Coordinated Multicast Beamforming for Multi-cell Massive MIMO Systems

by

Shiqi Yin

A Thesis Submitted to the
School of Graduate and Postdoctoral Studies in Partial
Fulfillment of the Requirements for the Degree of
Masters of Applied Science
in
Department of Electrical, Computer and Software Engineering
The Faculty of Engineering and Applied Science

University of Ontario Institute of Technology

September 2023

Thesis Examination Information
Submitted by: **Shiqi Yin**

Masters of Applied Science in **Electrical and Computer Engineering**

Thesis title: Coordinated Multicast Beamforming for Multi-cell Massive MIMO System

An oral defense of this thesis took place on September 21, 2023 in front of the following examining committee:

Examining Committtee:

Chair of Examining Committee	Dr. Jana Abou-Ziki
Research Supervisor	Dr. Min Dong
Examining Committee Member	Dr. Ying Wang
Thesis Examiner	Dr. Mehran Ebrahimi

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

Abstract

In this thesis, we consider the efficient coordinated multi-cell multicast beamforming design in massive massive multiple-input multiple output (MIMO) cellular networks. For the problem that aims to minimize the base station transmit power, we obtain the optimal multicast beamforming solution structure. Utilizing the optimal structure to reduce the computational complexity, we propose to determine the parameters in the optimal beamforming solution via semi-definite relaxtion (SDR) or successive concex approximation (SCA). To further design a fast algorithm applicable to massive MIMO systems, we apply the alternating direction method of multipliers (ADMM) technique and propose an ADMM-based first-order algorithm to solve the quality of service (QoS) problem, which decompose the QoS problem into subproblems with closed/semi-closed form updates.

Following this, we consider the max-min fair problem that aims to maximize the minimum signal to the interference and noise (SINR) of each user. We present the optimal max-min fair (MMF) coordinated multicast beamforming solution structure through the inverse relation between the QoS and MMF problem. We propose to use the projected subgradient algorithm (PSA) to obtain a solution to the MMF problem, which is a fast algorithm with closed-form updates and low computational complexity.

Simulation results show that our proposed methods can achieve near-optimal performance. Besides, with much lower computational complexity in a massive MIMO system than existing algorithms.

Keywords: multi-cell; multicast beamforming; MIMO; coordinated beamforming

Author's Declaration

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

_____ Shiqi Yin

Statement of Contributions

Part of this thesis is under preparation to be submitted for journal publication as:

S. Yin and M.Dong, "Low-complexity and structurally optimal coordinated multicast beamforming design for Multi-cell Massive MIMO Systems," to be submitted to IEEE Transactions on Wireless Communications.

I performed the majority of development of the algorithm design, simulation experiement and the writing of the manuscript under the supevison or Prof. Min Dong. I have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Min Dong. Throughout my two years at Ontario Tech University, her enormous knowledge and experience taught me how to do research systematically. I have learned many necessary skills from her kind instruction, patience, motivation, and enthusiasm. Without my supervisor, I would not have improved my studying, writing, and research skills during the last two years.

I would like to show my thanks to my group members, *i.e.,* Yong, Mohammad, and Reza. Their encouragement helped me keep pursuing my research. Also, I would like to thank my friends. Their company made me not feel alone and helped me adopt a new culture, especially Kevin, Lingli, and Nancy.

In the end, I would like to thank my family for their support and love. Their caring and deep love help me insist when I have to face every difficulty!

Contents

List of Figures

List of Tables

List of Abbreviation

5G	5th generation
ADMM	alternating direction method of multipliers
BS	base station
IoT	Internet of Things
i.i.d	independent and identically distributed
KKT	Karush-Kuhn-Tucker
MIMO	massive multiple-input multiple output
MMF	max-min fair
MMSE	minimum mean square error
MRT	matimum ratio transmit
NP-hard	non-deterministic polynomial-time hard
PSA	projected subgradient algorithm
QoS	quality of service
QCQP	quadrtically constrained quadratic program
SDR	semi-definite relaxtion
SCA	successive concex approximation
SOCP	second-order cone programming
SDP	semi-definite programming
SINR	signal to the interference and noise
ZF	zero-forcing

Chapter 1

Introduction

1.1 Overview

In recent decades, wireless communication technologies has developed dramatically. According to [1], the number of subscribers worldwide has increased from 300 million to more than 6 billion. The large amount mobile users leads to the vast development of wireless services, and the user demands significantly different as compared to 25 years ago. Nowadays, the new activities provided by the smartphone, *i.e.,* video/audio streaming, social media, gaming, and other host apps, require a higher data transmission rate. In addition, other non-human-operated devices like the Internet of Things (IoT) also need robust wireless communication. To satisfy these increasing requirements, many new wireless technologies have emerged in recent years [2–15]. Among them, massive multiple-input multiple output (MIMO) system has become a underpinning technology for the 5th generation (5G) and beyond wireless network. According to the *Cisco Visual Networking Index* Forecast for 2017-2022 [16], wireless data will grow at a compound annual growth rate of 46%, and video data is expected to account for 82% of all internet traffic by 2022, which grows from 75% in 2017. Moreover, users demand high quality of videos, such as or/and these types of videos

require high transmit data rate.

The massive MIMO system [2] is a MIMO system equipped with large scale antennas, which can be considered as an extension of multi-user MIMO system. Compared to point-to-point or multi-user MIMO, the massive MIMO has shown more benefits [3]: improving the energy efficiency, reducing interference between groups or cells. For energy efficiency, the massive MIMO comprise small and low cost antennas, these independent antennas make the system more robust and reduce the energy cost. In addition, the large scale antennas at the transmitter enables transmitting different signals to multiple users simultaneously, which leads to the massive MIMO achieve spatial multiplexing. Thus, both the signal to the interference and noise (SINR) at users and the network capacity can be increased significantly [17].

Besides the massive MIMO, another important technique to satisfy the nowadays increasing demanding is beamforming. Since 5G or beyond wireless networks and services require ultra high data rates, coverage and connectivity, the multi-antenna multicast beamforming has been playing a key role in recent years for supporting high-speed content distribution. This transmission technology is to send the common data to multiple users simultaneously. Based on the user distribution, multicast beamforming can be divided into several categories, *i.e.,* single-group, multi-group, and multi-cell multicast beamforming. For designing beamforming techniques, several major performance metrics, such as transmit power, quality of service (QoS) [18–24], max-min fair (MMF), [21, 25–30], are considered most recently. The objective of QoS problem is to minimize the transmit power, which is subject to the SINR among all the users satisfy the target. For the MMF problem, which is to maximize the worst

SINR among all users, subject to the transmit power satisfy a budget.

Since the family of the above problems (QoS and MMF) are non-deterministic polynomial-time hard (NP-hard) [18], most literature works are rely on designing iterative algorithms to solve them. Among the existing approaches, the semi-definite relaxtion (SDR) [18, 19, 21, 31, 32] is a typical method to get an approximate solution to solve the QoS or MMF problem. However, these SDR-based methods perform deteriorate in large scale antenna systems. To overcome this drawback, thesuccessive concex approximation (SCA) [33, 34] is proposed for multicast beamforming with a larger number of antennas. However, the computational complexity will still grow relatively fast as the number of antenna array increases, which is not suitable for massive MIMO systems. Therefore, more numerical algorithms are developed to reduce the computational complexity under massive MIMO systems.

For 5G wireless networks design, the combination of massive MIMO system with the multicast beamforming technique has been becoming a major trend to support high user data rate. In a massive MIMO cellular system, the base stations (BSs) can coordinate with each other to send user data, which is beneficial to reduce the interference [35]. For downlink transmission, there are two types of multi-cell multicast beamforming: coordinated beamforming and cooperative beamforming.

For coordinated beamforming [4, 5, 36, 37], the BSs coordinate their beamforming vectors to transmit data to reduce interference to each other. There is no data information exchange between the BSs. Therefore, the inter-cell interference can be reduced and the SINR at users can be improved effectively. Also, the BSs can improve the overall performance by jointly optimizing the beamformers. These benefits

make the coordinated multi-cell multicast beamforming one main technique in multi-antenna cellular systems. For coordinated multicast beamforming in conventional multi-antenna systems, both the QoS and MMF problems are considered in [31], and the SDR method is applied to find a solution. Also, the authors in [38, 39] designed a low-complexity multi-cell multicast beamforming method to obtain a suboptimal solution for massive MIMO system.

For cooperative beamforming, BS cluster is formed by several BSs to share data among BSs and transmit data to users simultaneously. This approach requires stringent symbol-level synchronization among BSs, which is difficult to achieve in practice.

There are limited literature works to design coordinated multicast beamforming for massive MIMO systems. In particular, no existing work study the optimal beamforming structure or propose other low-complexity algorithms for multi-cell multicast beamforming in massive MIMO systems.

1.2 Motivation and Objective

The optimal beamforming structure for both QoS and MMF problems under the multi-group multicast beamforming scenario with a single BS have been obtained in [40], which reveals a low-dimensional structure that can be exploited for the beamforming design. In particular, it shows that the computational method can be independent of the number of users. In a multi-cell multicast scenario, the weighted matimum ratio transmit (MRT), and zero-forcing (ZF) multicast beamforming structure are proposed to solve the MMF problem under [38]. However, these considered structures are suboptimal. The optimal coordinated mulitcast beamforming structure in the multi-

cell multicast scenario is still unknown, and whether it will entail a low-complexity beamforming design for massive MIMO multi-cell systems is still an open problem.

This thesis considers the multi-cell multicast downlink transmission for a large-scale massive MIMO system under a multi-cell multicast scenario. Motivated by the above challenges, we aim to obtain the optimal beamforming structure for the QoS and MMF problems. Furthermore, we aim to explore the optimal structure to design low-complexity algorithms for obtaining the coordinated multicast beamforming solution for both the QoS and MMF problems.

1.3 Thesis Contribution

This thesis mainly considers solving the QoS and MMF problem under a multi-cell multicast beamforming scenario. We first develop the optimal beamforming structure for the QoS problem based on [40]. Following this structure, we proposed several methods to obtain the optimal/suboptimal beamforming solution. Also, for the QoS problem, we adopted the alternating direction method of multipliers (ADMM) [41] combined with the optimal structure to decrease the complexity. Furthermore, a second-order iteratively algorithm and a fast-first order projected subgradient algorithm (PSA) [42], along with the optimal structure of QoS, are presented to solve the MMF problem.

1. The optimal QoS multicast beamforming solution

   Compared with the multi-group multicast scenario, the power inequality constraint in the multi-cell multicast scenario is individual, implying that the QoS becomes a min-max optimization problem with more constraints than the multi-

group case. We obtain that the optimal beamforming structure at each BS and show that it is only related to the channel covariance matrices and its own cell's user channels, we also prove the inverse relation between the QoS and MMF problem, which leads to the optimal beamforming structure for the MMF problem.

2. Efficient algorithm for QoS multicast beamforming

   Based on the optimal QoS beamforming structure, we apply several common numerical methods, *e.g.,* SDR, and SCA, to solve the QoS problem but in a much reduced problem dimension. To develop a computationally efficient algorithm suitable for massive MIMO, we develop an ADMM-based fast algorithm to obtain the multicast beamforming solution, where we obtain semi-closed form updates to iteratively solve each SCA subproblem. The simulation results show that our proposed algorithms can nearly achieve the stationary optimal solution. At the same time, the complexity is reduced substantially compared with directly utilizing SDR or SCA.

3. Efficient algorithm for MMF multicast beamforming

   We utilize the optimal beamforming structure and propose a first-order fast algorithm to solve the MMF problem. First, following the inverse relation between QoS and MMF, we iteratively solve the MMF problem by utilizing the optimal beamforming structure and the QoS beamforming solution. The reduced-dimension size structure can not only help to reduce the consumption cost but achieves near-optimal performance. Next, to avoid solving the QoS problem

multiple times and further reduce the complexity, we adopt a projected subgradient algorithm [42] to solve the MMF problem under the multi-cell scenario. The PSA-based algorithm directly solves the MMF problem. It only requires simple closed-form updates and has convergence guarantee. The simulation results show near-optimal performance and fast computation time as compared with existing algorithms, demonstrating its effectiveness.

1.4 Thesis Organization

The organization of this thesis is arranged as follows. In Chapter 2, we mainly present beamforming techniques for multi-group/multi-cell multicast scenarios, the standard form of QoS or MMF problem, and fast algorithm techniques. In Chapter 3, we mainly focused on deriving the optimal beamforming structure and following with the fast first-order algorithms to solve the QoS problem. In Chapter 4, the iteratively second-order algorithm and the fast first-order algorithm PSA are presented. In Chapter 5, we show the conclusion of this thesis.

1.5 Notation

In this thesis, transpose, Hermitian, trace, and conjugate of $\mathbf{A}$ are denoted by $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathrm{tr}(\mathbf{A})$ and $\mathbf{A}^*$ respectively. An identity matrix is denoted by $\mathbf{I}$. A semi-definite matrix $\mathbf{A}$ is denoted as $\mathbf{A} \succcurlyeq \mathbf{0}$. Notation $\mathbf{x} \sim \mathcal{CN}(a, \mathbf{Y})$ means $\mathbf{x}$ forms the complex Gaussian distribution with mean a and covariance matrix $\mathbf{Y}$. $\mathbb{E}[\mathbf{x}]$ stands for the expected value of variable $\mathbf{x}$. For the convergence study section, $a \to b$ represents a converges to b, and ϵ is defined as a fixed threshold for all the algorithms. The

Euclidean norm of a vector $\mathbf{a}$ is denoted by $\|\mathbf{a}\|$. The abbreviation i.i.d. stands for independent and identically distributed.

Chapter 2

Literature Review

2.1 Beamforming Techniques

Beamforming techniques have been studied for wireless communication a couple of decades ago since combining multiple signal paths can provide diversity and array gain. Nowadays, this technique has already become a crucial part of communication applications. Generally, there are two major beamforming techniques: unicast and multicast beamforming. In the following subsections, we present a brief overview of different techniques for beamforming.

2.1.1 Unicast Beamforming

Unicast beamforming is a typical technique for transmitting data from BS to users. It is used by the BS to send the private data stream dedicated to each user. For regular downlink unicast beamforming, QoS, MMF, or sum-rate maximization are typically considered as the beamforming design metrics [43–45]. For example, the author in [43] combined the diversity scheme and power allocation in a downlink cellular wireless system. In [44], a joint downlink beamforming and power control problem was considered in a frequency flat multi-antenna broadcast channel. In [45], a simple closed-form

solution for power minimization in the downlink unicast beamforming scenario was provided.

2.1.2 Multicast Beamforming

As the increasing demands for wireless multicast services and applications, *e.g.,* video conference, mobile commerce, intelligent transportation system, multicast beamforming techniques have been developed for this type of services. Multicast beamforming is a transmission technique used by the BS to send the common data to multiple users simultaneously. This technique provides better spectrum and power efficiency in sending the common data to multiple users, than using the conventional unicast beamforming to treat the common message as private data to send to each user. Based on the transmission scenarios, the multicast beamforming technique can be divided into single-group multicasting, multi-group multicasting, multi-cell multicasting.

Single-Group Multicast Beamforming

Single-group multicast beamforming is a technique that only considers a single group of users in the cell, and all the users in the group will receive the same message from the BS. The QoS and the MMF problem under single-cell single group scenario are considered in [18]. In [46], it is proven that both the QoS and MMF problem are NP-hard, and the SDR-based method is developed to solve these NP-hard problems.

The two major problems QoS and MMF are already proven to be NP-hard and non-convex [47]. The sub-optimal solutions obtained at the early stage mostly are through the SDR-based methods along with Gaussian randomization to recover a feasible solution [19,47]. Furthermore, [19] also developed a new strategy called stochas-

tic beamforming to avoid randomization, and at the same time improve the SDR approximation performance. Due to the high computational complexity faced by the SDR-based methods, [33] proposed to apply the SCA method for the QoS problem, which leads to a convergent iterative second-order cone programming (SOCP) solution. Different from the above works, [48] study a particular class of non-convex quadrtically constrained quadratic program (QCQP) problems. They construct a convex surrogate function and present a first-order method to solve the subproblems more efficiently than directly using the SCA method.

