_{k'=1}^{K} \mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} + \mathbf{n}_m\right)\right)^H \left(\hat{\mathbf{g}}_{mk}^H \left(\sqrt{\rho}\sum_{k'=1}^{K} \mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'} + \mathbf{n}_m\right)\right)\right\} \\
&= \underbrace{\rho\,\mathbb{E}\left\{\left(\hat{\mathbf{g}}_{mk}^H \left(\sum_{k'=1}^{K} \mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)\right)^H \left(\hat{\mathbf{g}}_{mk}^H \left(\sum_{k'=1}^{K} \mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)\right)\right\}}_{A} \\
&+ \mathbb{E}\left\{\left|\hat{\mathbf{g}}_{mk}^H \mathbf{n}_m\right|^2\right\}. \qquad (4.69)
\end{aligned}$$

For the second term of (4.69), we have

$$\mathbb{E}\left\{\left|\hat{\mathbf{g}}_{mk}^H \mathbf{n}_m\right|^2\right\} = N\gamma_{mk}, \qquad (4.70)$$

and the first term in (4.69) can be obtained as

$$
\begin{aligned}
A &= \mathbb{E}\left\{\left(\hat{\mathbf{g}}_{mk}^H\left(\sum_{k'=1}^{K}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)\right)^H\left(\hat{\mathbf{g}}_{mk}^H\left(\sum_{k'=1}^{K}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)\right)\right\} \\
&= \mathbb{E}\left\{\left(\sum_{k'=1}^{K}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)^H\hat{\mathbf{g}}_{mk}\hat{\mathbf{g}}_{mk}^H\left(\sum_{k'=1}^{K}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)\right\} \\
&= \mathbb{E}\left\{\left(\sum_{k'=1}^{K}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)^H\left(c_{mk}\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)\right)\right. \\
&\quad \left.\left(c_{mk}\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)\right)^H\left(\sum_{k'=1}^{K}\mathbf{g}_{mk'}\sqrt{q_{k'}}s_{k'}\right)\right\} \\
&= c_{mk}^2\sum_{k'=1}^{K}q_{k'}\mathbb{E}\left\{\left(\mathbf{g}_{mk'}s_{k'}\right)^H\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)\right. \\
&\quad \left.\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)^H\left(\mathbf{g}_{mk'}s_{k'}\right)\right\} \\
&= q_k c_{mk}^2\mathbb{E}\left\{\left(\mathbf{g}_{mk}s_{k}\right)^H\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)\right. \\
&\quad \left.\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)^H\left(\mathbf{g}_{mk}s_{k}\right)\right\} \\
&\quad + c_{mk}^2\sum_{k'\neq k}^{K}q_{k'}\mathbb{E}\left\{\left(\mathbf{g}_{mk'}s_{k'}\right)^H\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)\right. \\
&\quad \left.\left(\sqrt{\tau_p p_p}\mathbf{g}_{mk}+\sqrt{\tau_p p_p}\sum_{k''\neq k}^{K}\mathbf{g}_{mk''}\boldsymbol{\phi}_{k''}^H\boldsymbol{\phi}_k+\mathbf{W}_{p,m}\boldsymbol{\phi}_k\right)^H\left(\mathbf{g}_{mk'}s_{k'}\right)\right\}. \qquad (4.71)
\end{aligned}
$$

For the first bracket in last term in (4.71), we have:

$$
\begin{aligned}
I_1 &= c_{mk}^2 \mathbb{E}\left\{ (\mathbf{g}_{mk} s_k)^H \left(\sqrt{\tau_p p_p}\mathbf{g}_{mk} + \sqrt{\tau_p p_p} \sum_{k'' \neq k}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \right) \right. \\
&\quad \left. \left(\sqrt{\tau_p p_p}\mathbf{g}_{mk} + \sqrt{\tau_p p_p} \sum_{k'' \neq k}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \right)^H (\mathbf{g}_{mk} s_k) \right\} \\
&= c_{mk}^2 \tau_p p_p \mathbb{E}\left\{ \mathbf{g}_{mk}^H \mathbf{g}_{mk} \mathbf{g}_{mk}^H \mathbf{g}_{mk} \right\} + c_{mk}^2 \mathbb{E}\left\{ \mathbf{g}_{mk}^H \left(\sqrt{\tau_p p_p} \sum_{k'' \neq k}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \right) \right. \\
&\quad \left. \left(\sqrt{\tau_p p_p} \sum_{k'' \neq k}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \right)^H \mathbf{g}_{mk} \right\} \\
&= c_{mk}^2 \tau_p p_p N(N+1) \beta_{mk}^2 + N c_{mk}^2 \beta_{mk} \left(\tau_p p_p \sum_{k'' \neq k}^{K} \beta_{mk''} \left| \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k \right|^2 + 1 \right) \\
&= c_{mk}^2 \tau_p p_p N^2 \beta_{mk}^2 + c_{mk}^2 \tau_p p_p N \beta_{mk}^2 + N c_{mk}^2 \beta_{mk} \left(\tau_p p_p \sum_{k'' \neq k}^{K} \beta_{mk''} \left| \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k \right|^2 + 1 \right) \\
&= N^2 \gamma_{mk}^2 + N c_{mk}^2 \beta_{mk} \left(\tau_p p_p \sum_{k''=1}^{K} \beta_{mk''} \left| \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k \right|^2 + 1 \right) \\
&= N^2 \gamma_{mk}^2 + N c_{mk}^2 \beta_{mk} \frac{\sqrt{\tau_p p_p} \beta_{mk}}{c_{mk}} \\
&= N^2 \gamma_{mk}^2 + N \beta_{mk} \gamma_{mk}.
\end{aligned} \tag{4.72}
$$

Next, for the second bracket in last term in (4.71), we have

$$
\begin{aligned}
I_2 = c_{mk}^2 \mathbb{E}\Bigg\{ (\mathbf{g}_{mk'} s_{k'})^H \Bigg(\sqrt{\tau_p p_p} \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \sqrt{\tau_p p_p} \sum_{k'' \neq k'}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \Bigg) \\
\Bigg(\sqrt{\tau_p p_p} \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \sqrt{\tau_p p_p} \sum_{k'' \neq k'}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \Bigg)^H (\mathbf{g}_{mk'} s_{k'}) \Bigg\} \\
= c_{mk}^2 \Bigg[\tau_p p_p \mathbb{E}\left\{ \mathbf{g}_{mk'}^H \mathbf{g}_{mk'} \mathbf{g}_{mk'}^H \mathbf{g}_{mk'} \right\} \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 + \mathbb{E}\Bigg\{ \mathbf{g}_{mk'}^H \Bigg(\sqrt{\tau_p p_p} \sum_{k'' \neq k'}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \Bigg) \\
\Bigg(\sqrt{\tau_p p_p} \sum_{k'' \neq k'}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \Bigg)^H \mathbf{g}_{mk'} \Bigg\} \Bigg] \\
= c_{mk}^2 \Bigg[\tau_p p_p N(N+1) \beta_{mk'}^2 \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 + \mathbb{E}\Bigg\{ \mathbf{g}_{mk'}^H \Bigg(\sqrt{\tau_p p_p} \sum_{k'' \neq k'}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \Bigg) \\
\Bigg(\sqrt{\tau_p p_p} \sum_{k'' \neq k'}^{K} \mathbf{g}_{mk''} \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k + \mathbf{W}_{p,m} \boldsymbol{\phi}_k \Bigg)^H \mathbf{g}_{mk'} \Bigg\} \Bigg] \\
= c_{mk}^2 \Bigg[\tau_p p_p N(N+1) \beta_{mk'}^2 \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 + N \beta_{mk'} \Bigg(\tau_p p_p \sum_{k'' \neq k'}^{K} \beta_{mk''} \left| \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k \right|^2 + 1 \Bigg) \Bigg] \\
= c_{mk}^2 \Bigg[\tau_p p_p N^2 \beta_{mk'}^2 \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 + N \beta_{mk'} \Bigg(\tau_p p_p \sum_{k''=1}^{K} \beta_{mk''} \left| \boldsymbol{\phi}_{k''}^H \boldsymbol{\phi}_k \right|^2 + 1 \Bigg) \Bigg] \\
= c_{mk}^2 \tau_p p_p N^2 \beta_{mk'}^2 \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 + N c_{mk}^2 \beta_{mk'} \frac{\sqrt{\tau_p p_p} \beta_{mk}}{c_{mk}} \\
= N^2 \gamma_{mk'}^2 \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 + N \beta_{mk'} \gamma_{mk}. \qquad (4.73)
\end{aligned}
$$

Hence, using (4.72) and (4.73), the term A in (4.69) is obtained as

$$
A = N^2 \sum_{k'=1}^{K} \gamma_{mk'}^2 \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 q_{k'} + N \gamma_{mk} \sum_{k'=1}^{K} \beta_{mk'} q_{k'}. \qquad (4.74)
$$

Finally, using (4.69), (4.70), and (4.74), we have

$$
\mathbb{E}\left\{ |z_{mk}|^2 \right\} = N^2 \sum_{k'=1}^{K} \gamma_{mk'}^2 \left| \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k \right|^2 \rho q_{k'} + N \gamma_{mk} \sum_{k'=1}^{K} \beta_{mk'} \rho q_{k'} + N \gamma_{mk}. \qquad (4.75)
$$

As a result, using (4.75) and (4.68) we have:

$$\mathbb{E}\{|z_{mk}|^2\} = N^2 \sum_{k'=1}^{K} \gamma_{mk'}^2 \left|\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k\right|^2 \rho q_{k'} + N\gamma_{mk} \sum_{k'=1}^{K} \beta_{mk'} \rho q_{k'} + N\gamma_{mk}, \tag{4.76a}$$

$$\mathbb{E}\left\{(z_{mk})^H \times z_{nk}\right\} = N^2 c_{mk} c_{nk} \tau_p p_p \sum_{k'=1}^{K} \rho q_{k'} \beta_{mk'} \beta_{nk'} \left|\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k\right|^2. \tag{4.76b}$$

Next, for simplicity, let us consider the case of orthogonal pilots, where $\gamma_{mk} = \dfrac{\beta_{mk}^2}{\beta_{mk} + \frac{1}{\tau_p}} = \dfrac{\beta_{mk}^2}{a_{mk}}$. Therefore, we have

$$\rho_{mnk} = \tag{4.77}$$

$$\sqrt{\frac{\beta_{mk}^2 \beta_{nk}^2}{\beta_{mk}^2 \beta_{nk}^2 + a_{nk}\left(\frac{\beta_{mk}^2}{N} + \frac{a_{mk}}{N^2 \rho_d}\right) \sum_{k'=1}^{K} \beta_{nk'} + a_{mk}\left(\frac{\beta_{nk}^2}{N} + \frac{a_n}{N^2 \rho_d}\right) \sum_{k'=1}^{K} \beta_{mk'} + \frac{a_{mk} a_{nk}}{N^2} \sum_{k'=1}^{K} \beta_{mk'} \sum_{k'=1}^{K} \beta_{nk'} + \frac{a_{nk} \beta_{mk}^2 + a_{mk} \beta_{nk}^2}{N \rho_d} + \frac{a_{mk} a_{nk}}{N^2 \rho_d^2}}}.$$

Exploiting the definition of ρ_{mnk} in (4.77), one can conclude that under the conditions listed below, the correlation coefficient ρ_{mnk} between APs n and m is small enough, based on the results in Fig. 4.12, that the quantization distortions are approximately uncorrelated

1. As we have the term $\beta_{mk}^2 \beta_{nk}^2$ in both numerator and denominator of ρ_{mnk}, we need to have large path loss differences at APs n and m to avoid large ρ_{mnk},
2. As there are terms $\sum_{k'=1}^{K} \beta_{mk'}$, $\sum_{k'=1}^{K} \beta_{nk'}$ and $\sum_{k'=1}^{K} \beta_{mk'} \sum_{k'=1}^{K} \beta_{nk'}$ in the denominator of ρ_{mnk}, having a large number of users results in smaller ρ_{mnk},
3. Having small N results in small ρ_{mnk}.

The Actual Value of Correlation Between the Quantization Distortions Versus the Correlation Between the Inputs of the Quantizers

Fig. 4.12 demonstrates $\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mn}$ versus ρ_{mnk} for different number of quantization bits α, where the diagonal elements of the covariance matrix are obtained by setting $\rho_{mnk} = 1$. There are two important observations based on the results in Fig. 4.12, which are summarized as follows:

i) the elements of $\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mn}, \forall m \neq n, \alpha > 3$ are very small. Hence $\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}$ can be

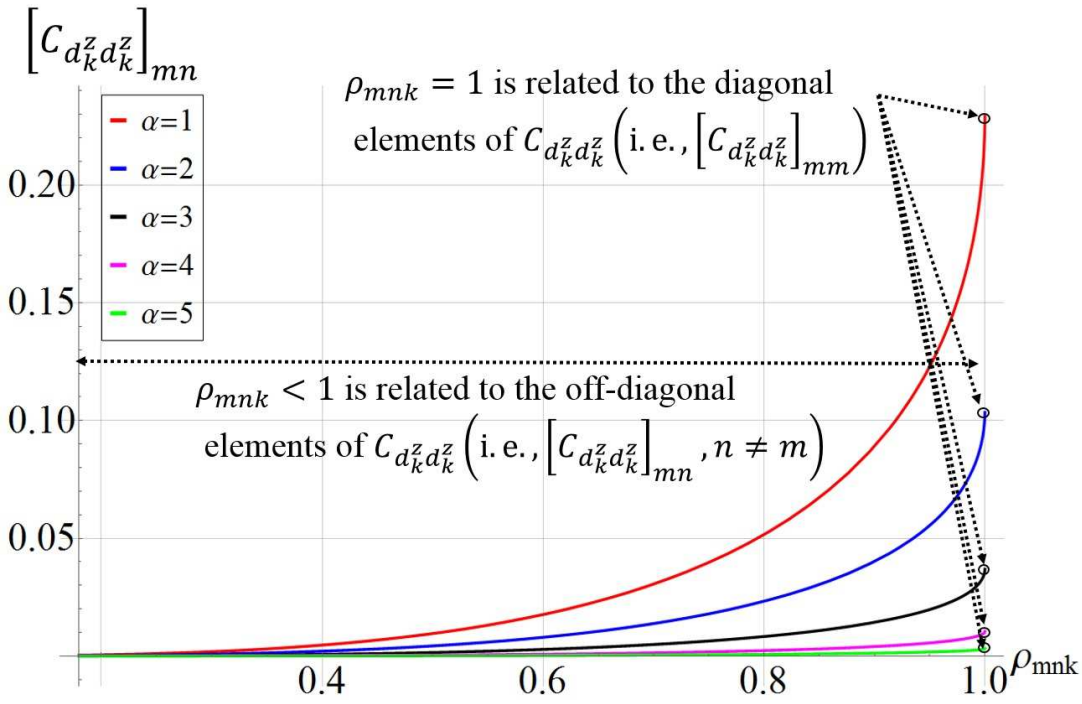

Figure 4.12: Figure presents the mnth element of the covariance matrix of the quantization distortion (i.e., $\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mn}$ in (4.46)) versus ρ_{mnk} for different number of quantization bits for a given user k. To remind, ρ_{mnk} refers to the correlation coefficient between the input of the quantizers at APs m and n. Note that the diagonal elements of the covariance matrix are obtained by setting $\rho_{mnk} = 1$.

approximated by a diagonal matrix for number of quantization bits $\alpha > 3$ (note that this verifies the assumption of a diagonal matrix for quantization distortions in [53–57]),

ii) Fig. 4.12 shows that the off-diagonal elements of the matrix $\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}$ are *small* compared to the diagonal elements of $\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}, \forall \alpha$ for $\rho_{mkn} \leq 0.75$. Here, by the term *small* we mean that $\dfrac{\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mm}}{\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mn}} \geq 10, \forall m, n, \alpha$ and with $\rho_{mkn} \leq 0.75$.

Below, we provide numerical results and discussion to address the effect of correlation between the inputs of the quantizers at different APs.