Multi-Group Multicast Beamforming

Extending the single-group multicast beamforming scenario to multiple groups [21, 25, 26, 49], the BS sends messages to multiple group, one message to one group. Thus, the users in the same group receive common data from the BS. This scenario can efficiently multiplex data to multiple groups but also bring inter-group interference. Thus, beamforming design is more challenging in this case.

For the multiple multicast group scenario, both QoS and the MMF problem are considered in [21]. Also, the authors in [25] proposed a suboptimal solution for the MMF problem under per-antenna power limit. To avoid the inter-group interference, dirty paper preceding is applied for minimizing the total power objective in [49]. In [26], a SCA-based algorithm is considered to solve the MMF problem for the multi-group multicasting scenario.

Multi-Cell Multicast Beamforming

Multicast beamforming design has also been studied for multi-cell networks. The early study in [50] mainly considered maxmizing the worst SINR in a single frequency network, in which the algorithm is based on spatio-temporal correlation knowledge, with a regular update of SINR. In [31], both the QoS and MMF beamforming are investigated under coordinated multi-cell multicast beamforming. A QoS-based algorithm for the QoS problem was proposed and the inverse relationship between the QoS and MMF problem. In [51], BS-clustering and local BS cache are applied jointly to minimize the transmit power for each multicast group.

Under different multicast beamforming scenarios, two types of problem formulations are typically considered: QoS problem and the MMF.

2.2 Massive MIMO Systems

2.2.1 Overview

As the wireless service demanding grows, *i.e.,* social media, intelligent system, IoT, the requirement for higher wireless communication technique has been increasing in recent decades. These increasing demandings make the massive MIMO system become an indispensable part. Compared to the traditional MIMO system *e.g.,* point-to-point MIMO system [52–54], multi-user MIMO system [55–59], massive MIMO system can improve the data rate or link reliability [2]. Furthermore, other benefits of the massive MIMO system are inexpensive low-power components, reduced latency, simplification of the MAC layer, and robustness against intentional jamming [3].

Generally, the massive MIMO system is combined with the beamforming tech-

nique. A single-group multicast beamforming problem with large antenna/user numbers is considered in [60]. They proposed linear programming-assisted sub gradient descent method, which is a first-order method for solving the MMF problem. In [61], the MMF problem under single-group in a massive MIMO setting is studied.

For multi-group multicasting in massive MIMO, the authors in [62] considered the MMF problem using a fixed maximum-ratio transmission or ZF beamforming structures. In [63], the ADMM technique was proposed for multi-group multicast beamforming design in large-scale wireless systems. The optimal beamforming structures for the QoS and the MMF problems are obtained in [40]. The obtained beamforming structure shows an inherent low-dimensional structure, which provides opportunity for low-complexity algorithm designs for massive MIMO systems. Following this, the author in [64] developed two fast-first order algorithms to solve the QoS problem, including the extragradient method and the ADMM-based method. Besides these computational methods, the authors in [65] consider applying jointly adaptive grouping and beamforming to improve the data rate in massive MIMO system.

Jointly unicast and multicast beamforming in massive MIMO has also been considered in the literature. HybridCast is designed in [66] for multi-user MIMO system to improve the efficiency. Since the pure multicasting is sent at a base rate, which would waste the link margin when delivering extra information. Different only considering the QoS and MMF problem, multi-objective optimization problem is studied in [27]. Furthermore, a null-space method based interference cancellation is proposed in [67]. In [68], the jointly problem was decomposed into two subproblems for the unicast users and multicast users respectively. For the unicast part, a closed-form

solution was presented in [68]. For the multicast part, based on the optimal multicast beamforming structure in [40], the authors combined it with an ADMM-based fast algorithm to get closed-form updates for the QoS problem beamforming design.

The multi-cell multicast beamforming in massive MIMO has been investigated in several studies [38, 39, 69, 70]. Coordinated multicast beamforming design was considered in [69], and the pilot scheme for eliminating contamination was proposed for coordinated multi-cell network. In [38] the MMF problem for both coordinated and cooperative multi-cell multicast beamforming with large-scale antennas was considered, and the suboptimal beamforming schemes, weighted maximum ratio transmission and zero-forcing beamforming, were used to reduce the computational complexity. In [39], with the same suboptimal beamforming schemes, the SDR-based approach was proposed for maximizing the minimum signal-to-leakage ratio. In [70], a constant envelope precoding construction was proposed for reducing the implementation cost and improving efficiency.

Besides the QoS and MMF problem mentioned above, other design objectives have also been considered under multicast beamforming. In [29, 71], the energy-efficient problem is studied under multi-cell multi-user multigroup multicast scenario, in which the joint coordinated beamforming and antenna selection are applied. For the sum-rate maximization problem, are considered in [72–74].

2.2.2 Challenges in Massive MIMO

Although the massive MIMO system can bring many benefits, there still exit challenges to design effective and efficient beamforming schemes for massive MIMO. In

recent years, one main focus is on developing low-complexity algorithms that are implementable in massive MIMO.

In [40], the optimal beamforming structure is proposed for solving the QoS problem. This low-dimension structure is independent with the antenna size, which yield a huge efficiency improvement. For the large number user case, the first-order algorithm can reduce the computational complexity. The ADMM-based algorithms [63,64,68,75] mainly focus on decomposing the QoS problem to a sequence of subproblems, and a closed/semi-closed form solution can be generated for each subproblem. Besides, an extragradient method with an adaptive step-size procedure is proposed in [64] to solve the QoS problem, which can also provides closed-form updates.

However, these low-complexity first-order algorithms are mostly designed for multi-group multicast beamforming with a single BS in massive MIMO. The optimal structure or algorithms cannot be directly applied to the multi-cell case, when multiple BSs exist with individual BS power requirement. In this thesis, we consider to design the optimal beamforming structure and low-complexity algorithms for multi-cellular systems.

For the MMF problem, the optimal beamforming structure derived in [40] is difficult to be used directly due to unknown parameters. The MMF problem is solved via solving the QoS problem iteratively [40]. Although this method can reduce the computational complexity at each iteration, the computational efficiency still needs to be improved. In [42], a first-order algorithm, the PSA is designed for overcoming this challenge, for multi-group mutlicast beamforming with a single BS. However, limited results exist for coordinated multicast beamforming for massive MIMO.

Chapter 3

The Optimal QoS Multi-cell Multicast Beamforming Design

In this chapter, we first consider deriving the optimal beamforming structure for the QoS problem under a multi-cell multicast beamforming scenario. Following this, we proposed numerical algorithms to solve the QoS problem. The optimal structure is applied for both the fast first-order and second-order algorithms.

3.1 System Model

We consider a downlink multicast transmission scenario in a multi-cell massive MIMO system consisting of J cells. The BS in each cell is equipped with M antennas. Each BS provides the multicast service to a group of K users in its cell, where each user is equipped with a single antenna. We assume that all BSs use the same spectrum bandwidth for transmission.

We study the coordinated multicast beamforming among the J coordinating cells for the multicast service. Each BS multicasts a message to the K users in its own cell, using the beamforming vector that is jointly designed among all the BSs. Define the cell index set $\mathcal{J} \triangleq \{1, \cdots, J\}$ and the user index set $\mathcal{K} \triangleq \{1, \cdots, K\}$. The serving

BS in cell j is denoted by BS j. Let $\mathbf{h}_{j,ik}$ denote the $M \times 1$ channel vector from BS j to user k in cell i, for $j, i \in \mathcal{J}$ and $k \in \mathcal{K}$. Let $\mathbf{w}_i$ denote the $M \times 1$ multicast beamforming vector at BS i. The received signal at user k in cell i is given by

$$y_{ik} = \mathbf{w}_i^H \mathbf{h}_{i,ik} s_i + \sum_{j=1, j\neq i}^{J} \mathbf{w}_j^H \mathbf{h}_{j,ik} s_j + n_{ik}, \ k \in \mathcal{K}, \ i \in \mathcal{J}. \tag{3.1}$$

where s_i is the data symbol transmitted from BS i with $\mathrm{E}[|s_i|^2] = 1$, and n_{ik} is the receiver additive white Gaussian noise at the user with zero mean and variance σ^2. The first term in (4.2) is the desired signal, and the second term is the interference from the other BSs of the coordinated neighboring cells.

In this chapter, we consider the QoS multicast beamforming problem. The objective is to minimize the BS transmit power while meeting the minimum SINR target at each user. For a multi-cell system, each BS may have its individual power budget target, denoted by p_i, $i \in \mathcal{J}$. Therefore, this problem is formulated as minimizing the maximum power margin $\|\mathbf{w}_i\|^2/p_i$ of all the BSs among the coordinated cells, given by

$$\begin{aligned} \mathcal{P}_o : & \min_{\mathbf{w}} \max_{i} \ \frac{1}{p_i} \|\mathbf{w}_i\|^2 \\ \text{s.t.} \ & \frac{|\mathbf{h}_{i,ik}^H \mathbf{w}_i|^2}{\sum_{j=1, j\neq i}^{J} |\mathbf{h}_{j,ik}^H \mathbf{w}_j|^2 + \sigma^2} \geqslant \gamma_{ik}, \ k \in \mathcal{K}, i \in \mathcal{J} \end{aligned}$$

where $\mathbf{w} \triangleq [\mathbf{w}_1^H, \cdots, \mathbf{w}_J^H]^H$ is the concatenated multicast beamforming vectors of all BSs, and γ_{ik} is the minimum SINR target at user k in group i.

3.2 The Optimal Coordinated Multicast Beamforming

We now consider the QoS problem $\mathcal{P}_o$ for multi-cell coordinated multicast beamforming. The problem has a min-max objective. To handle this, we introduce the auxiliary variable t and convert $\mathcal{P}_o$ into the following equivalent problem for $(\mathbf{w}, t)$:

$$\begin{aligned} \mathcal{P}_1: \ & \min_{\mathbf{w},t} \ t \\ \text{s.t.} \ & \frac{|\mathbf{h}_{i,ik}^H \mathbf{w}_i|^2}{\sum_{j=1,j\neq i}^{J} |\mathbf{h}_{j,ik}^H \mathbf{w}_j|^2 + \sigma^2} \geqslant \gamma_{ik}, \ k \in \mathcal{K}, i \in \mathcal{J} \\ & \frac{1}{p_i}\|\mathbf{w}_i\|^2 - t \leqslant 0, \ i \in \mathcal{J}. \end{aligned} \tag{3.2}$$

As mentioned earlier, $\mathcal{P}_1$ is an NP-hard problem, whose optimal solution is difficult to obtain directly in either primal or dual domain. Despite of this, we are able to obtain the structure of the optimal solution to $\mathcal{P}_1$. To find the optimal solution structure, we follow the same technique in [40] that is used to derive the optimal beamforming structure for the scenario of multi-group multicast beamforming under a single BS.

In particular, we consider the SCA method for solving $\mathcal{P}_1$. It iteratively solves a sequence of convex approximations of the original problem to obtain the sub-optimal solution that is guaranteed to converge to a stationary solution. We will analyze the solution of the SCA method to obtain the structure of the optimal beamforming solution to $\mathcal{P}_1$.

3.2.1 The SCA Method

For $\mathcal{P}_1$, consider $M \times 1$ auxiliary vector $\mathbf{z}_i, i \in \mathcal{J}$. Based on the property of positive semi-definite matrix, for any $\mathbf{z}_i, i \in \mathcal{J}$, we have

$$\mathbf{w}_i^H \mathbf{h}_{i,ik} \mathbf{h}_{i,ik}^H \mathbf{w}_i \geqslant 2\mathfrak{Re}\{\mathbf{w}_i^H \mathbf{h}_{i,ik} \mathbf{h}_{i,ik}^H \mathbf{z}_i\} + \mathbf{z}_i^H \mathbf{h}_{i,ik} \mathbf{h}_{i,ik}^H \mathbf{z}_i, \tag{3.3}$$

where the equality holds if and only if $\mathbf{w}_i = \mathbf{z}_i$. Thus, given $\mathbf{z} \triangleq [\mathbf{z}_1^H, \ldots, \mathbf{z}_J^H]^H$, applying (3.3) to the numerator of the SINR expression in the constraint in (3.2), we obtain a lower bound on the SINR. Replacing the SINR with this lower bound, we obtain the following jointly convex approximation optimization problem for $\mathcal{P}_1$:

$$\begin{aligned}
\mathcal{P}_{1\text{SCA}}(\mathbf{z}): \min_{\mathbf{w},t} \ & t \\
\text{s.t.} \ & \gamma_{ik} \sum_{j=1, j\neq i}^{J} |\mathbf{h}_{j,ik}^H \mathbf{w}_j|^2 - 2\mathfrak{Re}\{\mathbf{w}_i^H \mathbf{h}_{i,ik} \mathbf{h}_{i,ik}^H \mathbf{z}_i\} \\
& \quad + |\mathbf{z}_i^H \mathbf{h}_{i,ik}|^2 + \gamma_{ik}\sigma^2 \leqslant 0, \ k \in \mathcal{K}, i \in \mathcal{J} \qquad (3.4) \\
& \frac{1}{p_i}\|\mathbf{w}_i\|^2 - t \leqslant 0, \ i \in \mathcal{J}. \qquad (3.5)
\end{aligned}$$

where the non-convex SINR constraint in (3.2) is replaced by the convex constraint on the lower bound of SINR in (3.4). Let $(\mathbf{w}^\star(\mathbf{z}), t^\star(\mathbf{z}))$ be the optimal solution to $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$. As a result, this solution is feasible to $\mathcal{P}_1$. Replacing $\mathbf{z}$ with the optimal solution $\mathbf{w}^\star(\mathbf{z})$, we iteratively solve a sequence of convex approximation subproblems until convergence. This SCA method is guaranteed to converge to a stationary point $\mathbf{w}^\star$ of $\mathcal{P}_o$.

3.2.2 The optimal solution to SCA subproblem

For $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ in each SCA iteration under given $\mathbf{z}$, since it is a jointly convex problem and Slater's condition holds, we can obtain the optimal solution by solving the

Lagrange dual problem [76], similar as in [40] except that $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ now is a joint optimization problem with respect to (w.r.t.) $(\mathbf{w}, t)$. The Lagrangian for $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ is given by

$$\begin{aligned}
&\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}) \\
&= t + \sum_{i=1}^{J} \mu_i \left(\frac{\|\mathbf{w}_i\|^2}{p_i} - t \right) + \sum_{i=1}^{J} \sum_{k=1}^{K} \lambda_{ik} \Big[\gamma_{ik} \sum_{j \neq i} |\mathbf{w}_j^H \mathbf{h}_{j,ik}|^2 \\
&\quad - 2\mathfrak{Re}\{\mathbf{w}_i^H \mathbf{h}_{i,ik} \mathbf{h}_{i,ik}^H \mathbf{z}_i\} + |\mathbf{z}_i^H \mathbf{h}_{i,ik}|^2 + \gamma_{ik}\sigma^2 \Big]
\end{aligned} \tag{3.6}$$

where λ_{ik} and μ_i are the Lagrange multipliers associated with the QoS constraint in (3.4) for user k in cell i and the BS power constraint in (3.5), respectively, and we define $\boldsymbol{\lambda} \triangleq [\boldsymbol{\lambda}_1^T, \ldots, \boldsymbol{\lambda}_J^T]^T$ with $\boldsymbol{\lambda}_i \triangleq [\lambda_{i1}, \ldots, \lambda_{iK}]^T$ and $\boldsymbol{\mu} \triangleq [\mu_1, \ldots, \mu_J]^T$. After regrouping the terms w.r.t. t and $\mathbf{w}_i$ in (3.6), we rewrite the Lagrangian as

$$\begin{aligned}
&\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}) \\
&= (1 - \mathbf{1}^T \boldsymbol{\mu}) t + \sum_{i=1}^{J} \sum_{k=1}^{K} \lambda_{ik} \left(\sigma^2 \gamma_{ik} + |\mathbf{z}_i^H \mathbf{h}_{i,ik}|^2 \right) \\
&\quad + \sum_{i=1}^{J} \mathbf{w}_i^H \left(\frac{\mu_i}{p_i} \mathbf{I} + \sum_{j \neq i} \sum_{k=1}^{K} \lambda_{jk} \gamma_{jk} \mathbf{h}_{i,jk} \mathbf{h}_{i,jk}^H \right) \mathbf{w}_i \\
&\quad - 2 \sum_{i=1}^{J} \mathfrak{Re} \left\{ \mathbf{z}_i^H \left(\sum_{k=1}^{K} \lambda_{ik} \mathbf{h}_{i,ik} \mathbf{h}_{i,ik}^H \right) \mathbf{w}_i \right\} \\
&= (1 - \mathbf{1}^T \boldsymbol{\mu}) t + \sum_{i=1}^{J} \sum_{k=1}^{K} \lambda_{ik} \left(\sigma^2 \gamma_{ik} + |\mathbf{z}_i^H \mathbf{h}_{i,ik}|^2 \right) \\
&\quad + \sum_{i=1}^{J} \mathbf{w}_i^H \mathbf{R}_{i^-}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{w}_i - 2 \sum_{i=1}^{J} \mathfrak{Re} \left\{ \boldsymbol{\nu}_i^H \mathbf{w}_i \right\}
\end{aligned} \tag{3.7}$$

where

$$\mathbf{R}_{i^-}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \frac{\mu_i}{p_i} \mathbf{I} + \sum_{j=1, j \neq i}^{J} \sum_{k=1}^{K} \lambda_{jk} \gamma_{jk} \mathbf{h}_{i,jk} \mathbf{h}_{i,jk}^H, \tag{3.8}$$

$$\boldsymbol{\nu}_i \triangleq \left(\sum_{k=1}^{K} \lambda_{ik} \mathbf{h}_{i,ik} \mathbf{h}_{i,ik}^H \right) \mathbf{z}_i. \tag{3.9}$$

Note that $\mathbf{R}_{i^-}(\boldsymbol{\lambda}, \boldsymbol{\mu})$ contains the sample channel covariance matrix from BS i to all users in other cells.