How Large is the Correlation Between the Input of the Quantizers at Different APs in Cell-Free Massive MIMO?

The CDF of ρ_{mnk} in cell-free massive MIMO with different system parameters is plotted in Figs. 4.13-4.15. The figure shows that with the practical parameters used in Figs.

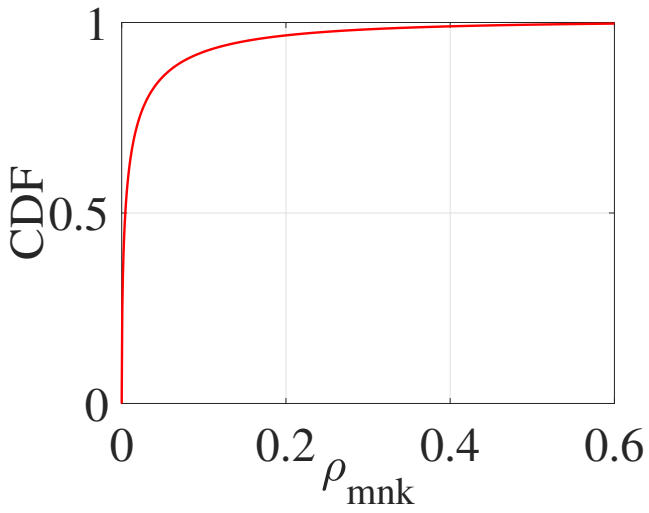

Figure 4.13: CDF of ρ_{mnk}, given in (4.60), in the cell-free massive MIMO system with different system parameters with $M = 60$, $N = 4$, $K = 20$, and $\tau_p = 20$.

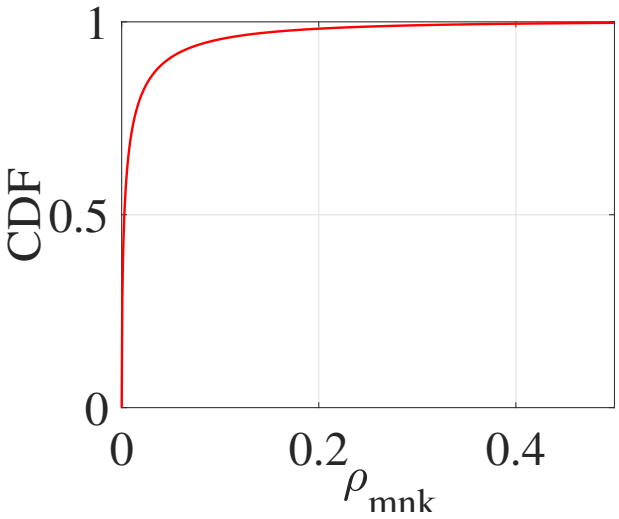

Figure 4.14: CDF of ρ_{mnk}, given in (4.60), in the cell-free massive MIMO system with different system parameters with $M = 80$, $N = 4$, $K = 30$, and $\tau_p = 30$.

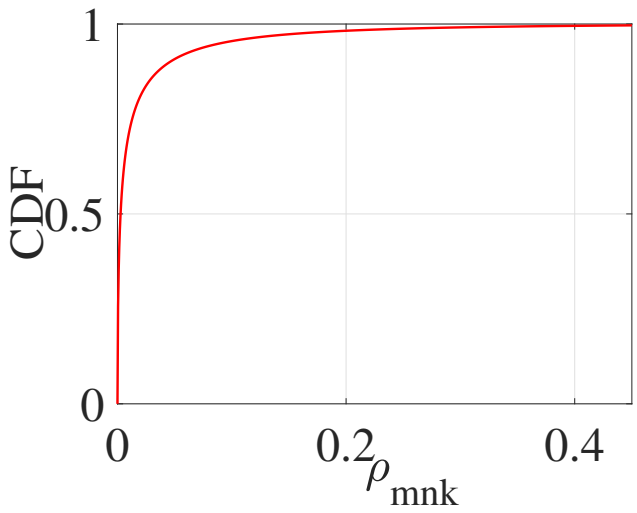

Figure 4.15: CDF of ρ_{mnk}, given in (4.60), in the cell-free massive MIMO system with different system parameters with $M = 60$, $N = 6$, $K = 40$, and $\tau_p = 30$..

4.13-4.15, and with probability of 95%, the correlation between the inputs of the quantizers at different APs is $\rho_{mn} \leq 0.4$. Next, based on Figs. 4.12, we can observe that with $\rho_{mnk} \leq 0.4$, the off-diagonal elements of the covariance matrix of the quantization distortion $\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mn}, \forall\, m \neq n$, are negligible compared to the diagonal elements, i.e., $\left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mm}$.

The Performance Gap Between the Exact Uplink Per-User Rate and the Uplink Per-User Rate While Ignoring the Correlation Between the Inputs of the Quantizers:

In this section, we present the uplink per-user rate with different system parameters for two different scenarios. To remind, the correlation between the quantization distortions at different APs is given by

$$\mathbb{E}\left\{|\mathrm{TQD}_k|^2\right\} = \sum_{m=1}^{M} \left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{mm} + \sum_{m=1}^{M} \sum_{n \neq m}^{M} \left[\mathbf{C}_{\mathbf{d}_k^z \mathbf{d}_k^z}\right]_{nm}, \tag{4.78}$$

The SINR obtained by the exact value of $\mathbb{E}\left\{|\mathrm{TQD}_k|^2\right\}$ in (4.79) is referred to as "Exact" in Fig. 4.16-4.18. Next, we exploit the results in Fig. 4.12 and Fig. 4.13-4.15, the covariance matrix of the quantization distortion is approximated with a diagonal

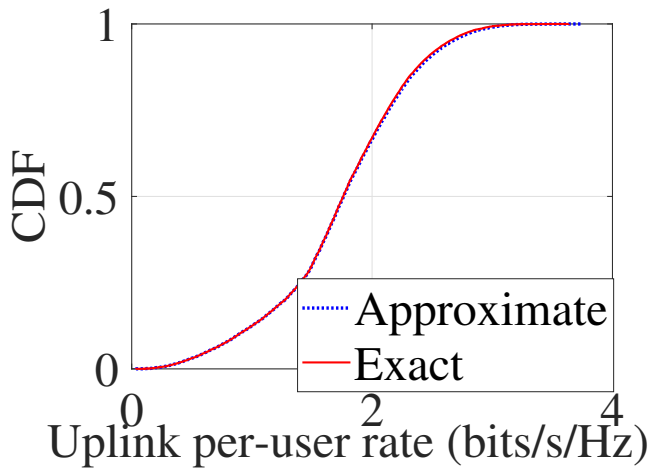

Figure 4.16: Uplink per-user rate of cell-free massive MIMO with. Here, the term "Exact" refers to the case where we include the correlation between the quantization distortions at different APs whereas the term "Approximate" refers to the case when we ignore the correlations between the error at different APs. In all figures, we set $\alpha = 2$ quantization bits, and use equal power allocation with $M = 60$, $N = 4$, $K = 20$, and $\tau_p = 20$.

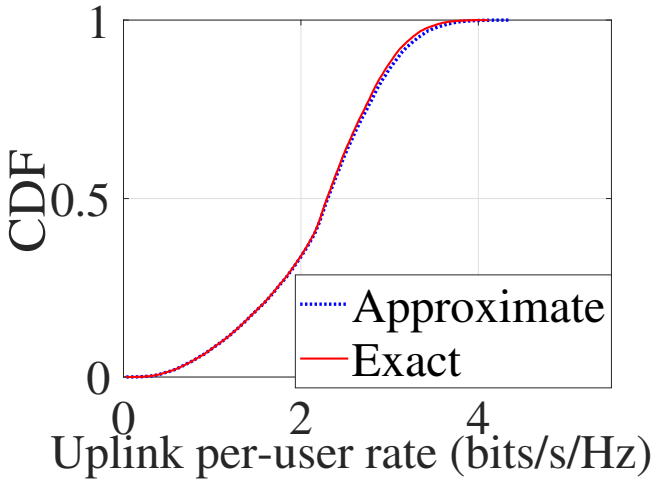

Figure 4.17: Uplink per-user rate of cell-free massive MIMO with. Here, the term "Exact" refers to the case where we include the correlation between the quantization distortions at different APs whereas the term "Approximate" refers to the case when we ignore the correlations between the error at different APs. In all figures, we set $\alpha = 2$ quantization bits, and use equal power allocation with $M = 80$, $N = 4$, $K = 30$, and $\tau_p = 30$.

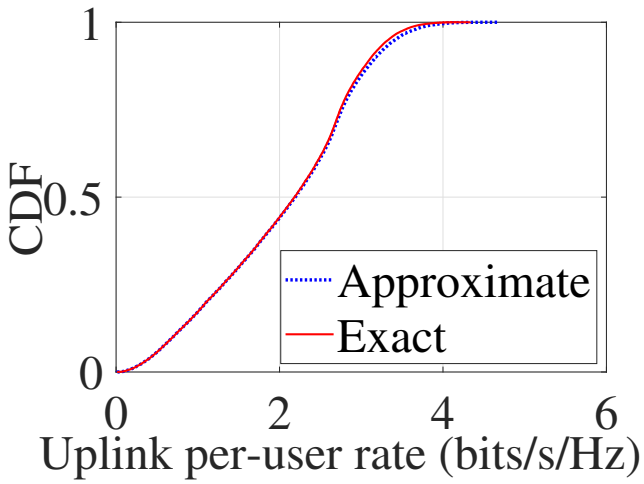

Figure 4.18: Uplink per-user rate of cell-free massive MIMO with. Here, the term "Exact" refers to the case where we include the correlation between the quantization distortions at different APs whereas the term "Approximate" refers to the case when we ignore the correlations between the error at different APs. In all figures, we set $\alpha = 2$ quantization bits, and use equal power allocation with $M = 60$, $N = 6$, $K = 40$, and $\tau_p = 30$.

matrix as follows:

$$\mathbb{E}\left\{|\mathrm{TQD}_k|^2\right\} = \sum_{m=1}^{M}\left[\mathbf{C}_{\mathbf{d}_k^z\mathbf{d}_k^z}\right]_{mm} + \sum_{m=1}^{M}\sum_{n\neq m}^{M}\left[\mathbf{C}_{\mathbf{d}_k^z\mathbf{d}_k^z}\right]_{nm} \approx \sum_{m=1}^{M}\left[\mathbf{C}_{\mathbf{d}_k^z\mathbf{d}_k^z}\right]_{mm}. \quad (4.79)$$

This scenario is given as "Approximate" in Fig. 4.16-4.18. As Fig. 4.16-4.18 shows, there is a negligible performance gap between the exact SINR and the approximate SINR. Note that exploiting the exact definition for SINR, the optimization problem cannot be approximated with a convex problem. Hence, here we consider the uncorrelated quantization distortions at different APs, and the total quantization distortion is obtained as follows:

$$\begin{aligned}\mathbb{E}\left\{|\mathrm{TQD}_k|^2\right\} &= \mathbb{E}\left\{\left|\sum_{m=1}^{M} u_{mk} n_{d,mk}\right|^2\right\} \quad (4.80)\\ &\approx \sum_{m=1}^{M} u_{mk}^2 \mathbb{E}\left\{|n_{d,mk}|^2\right\}\\ &= \left(\tilde{b}-\tilde{a}^2\right)\sum_{m=1}^{M} u_{mk}^2\left[N^2\sum_{k'=1}^{K}\gamma_{mk'}^2\left|\boldsymbol{\phi}_{k'}^H\boldsymbol{\phi}_k\right|^2\rho q_{k'} + N\gamma_{mk}\sum_{k'=1}^{K}\beta_{mk'}\rho q_{k'} + N\gamma_{mk}\right].\end{aligned}$$

By substituting (4.38), (4.39), (4.43), (4.44) and (4.81) into (4.6), the corresponding SINR of the kth user is obtained by (4.8), which completes the proof of Theorem 5. ■

Appendix 4.B: Details of Finding $\nu^\star$ in Remark 3

Assuming a total transmit power of $\sum_{k=1}^{K} q_k$, the power minimization problem can be defined as follows:

$$P_{13}: \min_{q_k} \sum_{k=1}^{K} q_k \tag{4.81a}$$

$$\text{s.t.} \quad S_k\left(q_k, \mathbf{u}_k, \alpha\right) \geq S_k^{(\text{r})}, \ \ \forall k, \tag{4.81b}$$

$$0 \leq q_k \leq p_{\text{max}}^{(k)}, \forall k. \tag{4.81c}$$

Problem P_{13} is a GP and can be efficiently solved. After solving Problem P_{13} and finding the optimal solution $q_k^+, \forall k$, the slack variable $\nu^\star$ is obtained as follows:

$$\nu^\star = \frac{\sum_{k=1}^{K} p_{\text{max}}^{(k)}}{\sum_{k=1}^{K} q_k^+}, \tag{4.82}$$

which completes the definition for Remark 3. ■

Appendix 4.C: Proof of Proposition 3

The standard form of GP is defined as follows [46, 50]:

$$P_{14}: \quad \min \quad f_0(\mathbf{x}), \tag{4.83a}$$

$$\text{subject to} \quad f_i(\mathbf{x}) \leq 1, \ \ i = 1, \cdots, m, \tag{4.83b}$$

$$g_i(\mathbf{x}) = 1, \ \ i = 1, \cdots, p, \tag{4.83c}$$

where f_0 and f_i are posynomial and g_i are monomial functions. Moreover, $\mathbf{x} = \{x_1, \cdots, x_n\}$ represents the optimization variables. The SINR constraint in (4.83) is not a posynomial function in its initial form, however it can be rewritten into the following posynomial function:

$$\frac{\mathbf{u}_k^H \left(N^2 \sum_{k' \neq k}^{K} q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \boldsymbol{\Delta}_{kk'} \boldsymbol{\Delta}_{kk'}^H + N \sum_{k'=1}^{K} q_{k'} \mathbf{D}_{kk'} + \frac{N}{\rho} \mathbf{R}_k\right) \mathbf{u}_k}{\mathbf{u}_k^H (N^2 q_k \boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_k^H) \mathbf{u}_k} \leq \frac{1}{t}, \ \forall k. \tag{4.84}$$

By applying a simple transformation, (4.84) is equivalent to the following inequality:

$$q_k^{-1}\left(\sum_{k'\neq k}^{K} a_{kk'}q_{k'}+\sum_{k'=1}^{K} b_{kk'}q_{k'}+c_k\right)\leq \frac{1}{t}, \tag{4.85}$$

where

$$a_{kk'}=\frac{\mathbf{u}_k^H\left(\left|\boldsymbol{\phi}_k^H\boldsymbol{\phi}_{k'}\right|^2\boldsymbol{\Delta}_{kk'}\boldsymbol{\Delta}_{kk'}^H\right)\mathbf{u}_k}{\mathbf{u}_k^H(\boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^H)\,\mathbf{u}_k}, \tag{4.86a}$$

$$b_{kk'}=\frac{\mathbf{u}_k^H\mathbf{D}_{kk'}\mathbf{u}_k}{\mathbf{u}_k^H\left(N\boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^H\right)\mathbf{u}_k}, \tag{4.86b}$$

$$c_k=\frac{\mathbf{u}_k^H\mathbf{R}_k\mathbf{u}_k}{\mathbf{u}_k^H\left(\rho N\boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^H\right)\mathbf{u}_k}. \tag{4.86c}$$

The transformation in (4.85) shows that the left-hand side of (4.84) is a posynomial function. Moreover, the spectral efficiency constraint in (4.24b) is not a posynomial function in its original form, however, through some mathematical manipulation, it can be written as:

$$\frac{\mathbf{u}_k^H\left(N^2\sum_{k'\neq k}^{K}q_{k'}|\boldsymbol{\phi}_k^H\boldsymbol{\phi}_{k'}|^2\boldsymbol{\Delta}_{kk'}\boldsymbol{\Delta}_{kk'}^H+N\sum_{k'=1}^{K}q_{k'}\mathbf{D}_{kk'}+\frac{N}{\rho}\mathbf{R}_k\right)\mathbf{u}_k}{\mathbf{u}_k^H(N^2q_k\boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^H)\mathbf{u}_k}\leq\frac{1}{\hat{S}_k^{(\mathrm{r})}},\ \forall k, \tag{4.87}$$

where $\hat{S}_k^{(\mathrm{r})}=2^{\frac{\tau_c S_k^{(\mathrm{r})}}{\tau_c-\tau_p}}-1$. By applying a simple transformation, (4.87) is equivalent to the following inequality:

$$q_k^{-1}\left(\sum_{k'\neq k}^{K} a_{kk'}q_{k'}+\sum_{k'=1}^{K} b_{kk'}q_{k'}+c_k\right)\leq \frac{1}{\hat{S}_k^{(\mathrm{r})}}, \tag{4.88}$$

where $a_{kk'}$, $b_{kk'}$, and c_k are given in (4.86a)-(4.86c). Therefore, the power allocation problem P_6 is a standard GP (convex problem), where the objective function and constraints are monomial and posynomial, respectively, which completes the proof of Proposition 3. ■

Chapter 5

Evaluation of Low Complexity Massive MIMO Techniques Under Realistic Channel Conditions

Contents

5.1 Introduction

To investigate the performance of massive multiple-input multiple-output (MIMO) systems, an accurate multi-user channel model is necessary. Most standardized MIMO channel models such as IEEE 802.11, the 3GPP spatial model, and the COST 273 model rely on clustering [71]. The GSCMs consider the physical reality of channels to

investigate the performance of MIMO systems using the concept of clusters [21]. The COST 2100 model is a well known GSCM [21].

The problem of correlation-based user scheduling and beamforming (CUSBF) in cluster-based channel models and its effect on the system performance of massive MIMO has not been well studied in the literature. In this chapter, we investigate the problem of joint user scheduling and beamforming design when only knowledge of the statistics of the channel is available at the BS. The second order statistics of the channel depend on the position of the users and the geometry of the system, including the relative position of clusters in the area with respect to the BS and users. The fixed positions of the users and clusters mean that channel variation is stationary. In the other words, if the geometry of the system is fixed, the channel covariance matrix remains constant over time. Moreover, changing the position of the users by a few meters will not affect the statistics of the channel [72]. In [72], the authors quantize the given interval for the angle of departure (AoD) of paths into M angular bins of size $\frac{1}{M}$; while exploiting the concept of bins, the current chapter uses an approximated version of the channel matrix for beamforming design. In general, multi-path components (MPCs) from common clusters cause correlation which reduces the rank of the channel. We therefore work in this chapter on the effect of common bins on the system performance.

Given the second order statistics of the channel, we perform low-complex user scheduling and precoding based only on the covariance matrix of the users. The behaviour of the eigenvalues of the channel covariance matrix for a large number of antennas at the BS is studied. When the number of antennas tends to infinity, based on Szego's theorem for large Toeplitz matrices [72, 73], the eigenvalue spectrum of the channel covariance matrix can be obtained by the discrete-time Fourier transform of the antenna correlation function. Assume M denotes the number of antennas in a $\frac{\lambda}{2d}$ at the BS (where λ and d refer to the wavelength and spaced linear array, respectively), $\frac{1}{M}$ is a rough estimate of the angular resolution. Although the BS does not require the instantaneous channel information of the users but only their second-order statistics, analysis and numerical results for the proposed user scheduling and beamforming scheme show throughput superiority over others. The proposed user selection scheme relies on a trade off between the number of occupied spectral bins for each user and the spectral overlap among the selected users. This trade off is shown by introducing the

variable ϵ to quantify the correlation between clusters.

We exploit the COST 2100 channel model, which is a more realistic multi-user channel model than that assumed in [74], as the parameters of the COST model are based on real-world measurements. In this chapter, we assume that the frequency is 2 GHz. Note that clustered channels are much sparser in the millimetre-wave range considered in [74] compared to 2 GHz. In the proposed scheme, we first build up the approximate eigenchannel matrix for the channels of users based on the channel covariance matrix. Next, we propose that the BS generates the approximate eigenchannel matrix. Note that this is the first work which employs this approximate eigenchannel matrix. Moreover, this is a low-complexity practical scheme to implement a massive MIMO system. Our results and contributions are summarized as follows:

1. Exploiting the eigenvalue spectrum of the channel covariance matrix, we propose to use the angular bins to build up an approximate eigenchannel, which can be used for linear precoding design. Next, a new user scheduling scheme is proposed *under the assumption that no instantaneous channel information is available at the BS, other than the channel correlation*.

2. The complexity of the proposed scheme is presented.

3. Numerical results show significant sum rate performance improvement of the massive MIMO system compared to the joint spatial division and multiplexing (JSDM)-based scheduling scheme presented in [72]. Moreover, in [75], the BS exploits knowledge of the estimated channel to design the beamformer. Hence it is very difficult to achieve the performance of the greedy weight clique (GWC) scheme [75] knowing only the correlation matrix. The numerical results confirm that there is only a small gap between the performance of the proposed correlation-based scheme and the GWC scheme (which relies on the availability of the estimated channel at the BS).

5.2 System Model

Consider downlink transmission in a single cell with M antennas at the BS and K single antenna user terminals on the same time-frequency resource. Here, we assume time division duplexing (TDD) mode where the uplink and downlink channels are the same.

5.2.1 Downlink Transmission

The transmitted signal when $K_s\,(K_s \ll M)$ users have been selected from the pool of K users, is given by $\mathbf{x} = \sum_{k=1}^{K_s} \sqrt{p_k}\mathbf{w}_k s_k$, where s_k denotes the data symbol of user k, $\mathbf{w}_k$ denotes the precoding vector of size M and p_k denotes the power assigned to user k. Then the received signal at user k is given by

$$y_k = \sqrt{p_k}\mathbf{h}_k^T\mathbf{w}_k s_k + \sum_{j=1, j\neq k}^{K_s} \sqrt{p_j}\mathbf{h}_k^T\mathbf{w}_j s_j + n_k, \tag{5.1}$$

where the vector $\mathbf{h}_k$ of size M denotes the downlink channel of the kth ($k = 1, \cdots, K_s$) user and $n_k \in \mathbb{C}(0, 1)$ is the complex additive white Gaussian noise (AWGN).

5.2.2 Uplink Channel Estimation in Single-Cell Massive MIMO with Correlated Channel

In this section, we investigate the problem of estimating the channel in the TDD mode. Suppose $\mathbf{H}^T \in \mathbb{C}^{M\times K}$ represents the uplink aggregate channel matrix between the users and the BS. The channel covariance matrix $\mathbf{R} \in \mathbb{C}^{MK\times MK}$ is given by

$$\mathbf{R} = \mathbb{E}\left\{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H\right\}, \tag{5.2}$$

where $\tilde{\mathbf{h}} = \text{vec}(\mathbf{H}^T)$. For MMSE estimation of the channel, we use a pilot sequence [76], [77]. Let us assume $\mathbf{\Phi}_p \in \mathbb{C}^{K\times\tau_p}$ denotes pilot matrix, where τ_p is the length of pilot

sequence for each user. The received pilot signal $\mathbf{Y} \in \mathbb{C}^{M \times \tau_p}$ is given by

$$\mathbf{Y} = \mathbf{H}^T \mathbf{\Phi}_p + \mathbf{N}, \tag{5.3}$$

where $\text{vec}(\mathbf{N}) \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{M\tau_p})$ denotes circularly symmetric complex Gaussian noise and $\mathbf{I} \in \mathbb{C}^{M\tau_p \times M\tau_p}$ denotes the identity matrix. The Bayesian MMSE estimator of the channel is given by [76]

$$\tilde{\mathbf{h}}_{MMSE} = \mathbf{R}\tilde{\mathbf{\Phi}}_p{}^H \left(\tilde{\mathbf{\Phi}}_p \mathbf{R} \tilde{\mathbf{\Phi}}_p{}^H + \sigma_n^2 \mathbf{I}\right)^{-1} \tilde{\mathbf{y}}, \tag{5.4}$$

where $\tilde{\mathbf{\Phi}}_p = \mathbf{\Phi}_p^T \otimes \mathbf{I}$, where $\otimes$ is the Kronecker product, and $\tilde{\mathbf{y}} = \text{vec}(\mathbf{Y})$.

5.3 Eigenvalue Spectrum of the Antenna Correlation Function

In the COST 2100 channel model, each entry of the channel matrix can be written as

$$h_{km} = \sum_{i=1}^{N_l} a_{ki}\, \delta(\phi - \phi_{ki}) \delta(\theta - \theta_{ki}) \delta(\tau - \tau_{ki}), \tag{5.5}$$

where $N_l = N_C \times N_p$, and it denotes the total number of paths and $\boldsymbol{\phi}_{ki}$ and $\boldsymbol{\theta}_{ki}$ represent the DoD and DoA respectively of path i to the kth user. The complex amplitude of the ith MPC in (5.5) is given by

$$a_{ki} = \underbrace{L_p\, A_{\text{VR}}\, \sqrt{A_C}}_{\text{geometry-based attenuation}} \times \underbrace{A_{\text{MPC}}}_{\text{small-scale fading}} = a_{ki}^{\text{ga}} \times a_{ki}^{\text{sf}}. \tag{5.6}$$

Note that the power of each path in (5.6) is scaled with respect to the small-scale fading and the attenuation due to the geometry of the system which we call *geometry-based attenuation*. Hence, assuming a linear array response at the BS side the $K \times M$ aggregate

channel of all K users is given by

$$\mathbf{H} = \begin{bmatrix} \sum_{i=1}^{N_l} a_{1i} & \sum_{i=1}^{N_l} a_{1i} e^{j\alpha \sin \phi_{1i}} & \dots & \sum_{i=1}^{N_l} a_{1i} e^{j\alpha(M-1) \sin \phi_{1i}} \\ \sum_{i=1}^{N_l} a_{2i} & \sum_{i=1}^{N_l} a_{2i} e^{j\alpha \sin \phi_{2i}} & \dots & \sum_{i=1}^{N_l} a_{2i} e^{j\alpha(M-1) \sin \phi_{2i}} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N_l} a_{Ki} & \sum_{i=1}^{N_l} a_{Ki} e^{j\alpha \sin \phi_{Ki}} & \dots & \sum_{i=1}^{N_l} a_{Ki} e^{j\alpha(M-1) \sin \phi_{Ki}} \end{bmatrix}, \tag{5.7}$$

where $\alpha = -2\pi \frac{d}{\lambda}$, d is the spacing between two antenna elements and λ denotes the wavelength (in m). The $M \times M$ channel spatial covariance of the kth user channel vector is given by $\mathbf{R}_k = \mathbb{E}\{\mathbf{h}_k \mathbf{h}_k^H\}$. Assuming that the positions of users and clusters are fixed, the expectation is taken over the power of MPCs which have the Rayleigh fading distribution. Assuming a linear array response for the AoD ϕ and wide sense stationary (WSS) over the array (ϕ is the same for all antenna elements confirming that you have planar wave), each (m, n)-th entry of the channel covariance matrix for the kth user, $\mathbf{R}_k$, is given by

$$[\mathbf{R}_k]_{m,n} = \sum_{i=1}^{N_l} (a_{ki}^{\text{ga}})^2 \, e^{j\alpha(n-m) \sin \phi_{ki}}, \tag{5.8}$$

where the second equality comes from the fact that $\mathbb{E}\left\{|a_{ki}|^2\right\} = (a_{ki}^{\text{ga}})^2$. [11]

5.3.1 Eigenvalue Spectrum with $M \to \infty$

In [74], the authors exploit Szego's theory for large Toeplitz matrices [73], and show that for massive MIMO systems, the eigenvalue spectrum of the antenna correlation function converges to the discrete-time Fourier transform of the antenna correlation function. In other words, in the limit of a large number of antennas, the empirical eigenvalue cumulative CDF of the empirical eigenvalues from the channel correlation matrix can be approximated by the samples of the discrete-time Fourier transform of the antenna correlation function [74]. The eigenvalue spectrum, $S_k(f)$, is obtained by the discrete-time Fourier transform of the autocorrelation function. Hence, we consider the spectrum over the range $f \in \left[\frac{-1}{2}, \frac{1}{2}\right]$ (the eigenvalue spectrum is wrapped around the

[1] [1]Note that the measurement results in [78] show that at the frequency of 2 GHz, to calculate the channel covariance matrix, the BS needs to average the channel samples over around 300-400 samples and 100-200 samples for the case of urban and rural environments, respectively.

interval $f \in \left[\frac{-1}{2}, \frac{1}{2}\right]$ by the periodicity of the discrete-time Fourier transform.). As the eigenvalue spectrum can take any positive real value, similar to [74], we write $S_k(f); f \in \left[\frac{-1}{2}, \frac{1}{2}\right] \rightarrow \mathbb{R}^+$, where $\mathbb{R}^+ = \{x \in \mathbb{R} | x > 0\}$ refers to the positive real values. Each entry of the channel correlation matrix for the kth user is given by $r_{k(mn)} = [\mathbf{R}_k]_{m,n}$, which with a change of notation, we rewrite as $r_{k(m)} = [\mathbf{R}_k]_{l,l-m}$. Hence, the general expression for the discrete-time Fourier transform of the antenna correlation function is given by the following Lemma.

Lemma 4. *The discrete-time Fourier transform of the antenna correlation for COST 2100 channel model with large number of antennas at the BS is obtained as:*

$$\begin{aligned} S_k(f) &= \sum_{m=-\infty}^{\infty} r_{k(m)} e^{-j2\pi fm} \\ &= \sum_{m=-\infty}^{\infty} \left(\sum_{i=1}^{N_l} (a_{ki}^{ga})^2 \, e^{-j2\pi \frac{d}{\lambda}(m) \sin \phi_{ki}} \right) e^{-j2\pi fm} \\ &= \sum_{i=1}^{N_t} (a_{ki}^{ga})^2 \sum_{m=-\infty}^{\infty} e^{-j2\pi fm(\frac{d}{\lambda} \sin \phi_{ki} + f)} \\ &\overset{(a)}{=} \sum_{i=1}^{N_l} (a_{ki}^{ga})^2 \sum_{m=-\infty}^{\infty} \delta \left(m - \left(\frac{d}{\lambda} \sin \phi_{ki} + f \right) \right) \\ &= \sum_{i=1}^{N_l} (a_{ki}^{ga})^2 \delta \left(f + \frac{d}{\lambda} \sin \phi_{ki} \right), \end{aligned} \tag{5.9}$$

where the step (a) we have

$$red \sum_{m=-\infty}^{\infty} e^{-j2\pi fm(\frac{d}{\lambda} \sin \phi_{ki} + f)} = \sum_{m=-\infty}^{\infty} \delta \left(m - \left(\frac{d}{\lambda} \sin \phi_{ki} + f \right) \right), \tag{5.10}$$

which is the property of sum of complex exponentials [79].

Equation (5.9) shows that the DoD of paths can be estimated perfectly from the eigenvalue spectrum in the case of $M \rightarrow \infty$. In the next section, we show that the eigenvalue spectrum of $\mathbf{R}_k$ can be used to build up an approximate eigenchannel matrix for precoding and user scheduling.