The Lagrange dual function is given by

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}) \triangleq \min_{\mathbf{w},t} \mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}), \tag{3.10}$$

and the dual problem for $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ is

$$\mathcal{D}_{1,\text{SCA}}(\mathbf{z}) : \max_{\boldsymbol{\lambda} \succcurlyeq \mathbf{0}, \boldsymbol{\mu} \succcurlyeq \mathbf{0},} g(\boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}).$$

By solving the the minimization of $\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z})$ in (3.10), we obtain the closed-form solution $\mathbf{w}_i^\star(\mathbf{z})$, $i \in \mathcal{J}$, to $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ as follows.

Proposition 3.1. *The optimal solution $\mathbf{w}_i^\star(\mathbf{z})$ to $\mathcal{P}_{1SCA}(\mathbf{z})$ is*

$$\mathbf{w}_i^\star(\mathbf{z}) = \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)\left(\sum_{k=1}^{K} \alpha_{ik}^\star \mathbf{h}_{i,ik}\right),\ i \in \mathcal{J} \tag{3.11}$$

where $(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ are the optimal Lagrange multipliers to the dual problem $\mathcal{D}_{1,SCA}(\mathbf{z})$ with $\mu_i^\star > 0$, $i \in \mathcal{J}$, and $\mathbf{1}^T \boldsymbol{\mu}^\star = 1$; and also, $\alpha_{ik}^\star \triangleq \lambda_{ik}^\star \mathbf{h}_{i,ik}^H \mathbf{z}_i$, $k \in \mathcal{K}, i \in \mathcal{J}$.

Proof. See Appendix A. □

Let $\mathbf{H}_i \triangleq [\mathbf{h}_{i,i1}, \ldots, \mathbf{h}_{i,iK}]$ denote the channel matrix between BS i and its own K users in cell i. Denote $\boldsymbol{\alpha}_i^\star \triangleq [\alpha_{i1}^\star, \ldots, \alpha_{iK}^\star]^T$. Then, we can rewrite the optimal solution $\mathbf{w}_i^\star(\mathbf{z})$ in (3.11) as

$$\mathbf{w}_i^\star(\mathbf{z}) = \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)\mathbf{H}_i \boldsymbol{\alpha}_i^\star,\ i \in \mathcal{J}. \tag{3.12}$$

3.2.3 The optimal solution to the QoS Problem

As mentioned earlier, we iteratively solve a sequence of SCA subproblems $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ by replacing $\mathbf{z}$ with $\mathbf{w}^\star(\mathbf{z})$ obtained from the previous subproblem, until $\mathbf{z}$ converges

to a stationary point $\mathbf{w}^\star$ of $\mathcal{P}_o$. In particular, if the stationary point is the global optimal solution $\mathbf{w}^\star = \mathbf{w}^o$, then $\mathbf{z}$ converges to the optimum $\mathbf{z} \to \mathbf{w}^o$, and we also obtain the maximum objective value t^o. Following this, first, we see that regardless of $\mathbf{z}$, the structure of $\mathbf{w}_i^\star(\mathbf{z})$ remains as in (3.12), while $\mathbf{w}_i^\star(\mathbf{z})$ depends on $\mathbf{z}$ only through $(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ for $\mathcal{D}_{1,\mathrm{SCA}}(\mathbf{z})$ and $\boldsymbol{\alpha}_i^\star$, Second, when $\mathbf{z} \to \mathbf{w}^o$, the optimal $(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ for the dual problem $\mathcal{D}_{1\mathrm{SCA}}(\mathbf{z})$ also converges to the optimal $(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ of $\mathcal{D}_{1,\mathrm{SCA}}(\mathbf{w}^o)$, *i.e.,* the dual problem of $\mathcal{P}_1$. Thus, we obtain the optimal solution to $\mathcal{P}_1$. The solution is stated in the following theorem.

Theorem 1. *The optimal beamforming solution to the QoS problem $\mathcal{P}_o$ for multi-cell coordinated multicast beamforming is given by*

$$\mathbf{w}_i^o = \mathbf{R}_i^{-1}(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)\mathbf{H}_i\mathbf{a}_i^o,\ i \in \mathcal{J} \tag{3.13}$$

where $(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ are the optimal dual solutions to $\mathcal{D}_{1,SCA}(\mathbf{w}^o)$ with $\mu_i^o > 0$, $i \in \mathcal{J}$, $\mathbf{a}_i^o \triangleq [a_{i1}^o, \ldots, a_{iK}^o]^T$ is the optimal weight vector with the weight for user k in cell i given by $a_{ik}^o \triangleq \lambda_{ik}^o(1+\gamma_{ik})(\mathbf{h}_{i,ik}^H\mathbf{w}_i^o)$, $k \in \mathcal{K}$, $i \in \mathcal{J}$, and

$$\mathbf{R}_i(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \frac{\mu_i}{p_i}\mathbf{I} + \sum_{j=1}^{J}\sum_{k=1}^{K}\lambda_{jk}\gamma_{jk}\mathbf{h}_{i,jk}\mathbf{h}_{i,jk}^H,\ i \in \mathcal{J}. \tag{3.14}$$

Furthermore, the optimal value of $\mathcal{P}_o$ is

$$\max_i\ \frac{1}{p_i}\|\mathbf{w}_i^o\|^2 = \sigma^2\boldsymbol{\lambda}^{oT}\boldsymbol{\gamma} \tag{3.15}$$

where $\boldsymbol{\gamma}$ is the SINR target vector of all users: $\boldsymbol{\gamma} \triangleq [\boldsymbol{\gamma}_1^T, \ldots, \boldsymbol{\gamma}_J^T]^T$ with $\boldsymbol{\gamma}_i \triangleq [\gamma_{i1}, \ldots \gamma_{iK}]^T$, $i \in \mathcal{J}$.

Proof. See Appendix B. □

Theorem 1 shows that the optimal multicast beamforming vector $\mathbf{w}_i^o$ for each BS i given in (3.13) is a weighted minimum mean square error (MMSE) beamformer. The matrix $\mathbf{R}_i(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is a per-BS noise-plus-weighted-channel-covariance matrix. The first term is the normalized receiver noise term scaled by μ_i/p_i, which reflects the BS transmit power used w.r.t. its power budget target. The second term contains the channels between BS i to all users in J cells. The term $\mathbf{H}_i\mathbf{a}_i^o$ is the weighted user channels in cell i, which can be viewed as a *group-channel direction* for cell i: $\hat{\mathbf{h}}_i \triangleq \mathbf{H}_i\mathbf{a}_i^o$. The optimal weight vector $\mathbf{a}_i^o$ shows the relative significance of the channel of each user in cell i in this group-channel direction and thus the beamformer $\mathbf{w}_i^o$.

The optimal multicast beamforming vector $\mathbf{w}_i^o$ for BS i is only a function of the channels from BS i to all users in J cells. Thus, $\mathbf{w}_i^o$ can be computed locally at each BS, without requiring the global knowledge of channel state information in other cells. However, note that the optimal solution $\mathbf{w}_i^o$ in (3.13) is shown in a semi-closed-form, where $\boldsymbol{\lambda}^o$, $\boldsymbol{\mu}^o$ and $\mathbf{a}_i^o$ need to be computed. Determining their optimal values require considering J cells jointly.

3.3 Numerical Algorithms

Note that even though we have the optimal structure for $\mathbf{w}_i^o$, since $\mathcal{P}_1$ is an NP-hard problem, obtaining the optimal values $(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ and $\mathbf{a}^o$ is difficult. Thus, in this section, we design algorithms to determine these parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$.

3.3.1 Obtaining Undetermined Parameters

From Proposition 3.1, $\mathbf{1}^H\boldsymbol{\mu} = 1$ and $\mu_i > 0$, $\forall i \in \mathcal{J}$. Thus, we have $\mu_i \in (0, 1)$. However, the value of each μ_i still can not be easily solved directly. Thus, to simplify the computation, we propose to set $\mu_i = 1/J, \forall i \in \mathcal{J}$, and denote $\tilde{\boldsymbol{\mu}} = [1/J, \ldots, 1/J]^T$.

After setting $\tilde{\boldsymbol{\mu}}$ in $\mathbf{R}_i(\boldsymbol{\lambda}, \tilde{\boldsymbol{\mu}})$, we need to determine $\boldsymbol{\lambda}$. From Theorem 1, we have $a_{ik}^o \triangleq \lambda_{ik}^o(1+\gamma_{ik})(\mathbf{h}_{i,ik}^H\mathbf{w}_i^o)$, $\forall k, i$. Let $\delta_{ik} \triangleq \mathbf{h}_{i,ik}^H\mathbf{w}^o$, $k \in \mathcal{K}, i \in \mathcal{J}$, and $\boldsymbol{\delta}_i = [\delta_{i1}, \ldots, \delta_{iK}]^T$. Then, we can express a_{ik}^o into the vector form $\mathbf{a}_i^o$ as

$$\mathbf{a}_i^o = \mathbf{D}_{\boldsymbol{\lambda}_i}\left(\mathbf{I} + \mathbf{D}_{\boldsymbol{\gamma}_i}\right)\boldsymbol{\delta}_i. \tag{3.16}$$

Also, based on the optimal solution in (3.13), we have

$$\begin{aligned}\delta_{ik} = \mathbf{h}_{i,ik}^H\mathbf{w}_i^o &= \mathbf{h}_{i,ik}^H\mathbf{R}_i^{-1}(\boldsymbol{\lambda}_i^o, \tilde{\boldsymbol{\mu}})\mathbf{H}_i\mathbf{a}_i^o \\ &= \mathbf{h}_{i,ik}^H\mathbf{R}_i^{-1}(\boldsymbol{\lambda}_i^o, \tilde{\boldsymbol{\mu}})\mathbf{H}_i\mathbf{D}_{\boldsymbol{\lambda}_i}(\mathbf{I} + \mathbf{D}_{\boldsymbol{\gamma}_i})\boldsymbol{\delta}_i. \end{aligned}\tag{3.17}$$

Writing (3.17) into a vector form for each $i \in \mathcal{J}$, we have

$$\boldsymbol{\delta}_i = \mathbf{H}_i^H\mathbf{R}_i^{-1}(\boldsymbol{\lambda}_i^o, \tilde{\boldsymbol{\mu}})\mathbf{H}_i\mathbf{D}_{\boldsymbol{\lambda}_i}\left(\mathbf{I} + \mathbf{D}_{\boldsymbol{\gamma}_i}\right)\boldsymbol{\delta}_i, \tag{3.18}$$

which leads to

$$\left(\mathbf{H}_i^H\mathbf{R}_i^{-1}(\boldsymbol{\lambda}_i^o, \tilde{\boldsymbol{\mu}})\mathbf{H}_i\mathbf{D}_{\boldsymbol{\lambda}_i}\left(\mathbf{I} + \mathbf{D}_{\boldsymbol{\gamma}_i}\right) - \mathbf{I}\right)\boldsymbol{\delta}_i = \mathbf{0}. \tag{3.19}$$

At the optimality, the optimal $\boldsymbol{\lambda}_i^o$ should satisfy (3.19) (assuming the optimal $\boldsymbol{\mu}^o$). Since directly solve (3.19) is difficult due to the unknown $\boldsymbol{\delta}_i$, similar to the approach in [40], we consider a sufficient condition for (3.19), below and use it, to develop a suboptimal algorithm obtaining $\boldsymbol{\lambda}$:

$$\mathbf{H}_i^H\mathbf{R}_i^{-1}(\boldsymbol{\lambda}_i^o, \tilde{\boldsymbol{\mu}})\mathbf{H}_i\mathbf{D}_{\boldsymbol{\lambda}_i}\left(\mathbf{I} + \mathbf{D}_{\boldsymbol{\gamma}_i}\right) = \mathbf{I},\ i \in \mathcal{J}. \tag{3.20}$$

Algorithm 1 Fixed-point Iterative method for $\boldsymbol{\lambda}_i$

1: **Initialization** : Set $m = 0$; Set ϵ; Initialize $\boldsymbol{\lambda} \succcurlyeq \mathbf{0}$.

2: **repeat**

3: Compute $\mathbf{R}_i(\boldsymbol{\lambda}_i^{(m)}, \tilde{\boldsymbol{\mu}})$ based on (3.14).

4: For each $i \in \mathcal{J}, k \in \mathcal{K}$, compute

$$\lambda_{ik}^{(m+1)} = \frac{1}{(1+\gamma_{ik})\,\mathbf{h}_{i,ik}^H \mathbf{R}_i^{-1}\left(\boldsymbol{\lambda}_i^{(m)}, \tilde{\boldsymbol{\mu}}\right)\mathbf{h}_{i,ik}} \geqslant 0.$$

5: $m \leftarrow m + 1$.

6: **until** $\max_{i,k}\left|\lambda_{ik}^{(m+1)} - \lambda_{ik}^{(m)}\right| \leqslant \epsilon$.

This sufficient condition can be described per-element as follows for each $i \in \mathcal{J}$:

$$\begin{cases} \lambda_{ik}\,(1+\gamma_{ik})\,\mathbf{h}_{i,ik}^H \mathbf{R}_i^{-1}(\boldsymbol{\lambda}_i^o, \tilde{\boldsymbol{\mu}})\mathbf{h}_{i,ik} = 1, & k \in \mathcal{K} \\ \lambda_{ik}\,(1+\gamma_{ik})\,\mathbf{h}_{i,ik}^H \mathbf{R}_i^{-1}(\boldsymbol{\lambda}_i^o, \tilde{\boldsymbol{\mu}})\mathbf{h}_{i,il} = 0, & l \neq k, l \in \mathcal{K}. \end{cases} \tag{3.21}$$

Note that there are typically more equations than variables in (3.21) to solve. Thus, λ_{ik} may not satisfy all the equations. Following the proposed method in [40], we obtain by only solving the first equation in (3.21) (*i.e.,* the diagonal components in (3.20), which can be solved using the fixed-point iterative method shown in Algorithm 1.