5.3.2 Eigenvalue Spectrum with Finite M

For the case of finite M, this chapter follows the methodology in [72]. In [72], Adhikari et al proposed quantizing the interval $[-\frac{1}{2}, \frac{1}{2}]$ into M disjoint intervals of size $\frac{1}{M}$. Using analysis in [72], each interval introduces an angular bin, where bin B_b is centred at $\frac{b}{M} - \frac{1}{2}$ with $b \in \{0, 1, \cdots, M-1\}$. Hence, based on [72], the kth user "occupies" bin B_b if the following condition holds:

$$-\frac{d}{\lambda}\sin\phi_{kp} \in B_b \equiv \frac{b}{M} - \frac{1}{2} - \frac{1}{2M} < -\frac{d}{\lambda}\sin\phi_{kp} \leq \frac{b}{M} - \frac{1}{2} + \frac{1}{2M}. \quad (5.11)$$

Let us assume, similar to [72], that $\pi(i)$ denotes the index of the bin occupied by the MPC i. Then, based on [72], $S_k(f)$ for the case of finite M can be written as

$$\begin{aligned} S_k(f) &= \sum_{i=1}^{N_l} (a_{ki}^{\text{ga}})^2 \sum_{m=-\infty}^{\infty} \delta\left(m - \left(\frac{d}{\lambda}\sin\phi_{ki} + f\right)\right) \\ &= \sum_{i=1}^{N_l} (a_{ki}^{\text{ga}})^2 \times 1\left\{f \in B_{\pi(i)}\right\}. \end{aligned} \quad (5.12)$$

As (5.12) shows, the discrete-time Fourier transform at a particular B_b, is summation of the paths with DoDs in the same bin, i.e. $-\frac{d}{\lambda}\sin\phi_{kp} \in B_b$. Hence, the estimated DoD based on the channel eigenvalue spectrum is not accurate for the case of finite M. However, as we show in the next section, (5.12) can still be used to build up an"approximate eigenchannels" matrix which can be used for beamforming and user scheduling.

5.4 Proposed User Scheduling and Beamforming

In this chapter, we aim to solve the problem of joint user scheduling and beamforming design assuming that only the second order statistics of the channel are available at the BS. The proposed user selection scheme relies on a trade off between the number of occupied spectral bins for each user and the spectral overlap among the selected users. For this case, the performance analysis are found in the next subsection. Once the set of active users has been determined, the BS exploits the covariance matrix of the selected users for

beamforming design and transmits data to the users.

5.4.1 Correlation-based User Scheduling

By using the discrete-time Fourier transform of the antenna correlation given in (5.12), we generate the $K \times B$ matrix $\mathbf{U}$ as follows:

$$\mathbf{U}=\begin{bmatrix} \sum_{i,-\frac{d}{\lambda}\sin\phi_{1i}\in B_1}(a_{1i}^{\mathrm{ga}})^2 & \sum_{i,-\frac{d}{\lambda}\sin\phi_{1i}\in B_2}(a_{1i}^{\mathrm{ga}})^2 & \cdots & \sum_{i,-\frac{d}{\lambda}\sin\phi_{1i}\in B_M}(a_{2i}^{\mathrm{ga}})^2 \\ \sum_{i,-\frac{d}{\lambda}\sin\phi_{2i}\in B_1}(a_{2i}^{\mathrm{ga}})^2 & \sum_{i,-\frac{d}{\lambda}\sin\phi_{2i}\in B_2}(a_{2i}^{\mathrm{ga}})^2 & \cdots & \sum_{i,-\frac{d}{\lambda}\sin\phi_{2i}\in B_M}(a_{2i}^{\mathrm{ga}})^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i,-\frac{d}{\lambda}\sin\phi_{Ki}\in B_1}(a_{Ki}^{\mathrm{ga}})^2 & \sum_{i,-\frac{d}{\lambda}\sin\phi_{Ki}\in B_2}(a_{Ki}^{\mathrm{ga}})^2 & \cdots & \sum_{i,-\frac{d}{\lambda}\sin\phi_{Ki}\in B_M}(a_{Ki}^{\mathrm{ga}})^2 \end{bmatrix}, \quad (5.13)$$

where each (k, b)-th entry of the matrix $\mathbf{U}$ denotes the discrete-time Fourier transform of the antenna correlation function of the kth user at the bth bin, i.e. $\sum_{i,-\frac{d}{\lambda}\sin\phi_{ki}\in B_b}(a_{ki}^{\mathrm{ga}})^2$. The BS uses the functions $f_1(\mathbf{u}_k)$ and $f_2(\mathbf{u}_k)$ to perform user scheduling, where $\mathbf{u}_k$ is the kth row of matrix $\mathbf{U}$ and we define the functions $f_1(\mathbf{u})$ and $f_2(\mathbf{u})$ in the following. As described in step 4.1 in Algorithm 7, the algorithm starts by calculating the summation over all area in terms of eigenvalue spectrum for all users, i.e. $f_1(\parallel \mathbf{u}_k \parallel) = \parallel \mathbf{u}_k \parallel, \forall\, k$, and selects the user which has the largest value among the users. Then in the next step, the proposed algorithm finds a set of ϵ-orthogonal users to the selected users. Here, ϵ-orthogonality among the user k and the user j means that $f_2(\mathbf{u}_k, \mathbf{u}_j) = \frac{|\mathbf{u}_k \mathbf{u}_j^*|}{||\mathbf{u}_k||||\mathbf{u}_j||} < \epsilon$. Note that if the user k and the user j do not have spectral overlap, which means they do not have any shared bins, we have $\frac{|\mathbf{u}_k \mathbf{u}_j^*|}{||\mathbf{u}_k||||\mathbf{u}_j||} = 0$. Hence, increasing the value of ϵ allows the users to have a bigger spectral overlap area. If the value of ϵ is too small, the area of spectral overlap between the selected users decreases and Algorithm 7 selects a small number of users. If the value of ϵ is too big, Algorithm 7 selects users with a large spectral overlap which can reduce the throughput due to interference. It is well known that in GSCMs, MPCs from shared clusters cause high correlation which reduces the rank of the channel [80, 81]. However, selecting users with no spectral overlap does not necessarily result in a higher throughput. So, to find the optimum value of ϵ, we draw the sum rate versus ϵ and set the optimum value as ϵ in Algorithm 7. Note that, $\mathcal{S}_0$ contains $K_s = |\mathcal{S}_0|$ indices of the selected users.

Algorithm 7 Correlation-based user scheduling and beamforming (CUSBF):

Step 1) Initialization: $\Upsilon_0 = [1, \cdots, K]$, $\mathcal{S}_0 = \emptyset$, $i = 1$,
Step 2) Calculate the eigenvalue spectrum of $\mathbf{R}_k$ by means of the discrete-time Fourier transform of the antenna correlation function,
Step 3) Generate matrix $\mathbf{U}$ given by (5.13),
Step 4) Greedy Algorithm:

- **4.1** $\pi(i) = \arg\max_{k \in \Upsilon_0} f_1(\| \mathbf{u}_k \|)$
  $= \arg\max_{k \in \Upsilon_0} \| \mathbf{u}_k \|$, $\mathcal{S}_0 \leftarrow \mathcal{S}_0 \cup \{k\}$, $\mathbf{u}_{(i)} = \mathbf{u}_{(\pi(i))}$,

- **4.2** If $|\Upsilon_0| < K_s$, $\Upsilon_i = \{k \in \Upsilon_{i-1}, k \neq \pi(i) \mid f_2(\mathbf{u}_k, \mathbf{u}_{(i)}) = \frac{|\mathbf{u}_k \mathbf{u}_{(i)}^*|}{\|\mathbf{u}_k\| \|\mathbf{u}_{(i)}\|} < \epsilon\}$,

- **4.3** If $|\Upsilon_0| < K_s$ and $\Upsilon_i \neq \emptyset$, then $i \leftarrow i + 1$, and go to step 4.1, else, go to step 5,

Step 5) Generate matrix $\mathbf{G}$ given by (5.16). BS does not require the instantaneous channels of the users and uses matrix $\mathbf{G}$ for beamforming design.

5.4.2 Correlation-based Beamforming

Once the set of users is fixed, the BS can design the precoding matrix based on the knowledge of $\mathbf{R}_k$, $\forall k$. If $\mathbf{R}_k$, $\forall k$, is available at the BS, it is possible to find an approximated version for the channel matrix $\mathbf{G}$. So, at step 5 of Algorithm 7, we propose to build up the approximate eigenchannel matrix for the channels of users based on the channel covariance matrix given by eq. (5.12) as follows:

$$g_{km} = \sum_{b=1}^{M} \left(\sum_{i, -\frac{d}{\lambda} \sin \phi_{ki} \in \mathrm{B}_b} (a_{ki}^{ga})^2 \right)^{\frac{1}{2}} e^{j2\pi(m-1)\left(\frac{b}{M} - \frac{1}{2}\right)}, \tag{5.14}$$

where the approximate eigenchannel g_{km} is a superposition of B approximated paths, where $B = M$ (denotes the total number of angular bins) and the bth approximated path is centred at $\frac{b}{M} - \frac{1}{2}$. We propose that the BS uses equation (5.14) to build up the approximate eigenchannel matrix $\mathbf{G}$ defined as follows:

Finite M : (5.15)

$$\mathbf{G} = \begin{bmatrix} \sum_{b=1}^{M} (\sum_{i, -\frac{d}{\lambda} \sin \phi_{1i} \in \mathrm{B}_b} (a_{1i}^{\mathrm{ga}})^2)^{\frac{1}{2}} & \cdots & \sum_{b=1}^{M} (\sum_{i, -\frac{d}{\lambda} \sin \phi_{1i} \in \mathrm{B}_b} (a_{1i}^{\mathrm{ga}})^2)^{\frac{1}{2}} e^{j2\pi(M-1)(\frac{b}{M} - \frac{1}{2})} \\ \sum_{b=1}^{M} (\sum_{i, -\frac{d}{\lambda} \sin \phi_{2i} \in \mathrm{B}_b} (a_{2i}^{\mathrm{ga}})^2)^{\frac{1}{2}} & \cdots & \sum_{b=1}^{M} (\sum_{i, -\frac{d}{\lambda} \sin \phi_{2i} \in \mathrm{B}_b} (a_{2i}^{\mathrm{ga}})^2)^{\frac{1}{2}} e^{j2\pi(M-1)(\frac{b}{M} - \frac{1}{2})} \\ \vdots & \ddots & \vdots \\ \sum_{b=1}^{M} (\sum_{i, -\frac{d}{\lambda} \sin \phi_{Ki} \in \mathrm{B}_b} (a_{Ki}^{\mathrm{ga}})^2)^{\frac{1}{2}} & \cdots & \sum_{b=1}^{M} (\sum_{i, -\frac{d}{\lambda} \sin \phi_{Ki} \in \mathrm{B}_b} (a_{Ki}^{\mathrm{ga}})^2)^{\frac{1}{2}} e^{j2\pi(M-1)(\frac{b}{M} - \frac{1}{2})} \end{bmatrix}.$$

Table 5.1: Computational Complexity of Different Schemes

Schemes	Channel estimation	Scheduling	Beamforming
[75]	$\mathcal{O}(K^3M^3)$	$\mathcal{O}(K)$	$\mathcal{O}(M^3)$
[72]	$\mathcal{O}(K_s^3M^3)$	$\mathcal{O}(K)$	$K_s\mathcal{O}\big(M^3{+}M\log^2 M\log b\big)$
Algorithm 1	–	$\mathcal{O}(K)$	$\mathcal{O}(M^3)$

The approximate eigenchannel matrix $\mathbf{G}$ can be used for user scheduling and precoding design. Note that only for the case of $M \to \infty$, the DoD of each single MPC is resolvable and are available at the BS.

5.5 Complexity Analysis

Without loss of generality the complexity, computation of the MMSE estimator is given by $\mathcal{O}(\tau^3M^3)$, where $\tau = K$ is sufficient to remove the effect of pilot contamination [76]. Hence, the complexity of the MMSE estimator scales as $\mathcal{O}(K^3M^3)$, which indicates the complexity of inverting of matrix size $KM \times KM$ to estimate the channel in equation (5.16). The proposed Algorithm 1 and the scheme in [72] do not exploit the knowledge of channel for user scheduling and beamforming design. For a given $M \times M$ matrix, the required operations to determine the eigenvectors is given by $\mathcal{O}\big[M^3 + (M\log^2 M)\log b\big]$, where b is the relative error bound [82]. Moreover, the complexity to search the user for the scheme in [75] is linear with the number of users [83]. Note that the complexity of user scheduling in the proposed Algorithm 1 and the scheme in [72] is linear in terms of the number of users. The number of arithmetic operations required for Algorithm 7 is shown in Table 5.1. The authors in [78] define the spatial WSS quality which is given by

$$Q_{\text{WSS}} = \frac{\tau_{\text{LT}}}{\tau_c}, \tag{5.16}$$

where τ_{LT} refers to the long-term time, where the statistics of the channel may be considered constant within this interval whereas τ_c is the channel coherence time. The measurement results for the outdoor scenario at a center frequency of 2 GHz shows that $Q_{\text{WSS}} =$

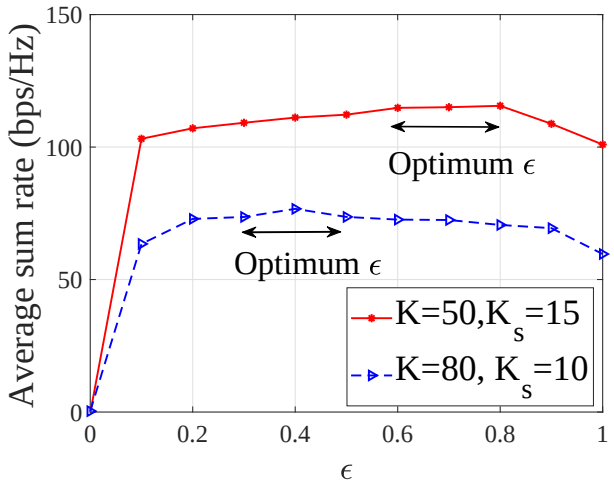

Figure 5.1: Average sum rate versus ϵ,$p_k = 10$ dBm and $R = 500$ meters.

120. As a result, every $120 \times \tau_c$, the correlation based schemes (the proposed Algorithm 1 and the scheme in [72]) need to be run, while the scheme in [75] need to be run at the beginning of each coherence time.

5.6 Numerical Results and Discussion

A square cell with a side length of $2 \times R$ has been considered; we call R the cell size and also assume users are uniformly distributed in the cell. As in [84], we assume that there is no user closer than $R_{th} = 0.1 \times R$ to the BS. We simulate a micro-cell environment for the NLoS case and set the operating frequency $f_C = 2$ GHz. The external parameters and stochastic parameters are extracted from chapter 3 of [32]. The BS and user heights are assumed to be $h_{BS} = 5$ m and $h_{MS} = 1.5$ m, respectively. The noise power is given by $P_n =$ BW $k_B\, T_0\, W$, where BW $= 20$ MHz denotes the bandwidth, $k_B = 1.381 \times 10^{-23}$ represents the Boltzmann constant, $T_0 = 290$ (Kelvin) denotes the noise temperature, and $W = 9$ dB is the noise figure. For this network setup, the average sum rate is evaluated for the three scenarios. First, we evaluate the average throughput of the proposed CUSBF scheme, given by Algorithm 1. In Fig. 5.1, the sum rate of users under the proposed scheme is plotted as a function of ϵ in Algorithm 1. If ϵ is too large, the spectral overlap (number of shared bins) is big, while if is too small, the multiuser diversity gain decreases and users with shared bins cannot be selected. As a result, there should be a trade off

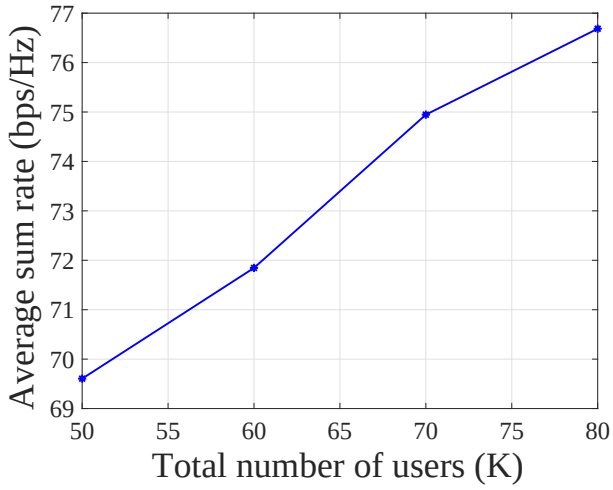

Figure 5.2: Average sum rate versus total number of users with $K_s = 10$, $p_k = 10$ dBm and $R = 500$ meters.

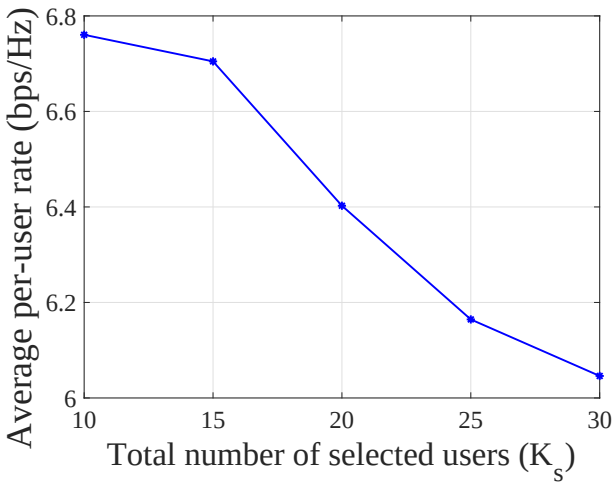

Figure 5.3: Average per-user rate versus total number of selected antennas with $K = 50$, $p_k = 10$ dBm and $R = 500$ meters.

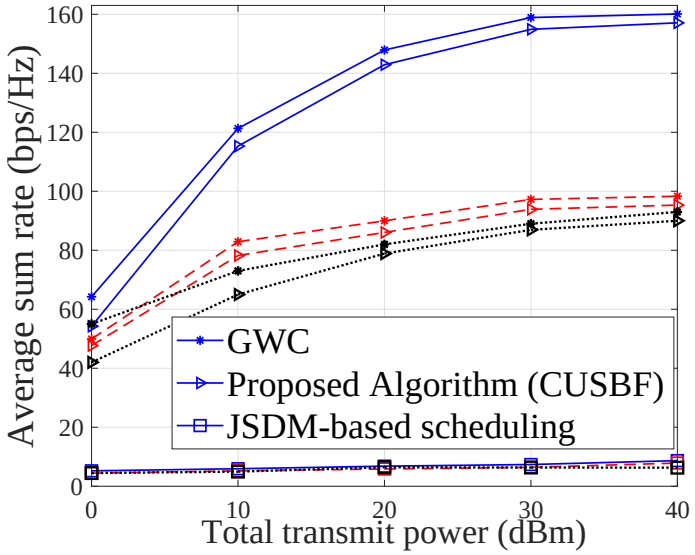

Figure 5.4: The average sum rate vs. transmit power. Solid (blue), dashed (red) and dotted (black) lines refer to $\{M = 300, K = 70, K_s = 20\}$, $\{M = 300, K = 50, K_s = 10\}$ and $\{M = 200, K = 50, K_s = 10\}$, respectively.

between total number of shared bins and summation over all area in terms of eigenvalue spectrum, which is explained in Subsection 5.4.1. The optimal value of ϵ is shown in Fig. 5.1. Next, we plot the average sum rate versus the total number of users in the system in Fig. 5.2. As the figure shows, by increasing the total number of users, the average sum rate increases, as a result of multi-user diversity gain. Fig. 5.3 demonstrates the average per-user rate versus the total number of users in the system. Note that the analysis in [85] demonstrate that in the limit of massive MIMO ($M, K_s \to \infty$ and $\alpha = \frac{M}{K_s}$), by increasing K_s the average per-user rate decreases.

Finally, we evaluate the average throughput of the proposed CUSBF scheme, given by Algorithm 1, and GWC [75, 86] with an MMSE estimate of the channel. For the case of GWC, similar to [86], we set the optimal channel direction constraint to achieve the best performance for GWC. Moreover, the comparison with the scheme proposed in [72] is provided. In [72], Adhikari et al propose to select users which occupy a larger number of bins and find users having a smaller spectral overlap with the selected users. This scheme is referred to JSDM-based scheduling. Fig. 5.4 depicts the average sum rate versus the total transmit power for three cases of $\{M = 300, K = 70, K_s = 20\}$, $\{M = 300, K = 50, K_s = 10\}$ and $\{M = 200, K = 50, K_s = 10\}$, while adopting the currently proposed scheme with ZFBF. As expected, since GWC exploits the estimated instantaneous CSI, it has the best throughput. As the figures show, the performance of the

proposed Algorithm 1 is slightly poorer than the case in which the BS has the knowledge of the estimated instantaneous channel to perform user scheduling and beamforming as in [75], i.e., GWC. Interestingly, for a larger number of antennas at the BS, the superiority of the proposed scheme is more obvious in terms of achieving performance close to that of the GWC scheme. Moreover, the performance of the proposed algorithm is several times higher than for the scheme in [72], i.e., JSDM-based scheduling. In addition, the figure demonstrate that the performance of the scheme in [72] is quite poor for the case of the COST 2100 channel model. This is because of the large number of clusters in the area, which means that the performance of eigen-beamforming is not as good as ZFBF. Note that the JSDM in [72] is designed to work well with the angularly-sparse multipath channels typically observed at mm-waves.

5.7 Summary

We proposed to use the angular bins of the eigenvalue spectrum of the channel covariance matrix to build up an approximate eigenchannel for the users. Using the discrete-time Fourier transform of the antenna correlation function, a novel user scheduling scheme and linear precoding design has been proposed and tested with the COST 2100 channel model. The results show that while the average throughput slightly decreases due to absence of instantaneous channel, the computational complexity of the system reduces significantly. As a result, the proposed scheme can be considered as a superior practical approach for massive MIMO systems.

Chapter 6

Cluster Parametrization at 60 GHz in a Large Indoor Environment

Contents

6.1 Introduction

Most standardized MIMO channel models such as IEEE 802.11 [21] and the most recent 3GPP channel model [87] rely on clustering [21]. The same applies to the recent COST channel models, e.g., the COST 2100 model [88]. These models are geometry-based stochastic channel models (GSCMs) that are mathematically tractable, though to a limited extent, to investigate the performance of MIMO systems [89]. The concept of clustering is an essential basis of GSCMs to characterize scatterers in the cell environments. In [90–94], the authors use clusters to characterize measured multipath channels for a GSCM in mmWave bands. The available GSCMs at mmWave do not necessarily retain the spatial

consistency of simulated channels due to lack of cluster dynamics, which is essential for small cells with ultra-dense users. In this chapter, we work on cluster parameterization to investigate the spatial consistency life-time of clusters, using a ray-tracer which is adjusted to produce results consistent with measurements.

Unlike previously available clustering algorithms in [95, 96] (which are done based on the angle and delay), in this chapter the coordinates are exploited for which the multipath components (MPCs) interact with surrounding objects for a fixed position of mobile station (MS) and BS. To the best of our knowledge, previously clustering has been performed in a double-directional setting, i.e., considering both angle of arrival (AoA) and angle of departure (AoD). A consistent scheme to identify and track clusters based on the spatial coordinates of the MPCs (the $[x, y, z]$-coordinates of the MPCs) is presented. To investigate the performance of the proposed clustering scheme we exploit a set of ray-tracer results in Helsinki's airport described in [97], which is very accurate to present the propagation properties such as specular reflections, diffraction, diffuse scattering [98]. The contributions of the chapter are summarized as follows:

1. We study whether clusters exist or not.

2. For the first time, we perform clustering of dynamic multipath channels.

3. $[x, y, z]$ coordinate-based clustering.

6.1.1 Outline

The chapter is organized as follows; Section II describes the ray-tracer and simulation area, and Section III provides the MPC clustering-and-tracking framework. The simulation results and discussion are presented in Section IV while Section V concludes the chapter.

6.2 The Ray-Tracer and Simulation Area

The ray-tracer simulates multipath channels for a large number of links between BS and MS [97]. Note that the ray-tracer works with accurate descriptions of the environment in the form of point clouds, obtained by laser scanning, and has the ability of simulating relevant propagation properties such as specular reflections, diffraction, diffuse scattering and shadowing [98]. For more details on the ray-tracer refer to [97,98]. A check-in hall of Helsinki airport as a representative small-cell scenario is considered as shown in Fig. 6.1. Exploiting the ray-tracer parameters in Fig. 6.1, we obtain the MPCs for links defined by BS and MS locations as in Fig. 6.1. The BS is located 1 m from a wall at a height of 5.7 m whereas the MS is placed at a height of 1.5 m at every 5 cm over a route. In total, 2639 links including 1816 LOS and 823 obstructed LOS (OLOS) are simulated. As the ray-tracer calculates interactions of MPC with physical objects in the environment, we save the first and last MPC interacting coordinates $[x, y, z]$ (in case of multi-bounce clusters, which is explained in Chapter 2) instead of the angle of departure and arrival of each MPC. We assume downlink where BS transmits and MS receives radio signals. The first and last interacting coordinates are the same for a single-bounce path ,related to single-bounce cluster (please refer to the definition of clusters given in Chapter 2), and are different for a multiple-bounce path. The ray-tracer also derives a complex gain for each MPC.

6.3 Clustering-and-Tracking Framework

Similar to standard clustering algorithms [95,96], we independently perform clustering at each snapshot and thereafter the clusters are tracked. Note that the snapshot refers to the position of MS. Consider $n = 1, \cdots, N$ data windows, where at each data window we have $L^{(n)}$ MPCs. Next, we define for each MPC $\mathbf{v}_{1,l}^{(n)} = [x_{MS,l}^{(n)}, y_{MS,l}^{(n)}, z_{MS,l}^{(n)}]$ (the position of MPCs from MS side) and $\mathbf{v}_{2,l}^{(n)} = [x_{BS,l}^{(n)}, y_{BS,l}^{(n)}, z_{BS,l}^{(n)}]$ (the position of MPCs from BS side), and finally we have

$$\chi_l^{(n)} = \left[\mathbf{v}_{1,l}^{(n)}\right] = \left[x_{MS,l}^{(n)}, y_{MS,l}^{(n)}, z_{MS,l}^{(n)}\right]. \tag{6.1}$$

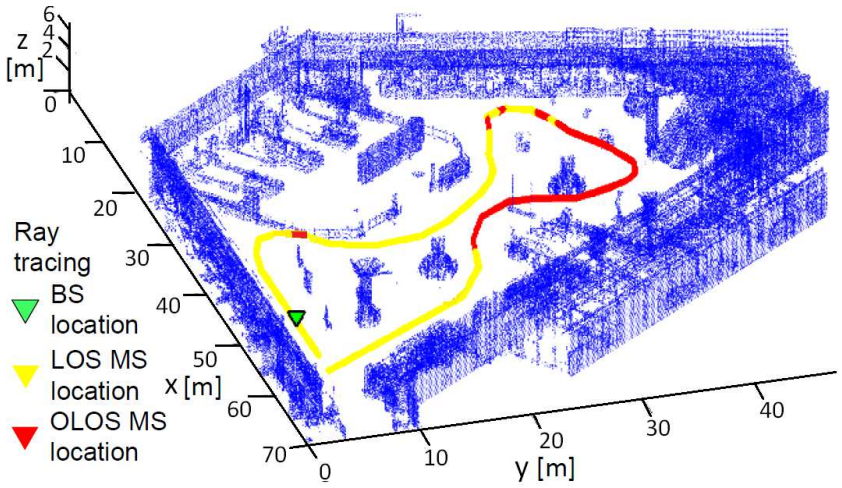

Figure 6.1: Floor plan of the small-cell site in Helsinki airport. For this simulation set-up $f_c = 61$ GHz, $BW = 2$ GHz refer to the carrier frequency and bandwidth, respectively. Moreover, the position of BS is fixed (the green triangle), while we investigate 2639 positions for MS (the yellow and red points demonstrate the LOS and OLOS, respectively. The total MS route is 132 m, and channels simulated at every 5 cm.

The same equality hold for the BS-side components. This enables us after visualising clusters to plot clusters separately for $v_{1,l}^{(n)}$ and $v_{2,l}^{(n)}$ in physical three-dimensional space as well as defining the matrix $\boldsymbol{\chi}^{(n)} = [\chi_1^{(n)}, \cdots, \chi_L^{(n)}]$ Moreover, the lth MPC in window n has a power represented by $p_l^{(n)}$ which enables us to define the power vector $\mathbf{p}^{(n)} = [p_1^{(n)}, \cdots, p_L^{(n)}]$.

6.3.1 Cluster Parameters

In next step, we define the following parameters for each cluster:

1. Cluster ID c.
2. Cluster power at time n: $\gamma_c^{(n)} = \sum_{l \in I_c^{(n)}} p_l^n$, where $I_c^{(n)}$ denotes the set of MPCs belonging to cluster c at time n.
3. Total number of MPCs in cluster c at time n: $L_c^{(n)} = |I_c^{(n)}|_{\text{size}}$.

4. Cluster centroid position:

$$\begin{aligned}\boldsymbol{\mu}_c^{(n)} &= \left[x_{MS,c}^{(n)}, y_{MS,c}^{(n)}, z_{MS,c}^{(n)}\right]^T \\ &= \frac{1}{\gamma_c^{(n)}} \left[\sum_{l \in I_c^{(n)}} p_l^n x_{MS,l}^{(n)}, \sum_{l \in I_c^{(n)}} p_l^n y_{MS,l}^{(n)}, \sum_{l \in I_c^{(n)}} p_l^n z_{MS,l}^{(n)}\right]^T . \end{aligned} \tag{6.2}$$

5. Combined cluster centroid position and speed:

$$\boldsymbol{\theta}_c^{(n)} = \left[x_{MS,c}^{(n)}, \Delta x_{MS,c}^{(n)}, y_{MS,c}^{(n)}, \Delta y_{MS,c}^{(n)}, z_{MS,c}^{(n)}, \Delta z_{MS,c}^{(n)}\right]^T . \tag{6.3}$$

6. Cluster spread matrix:

$$\mathbf{C}_c^{(n)} = \frac{\sum_{l \in I_c^{(n)}} p_l^{(n)} \left(\boldsymbol{\chi}_l^n - \boldsymbol{\mu}_c^n\right) \left(\boldsymbol{\chi}_l^n - \boldsymbol{\mu}_c^n\right)^T}{\gamma_c^{(n)}}. \tag{6.4}$$

Next, similar to terminology in [95], a Kalman filter [77] is used to both track and predict the cluster positions over time. Moreover, an initial-guess process introduces an appropriate initial guess for cluster centroids, and finally the clustering algorithm determines the clusters in the ray-tracer results exploiting the initial guess.

6.3.2 Kalman Filter to Track and Predict Cluster Positions

We exploit the cluster centroid positions and cluster centroid speeds for the Kalman tracking [77]. The following state equations are used:

$$\begin{cases} \boldsymbol{\theta}_c^{(n)} = \mathbf{A}\boldsymbol{\theta}_c^{(n-1)} + \mathbf{B}^{(n)}, & \text{(6.5a)} \\ \mathbf{A} = \mathbf{I}_3 \otimes \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} & \text{(6.5b)} \\ \boldsymbol{\mu}_c^{(n)} = \mathbf{D}\boldsymbol{\theta}_c^{(n)} + \mathbf{E}^{(n)}, & \text{(6.5c)} \\ \mathbf{D} = \mathbf{I}_3 \otimes \begin{bmatrix} 1 & 0 \end{bmatrix}, & \text{(6.5d)} \end{cases}$$

where $\mathbf{B}^{(n)}$ and $\mathbf{E}^{(n)}$ refer to the state-noise with covariance matrix $\mathbf{Q}$ and the observation-noise with covariance matrix $\mathbf{R}$, respectively. Note that $\boldsymbol{\mu}_c^{(n)}$ introduces the observed

cluster centroid position. The prediction and update equations are given by

$$\text{Prediction} \begin{cases} \boldsymbol{\theta}_c^{(n|n-1)} = \mathbf{A}\boldsymbol{\theta}_c^{(n-1|n-1)}, & (6.6a) \\ \mathbf{M}^{(n|n-1)} = \mathbf{A}\mathbf{M}_c^{(n-1|n-1)} + \mathbf{Q}, & (6.6b) \end{cases}$$

and update

$$\mathbf{K}^{(n|n)} = \mathbf{M}_c^{(n|n-1)}\mathbf{D}^T \left(\mathbf{D}\mathbf{M}^{(n|n-1)}\mathbf{D}^T + \mathbf{R}\right)^{-1}, \quad (6.7a)$$

$$\boldsymbol{\theta}_c^{(n|n)} = \boldsymbol{\theta}_c^{(n|n-1)} + \mathbf{K}^{(n|n)} \left(\boldsymbol{\mu}_c - \mathbf{D}\boldsymbol{\theta}_c^{(n|n-1)}\right), \quad (6.7b)$$

$$\mathbf{M}^{(n|n)} = \left(\mathbf{I} - \mathbf{K}^{(n|n)}\mathbf{D}\right)\mathbf{M}^{(n|n-1)} \quad (6.7c)$$

6.3.3 Association of Clusters

Association of predicted targets to identified targets is a substantial challenge in any multi-target tracking [95]. Based on [95], the distance between a cluster with parameters $(\boldsymbol{\mu}_c, C_c)$ and a cluster with centroid $\tilde{\boldsymbol{\mu}}$ is called the closeness function and is given by

$$d_c\left(\tilde{\boldsymbol{\mu}}|\boldsymbol{\mu}_c, \mathbf{C}_c\right) = \frac{1}{(2\pi)^{\frac{3}{2}}\left|\mathbf{C}_c\right|_{\text{det}}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_c\right)^T \mathbf{C}_c^{-1}.\left(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_c\right)\right), \quad (6.8)$$

First, the closeness function between the old clusters (with the old covariance matrix) and new centroids and the closeness function between the new clusters (with the old covariance matrix) and old centroids are calculated. Next, for each new cluster the closest old cluster and for each old cluster the closest new cluster is determined. Note that the closest cluster is determined by finding the maximum value of the closeness function. If the closeness function from both directions are exactly the same, these two clusters are associated and assumed to be one cluster. The clusters which are not associated are assumed to be new ones.