3.3.2 Conventional Methods for Computing Weight Vectors

The SDR Method

Based on the optimal beamforming structure in (3.13) and $(\boldsymbol{\lambda}, \tilde{\boldsymbol{\mu}})$ obtained from the above, we convert the original problem $\mathcal{P}_1$ w.r.t. $(\mathbf{w}, t)$ into a joint optimization of $(\mathbf{a}, t)$, given by

$$\begin{aligned} \mathcal{P}_2 : \min_{\mathbf{a},t}\ & t \\ \text{s.t. } & \frac{|\mathbf{a}_i^H \mathbf{H}_i^H \mathbf{R}_i^{-1H}(\boldsymbol{\lambda}^o, \tilde{\boldsymbol{\mu}})\mathbf{h}_{i,ik}|^2}{\sum_{j=1,j\neq i}^{J} |\mathbf{a}_j^H \mathbf{H}_j^H \mathbf{R}_j^{-1H}(\boldsymbol{\lambda}^o, \tilde{\boldsymbol{\mu}})\mathbf{h}_{j,ik}|^2 + \sigma^2} \geqslant \gamma_{ik}, \end{aligned}$$

$$k \in \mathcal{K}, i \in \mathcal{J} \tag{3.22}$$

$$\frac{1}{p_i}\|\mathbf{R}_i^{-1}(\boldsymbol{\lambda}^o, \tilde{\boldsymbol{\mu}})\mathbf{H}_i\mathbf{a}_i\|^2 - t \leqslant 0,\ i \in \mathcal{J}.$$

Note that by this problem conversion, we now have a much smaller problem $\mathcal{P}_2$ than $\mathcal{P}_1$, since $\mathbf{a}$ is a $JK \times 1$ vector, $\mathbf{w}$ is a $JM \times 1$ vector, and $K \ll M$ in a massive MIMO system. We can apply the conventional SDR method along with the Gaussian Randomization procedure to solve $\mathcal{P}_2$. Specifically, define $\mathbf{X}_i \triangleq \mathbf{a}_i\mathbf{a}_i^H$, $\mathbf{G}_i \triangleq \mathbf{R}_i^{-1}(\boldsymbol{\lambda}, \tilde{\boldsymbol{\mu}})\mathbf{H}_i$, $\mathbf{f}_{j,ik} \triangleq \mathbf{G}_j^H\mathbf{h}_{j,ik}$, for $j, i \in \mathcal{J}, k \in \mathcal{K}$, and drop the rank-1 constraint, $\mathcal{P}_2$ is relaxed to the following SDP problem w.r.t. $(\mathbf{X}, t)$

$$\begin{aligned}
\mathcal{P}_{2\text{SDR}}: &\min_{\{\mathbf{X}_i\},t}\ t \\
\text{s.t. } &\text{tr}(\mathbf{f}_{i,ik}\mathbf{f}_{i,ik}^H\mathbf{X}_i) \geqslant \sum_{j=1,j\neq i}^{J}\text{tr}(\mathbf{f}_{j,ik}\mathbf{f}_{j,ik}^H\mathbf{X}_j) + \sigma^2, k \in \mathcal{K}, i \in \mathcal{J} \\
&\frac{1}{p_i}\text{tr}(\mathbf{G}_i^H\mathbf{G}_i\mathbf{X}_i) \leqslant t,\ i \in \mathcal{J} \\
&\mathbf{X}_i \succcurlyeq 0,\ i \in \mathcal{J}.
\end{aligned}$$

Note that $\mathcal{P}_{2\text{SDR}}$ is a jointly SDP problem for $\mathbf{X}$ and t, which can be solved using standard CVX solvers. Recovering the solution $\mathbf{a}_i$ from the optimal solution $\mathbf{X}_i^\star$ depends on the rank of $\mathbf{X}_i^\star$. If $\text{rank}(\mathbf{X}_i^\star) = 1$, $\mathbf{X}_i = \mathbf{a}_i^\star\mathbf{a}_i^{\star H}$, then $\mathbf{a}_i^\star$ can be extracted directly from $\mathbf{X}_i^\star$. Otherwise, we use the Gaussian randomization method [77] to extract weight vector $\mathbf{a}_i^\star$, which is briefly described below: Denote $\tilde{\mathcal{J}} \triangleq \{i : \text{rank}(\mathbf{X}_i) > 1, i \in \mathcal{J}\}$.

i) Generate L independent and identically distributed (i.i.d) random vectors $\mathbf{a}_i^{(l)} \sim \mathcal{CN}(\mathbf{0}, \mathbf{X}_i)$, $l = 1, \ldots, L$, $i \in \tilde{\mathcal{J}}$.

ii) For each l, check whether $\{\mathbf{a}_i^{(l)}\}_{i=1}^J$ do not satisfy the SINR constraint (3.22); If

not, regenerate $\{\mathbf{a}_i^{(l)}\}_{i\in\mathcal{J}}$ until they meet the SINR constraint .

iii) Find $l^\star = \arg\min_l \max_i \left\{\frac{\|\mathbf{w}_i^{(l)}\|^2}{p_i}\right\}$, where $\mathbf{w}_i^{(l)} = \mathbf{R}_i^{-1}(\boldsymbol{\lambda}^o, \tilde{\boldsymbol{\mu}})\mathbf{H}_i\mathbf{a}_i^{(l)}$ and set $\mathbf{w}_i^\star = \mathbf{w}_i^{(l)}, i \in \mathcal{J}$.

SCA method

The computational complexity of SDR is relatively high, and its performance suffers as the problem size grows. SCA is a more computationally efficient method that overcomes these drawbacks of SDR. We can use SCA similar as $\mathcal{P}_{1\text{SCA}}$ to iteratively solve $\mathcal{P}_2$ for $\mathbf{a}$.

Specifically, denote $\mathbf{u} \triangleq [\mathbf{u}_1^H, \ldots, \mathbf{u}_J^H]^H$, where $\mathbf{u}_i$ is $K \times 1$ auxiliary vector for $i \in \mathcal{J}$. Given $\mathbf{u}$, we apply the convex approximation to the constraint in (3.22) The above problem can be solved using the standard SDP solvers., we have the following joint optimization problem over $(\mathbf{a}, t)$ at each SCA iteration: $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$:

$$
\begin{aligned}
\mathcal{P}_{2\text{SCA}}(\mathbf{u}) : \min_{\mathbf{a},t}\ & t \\
\text{s.t. } & 2\mathfrak{Re}\left\{\mathbf{a}_i^H\mathbf{f}_{i,ik}\mathbf{f}_{i,ik}^H\mathbf{u}_i\right\} - |\mathbf{u}_i^H\mathbf{f}_{i,ik}|^2 \\
& \qquad \geqslant \sum_{j=1,j\neq i}^{J} |\mathbf{a}_j^H\mathbf{f}_{j,ik}|^2 + \sigma^2,\ k \in \mathcal{K}, i \in \mathcal{J} \\
& \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2 - t \leqslant 0,\ i \in \mathcal{J}.
\end{aligned}
$$

The iterative procedure of SCA is the same as that described before Section 3.2.1, where we update $\mathbf{u}$ with the solution to $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ and solve $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ iteratively until convergence.

Initialization: The SCA method requires the initial $\mathbf{u}^{(0)}$ should to be feasible to $\mathcal{P}_2$. We can use the solution of the SDR method $\mathbf{a}^{\text{SDR}}$ as the initial point. Note

that since we only need a feasible point, we can limit the number of trials L in the Gaussian randomization procedure to further reduce the computational complexity in generating the initial point. Besides ensuring the feasibility, $\mathbf{a}^{\text{SDR}}$ is a good initial point that can accelerate the convergence for SCA.

3.3.3 Fast ADMM-SCA Method for Computing Weight Vectors

In the SCA method, each SCA update needs to solve the convex problem $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$, which typically is computed using the interior-point algorithm [76] by the standard convex solvers, such as CVX [78]. Since the interior-point algorithm is a second-order algorithm, its computational complexity is still relatively high, especially when the problem size grows. Thus, we consider using the first-order fast algorithm to provide fast computation of the solution. In particular, we propose to apply ADMM technique [64] to solve $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ at each SCA iteration. ADMM is a robust numerical method that can provide fast computation to solve large-scale problems. Depending on the problem structure and the specific ADMM construction design in the problem, it can be used to break down a large problem into small subproblems to be solved individually with lower computational complexity.

In order to apply the ADMM technique, we first introduce the auxiliary variables $v \in \mathbb{R}$ and $d_{j,ik} \in \mathbb{C}$, $k \in \mathcal{K}$, $i, j \in \mathcal{J}$, and equivalently transform $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ into the following problem:

$$\begin{aligned} \mathcal{P}_{\text{ADMM}}(\mathbf{u}): \ & \min_{\mathbf{a},\mathbf{d},t,v} \ t \\ & \text{s.t.} \ \ d_{j,ik} = \mathbf{a}_j^H \mathbf{f}_{j,ik}, \ k \in \mathcal{K}, \ i, j \in \mathcal{J}, \end{aligned} \tag{3.23}$$

$$v = t, \tag{3.24}$$

$$\gamma_{ik} \sum_{j=1, j\neq i}^{J} |d_{j,ik}|^2 + |\mathbf{u}_i^H \mathbf{f}_{i,ik}|^2 + \gamma_{ik}\sigma^2 - 2\mathfrak{Re}\{d_{i,ik}\mathbf{f}_{i,ik}^H \mathbf{u}_i\} \leqslant 0,$$

$$k \in \mathcal{K}, i \in \mathcal{J}, \tag{3.25}$$

$$\frac{1}{p_i}\|\mathbf{G}_i \mathbf{a}_i\|^2 - v \leqslant 0,\ i \in \mathcal{J} \tag{3.26}$$

where $\mathbf{d} \triangleq [\mathbf{d}_{11}^H, \ldots, \mathbf{d}_{JK}^H]^H \in \mathbb{C}^{JK_{\text{tot}}}$, with $\mathbf{d}_{ik} \triangleq [d_{1,ik}, \ldots, d_{j,ik}]^T \in \mathbb{C}^J$.

Denote the feasible set for $\mathbf{d}$ satisfying the constraint (3.25) as $\mathcal{F}$, and that for $(\mathbf{a}, v)$ satisfying the constraint (3.26) as $\mathcal{C}$. Define the indicator functions for $\mathcal{F}$ and $\mathcal{C}$ respectively as

$$I_{\mathcal{F}}(\mathbf{d}) \triangleq \begin{cases} 0 & \mathbf{d} \in \mathcal{F} \\ \infty & \text{o.w.} \end{cases}, \quad I_{\mathcal{C}}(\mathbf{a}, v) \triangleq \begin{cases} 0 & (\mathbf{a}, v) \in \mathcal{C} \\ \infty & \text{o.w.} \end{cases}. \tag{3.27}$$

Then, we can equivalently transform $\mathcal{P}_{\text{ADMM}}(\mathbf{u})$ into the following equality-constrained problem:

$$\begin{aligned} \mathcal{P}'_{\text{ADMM}}(\mathbf{u}) : \min_{\mathbf{a},\mathbf{d},t,v}\ & t + I_{\mathcal{F}}(\mathbf{d}) + I_{\mathcal{C}}(\mathbf{a}, v) \\ \text{s.t. } & d_{j,ik} = \mathbf{a}_j^H \mathbf{f}_{j,ik},\ k \in \mathcal{K},\ i, j \in \mathcal{J} \\ & v = t. \end{aligned}$$

Based on the ADMM technique, the augmented Lagrangian of $\mathcal{P}'_{\text{ADMM}}(\mathbf{u})$ is given by

$$\begin{aligned} &\mathcal{L}_\rho(\mathbf{a}, \mathbf{d}, t, v, \mathbf{q}, z) = t + I_{\mathcal{F}}(\mathbf{d}) + I_{\mathcal{C}}(\mathbf{a}, v) + \\ &\frac{\rho}{2} \sum_{j=1}^{J} \sum_{i=1}^{J} \sum_{k=1}^{K} |d_{j,ik} - \mathbf{a}_j^H \mathbf{f}_{j,ik} + q_{j,ik}|^2 + \frac{\rho}{2}(v - t + z)^2 \end{aligned} \tag{3.28}$$

where $\rho > 0$ is the penalty parameter, and $\{q_{j,ik} \in \mathbb{C}, k \in \mathcal{K},\ i, j \in \mathcal{J}\}$ and $z \in \mathbb{R}$ are the dual variables associated with the respective equality constraints in $\mathcal{P}'_{\text{ADMM}}(\mathbf{u})$. Also, we denote $\mathbf{q} \triangleq [\mathbf{q}_{11}^H, \ldots, \mathbf{q}_{JK}^H]^H$, with $\mathbf{q}_{ik} \triangleq [q_{1ik}, \ldots, q_{Jik}]^T$.

Note that our particular design of auxiliary variables v and $\mathbf{d}$ and their respective equivalency constraints in (3.23) and (3.24) enables us to break the minimize of $\mathcal{L}_\rho(\mathbf{a}, \mathbf{d}, t, v, \mathbf{q}, z)$ into smaller subproblems. Specifically, we note that the terms in (3.28) for the optimization variables $(\mathbf{d}, v)$ and $(\mathbf{a}, t)$ are separate. Thus, we can decompose the optimization of $\mathcal{L}_\rho(\mathbf{a}, \mathbf{d}, t, v, \mathbf{q}, z)$ into two subproblems for $(\mathbf{d}, v)$ and $(\mathbf{a}, t)$ separately, and solve them alternatively. Our proposed ADMM-based algorithm for $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ is summarized in Algorithm 2. The ADMM procedure constrains three updating steps in each iteration. The two ADMM blocks in the first two steps involve solving two minimization subproblems w.r.t. $(\mathbf{d}, v)$ and $(\mathbf{a}, t)$ in (3.29) and (3.30), respectively. In particular, we will show below that we are able to derive a closed-form solution for each of these subproblems. As a result, the above specific ADMM construction leads to a fast algorithm to compute the solution for $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ at each SCA iteration. Finally, since $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ is convex, the ADMM procedure is guaranteed to converge to the optimal solution.

Next, we present the details of the solution to each subproblem.

Closed-Form $(\mathbf{d}, v)$-Update

Given $\mathbf{a}_i^{(l)}$, $t^{(l)}$ and $\mathbf{q}_i^{(l)}$, from (3.28), the minimization of $(\mathbf{d}, v)$ in (3.29) can be separated, which is equivalent to solve the following two subproblems

$$\begin{aligned}
&\mathcal{P}_{\mathbf{d}}(\mathbf{u}) : \min_{\mathbf{d}} \frac{\rho}{2} \sum_{j=1}^{J} \sum_{i=1}^{J} \sum_{k=1}^{K} |d_{j,ik} - \mathbf{a}_j^{H(l)} \mathbf{f}_{j,ik} + q_{j,ik}^{(l)}|^2 \\
&\text{s.t. } \gamma_{ik} \sum_{j=1, j\neq i}^{J} |d_{j,ik}|^2 + |\mathbf{u}_i^H \mathbf{f}_{i,ik}|^2 + \gamma_{ik}\sigma^2 - 2\mathfrak{Re}\{d_{i,ik} \mathbf{f}_{i,ik}^H \mathbf{u}_i\} \leqslant 0, \\
&\qquad\qquad k \in \mathcal{K}, i \in \mathcal{J}. \qquad (3.33)
\end{aligned}$$

Algorithm 2 ADMM-based Algorithm for $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$

1: **Initialization**: Set ρ; Set $\mathbf{a}^{(0)} = \mathbf{u}^{(l)}, \mathbf{q}^{(0)} = \mathbf{0}, z^{(0)} = 0, t^{(0)} = 0$; Set $l = 0$.

2: **repeat**

3: Update the auxiliary vectors $\mathbf{d}^{(l+1)}$ and variable $v^{(l+1)}$

$$\{\mathbf{d}^{(l+1)}, v^{(l+1)}\} = \arg\min_{\mathbf{d},v} \mathcal{L}_\rho(\mathbf{a}^{(l)}, t^{(l)}, v, \mathbf{d}, \mathbf{q}^{(l)}, z^{(l)}) \tag{3.29}$$

4: Update $\mathbf{a}^{(l+1)}$ and $t^{(l+1)}$

$$\{\mathbf{a}^{(l+1)}, t^{(l+1)}\} = \arg\min_{\mathbf{a},t} \mathcal{L}_\rho(\mathbf{a}, t, v^{(l+1)}, \mathbf{d}^{(l+1)}, \mathbf{q}^{(l)}, z^{(l)}) \tag{3.30}$$

5: Update dual variables $\mathbf{q}^{(l+1)}$ and $z^{(l+1)}$

$$q_{j,ik}^{(l+1)} = q_{j,ik}^{(l)} + \left(d_{j,ik}^{(l+1)} - \mathbf{a}_j^{(l+1)H}\mathbf{f}_{j,ik}\right), \forall i,j,k \tag{3.31}$$

$$z^{(l+1)} = z^{(l)} + (v^{(l+1)} - t^{(l+1)}). \tag{3.32}$$

6: Set $l \leftarrow l + 1$.

7: **until** convergence.

and

$$\begin{aligned} \mathcal{P}_v : \min_v \ & t + \frac{\rho}{2}|v - t + z|^2 \\ \text{s.t. } & \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2 \leqslant v, \ i \in \mathcal{J}. \end{aligned} \tag{3.34}$$

We can decompose $P_{\mathbf{d}}(\mathbf{u})$ into K^{tot} subproblems, one for each user k in cell i, as

$$\begin{aligned} \mathcal{P}_{\text{dsub}}(\mathbf{u}) : \ & \min_{\mathbf{d}_{ik}} \frac{\rho}{2}\sum_{j=1}^{J} |d_{j,ik} - \mathbf{a}_j^{(l)H}\mathbf{f}_{j,ik} - q_{j,ik}^{(l)}|^2 \\ & \text{s.t. } \gamma_{ik} \sum_{j=1, j\neq i}^{J} |d_{j,ik}|^2 + |\mathbf{u}_i^H\mathbf{f}_{i,ik}|^2 + \gamma_{ik}\sigma^2 - 2\mathfrak{Re}\{d_{i,ik}\mathbf{f}_{i,ik}^H\mathbf{u}_i\} \leqslant 0, \\ & \qquad\qquad k \in \mathcal{K}, i \in \mathcal{J}. \end{aligned} \tag{3.35}$$

Note that $P_{\mathbf{d}}(\mathbf{u})$ is a convex quadrtically constrained quadratic program (QCQP)-

1 problem, for which a closed-form solution can be obtained. In particular, such QCQP-1 problem has been considered in [64], where the closed-form solution is shown. Specifically, Define $e_{1,j,ik}^{(l)} \triangleq \mathbf{a}_j^{(l)H}\mathbf{f}_{j,ik} - q_{j,ik}^{(l)}$, $e_{2,ik} \triangleq |\mathbf{u}_i^H\mathbf{f}_{i,ik}|^2 + \gamma_{ik}\sigma^2$, $e_{3,ik} \triangleq \mathbf{f}_{i,ik}^H\mathbf{u}_i$, then we have

$$\begin{aligned}\mathcal{P}_{\text{dsub}}(\mathbf{u}):\ &\min_{\mathbf{d}_{ik}} \frac{\rho}{2}\sum_{j=1}^{J}|d_{j,ik} - e_{1,j,ik}^{(l)}|^2 \\ &\text{s.t. } e_{2,ik} + \gamma_{ik}\sum_{j=1,j\neq i}^{J}|d_{j,ik}|^2 - 2\mathfrak{Re}\{d_{i,ik}e_{3,ik}\} \leqslant 0\end{aligned} \tag{3.36}$$

the optimal solution $\mathbf{d}_{ik}^o$ for $\mathcal{P}_{\text{dsub}}(\mathbf{u})$ is given by

$$d_{j,ik}^o = \begin{cases} e_{1,i,ik}^{(l)} + \nu_{ik}^o e_{3,ik}^*, & j = i; \\ \frac{e_{1,j,ik}^{(l)}}{1+\nu_{ik}^o\gamma_{ik}}, & j \neq i. \end{cases} \tag{3.37}$$

Substituting the closed-form expression of $d_{j,ik}^o$ in (3.37) into the inequality constraint (3.36), which leads to

$$f(\nu_{ik}^o) = e_{2,ik} + \gamma_{ik}\frac{\sum_{j\neq i}^{J}|e_{1,j,ik}^{(l)}|^2}{(1+\nu_{ik}^o\gamma_{ik})^2} - 2\mathfrak{Re}\{e_{1,i,ik}^{(l)}e_{3,ik}\} - 2\nu_{ik}^o|e_{3,ik}|^2 \leqslant 0. \tag{3.38}$$

and ν_{ik}^o is the unique real positive root of the cubic equation (3.38), which can be obtained by applying the cubic equation formula.

Now we solve $\mathcal{P}_v$ for v. We first rewrite the constraint in (3.34), and the problem is equivalent to

$$\begin{aligned}\mathcal{P}_v:&\min_{v}\ t + \frac{\rho}{2}|v - t + z|^2 \\ &\text{s.t. } v \geqslant \max\frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2,\ i \in \mathcal{J}.\end{aligned} \tag{3.39}$$

Note that P_v is a convex problem with linear constraints. $f(v) \triangleq t + \frac{\rho}{2}|v - t + z|^2$, the domain of $f(v)$ is $[\max_i \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2, \infty)$. Thus, $\mathcal{P}_v$ is equivalent to find the minimum

point of $f(v)$ over the domain of $f(v)$. Taking the derivative of $f(v)$, and let it equal to 0,

$$f^{'}(v) = 2v + 2(z - t) = 0 \Rightarrow v^{\star} = t - z. \tag{3.40}$$

If $v^* \in [\max_i \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2, \infty)$, then v^* is the optimal solution for $\mathcal{P}_t$. Otherwise, $f(v)$ monotonically increases in the domain $[\max_i \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2, \infty)$, and thus $f(v)$ achieves its minimum value at $v = \max_i \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2$. Thus, we obtain the closed-form solution of v^o as:

$$v^o \triangleq \max\left(\max_i \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2, t - z\right). \tag{3.41}$$

Semi-closed-form $(\mathbf{a}, t)$-Update

Given $\mathbf{d}^{(l+1)}$ and $\mathbf{q}^{(l)}$, optimizing $\mathbf{a}$ and t jointly for the minimization problem in (3.30) is equivalent to solve the following problem.

$$\begin{aligned} &\min_{\mathbf{a},t} \Big(t + \frac{\rho}{2}\sum_{j=1}^{J}\sum_{i=1}^{J}\sum_{k=1}^{K}|d_{j,ik}^{(l+1)} - \mathbf{a}_j^H\mathbf{f}_{j,ik} + q_{j,ik}^{(l)}|^2 + \frac{\rho}{2}(v^{(l+1)} - t + z^{(l)})^2\Big) \\ &\text{s.t. } \frac{1}{p_j}\|\mathbf{G}_j\mathbf{a}_j\|^2 - v^{(l+1)} \leqslant 0,\ j \in \mathcal{J}. \end{aligned} \tag{3.42}$$

Since the terms in the objective function and constraints for $\mathbf{a}$ and t are separate, the above joint optmization problem can be decomposed into two subproblems for t and $\mathbf{a}$ separately. The subproblem for t is given by:

$$\mathcal{P}_t : \min_t\ t + \frac{\rho}{2}\left(v^{(l+1)} - t + z^{(l)}\right)^2. \tag{3.43}$$

The subproblem for $\mathbf{a}$ can be further decomposed into J subproblems, one for each $\mathbf{a}_j$

$$\mathcal{P}_{\mathbf{a}_j}(\mathbf{u}) : \min_{\mathbf{a}_j} \frac{\rho}{2}\sum_{i=1}^{J}\sum_{k=1}^{K}|d_{j,ik}^{(l+1)} - \mathbf{a}_j^H\mathbf{f}_{j,ik} + q_{j,ik}^{(l)}|^2$$

$$\text{s.t.}\frac{1}{p_j}\|\mathbf{G}_j\mathbf{a}_j\|^2 \leqslant v, \tag{3.44}$$

The problem $\mathcal{P}_t$ is a convex quadratic unconstrained optimization problem for t. Let $f(t) = t + \frac{\rho}{2}(v^{(l+1)} - t + z^{(l)})^2$. taking the derivative of $f(t)$ and set it to zero, we obtain

$$\begin{aligned} f^{'}(t) &= 1 - \rho(v^{(l+1)} - t + z^{(l)}) = 0 \\ \Rightarrow t^{\star} &= v^{(l+1)} + z^{(l)} - \frac{1}{\rho}. \end{aligned} \tag{3.45}$$

The subproblem $P_{\mathbf{a}_j}(\mathbf{u})$ is again a convex QCQP-1 problem. We can solve it by the Karush-Kuhn-Tucker (KKT) conditions. The Lagrangian of $\mathcal{P}_{\mathbf{a}_j}(\mathbf{u})$ is given by

$$\mathcal{L}(\mathbf{a}_j, \tilde{\lambda}_j) = \frac{\rho}{2}\sum_{i=1}^{J}\sum_{k=1}^{K}|d_{j,ik} - \mathbf{a}_j^H\mathbf{f}_{j,ik} + q_{j,ik}|^2 + \tilde{\lambda}_j\left(\frac{1}{p_j}\|\mathbf{G}_j\mathbf{a}_j\|^2 - v\right), \tag{3.46}$$

where $\tilde{\lambda}_j$ is the Lagrangian multiplier associated with the constraint in (3.44). The KKT conditions are

$$\frac{1}{p_j}\|\mathbf{G}_j\mathbf{a}_j\|^2 - v \leqslant 0; \tag{3.47}$$

$$\tilde{\lambda}_j \geqslant 0; \tag{3.48}$$

$$\tilde{\lambda}_j\left(\frac{1}{p_j}\|\mathbf{G}_j\mathbf{a}_j\|^2 - v\right) = 0; \tag{3.49}$$

$$\rho\left(\sum_{i=1}^{J}\sum_{k=1}^{K}\mathbf{f}_{j,ik}\mathbf{f}_{j,ik}^H\right)\mathbf{a}_j - \rho\sum_{i=1}^{J}\sum_{k=1}^{K}(d_{j,ik} + q_{j,ik})^*\mathbf{f}_{j,ik} + \frac{2\tilde{\lambda}_j}{p_j}\mathbf{G}_j^H\mathbf{G}_j\mathbf{a}_j = 0. \tag{3.50}$$

From (3.50), $\mathbf{a}_j$ is obtained as

$$\mathbf{a}_j = \left(\frac{2}{\rho}\cdot\frac{\tilde{\lambda}_j}{p_j}\mathbf{G}_j^H\mathbf{G}_j + \sum_{i=1}^{J}\sum_{k=1}^{K}\mathbf{f}_{j,ik}\mathbf{f}_{j,ik}^H\right)^{-1}\sum_{i=1}^{J}\sum_{k=1}^{K}(d_{j,ik} + q_{j,ik})^*\mathbf{f}_{j,ik}. \tag{3.51}$$

Based on the complementary slackness condition in (3.49), the following is possible:

$$\tilde{\lambda}_j = \begin{cases} 0, & \text{if } \frac{1}{p_j}\|\mathbf{G}_j\mathbf{a}_j\|^2 - v < 0, \\ \geqslant 0, & \text{if } \frac{1}{p_j}\|\mathbf{G}_j\mathbf{a}_j\|^2 - v = 0. \end{cases} \tag{3.52}$$

Thus, if $\mathbf{a}_j$ in (3.51) with $\tilde{\lambda}_j = 0$ satisfies the constraint in (3.44) (*i.e.,* $\mathcal{P}_{\mathbf{a}_j}(\mathbf{u})$ becomes an unconstrained optimization problem), then, the optimal $\tilde{\lambda}_j^\star = 0$. Otherwise, $\tilde{\lambda}_j^\star$ should be such that (3.47) holds with equality. In the later, we can use bi-section method to find $\tilde{\lambda}_j$.

3.4 Simulation Results

We consider a coordinated multi-cell multicast beamforming scenario with $J = 3$ BSs and one group per cell. Each cell has a unit cell radius, and users in the cell are randomly located with a uniform distribution. User channels are generated independently with $\mathbf{h}_{j,ik} \sim \mathcal{CN}(\mathbf{0}, \beta_{j,ik}\mathbf{I})$, $\forall j, i \in \mathcal{J}, k \in \mathcal{K}$. We use simplified path loss model for channel variance $\beta_{j,ik}$ as $\beta_{j,ik} = \xi_0 d_{j,ik}^{-\kappa}$, where $d_{j,ik}$ is the distance between BS j the the user k in cell i. κ is the pathloss exponent, which is set to $\kappa = 3.5$, and ξ_0 is the path loss constant. For the power budget of each BS, we set the power constraint of all the BSs are equal to 10W. The value of ξ_0 is determined by setting the nominal averaged SNR at the cell boundary to be -5 dB, *i.e.,* $P\xi_0/\sigma^2 = -5$[dB]. The performance results are averaged over 100 channel realizations and over 10 realizations of user locations.

3.4.1 Convergence Behaviour Study

In this section, we consider to study the convergence behaviour of the Lagrangian multiplier $\boldsymbol{\lambda}$ and our proposed algorithms. The details and analysis are shown in the following sections.

The Performance of Lagrangian Multiplier $\boldsymbol{\lambda}$

Before comparing the performance on our propose algorithms, we study the convergence behaviour on the Lagrangian multiplier in the covariance matrix $\mathbf{R}_i(\boldsymbol{\lambda}, \boldsymbol{\mu})$ first. The convergence behaviour over $\boldsymbol{\lambda}$ over different antenna size M (From 16 to 64) is shown in Fig. 3.1. For this convergence study, we utilize the default system setup: $J = 3, K = 5$. From the figure, we can tell all of the λ_{ik}s can converge in less than 45 iterations. Besides, we can also conclude from Fig. 3.1 that it needs less iterations to converge when antenna size M increases.

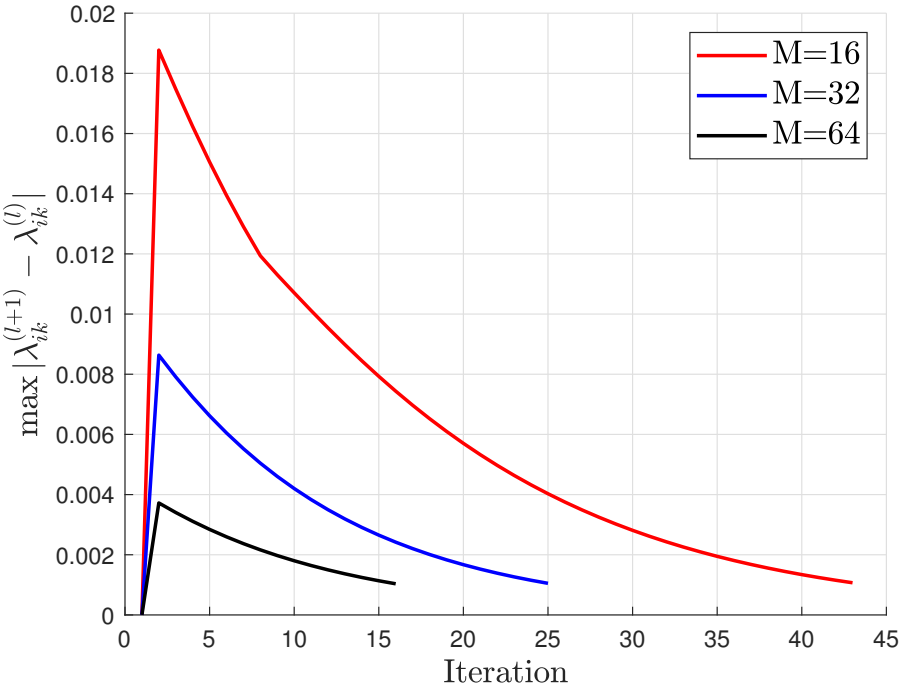

Figure 3.1: The convergence of $\widetilde{\boldsymbol{\lambda}}$ by using Algorithm 1 ($J = 3$, $K = 5$).

Moreover, we present the cumulative distribution function (CDF) curve over the iterations need for $\boldsymbol{\lambda}$ convergence under different M in Fig. 3.2, which is generated over 100 channel realizations. This CDF also indicates the $\boldsymbol{\lambda}$ converges in less iterations

when the M is large.