6.3.4 Initial Guess for Clusters

The initial guess of the cluster centroids is a challenging task in clustering algorithms. In [95], the authors propose a novel initial guess to maximize the distances between the cluster centroids. If there is no cluster prediction available, the path having the strongest power is selected as the first centroid $\hat{\boldsymbol{\mu}}_1$ whereas for the case of available cluster prediction, the initial-guess centroid from the prediction is to be as the current initial guess. Note that the multipath component distance (MCD) in this chapter is different from the one used in [95, 99]. The distance measure between MPCs i and j is given by

$$\text{MCD}_{ij} = \sqrt{||\text{MCD}_{x_{MS},ij}||^2 + ||\text{MCD}_{y_{MS},ij}||^{+}||\text{MCD}_{z_{MS},ij}||^2}.$$

Note that in (6.9) we have

$$\text{MCD}_{x_{MS},ij} = \frac{|x_{MS,i} - x_{MS,j}|}{\Delta x_{MS,\text{max}}}, \tag{6.9}$$

where $\Delta x_{MS,\text{max}} = \max\{|x_{MS,i} - x_{MS,j}|\}$, and the other terms in (6.9) are evaluated is a the similar way to (6.9). Next, the weighted distance matrix $\boldsymbol{\Upsilon} \in C^{l\times c}$ between all paths and all initial-guess centroids is evaluated as follows:

$$\boldsymbol{\Upsilon}\left(\boldsymbol{\chi}_l^n - \hat{\boldsymbol{\mu}}_c\right) = \log_{10}\left(p_l^{(n)}\right)\text{MCD}\left(\boldsymbol{\chi}_l^n - \hat{\boldsymbol{\mu}}_c\right). \tag{6.10}$$

Following the terminology in [95], we select the path with the maximum minimum distance to any centroid as follows:

$$l_{\text{sel}} = \max_l\left\{\min_c\left\{\boldsymbol{\Upsilon}\right\}\right\}. \tag{6.11}$$

We then assign all MPCs to their closest centroid and cluster power is evaluated. If we do not achieve the maximum number of clusters, and centroid powers are larger than 0.01% of the total snapshot power, we repeat the calculation of the weighted distance matrix $\boldsymbol{\Upsilon} \in \mathbb{C}^{l\times c}$ in (6.10). Otherwise, the last centroid is ignored and the algorithm is stopped.

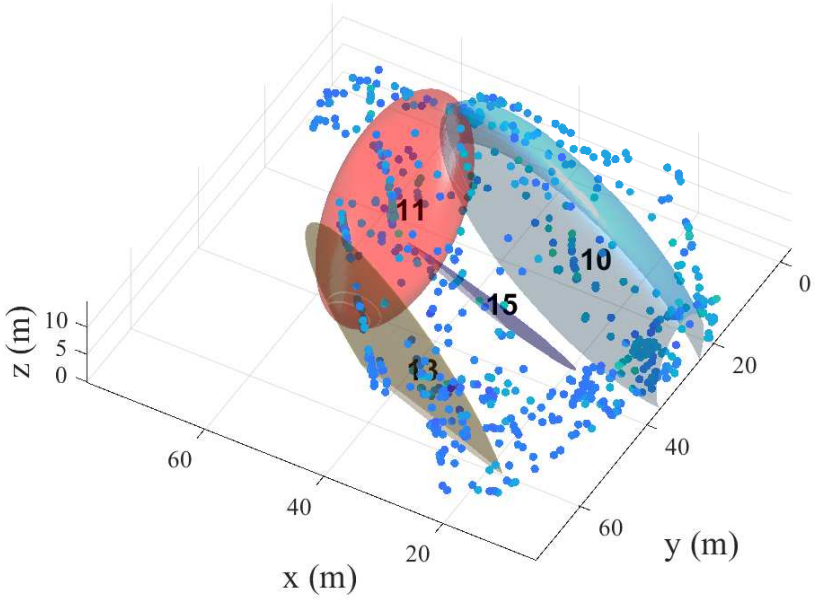

Figure 6.2: Tracked Rx-side clusters in Helsinki airport in snapshot 3.

6.3.5 Clustering Algorithm

The KPowerMeans clustering algorithm is investigated in [100], and it performs as follows: the initial-guess algorithm is applied, and the KPowerMeans clustering algorithm is run only once as the initial guess as are constant. For more details on the KPowerMeans clustering algorithm refer to [100]. Note that if any cluster occupies less than 1% of total cluster power, we re-start the clustering algorithm with the initial guess, with the number of clusters is reduced by one. Therefore, it is possible that the algorithm ends with a single cluster.

6.4 Results and Discussion

The joint clustering-and-tracking algorithm is applied to the ray-tracer results at Helsinki airport, explained in Section 6.2, where we have 2639 links. Figs. 6.2- and 6.6 present the exemplary plots for different snapshots. The MPCs are shown by dots, where their power is shown by light blue (weak power) and violet (strong power). The clusters are shown by ellipsoids and always 99% of the total power is carried by the MPCs within clusters.

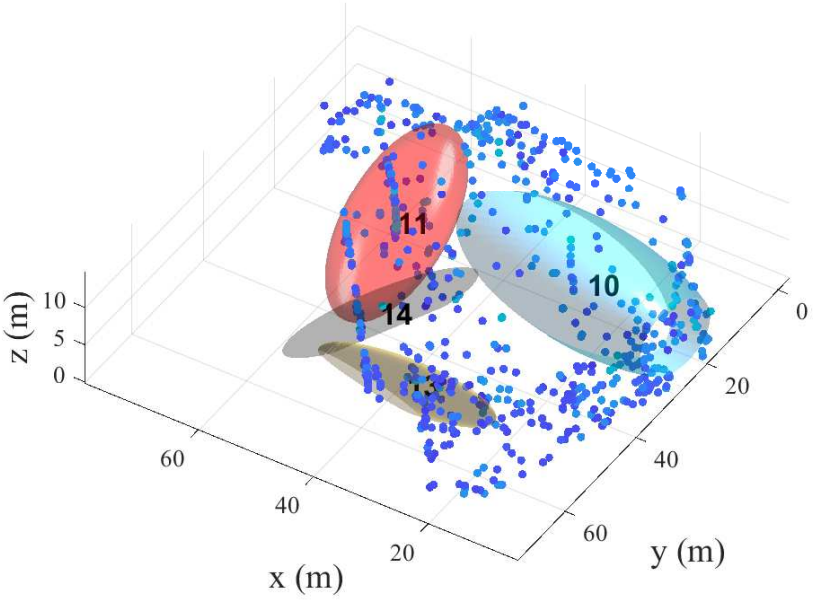

Figure 6.3: Tracked Rx-side clusters in Helsinki airport in snapshot 4.

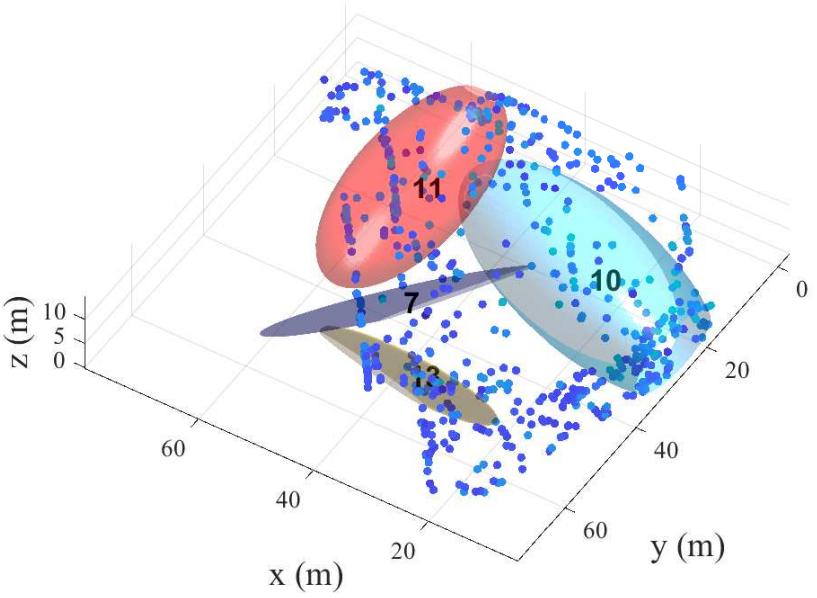

Figure 6.4: Tracked Rx-side clusters in Helsinki airport in snapshot 5.

We use different colors for ellipsoids just to make the cluster recognition easier. Each cluster is identified by a cluster ID which is written on each cluster. As these exemplary figures show for snapshots 2,3 and 4, cluster 2 is always tracked while the other clusters are determined as new clusters.

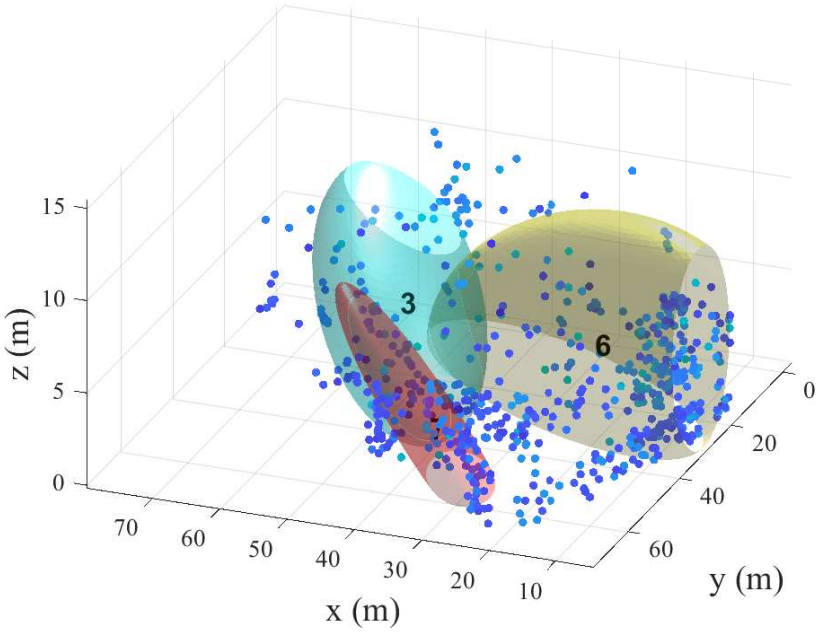

Figure 6.5: Tracked Tx-side clusters in Helsinki airport in snapshot 12.

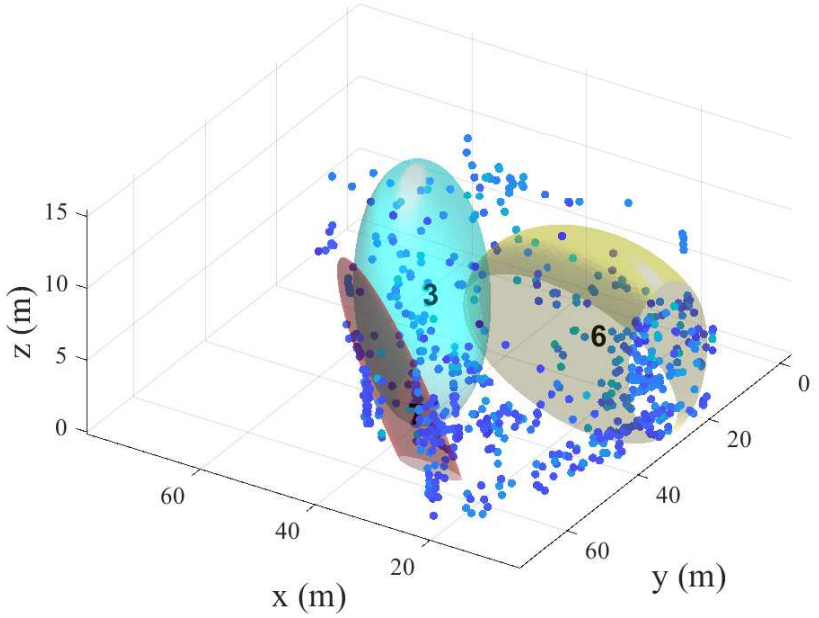

Figure 6.6: Tracked Tx-side clusters in Helsinki airport in snapshot 13.

Next, the lifetime of clusters for the available sets of ray-tracer results is investigated, for Tx-side clusters and Rx-side clusters separately. Figs. 6.7 and 6.8 show the histograms of cluster lifetimes for Rx-side (BS-side) and Tx-side (MS-side) scenarios, respectively. The figures show that in most cases clusters are active only for a few

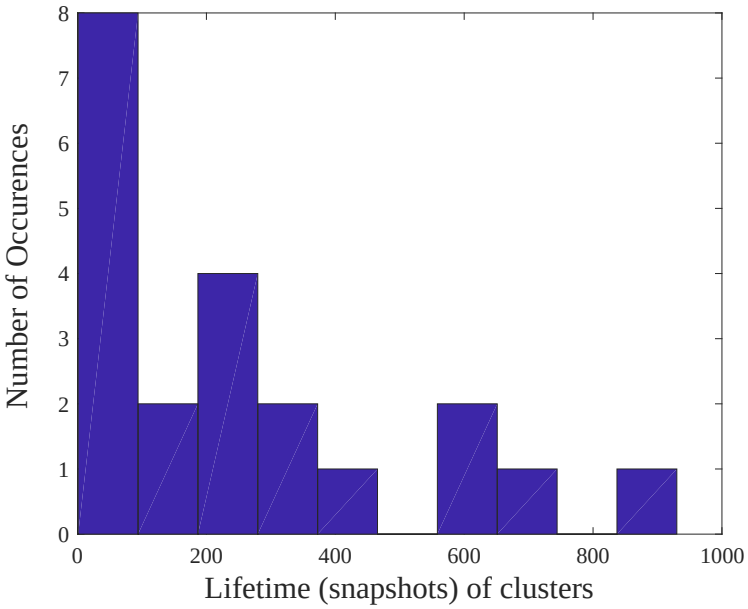

Figure 6.7: Histogram of Rx-side clusters cluster lifetimes (snapshots).