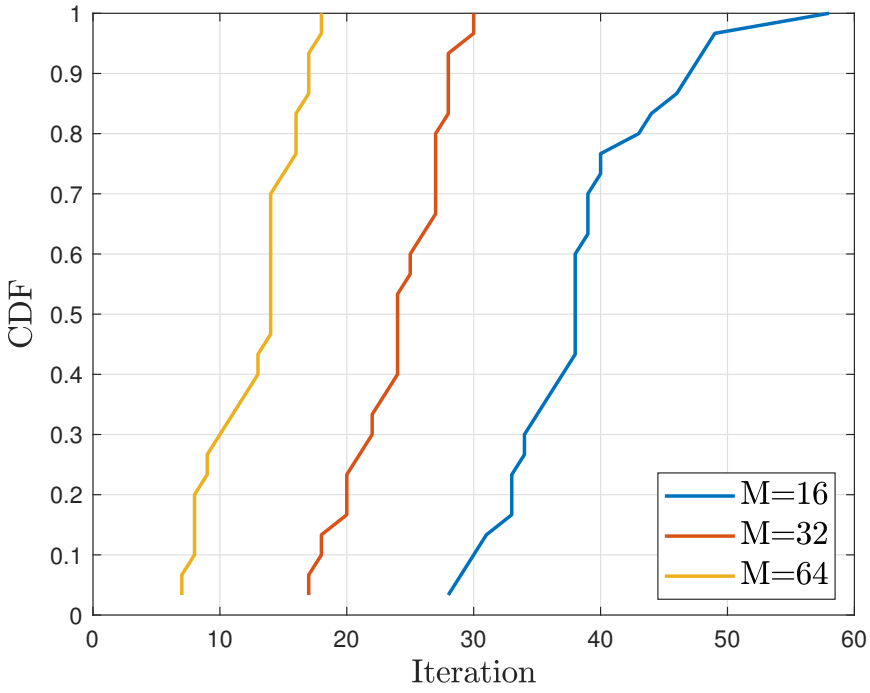

Figure 3.2: CDF of the iterations need for $\boldsymbol{\lambda}$ convergence under different M ($J = 3$, $K = 5$).

Convergence Behaviour for the OptSCA-ADMM Algorithm

In this section, we mainly study the convergence behaviour of OptSCA-CVX and OptSCA-ADMM, based on the optimal beamforming structure. (Theorem 1).

Before studying the convergence behavior, we need to determine the penalty parameter ρ first. Since the performance may be sensitive to the picked penalty parameter ρ, we set a range of value of ρ to observe the influence on different number of antennas M to the change of objective value. To ensure the accuracy of this simulation study, we use the same set of channel realizations for different parameter ρ under the same number of antenna M. Fig. 3.3 shows the influence of the objective value $\max_i \frac{1}{p_i}\|\mathbf{w}\|^2$ affected by the penalty parameter ρ under different number of antennas.

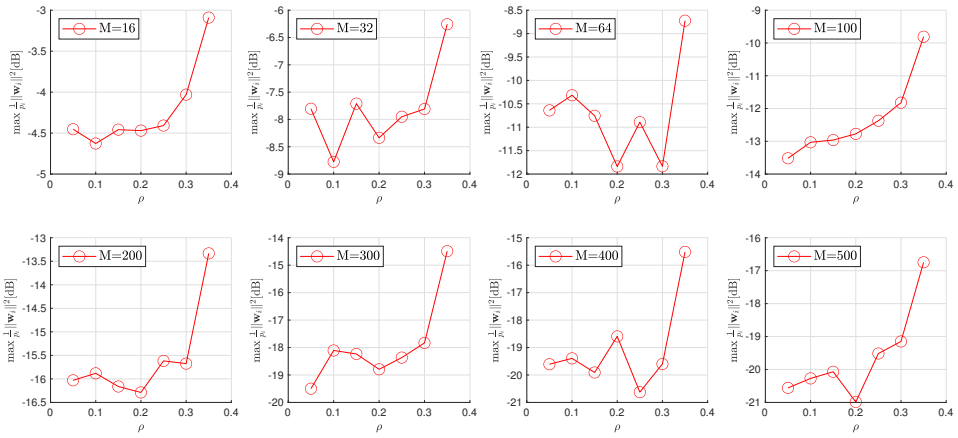

Figure 3.3: Performance affected by different ρ over number of antennas M ($J = 3, K = 5$).

From Fig. 3.3, we can conclude that when $\rho = 0.1$ or $\rho = 0.2$, the objective value $\max_i \frac{1}{p_i}\|\mathbf{w}\|^2$ stays relatively steadily. For the consideration of performance, $\rho = 0.2$ can make most cases achieve a lower objective value. Therefore, we set the penalty parameter ρ in the OptSCA-ADMM algorithm as 0.2 based on this simulation observation.

The fast first-order algorithm OptSCA-ADMM consists of outer-layer SCA iteration over $\mathbf{u}$; and the inner-layer Algorithm 2 for solving $\mathcal{P}_{\text{2SCA}}(\mathbf{u})$ in each SCA iteration. We first study the convergence behavior of the inner-layer ADMM (Algorithm 2). We set $M = 50, 100, 200$ and $K = 5$. For Algorithm 2, we consider the maximum relative difference $\max_i \frac{\|\mathbf{a}_i^{(l)} - \mathbf{a}_i^{(l-1)}\|}{\|\mathbf{a}_i^{(l-1)}\|}$ between two consecutive iterations. Fig. 3.4 shows the convergence behaviour of the maximum relative difference of $\mathbf{a}$ over iterations under Algorithm 2 in the first outer-layer SCA iteration. We observe that the trajectory decreases fast and reach the threshold in less than 10 iterations when M is large. While compared to the large-scale cases, it may need more iterations to

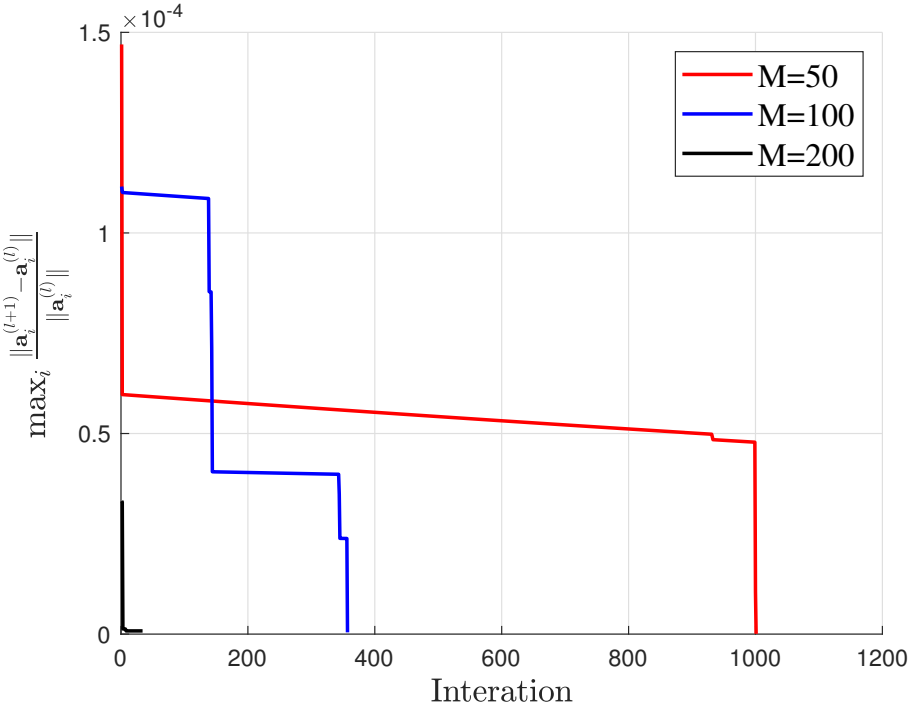

Figure 3.4: QoS problem for solving $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$: the maximum relative difference of $\mathbf{a}$ over the inner-layer iterations at the first outer-layer SCA iteration. ($J = 3$, $K = 5$).

converge when the M is small.

In Fig. 3.5, we also plot the trajectory of the objective value of $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ over iterations under Algorithm 2 in the first outer-layer of the SCA iteration. The system setting for Fig. 3.5 is $M = 50, J = 3, K = 5$, from which we can see the objective power ratio can be reduced more than 1 dB in the first outer-layer SCA iteration.

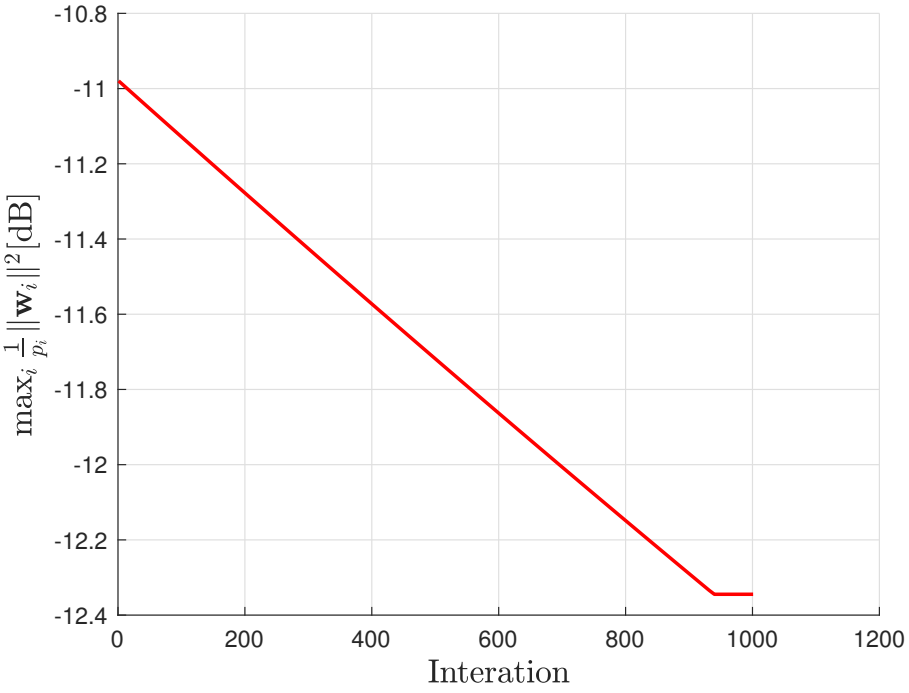

Figure 3.5: QoS problem for solving $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$: the maximum individual power ratio $\max_i \frac{1}{p_i}\|\mathbf{w}_i\|^2$ over the inner-layer iterations at the first outer-layer SCA iteration ($M = 50, J = 3, K = 5$).

In Fig. 3.6, we show the trajectory of the objective value of $\mathcal{P}_2$ over the outer-layer SCA iterations under OptSCA-ADMM and OptSCA-CVX. As twe see, since both algorithms solve $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ optimally, they achieve the same objective value outer iteration and converge fast in less than 10 iterations.

In Fig. 3.7 , we also show the CDF of the iterations need by the first outer iteration of the OptSCA-ADMM algorithm. As the CDF indicates that over 90% channel realizations can converge less than 500 iterations, and the needed iterations remains rough steadily over the M increases.

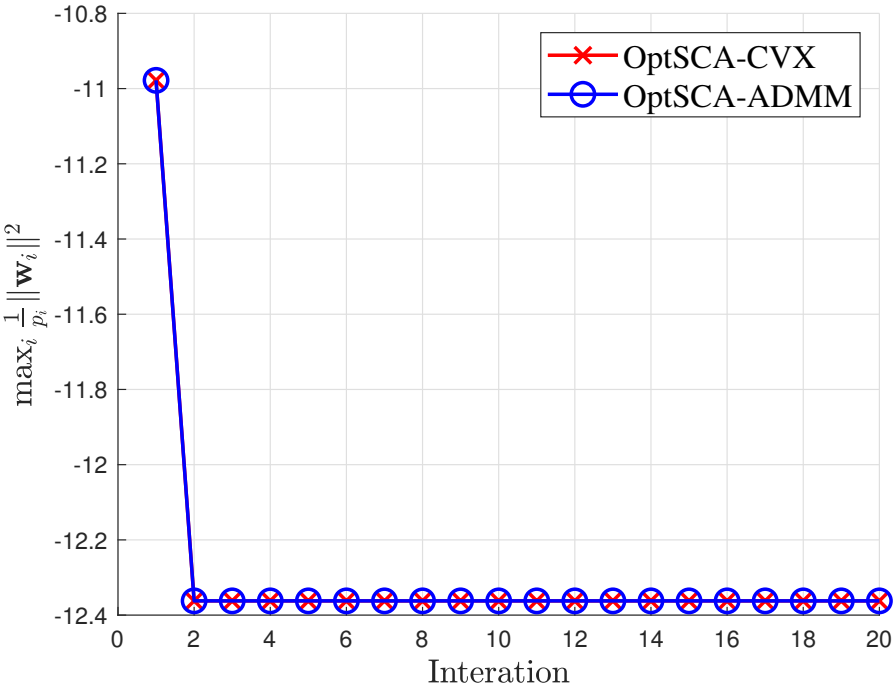

Figure 3.6: QoS problem for solving $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$: the maximum individual power ratio $\max_i \frac{1}{p_i}\|\mathbf{w}_i\|^2$ over the outer-layer iterations. $(M = 50, K = 5)$.

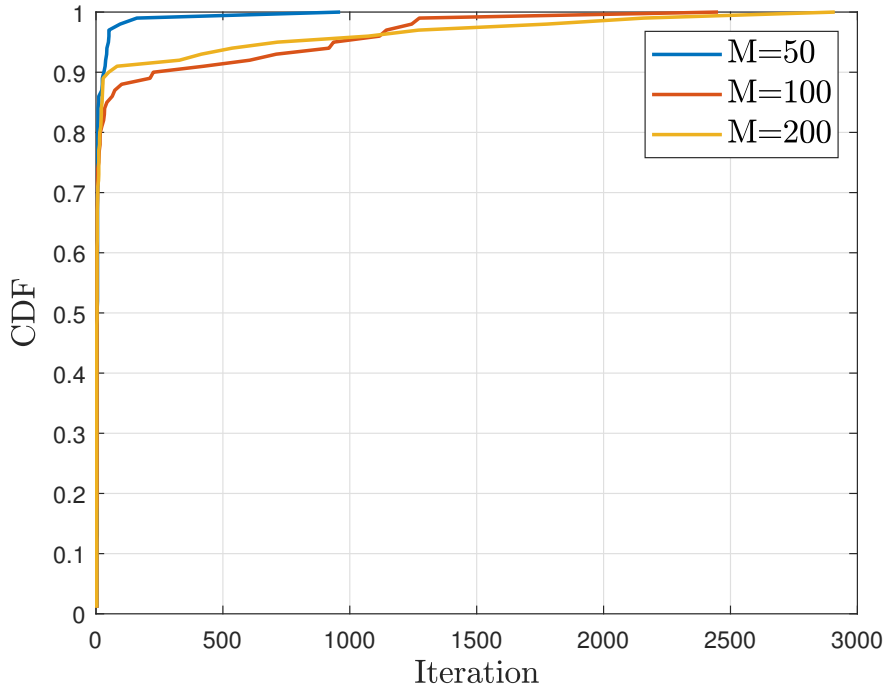

Figure 3.7: The CDF curve of inner-layer $\mathcal{P}_{2\text{SCA}}(\mathbf{u})$ $(J = 3, K = 5)$.

3.4.2 Performance Comparison QoS Problem

We now show the performance of our proposed algorithms for the QoS problem $\mathcal{P}_o$ and compare them with other methods. We set $K = 5$ and SINR target $\gamma_{ik} = 10$ dB, $\forall i, k$, Fig. 3.8 shows the maximum power margin $\max_i \|\mathbf{w}_i\|^2/p_i$ over the number of antennas M.

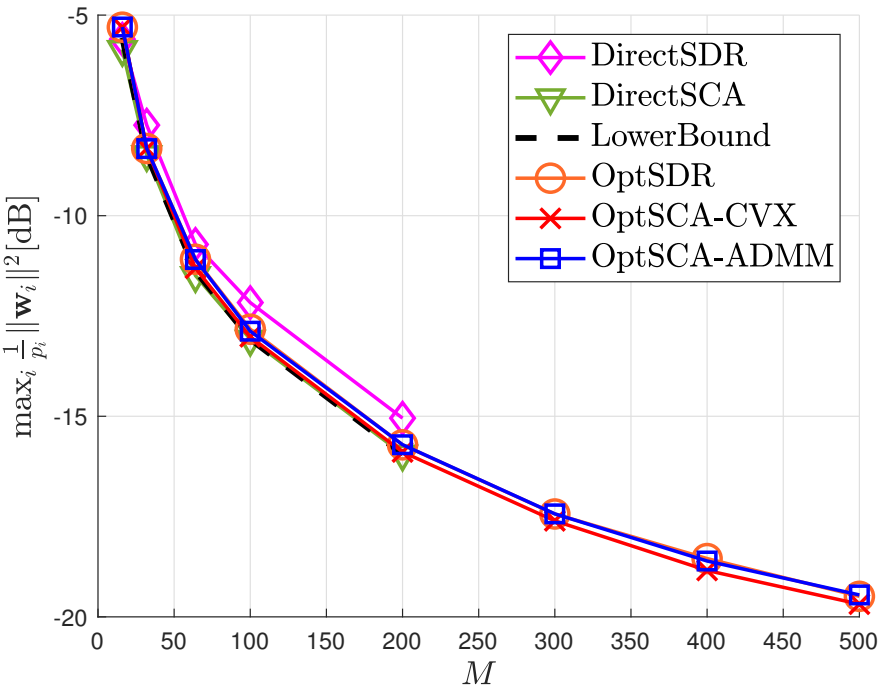

Figure 3.8: QoS problem: the maximum individual power ratio $\max_i \frac{1}{p_i}\|\mathbf{w}_i\|^2$ vs. the number of antenna size M $(K = 5)$.

We see that all of our proposed algorithms, OptSDR, OptSCA-CVX, and OptSCA-ADMM, perform very close to the lower bound, which means our proposed algorithms achieve nearly optimal performance for $K = 5$. To compare the computational complexity, we show the average computation time of the algorithms in Table 3.1. The first row shows the computational time for determining $\boldsymbol{\lambda}$, which is increasing as the antenna size grows. From Table 3.1, we see that

Table 3.1: Comparison of average computation time (sec) for QoS problem $\mathcal{P}_o$ ($J = 3, K = 5$).

M	100	200	300	400	500
$\boldsymbol{\lambda}$	0.0889	0.0573	0.1496	0.2433	0.3648
OptSDR	0.0199	0.0158	0.0174	0.0131	0.0108
OptSCA-CVX	2.5859	2.2480	2.6264	2.4451	2.3545
OptSCA-ADMM	0.0889	0.0573	0.0496	0.0433	0.0648
DirectSDR	33.8828	239.4786	-	-	-

1. The computation time of OptSDR, OptSCA-CVX and OptSCA-ADMM remain roughly unchanged as M increases, This is because by using the optimal beamforming structure in (3.13), we reduce the original problem $\mathcal{P}_o$ to weight optimization problem w.r.t. $\mathbf{a}$, and the dimension of the optimization variable is reduced from JM to JK, which only depend on the group size and no longer depend on the number of antennas M. In contrast, the computational complexity of DirectSDR increases with M significantly and is not practical for implementation as M becomes large.

2. The computational time of OptSCA-ADMM is only 1.5% $\sim$ 4% of that of OptSCA-CVX. This demonstrates that OptSCA-ADMM has significantly lower computational complexity than OptSCA-CVX in solving the SCA subproblem $\mathcal{P}_{\text{2SCA}}(\mathbf{u})$. n particular, OptSCA-ADMM uses Algorithm 2 to solve $\mathcal{P}_{\text{2SCA}}(\mathbf{u})$, which uses the closed-form or semi-closed form updates. This significantly reduces the computation time as compared with the conventional convex solver.

In Fig. 3.9, we show $\max_i \|\mathbf{w}_i\|^2/p_i$ vs. the number of users K per cell for $M = 50, 100, 200$. OptSCA-ADMM and OptSCA-CVX can nearly attain the lower bound

for all values of K and M. However, OptSDR performance deteriorates as K increases, with a noticeable gap to the lower bound. This is expected for the SDR-based method, as it is an approximate method and is known to be less accurate for problems of larger-size.

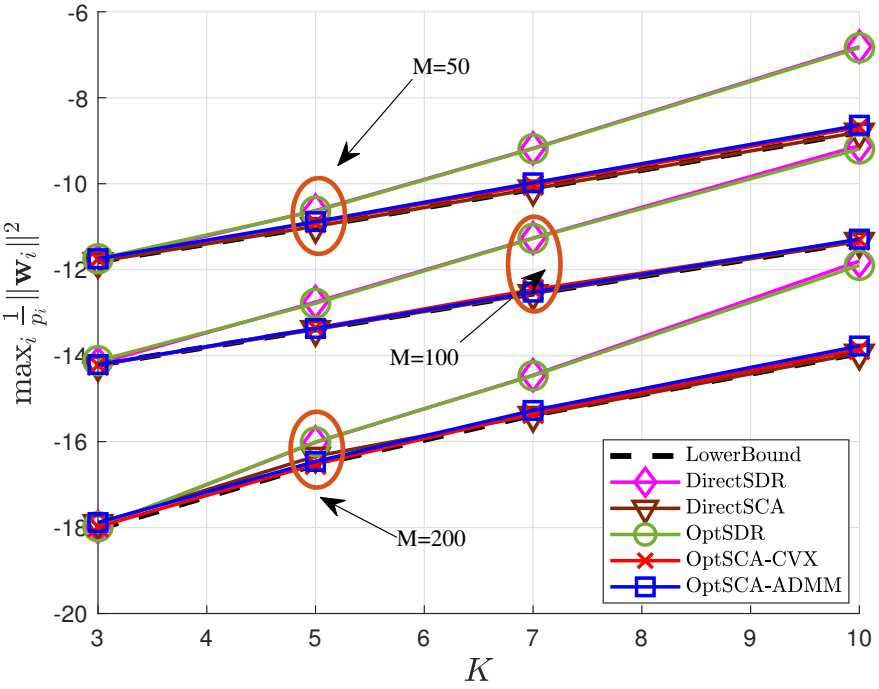

Figure 3.9: Performance for $\mathcal{P}_{\text{2SCA}}(\mathbf{u})$: the maximum individual power ratio $\max_i \frac{1}{p_i}\|\mathbf{w}_i\|^2$ over the number of users K. $(M = 100)$.

The average computation times of different methods used to generate Fig. 3.9 are shown in Table 3.2. Compared with our proposed algorithms, DirectSDR method has high computational complexity even when K is small. For the large problem size, *i.e.*, $K = 10$, the computational cost of utilizing ADMM technique is two magnitudes less than solving SCA via the conventional convex solver CVX solver. This again demonstrates the computational advantage of OptSCA-ADMM as a fast algorithm over other algorithms.

Table 3.2: Comparison of average computation time (sec) for QoS problem $\mathcal{P}_o$ ($J = 3, M = 100$).

K	3	5	7	10
OptSDR	0.0387	0.0533	0.0580	0.0784
OptSCA-CVX	2.2869	3.6130	6.1978	14.0025
OptSCA-ADMM	0.0100	0.0140	0.0207	0.1018
DirectSDR	45.78	217.11	374.65	1055.63

3.5 Summary

In this chapter, we focused on the QoS problem under multi-cell multicast scenario firstly. By applying the Lagrangian, we drive the optimal beamforming structure, which is a reduced-dimension size form compared to the original beamformer. Moreover, the optimal multicast beamforming vector $\mathbf{w}_i^o$ for BS i is only a function of the channels from BS i to all users in J cells. Thus, $\mathbf{w}_i^o$ can be computed locally at each BS, without requiring the global knowledge of channel state information in other cells.

Following the optimal structure, we proposed several numerical algorithms to solve the QoS problem. These algorithms are all based on a fixed-point iterative method to compute the Lagrange parameter $\boldsymbol{\lambda}$. All of the SCA based approaches, *i.e.,* OptSCA-CVX and OptSCA-ADMM can achieve lower bound nearly, while the OptSDR algorithm has a deteriorate performance when the number of users increases. For the computation complexity, the fast first order Algorithm 2 shows the best performance.

Chapter 4

Multicast Beamforming For the MMF Problem

The previous chapter focuses on the QoS problem for the coordinated multicast beamforming design. We have derived the optimal structure of the multicast beamforming vector $\mathbf{w}_i$ at each BS i. Utilizing this optimal structure, Several computational methods, including the second-order algorithms such as SDR and SCA and the first-order fast algorithm based on the ADMM technique, are proposed to compute the weight parameters with reduced computational complexity for large-scale systems. In this chapter, we consider the weighted MMF problem for multi-cell coordinated multicast beamforming.

4.1 System Model

We consider a downlink multicast transmission scenario in a multi-cell massive MIMO system consisting of J cells. The BS in each cell is equipped with M antennas. Each BS provides the multicast service to a group of K users in its cell, where each user is equipped with a single antenna. We assume that all BSs use the same spectrum bandwidth for transmission.

We study the coordinated multicast beamforming among the J coordinating cells for the multicast service. Each BS multicasts a message to the K users in its own cell, using the beamforming vector that is jointly designed among all the BSs. Define the cell index set $\mathcal{J} \triangleq \{1, \cdots, J\}$ and the user index set $\mathcal{K} \triangleq \{1, \cdots, K\}$. The serving BS in cell j is denoted by BS j. Let $\mathbf{h}_{j,ik}$ denote the $M \times 1$ channel vector from BS j to user k in cell i, for $j, i \in \mathcal{J}$ and $k \in \mathcal{K}$. Let $\mathbf{w}_i$ denote the $M \times 1$ multicast beamforming vector at BS i. The received signal at user k in cell i is given by

$$y_{ik} = \mathbf{w}_i^H \mathbf{h}_{i,ik} s_i + \sum_{j=1, j \neq i}^{J} \mathbf{w}_j^H \mathbf{h}_{j,ik} s_j + n_{ik}, \ k \in \mathcal{K}, \ i \in \mathcal{J}. \tag{4.1}$$

where s_i is the data symbol transmitted from BS i with $\mathrm{E}[|s_i|^2] = 1$, and n_{ik} is the receiver additive white Gaussian noise at the user with zero mean and variance σ^2. The first term in (4.2) is the desired signal, and the second term is the interference from the other BSs of the coordinated neighboring cells.

We study the coordinated multicast beamforming among the J coordinating cells for the multicast service. Each BS multicasts a message to the K users in its own cell, using the beamforming vector that is jointly designed among all the BSs. Define the cell index set $\mathcal{J} \triangleq \{1, \cdots, J\}$ and the user index set $\mathcal{K} \triangleq \{1, \cdots, K\}$. The serving BS in cell j is denoted by BS j. Let $\mathbf{h}_{j,ik}$ denote the $M \times 1$ channel vector from BS j to user k in cell i, for $j, i \in \mathcal{J}$ and $k \in \mathcal{K}$. Let $\mathbf{w}_i$ denote the $M \times 1$ multicast beamforming vector at BS i. The received signal at user k in cell i is given by

$$y_{ik} = \mathbf{w}_i^H \mathbf{h}_{i,ik} s_i + \sum_{j=1, j \neq i}^{J} \mathbf{w}_j^H \mathbf{h}_{j,ik} s_j + n_{ik}, \ k \in \mathcal{K}, \ i \in \mathcal{J}. \tag{4.2}$$

where s_i is the data symbol transmitted from BS i with $\mathrm{E}[|s_i|^2] = 1$, and n_{ik} is the receiver additive white Gaussian noise at the user with zero mean and variance σ^2.

The first term in (4.2) is the desired signal, and the second term is the interference from the other BSs of the coordinated neighboring cells. The transmit power at BS i is given by $\|\mathbf{w}_i\|^2$, for $i \in \mathcal{J}$.

In this chapter, we consider the QoS multicast beamforming problem $\mathcal{S}_o$ (4.3). The objective of MMF problem is to maximize the worst-case performance among all cells, subject to each BS transmit power limit p_i, $i \in \mathcal{J}$. The worst-case performance in each cell is represented by the minimum weighted SINR among K users. Thus, the problem is formulated as

$$\begin{aligned} \mathcal{S}_o: \ & \max_{\mathbf{w}} \min_{i,k} \ \frac{1}{\gamma_{ik}} \frac{|\mathbf{h}_{i,ik}^H \mathbf{w}_i|^2}{\sum_{j=1,j\neq i}^{J} |\mathbf{h}_{j,ik}^H \mathbf{w}_j|^2 + \sigma^2} \\ & \text{s.t. } \|\mathbf{w}_i\|^2 \leqslant p_i, \ i \in \mathcal{J}. \end{aligned} \tag{4.3}$$

In particular, we solve $\mathcal{S}_o$ based on the optimal results obtained from the QoS problem. Firstly, $\mathcal{S}_o$ in (4.3) can be equivalently expressed by

$$\begin{aligned} \mathcal{S}_1: & \max_{\mathbf{w},t} \ t \\ \text{s.t. } & \frac{|\mathbf{h}_{i,ik}^H \mathbf{w}_i|^2}{\sum_{j\neq i} |\mathbf{h}_{j,ik}^H \mathbf{w}_j|^2 + \sigma^2} \geqslant t\gamma_{ik}, \ k \in \mathcal{K}, i \in \mathcal{J} \\ & \frac{1}{p_i}\|\mathbf{w}_i\|^2 \leqslant 1, \ i \in \mathcal{J}. \end{aligned} \tag{4.4}$$

4.2 The Optimal Solution Structure for MMF Problem

For the scenario of a single BS with multiple multicast groups, the inverse relationship between the QoS problem and MMF problem are two inverse problems has been shown [40]. For multi-cell coordinated multicast beamforming, we show that such inverse relationship also holds for QoS problem $\mathcal{P}_1$ in (3.2) and MMF problem $\mathcal{S}_1$ in

(55). Specifically, we parameterize the weighted MMF problem $\mathcal{S}_1$ as $\mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p})$ for a given individual transmit power budget $\mathbf{p} \triangleq [p_N, \ldots, p_J]^T$, and an SINR target vector $\boldsymbol{\gamma} \triangleq [\boldsymbol{\gamma}_1^T, \cdots, \boldsymbol{\gamma}_J^T]^T$ with $\boldsymbol{\gamma}_i \triangleq [\gamma_{i1}, \ldots, \gamma_{iK}]^T$. Also, we denote objective value as $t^o = \mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p})$. Similarly, the QoS problem $\mathcal{P}_1(\mathbf{p})$ is parameterized as $\mathcal{P}_1(\mathbf{p}, \boldsymbol{\gamma})$.Then we have the following proposition.

Proposition 4.1. *The QoS problem $\mathcal{P}_1$ and the MMF problem $\mathcal{S}_1$ is stated below are inverse problems with the following inverse relation:*

$$t^o = \mathcal{S}_1(\boldsymbol{\gamma}, \mathcal{P}_1(\mathbf{p}, t^o\boldsymbol{\gamma}) \cdot \mathbf{p}), \tag{4.5}$$

$$1 = \mathcal{P}_1(\mathbf{p}, \mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p}) \cdot \boldsymbol{\gamma}). \tag{4.6}$$

Proof. We first prove (4.5) by contradiction. Let $\{\mathbf{w}\}_{i=1}^J$ be the optimal solution of $\mathcal{P}_1(\mathbf{p}, \boldsymbol{\gamma})$ with optimal value 1. Consider the MMF problem $\mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p})$, the set $\{\mathbf{w}\}_{i=1}^J$ is a feasible solution since $\max_i \frac{1}{p_i}\|\mathbf{w}_i\|^2 = 1$, which yields $\frac{1}{p_i}\|\mathbf{w}_i\|^2 \leqslant 1, i \in \mathcal{J}$. Let t^o be the corresponding optimal value with $\{\mathbf{w}_i\}_{i=1}^J$. Assume the existence of another feasible solution $\{\tilde{\mathbf{w}}_i\}_{i=1}^J$ with associated optimal value $\tilde{t}^o > t^o$. Then, it is possible to find a constant $c < 1$ to scale down this solution set $\{\tilde{\mathbf{w}}_i\}_{i=1}^J$, while still fulfilling the SINR constraints of problem $\mathcal{P}_1(\mathbf{p}, \boldsymbol{\gamma})$. The resulting set $\{c\tilde{\mathbf{w}}_i\}_{i=1}^J$ has a smaller objective value (minimum individual transmitted power) than 1, which contradicts optimality of $\{\mathbf{w}_i\}_{i=1}^J$.