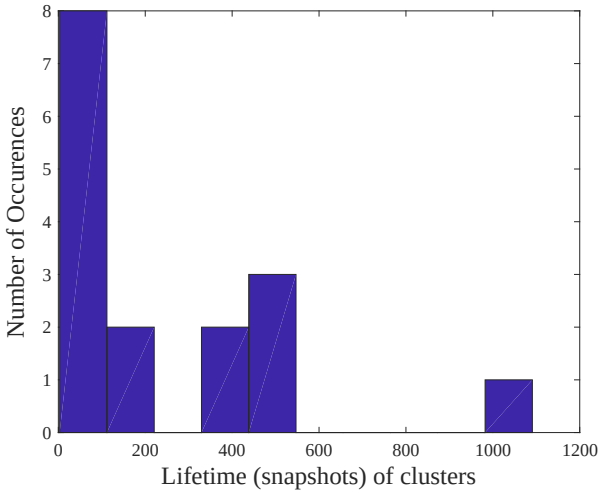

Figure 6.8: Histogram of Tx-side cluster lifetimes (snapshots).

snapshots for this set of ray-tracer results. This requires more investigation. Moreover, the number of clusters per snapshot is presented in Figs. 6.9 and 6.10 for Rx-side and Tx-side clusters, respectively.

The other interesting phenomenon is the movement of the tracked cluster centroids, which is shown in Fig. 6.11. Based on these figures the cluster centroid moves rapidly in the x or y direction while its speed is very low in other direction. Moreover, the figure shows for these clusters that the centroid's speed is very low in the z direction. Finally,

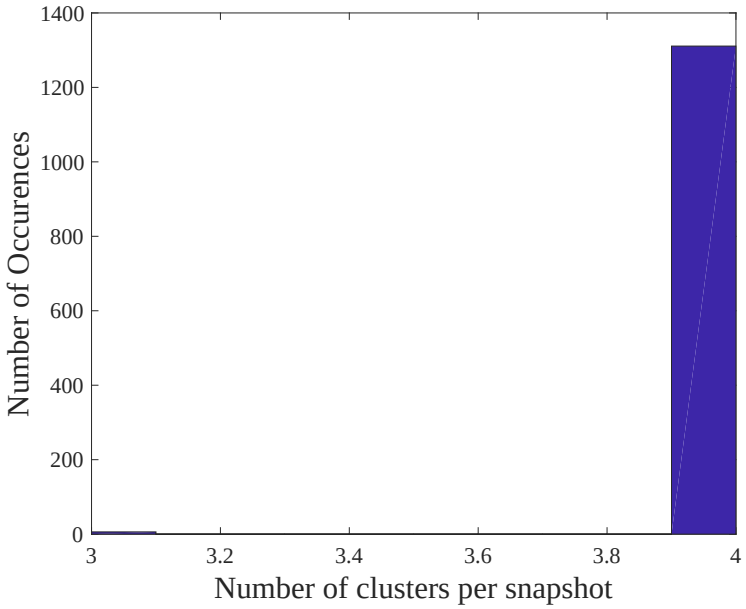

Figure 6.9: Histogram of total number of Rx-side clusters.

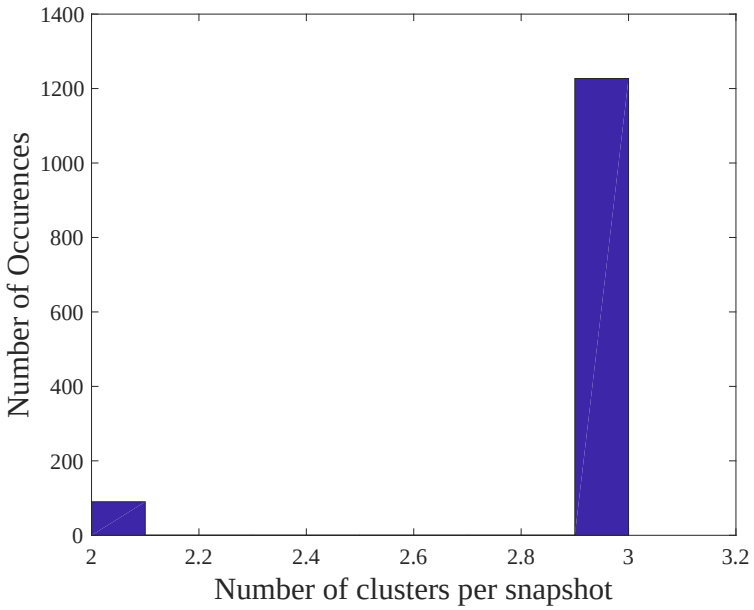

Figure 6.10: Histogram of total number of Tx-side clusters.

Figs. 6.12 and 6.13 investigate the distribution of the percentage of power in Tx-side and Rx-side clusters.

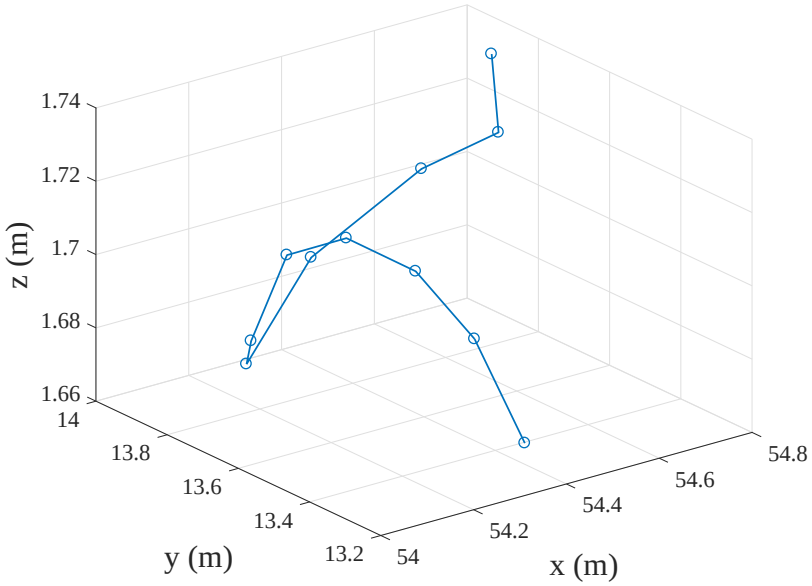

Figure 6.11: Tracked centroid of exemplary moving cluster.

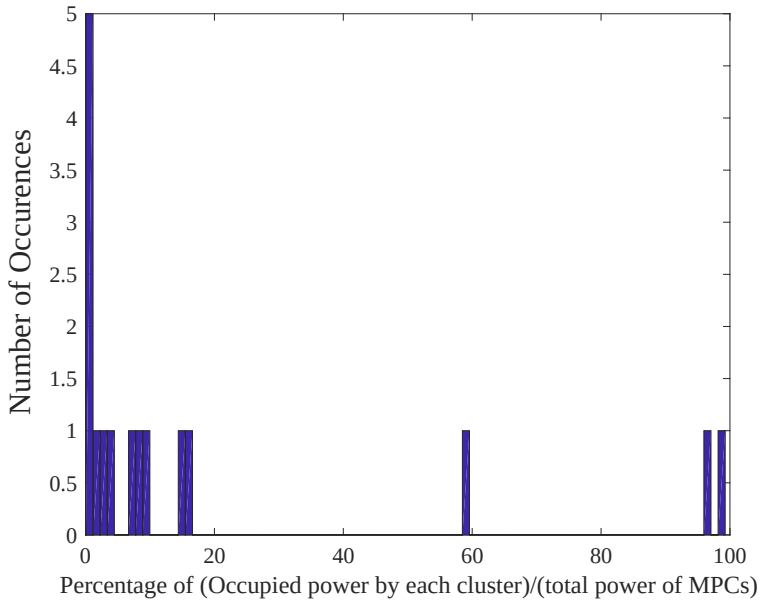

Figure 6.12: Histogram of percentage of occupied power by each Tx-side cluster.

6.5 Summary

In this chapter, we have worked on parameterization for the COST 2100 channel model at 60 GHz band. We have worked on a ray-tracer, which has been optimized to match measurements, to get double-directional channels at mmWaves. We have combined

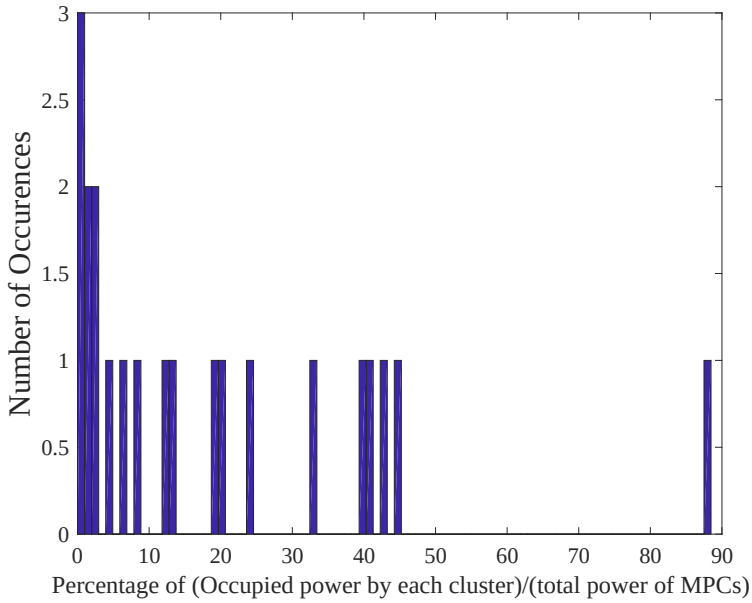

Figure 6.13: Histogram of percentage of occupied power by each Rx-side cluster.

clustering and tracking to improve the performance of consistent clustering. The results showed that the joint clustering-and-tracking allows for cluster identification and tracking for the ray-tracer results. Cluster lifetime and number of clusters per snapshot have been investigated.

Chapter 7

Conclusions and Future Work

Contents

7.1 Summary of the Work

The conclusions of the thesis can be itemized as follows:

- In Chapter 1, a introduction to the research area has been presented, followed by outline of thesis and the publication list.
- A literature review was presented in Chapter 2, and the required and general basics of thesis were presented. These includes: the basics of multi-user multiple-input multiple-output (MIMO), channel estimation, the basics of massive MIMO, and alternative ways to perform massive MIMO. Next, the basics of COST 2100 channel model and general basics of optimum uniform quantization have been discussed. Finally, achievable rate with unknown gain at the receiver was reviewed.

- In Chapter 3, we have studied the uplink max-min rate problem in cell-free massive MIMO with the realistic assumption of limited-capacity fronthaul links. Next, we have proposed an optimal solution to maximize the smallest user rate. The max-min problem was divided into two sub-problems which were iteratively solved by formulating them into generalized eigenvalue problem and GP. We have validated the optimality of the proposed solution through presenting an uplink-downlink duality. Numerical results have been provided to demonstrate the optimality of the proposed scheme. In addition, these results confirmed that the proposed max-min rate algorithm can increase the median of the CDF of the minimum uplink rate of the users by more than two times, compared to existing algorithms. Finally, we presented a user assignment algorithm to further improvement in minimum rate of the users.

- In Chapter 4, we have considered cell-free massive MIMO when the quantized version of the weighted signals are available at the CPU. Bussgang decomposition has been used to model the quantization effects. A closed-form expression for spectral efficiency has been derived. We have then studied the problem of the energy efficiency maximization with per-user power constraints, fronthaul capacity constraints and throughput requirements. We have developed an SCA to efficiently solve this non-convex problem. Next a low-complexity sub-optimal scheme is proposed. In addition, complexity and convergence of the proposed schemes have been investigated. Numerical results confirmed that the limited-fronthaul cell-free massive MIMO system with the proposed algorithm can reach almost twice the uplink total energy efficiency compared to the case of equal power allocation. In addition, a trade-off between the total number of APs and the number of antennas at the APs has been shown. Moreover, we investigated the optimal number of AP antennas along with the optimal number of quantization bits to maximize the uplink total energy efficiency of cell-free massive MIMO.

- In Chapter 5, a collocated single-cell massive MIMO was considered with realistic geometry-based COST 2100 channel model. Exploiting the eigenvalue spectrum of the correlation of the channel at the BS, a correlation-based user scheduling and beamforming scheme has been proposed. Complexity of the proposed scheme has been investigated and finally we presented the numerical results.

- In Chapter 6, we assumed ray-tracer results in Helsinki airport at 61 GHz. A

clustering-and-tracking framework was used to find the cluster parameters, and a Kalman filter has been exploited to track and predict cluster positions.

7.2 General Conclusion

The thesis investigates mainly massive MIMO and COST 2100 channel model. The benefits and challenges of different schemes to implement a massive MIMO system have been presented. Thanks to the distributed APs, cell-free massive MIMO has the ability to provide a great performance to all users. In practice the links from the APs to the CPU cannot have infinite capacity. Therefore, a cell-free massive MIMO with limited-capacity fronthaul links have been investigated. The problem of max-min signal-to-interference-plus-noise ratio (SINR) has been considered which provides all users with the same throughput. A user assignment technique was proposed. Next, energy efficiency optimization problem has been considered.

In this thesis, we provided the basics of real geometry-based COST channel model and draw an elegant connection between collocated massive MIMO and COST 2100 channel model. A correlation-based user scheduling and beamforming technique has been presented.

Finally, a channel parameterization scheme at 61 GHz was proposed which exploits Kalman filter to find the cluster parameters.

7.3 Future Work

Exploiting the proposed schemes, algorithms, performance analysis, and the presented results, the future direction of related research are summarized as follows:

- In Chapters 3 and 4, the performance of cell-free massive MIMO with uncorrelated Rayleigh fading channel were provided. However, note that this is not a realistic

channel model and in practice the channel between different APs and different users could be correlated. A proper investigation on the performance of cell-free massive with correlated channel model could be an interesting research direction.

- In Chapter 3, a user assignment scheme was proposed to improve the max-min throughput of the system. Although the proposed user assignment scheme improves the performance, however, this is not the optimal assignment scheme. Future work is needed to investigate the optimal user assignment scheme with different linear receivers.

  Further investigation to find the optimal user assignment scheme is required.

- In Chapter 5, a correlation-based user scheduling and beamforming design has been proposed. The extension of the proposed scheme to distributed massive MIMO is a future research direction.

- In Chapter 6, we found the cluster parameters. However, the other channel parameters like path loss and cluster delay spreads have not been investigated. This provides a future research direction.

- Deep learning can be used to decrease the effect of pilot contamination in cell-free massive MIMO. Pilot contamination is one of the difficult challenges in cell-free massive MIMO. Investigating optimal pilot assignment schemes is an interesting future direction.

- In this thesis, we provided the performance analysis of cell-free massive MIMO with MRC. However, other signal processing techniques (e.g., zero-forcing processing) can be implemented to improve the system performance and can be considered in future work.

- The original idea of additive quantization noise model (AQNM) comes from [101], where the authors exploit a scalar non-uniform MMSE quantizer at each ADC and compute the variance quantization distortion using Llyod-Max algorithm [58]. However, in this thesis, we consider the uniform quantizer which has low complexity. This can be regarded as future work.

- The performance with Ricean channels and the effect of LoS have not been considered in this thesis. The case of Ricean channels is kept for future work.

Glossary

5G	**5th Generation**
4G	**4th Generation**
AP	**Access Point**
AWGN	**Additive White Gaussian Noise**
ADC	**Analog**-to-**Digital Converter**
BS	**Base Station**
BU	**Beamforming Uncertainty**
CPU	**Central Processing Unit**
CP	**Circuit Power**
CUSBF	**Correlation**-Based **User** **Scheduling** and **Beamforming**
CoMP	**Coordinated Multipoint Processing**
CDF	**Cumulative Distribution Function**
DS	**Desired Signal**
DoA	**Direction** of **Arrival**
DoD	**Direction** of **Departure**
eCOST	**European Cooperation** in **Science** and **Technology**
FDD	**Frequency Division Duplexing**
FT	**Fronthaul Traffic**
GWC	**Greedy Weight Clique**
GP	**Geometric Programming**
GSCMs	**Geometry**-based **Stochastic Channel Models**
IUI	**Inter-User-Interference**
i.i.d.	**Independent** and **Identically Distributed**

JSDM	**Joint Spatial Division** and **Multiplexing**
LOS	**Line**-of-**Sight**
ML	**Maximum-Likelihood**
MRC	**Maximum Ratio Combining**
MRT	**Maximum Ratio Transmission**
MIMO	**Multiple Input Multiple**
Output	
MM-Wave	**Millimetre-Wave**
MMSE	**Minimum Mean-Square Error**
MS	**Mobile Stations**
MPC	**Multi-Path Component**
MCD	**Multi-Path Component Distance**
NLoS	**Non-Line**-of-**Sight**
OLOS	**Obstructed Line**-of-**Sight**
PA	**Power Amplifier**
RV	**Random Variable**
SCA	**Successive Convex Approximation**
SDNR	**Signal**-to-**Distortion Noise Ratio**
SIC	**Successive**-to-**Interference Cancellation**
SNR	**Signal**-to-**Noise Ratio**
SINR	**Signal**-to-**Interference** plus **Noise Ratio**
SOCP	**Second Order Cone Programming**
TDD	**Time Division Duplexing**
TN	**Total Noise**
TQD	**Total Quantization Distortion**
UaF	**Use**-and-**Then-Forget**
VR	**Visibility Region**
WSS	**Wide Sense Stationary**
ZF	**Zero-Forcing**
ZFBF	**Zero-Forcing Beamforming**

Bibliography

[1] H. Q. Ngo, *Massive MIMO: Fundamentals and System Designs*, Ph.D. dissertation, Dept. Electical Eng., Linköping University, Linköping, Sweden, 2015.

[2] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[3] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *[online]. Available: https://arxiv.org/abs/1903.10611*, submitted.

[4] H. Yang and T. L. Marzetta, "Capacity performance of multicell large-scale antenna systems," in *Proc. IEEE Allerton Conference*, Oct. 2013, pp. 668–675.

[5] Y. Xu, G. Yue, and S. Mao, "User grouping for massive MIMO in FDD systems: New design methods and analysis," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6827–6842, July 2015.

[6] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, July 2017.

[7] S. Buzzi and C. DAndrea, "Cell-free massive MIMO: user-centric approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 1–4, Aug. 2017.

[8] G. Interdonato, E. Bjornson, H. Q. Ngo, P. Frenger, and Erik G. Larsson, "Ubiquitous cell-free massive MIMO communications," *IEEE Commun. Mag.*, pp. 1–19, submitted.

[9] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Trans. Wireless Commun. Mag.*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[10] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, Aug. 2010.

[11] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[12] A. Ashikhmin, T. L. Marzetta, and L. Li, "Interference reduction in multi-cell massive MIMO systems I: Large-scale fading precoding and decoding," *Available: http:https://arxiv.org/abs/1411.4182*, Submitted to IEEE Trans. Inf. Theory.

[13] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[14] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[15] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[16] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. and Net.*, vol. 2, no. 1, pp. 25–39, Mar. 2017.

[17] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "MmWave massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 13–21, Oct. 2015.

[18] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user mimo systems: Is massive MIMO the answer?," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, June 2015.

[19] G. Lee and Y. Sung, "A new approach to user scheduling in massive multi-user mimo broadcast channels," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1481–1495, Apr. 2018.

[20] 3GPP Specification 38.900, *Study on channel model for frequency spectrum above 6 GHz*, IEEE802.11-09/0334r3, 2016.

[21] A. F. Molisch and F. Tufvesson, "Propagation channel models for next-generation wireless communications systems," *IEEE Trans. Commun.*, vol. E97-B, no. 10, pp. 2022–2034, Oct. 2014.

[22] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[23] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*, Cambridge University Press, 2016.

[24] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*, Cambridge University Press, 2016.

[25] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *Proc. IEEE EUSIPCO*, Sept. 2014, pp. 1–6.

[26] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[27] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[28] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. IEEE VTC*, May 2001, pp. 1745–1749.

[29] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "On the performance of cell-free massive MIMO with short-term power constraints," in *Proc. IEEE CAMAD*, Oct. 2016, pp. 1–6.

[30] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.

[31] H. Ahmadi, A. Farhang, N. Marchetti, and A. MacKenzie, "A game theoretic approach for pilot contamination avoidance in massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 12–15, Feb. 2016.

[32] R. Verdone and Eds. A. Zanella, *Pervasive Mobile and Ambient Wireless Communications: COST Action 2100*, Springer, 2012.

[33] L. M. Correia, *Mobile Broadband Multimedia Networks*, Academic Press, San Diego, 2006.

[34] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Baltimore: Johns Hopkins, 1996.

[35] M. J. Feuerstein, K. L. Blackard, T. S. Rappaport, S. Y. Seidel, and H. H. Xia, "Path loss, delay spread, and outage models as a function of antenna height for microcellular system design," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp. 487–498, Aug. 1994.

[36] P. Zillmann, "Relationship between two distortion measures for memoryless nonlinear systems," *IEEE Signal Process. Lett.*, vol. 17, no. 11, pp. 917–920, Feb. 2010.

[37] L. C. Andrews, *Special Functions of Mathematics for Engineers*, Oxford University Press, second edition, 1997.

[38] J. Max, "Quantizing for minimum distortion," *IEEE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.

[39] K. Cumanan, R. Krishna, L. Musavian, and S. Lambotharan, "Joint beamforming and user maximization techniques for techniques for cognitive radio networks based on branch and bound method," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3082–3092, Oct. 2010.

[40] K. Cumanan, L. Musavian, S. Lambotharan, and A. B. Gershman, "SINR balancing technique for downlink beamforming in cognitive radio networks," *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 133–136, Feb. 2010.

[41] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161 – 176, Jan. 2006.

[42] D. W. H. Cai, T. Q. S. Quek, and C. W. Tan, "A unified analysis of max-min weighted SINR for MIMO downlink system," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3850–3862, Aug. 2011.

[43] M. Schubert and H. Boche, "Solution of the multiuser downlink beam-forming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18 – 28, Jan. 2004.

[44] D. N. C. Tse and P. Viswanath, "Downlink-uplink duality and effective bandwidths," in *Proc. IEEE ISlT*, July 2002.

[45] G. Golub and C. V. Loan, *Matrix Computations*, Baltimore, MD: The Johns Hopkins Univ. Press, second edition, 1996.

[46] S. P. Boyd, S. J. Kim, A. Hassibi, and L. Vandenbarghe, "A tutorial on geometric programming," *Optim. Eng.*, vol. 8, no. 1, pp. 67–128, 2007.

[47] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, *Power Control by Geometric Programming, in Resource Allocation in Next Generation Wireless Networks*, W. Li, Y. Pan, Editors, Nova Sciences Publishers, 2006.

[48] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-time signal processing*, Prentice-hall Englewood Cliffs, 1989.

[49] A. Pizzo, D. Verenzuela, L. Sanguinetti, and E. Björnson, "Network deployment for maximal energy efficiency in uplink with multislope path loss," *IEEE Trans. Green Commun. and Net.*, pp. 1–30, submitted.

[50] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, UK: Cambridge University Press, 2004.

[51] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under sinr constraints," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2324–2334, July 2005.

[52] A. J. Fehske, P. Marsch, and G. P. Fettweis, "Bit per joule efficiency of cooperating base stations in cellular networks," in *Proc. IEEE Globecom Workshops*, Dec. 2010, pp. 1406–1411.

[53] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[54] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *Proc. IEEE ISIT*, Aug. 2012, pp. 1–5.

[55] A. Mezghani, *Information theoretic Analysis and Signal Processing Techniques for Quantized MIMO Communications*, Ph.D. dissertation, Technical University of Munich, Germany, 2015.

[56] A. Kakkavas, J. Munir, A. Mezghani, H. Brunner, and J. A. Nossek, "Weighted sum rate maximization for multiuser MISO systems with low resolution digital to analog converters," in *Proc. IEEE WSA*, Mar. 2016.

[57] J. Munir, D. Plabst, and J. A. Nossek, "Efficient equalization method for cyclic prefix-free coarsely quantized massive MIMO systems," in *Proc. IEEE ICC*, May 2018, pp. 1–6.

[58] J. Max, "The worst additive noise under a covariance constraint," *IEEE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Nov. 1960.

[59] J. G. Proakis, *Digital Communications*, McGraw-Hill, 1995.

[60] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, June 2015.

[61] O. Onireti, F. Heliot, and M. A. Imran, "On the energy efficiency-spectral efficiency trade-off of distributed MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3741–3753, Sept. 2013.

[62] O. Onireti, F. Heliot, and M. A. Imran, "On the energy efficiency-spectral efficiency trade-off in the uplink of CoMP system," *IEEE Trans. Commun.*, vol. 11, no. 2, pp. 556–561, Feb. 2012.

[63] L. Falconetti and E. Yassin, "Towards energy efficiency with uplink cooperation in heterogeneous networks," in *Proc. IEEE WCNC*, Apr. 2014, pp. 1649–1654.

[64] S. He, Y. Huang, L. Yang, B. Ottersten, and W. Hong, "Energy efficient coordinated beamforming for multicell system: duality-based algorithm design and massive MIMO transition," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4920–4935, Dec. 2013.

[65] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, Jan. 2010.

[66] P. C. Weeraddana, M. Codreanu, M. Latva-aho, and A. Ephremides, "Resource allocation for cross-layer utility maximization in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2790–2809, July 2011.

[67] E. Björnson and E. Jorswieck, *Optimal Resource Allocation in Coordinated Multi-Cell Systems*, Foundations and Trends in Communications and Information Theory, 2012.

[68] P. Arbenz, *Lecture Notes on Solving Large Scale Eigenvalue Problems*, Computer Science Department, ETH Zurich, [online]. available:https:https://people.inf.ethz.ch/arbenz/ewp/lnotes/lsevp.pdf edition, 2016.

[69] Y. Nesterov and A. Nemirovsky, *Interior-point polynomial methods in convex programming*, Studies in Applied Mathematics, SIAM, Philadelphia, 1994.

[70] R. Price, "A useful theorem for nonlinear devices having gaussian inputs," *IEEE Trans. Inf. Theory*, vol. 4, no. 2, pp. 69–72, June 1958.

[71] H. Hofstetter, A. F. Molisch, and N. Czink, "A twin-cluster MIMO channel model," in *Proc. IEEE EuCAP*, Nov. 2006.

[72] A. Adhikary, E. Al Safadi, M. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, "Joint spatial division and multiplexing for mm-wave channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1239–1255, June 2014.

[73] U. Grenander and G. Szegõ, *Toeplitz forms and their applications*, London, UK: Chelsea, 1984.

[74] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[75] T. Yoo and A. Goldsmith, "Sum rate optimal multi-antenna downlink beamforming strategy based on clique search," in *Proc. IEEE Globecom*, Dec. 2005.

[76] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process*, vol. 14, no. 5, pp. 2868–2882, Jan. 2015.

[77] S. Kay, *Fundamentals of statistical signal processing: Estimation theory*, Englewood Cliffs, NJ, USA, Prentice-Hall, 1993.

[78] I. Viering, H. Hofstetter, and W. Utschick, "Spatial long-term variations in urban, rural and indoor environments," in *Proc. the 5th Meeting of COST273, Lisbon, Portugal*, Sept. 2002.

[79] D. Mitrovic and D Zubrinic, *Fundamentals of Applied Functional Analysis*, CRC Press, 1997.

[80] A. G. Burr, "Capacity bounds and estimates for the finite scatterers MIMO wireless channel," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 812–818, June 2003.

[81] A. G. Burr, "Multiplexing gain of multiuser MIMO on finite scattering channels," in *Proc. IEEE ISWCS*, May 2010, pp. 1–5.

[82] V. Y. Pan and Z. Q. Chen, *The complexity of the matrix eigenproblem*, 31st Annual ACM Symp. on Theory of Computing, New York, 1999.

[83] T. Yoo, *Sum-capacity, scheduling, and multi-user diversity in MIMO broadcast systems*, Ph.D. dissertation, Stanford University, United states, 2007.

[84] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[85] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 40–60, Jan. 2013.

[86] P. Lu and H. C. Yang, "Sum rate analysis of multiuser MIMO system with zero-forcing transmit beamforming," *IEEE Trans. on Commun.*, vol. 57, no. 9, pp. 2585–2589, Sept. 2009.

[87] LTE, *Study on channel model for frequency spectrum above 6 GHz*, 3GPP Specification 38.900, May 2016.

[88] L. Liu, J. Poutanen, F. Quitin, K. Haneda, F. Tufvesson, P. D. Doncker, P. Vainikainen, and C. Oestges, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.

[89] M. K. Samimi and Theodore S. Rappaport, "Statistical channel model with multi frequency and arbitrary antenna beamwidth for millimetre-wave outdoor communications," in *Proc. IEEE Globecom*, Dec. 2015.

[90] A. Karttunen, J. Järveläinen, A. Khatun, and K. Haneda, "Radio propagation measurements and WINNER II parametrization for a shopping mall at 61-65 GHz," in *Proc. IEEE VTC*, May 2015, pp. 1–6.

[91] K. Haneda., L. Tian, H. Asplund, J. Li, Y. Wnag, D. Steer, C. Li, T. Balercia, S. Lee, Y.-S. Kim, A. Ghosh, T. Tomas, T. Nakamura, Y. Kakishima, T. Imai, H. Papadopulas, T. S. Rappaport, G.-R. McCartney, M. K. Samimi, S. Sun, O. Koymen, S. Hur, J. Park, J. Zhang, E. Mellios, A. F. Molisch, S. S. Ghassamzadeh, and A. Ghosh, "Indoor 5G 3GPP-like channel models for office and shopping mall environments," in *Proc. IEEE ICC Workshop*, May 2016, pp. 694–699.

[92] C. Schneider, J. Gedschold, M. Kaske, R. S. Thoma, and G. D. Galdo, "Estimation and characterization of multipath clusters in urban scenarios," in *Proc. IEEE EuCAP*, Apr. 2018, pp. 1–5.

[93] N. Iqbal, D. Dupleich, C. Schneider, J. Luo, R. Muller, S. Hafner, G. D. Galdo, and R. S. Thoma, "Tmodeling of intra-cluster multipaths for 60 GHz fading channels," in *Proc. IEEE EuCAP*, Apr. 2018, pp. 1–5.

[94] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-wave multipath clustering and channel modeling," *IEEE Trans. Ant. Prop.*, vol. 62, no. 3, pp. 1445–1455, Mar. 2014.

[95] N. Czink, R. Tian, S. Wyne, F. Tufvesson, J. P. Nuutinen, J. Ylitalo, E. Bonek, and A. F. Molisch, "Tracking time-variant cluster parameters in MIMO channel measurements," in *Proc. CHINACOM*, Aug. 2007, pp. 1147–1151.

[96] N. Czink, E. Bonek, L. Hentila, J. P. Nuutinen, and J. Ylitalo, "Cluster-based MIMO channel model parameters extracted from indoor time-variant measurements," in *Proc. IEEE Globecom*, Nov. 2006, pp. 1–5.

[97] K. Haneda, J. Järveläinen, and A. Karttunen, "[online]. available: https://arxiv.org/pdf/1802.08591.pdf," in *Proc. IEEE VTC*, 2018, pp. 1–6.

[98] J. Järveläinen, K. Haneda, and A. Karttunen, "Indoor propagation channel simulations at 60 GHz using point cloud data," *IEEE Trans. Ant. Prop.*, vol. 64, no. 8, pp. 4457–4467, Aug. 2016.

[99] N. Czink, P. Cera, J. Salo, E. Bonek, J. P. Nuutinen, and J. Ylitalo, "Improving clustering performance using multipath component distance," *Electronics Letters*, vol. 42, no. 1, pp. 1–2, Jan. 2006.

[100] N. Czink, P. Cera, J. Salo, E. Bonek, J. p. Nuutinen, and J. Ylitalo, "A framework for automatic clustering of parametric MIMO channel data including path powers," in *Proc. IEEE VTC*, Sept. 2006, pp. 1–5.

[101] Q. Bai and J. A. Nossek, "Energy efficiency maximization for 5G multiantenna receivers," *Trans. Emerg. Telecommun. Technol.*, vol. 26, no. 1, pp. 3–4, Jan. 2015.