Similarly, this procedure can be also utilized to prove (4.6). Specifically, let $\{\tilde{\mathbf{w}}_i\}_{i=1}^J$ and $\tilde{t}^o$ denote an optimal solution and the associated optimal value to the MMF problem $\mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p})$. Consider the QoS problem $\mathcal{P}_1(\tilde{t}^o\boldsymbol{\gamma}, \mathbf{p})$, the set $\{\tilde{\mathbf{w}}_i\}_{i=1}^J$ is a feasible solution since $\min_{i,k} \frac{1}{\gamma_{ik}}\mathrm{SINR}_{ik} = \tilde{t}^o$, then $\mathrm{SINR}_{ik} \geqslant \tilde{t}^o\gamma_{ik}$. From the individual

power inequality constraint in (4.4), we can see that $\frac{1}{p_i}\|\tilde{\mathbf{w}}_i\|^2 \leqslant 1$. In this case, the objective value associated with $\{\tilde{\mathbf{w}}_i\}_{i=1}^J$ is $\frac{1}{p_i}\|\tilde{\mathbf{w}}_i\|^2 = 1$. Assume there exists another feasible solution $\{\mathbf{w}_i\}_{i=1}^J$ with associated optimal value $\rho_i(\mathbf{w}_i) < 1$. This contradicts optimality of $\{\tilde{\mathbf{w}}_i\}_{i=1}^J$ for $\mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p})$, since the power budget $1-\rho_i(\mathbf{w}_i)$ can be distributed evenly to yield an objective value t^o larger than $\tilde{t}^o$. □

The inverse relation shown in Proposition 4.1 allows us to solve the MMF problem $\mathcal{S}_1$ by iteratively solving $\mathcal{P}_1$ along with a bi-section search over t until the maximum transmit power margin among the BSs is equal to 1. From this approach, we see that the solution of the MMF problem has a structure similar to that of the QoS problem shown in Theorem 1. We state the optimal beamforming vector for MMF problem $\mathcal{S}_1$ (4.4) below.

Theorem 2. *The optimal beamforming solution for the multi-cell coordinated multicast beamforming MMF problem $\mathcal{S}_1$ is given by*

$$\mathbf{w}^o_{MMF,i} = \mathbf{R}^{-1}_{MMF,i}(\boldsymbol{\lambda}^o_{QoS}, \boldsymbol{\mu}^o_{QoS})\mathbf{H}_i\mathbf{a}^o_{MMF,i},\ i \in \mathcal{J} \tag{4.7}$$

where

$$\mathbf{R}^{-1}_{MMF,i}(\boldsymbol{\lambda}^o_{QoS}, \boldsymbol{\mu}^o_{QoS}) \triangleq \frac{\mu^o_{QoS,i}}{p_i}\mathbf{I} + \frac{1}{\sigma^2}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\lambda^o_{QoS,jk}\gamma_{jk}}{\boldsymbol{\lambda}^{oT}_{QoS}\boldsymbol{\gamma}}\mathbf{h}_{i,jk}\mathbf{h}^H_{i,jk}, \tag{4.8}$$

with $\boldsymbol{\lambda}^o_{QoS}$ and $\boldsymbol{\mu}^o_{QoS}$ being the optimal Lagrangian multipliers from $\mathbf{w}_{QoS}$ in (3.13) *from the QoS problem $\mathcal{P}_1(\mathbf{p}, t^o\boldsymbol{\gamma})$ in Theorem 1.*

In the above theorem, $\mathbf{a}_{MMF,i} \triangleq [a_{MMF,i1}, \ldots, a_{MMF,iK}]^T$ is the weight vector for BS i with given by

$$a^o_{MMF,ik} \triangleq \lambda^o_{QoS,ik}\delta_{ik}\left(1 + \frac{\gamma_{ik}}{\sigma^2\boldsymbol{\lambda}^{oT}_{QoS}\boldsymbol{\gamma}}\right) \tag{4.9}$$

where $\delta_{ik} \triangleq \mathbf{h}_{i,ik}^H \mathbf{w}_{MMF,i}^o, i \in \mathcal{J}, k \in \mathcal{K}$.

The optimal objective value t^o *for problem* $\mathcal{S}_1$ *is given by*

$$t^o = \frac{1}{\sigma^2 \boldsymbol{\lambda}_{QoS}^{oT} \boldsymbol{\gamma}}. \tag{4.10}$$

Proof. From the inverse relation in (4.5) and (4.6), we can obtain the optimal $\mathbf{w}_{\mathrm{MMF},i}^o$ from the QoS problem $\mathcal{P}_1(t^o\boldsymbol{\gamma}, \mathbf{p})$. Note according to the proof of Theorem 1 in Appendix A at the optimality, we have (B.10). Based on (4.6), the minimum objective value of $\mathcal{P}_1(t^o\boldsymbol{\gamma}, \mathbf{p})$, *i.e.,* the maximum individual power ratio, is 1,

$$1 = t^o \sigma^2 \boldsymbol{\lambda}_{\mathrm{QoS}}^{oT} \boldsymbol{\gamma}, \tag{4.11}$$

where $\boldsymbol{\lambda}_{\mathrm{QoS}}^o$ is obtained by (3.21) of $\mathcal{P}_1(t^o\boldsymbol{\gamma}, \mathbf{p})$. Thus,

$$t^o = \frac{1}{\sigma^2 \boldsymbol{\lambda}_{\mathrm{QoS}}^{oT} \boldsymbol{\gamma}} \tag{4.12}$$

Also, since the optimal $\mathbf{w}_{\mathrm{MMF},i}^o$ for $\mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p})$ has the similar optimal beamforming structure as in QoS problem, except that γ_{ik} is now replaced by $t^o\gamma_{ik}$, which based on (4.12) is given by

$$t^o \gamma_{ik} = \frac{\gamma_{ik}}{\sigma^2 \boldsymbol{\lambda}_{\mathrm{QoS}}^{oT} \boldsymbol{\gamma}}. \tag{4.13}$$

Thus, substituting (4.13) into the covariance matrix $\mathbf{R}_i(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ in (3.14), we have $\mathbf{R}_{\mathrm{MMF},i}^{-1}(\boldsymbol{\lambda}_{\mathrm{QoS}}^o, \boldsymbol{\mu}_{\mathrm{QoS}}^o)$ in (4.8). Similarly, using the same relation, a_{ik}^o now becomes $a_{\mathrm{MMF},ik}^o$ shown in (4.9). Hence, we have the optimal beamforming vectors $\mathbf{w}_{\mathrm{MMF},i}^o$ in (4.7). □

However, it is still challengeable to obtain the solution through directly applying the optimal structure in Theorem 2. Since the parameters $\boldsymbol{\lambda}_{\mathrm{QoS}}^o$ and $\boldsymbol{\mu}_{\mathrm{QoS}}^o$ need to be determined in the covariance matrix $\mathbf{R}_{\mathrm{MMF},i}^{-1}(\boldsymbol{\lambda}_{\mathrm{QoS}}^o, \boldsymbol{\mu}_{\mathrm{QoS}}^o)$ are related to the QoS

problem, which are involved and impossible to compute at a time. Thus, we consider to apply the inverse relationship of QoS and MMF problem in Proposition 4.1 to obtain $\mathbf{w}_{\mathrm{MMF},i}$, which is named as QoS2MMF algorithm. This method focuses on iteratively solving the QoS problem to find $\mathbf{w}_i$ for $\mathcal{P}_1(\mathbf{p}, t\boldsymbol{\gamma})$ with a bisection search over t until the maximum transmit power margin $\frac{1}{p_i}\|\mathbf{w}_i\|^2$ among all the BSs is equal to 1.

The main steps for this QoS2MMF iterative method is summarized as follows.

1. Set the upper bound t_{high} and lower bound t_{low} of the SINR target for the QoS problem $\mathcal{P}_1(\mathbf{p}, t\boldsymbol{\gamma})$, since $t = \mathcal{S}_1(\boldsymbol{\gamma}, \mathbf{p})$.

2. Set $t = \frac{t_{\mathrm{high}}+t_{\mathrm{low}}}{2}$.

3. Solve $\mathcal{P}_1(\mathbf{p}, t\boldsymbol{\gamma})$ and obtain the solution $\{\mathbf{w}_i(t)\}_{i=1}^{J}$.

4. Update $t_{\mathrm{high}} = t$ if $\rho \geqslant 1$, or $t_{\mathrm{low}} = t$, otherwise.

5. Repeat Step 2 to 4 until $t_{\mathrm{high}} \to t_{\mathrm{low}}$.

Compared to directly solving $\mathcal{S}_o$ by the SDR method, which has computational complexity of $\mathcal{O}((MJ)^3)$, the complexity of the proposed QoS2MMF algorithm is $\mathcal{O}((KJ)^3)$ per bi-section iteration, the computational complexity for each iteration is reduced from $\mathcal{O}((MJ)^3)$ to $\mathcal{O}((KJ)^3)$.

4.2.1 Fast First-Order Algorithm for MMF Problem

The proposed QoS2MMF method requires to solve the QoS problems multiple times during the bi-section search over t, which adds computational complexity and is not

the most efficient algorithm. To further reduce the computational complexity, we propose a first-order algorithm based on the PSA [42] to solve the MMF problem $\mathcal{S}_o$. The PSA-based algorithm was originally proposed for the single-BS multi-group multicast beamforming MMF problem in [42]. We show below that we can adopt this approach to solve the multi-cell coordinated multicast beamforming MMF problem.

4.2.2 Asymtotic Structure of the Channel Covariance Matrix

From the analysis in Section 4.2, the difficulty of directly computing $\mathbf{w}_{\mathrm{MMF},i}$ (4.7) is in the determination of $\mathbf{R}_{\mathrm{MMF},i}(\boldsymbol{\lambda}^o_{\mathrm{QoS}}, \boldsymbol{\mu}^o_{\mathrm{QoS}})$, since it requires the knowledge of t^o. Observing the structure of covariance matrix $\mathbf{R}_{\mathrm{MMF},i}(\boldsymbol{\lambda}^o_{\mathrm{QoS}}, \boldsymbol{\mu}^o_{\mathrm{QoS}})$ in (4.8), we notice that the contribution from each user channel $\mathbf{h}_{ik}$ is weighted by $\frac{1}{\sigma^2}\frac{\lambda^o_{\mathrm{QoS},jk}\gamma_{jk}}{\boldsymbol{\lambda}^{oT}_{\mathrm{QoS}}\boldsymbol{\gamma}}$, which represents the portion of transmit power allocated to each user. As $M \to \infty$, we can derive the asymptotic expression for $\mathbf{R}_{\mathrm{MMF},i}(\boldsymbol{\lambda}^o_{\mathrm{QoS}}, \boldsymbol{\mu}^o_{\mathrm{QoS}})$ and use this asymptotic expression as an approximation to design a low-complexity approach to compute the solution.

We express each channel as $\mathbf{h}_{i,jk} \triangleq \sqrt{\beta_{i,jk}}\mathbf{g}_{i,jk}$, where $\beta_{i,jk}$ is the channel variance, and $\mathbf{g}_{i,jk}$ is the normalized channel vector representing the small-scale fading whose elements are i.i.d and zero mean. Then, we obtain the asymptotic expression of $\mathbf{R}_{\mathrm{MMF},i}(\boldsymbol{\lambda}^o_{\mathrm{QoS}}, \boldsymbol{\mu}^o_{\mathrm{QoS}})$ as M becomes large by following the techniques and results in [40]. Using this expression, $\mathbf{R}_{\mathrm{MMF},i}(\boldsymbol{\lambda}^o_{\mathrm{QoS}}, \boldsymbol{\mu}^o_{\mathrm{QoS}})$ can be approximated by

$$\mathbf{R}_{\mathrm{MMF},i}(\boldsymbol{\lambda}_{\mathrm{QoS}}, \boldsymbol{\mu}_{\mathrm{QoS}}) \approx \frac{\mu_i}{p_i}\mathbf{I} + \frac{\bar{\beta}_{\mathbf{h}}}{\sigma^2 K}\sum_{j=1}^{J}\sum_{k=1}^{K}\mathbf{g}_{i,jk}\mathbf{g}^H_{i,jk} \triangleq \mathbf{R}^{\infty}_{\mathrm{MMF},i}, \tag{4.14}$$

where $\bar{\beta}_{\mathbf{h}}$ is the harmonic mean of the large-scale channel variances of all users, is

represented by

$$\bar{\beta}_{\mathbf{h}} \triangleq \frac{1}{\frac{1}{K}(\sum_{j=1}^{J}\sum_{i=1}^{J}\sum_{k=1}^{K}\frac{1}{\beta_{i,jk}})}. \tag{4.15}$$

Since channel covariance $\mathbf{R}^{\infty}_{\mathrm{MMF},i}$ in (4.14) is expressed in closed-form, the only unknown variable in (4.7) to be computed is weight vectors $\mathbf{a}_i$. Similar to Section **??**, we apply the structure of $\mathbf{w}^{o}_{\mathrm{MMF},i}$ in (4.7) to $\mathcal{S}_1$ to convert it into a weight minimization problem w.r.t. $\mathbf{a}$:

$$\mathbf{w}_{\mathrm{MMF},i} = \mathbf{R}^{-1}_{\mathrm{MMF},i}(\boldsymbol{\lambda}_{\mathrm{QoS}}, \boldsymbol{\mu}_{\mathrm{QoS}})(4.14)\mathbf{H}_i\mathbf{a}_{\mathrm{MMF},i} \tag{4.16}$$

$$\begin{aligned}\mathcal{S}_2 : \max_{\mathbf{a}} \min_{i,k} \frac{1}{\gamma_{ik}} \frac{\mathbf{a}_i^H \mathbf{B}_{i,ik}\mathbf{a}_i}{\sum_{j=1,j\neq i}^{J}\mathbf{a}_j^H\mathbf{B}_{j,ik}\mathbf{a}_j + \sigma^2} \; k \in \mathcal{K}, i \in \mathcal{J} \\ \text{s.t. } \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2 \leqslant 1, i \in \mathcal{J}\end{aligned} \tag{4.17}$$

where $\mathbf{a} \triangleq [\mathbf{a}_1^H, \dots \mathbf{a}_J^H]^H$, $\mathbf{G}_i \triangleq \mathbf{R}^{\infty}_{\mathrm{MMF},i}\mathbf{H}_i$, for $i \in \mathcal{J}$, and $\mathbf{B}_{j,ik} \triangleq \mathbf{G}_j^H\mathbf{h}_{j,ik}\mathbf{h}_{j,ik}^H\mathbf{G}_j$, for $k \in \mathcal{K}, j, i \in \mathcal{J}$.

Note that the dimension of weight vector $\mathbf{a}$ is KJ, which is much lower than that of the beamforming vector $\mathbf{w}$, which is MJ for massive MIMO systems with $K \ll M$. Thus, $\mathcal{S}_2$ has much smaller size than $\mathcal{S}_1$.

4.2.3 Problem Reformulation

Since $\mathcal{S}_2$ is still a non-convex and NP-hard problem, we consider to adopt the PSA method [42] to compute a near-stationary solution to $\mathcal{S}_1$ efficiently.

Denote the weighted SINR for user k in cell i as

$$\psi_{ik}(\mathbf{a}) = -\frac{1}{\gamma_{ik}} \frac{\mathbf{a}_i^H \mathbf{B}_{i,ik}\mathbf{a}_i}{\sum_{j=1,j\neq i}^{J}\mathbf{a}_j^H\mathbf{B}_{j,ik}\mathbf{a}_j + \sigma^2}, \tag{4.18}$$

and denote the feasible set of $\mathcal{S}_2$ as

$$\mathcal{A} \triangleq \{\mathbf{a} : \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2 \leqslant 1\}. \tag{4.19}$$

Then, we can equivalently express the optimization objective in $\mathcal{S}_2$ in a min-max form as $\min_{\mathbf{a}} \max_{i,k} \psi_{i,k}(\mathbf{a})$. Furthermore, define a probability vector $\mathbf{y} : y \geqslant 0$ and $\mathbf{1^T y} = \mathbf{1}$. Define

$$f(\mathbf{a}, \mathbf{y}) \triangleq \boldsymbol{\psi}^T(\mathbf{a})\mathbf{y}, \tag{4.20}$$

where $\boldsymbol{\psi}(\mathbf{a}) \in \mathbb{R}^{K_{\text{tot}}}$ containing all $\psi_{ik}(\mathbf{a})$'s.

Thus, we can further equivalently rewrite $\mathcal{S}_2$ as

$$\mathcal{S}_3 : \min_{\mathbf{a}\in\mathcal{A}} \max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{a}, \mathbf{y}). \tag{4.21}$$

From the definition of weighted SINR, we can tell an optimal solution of $\mathbf{y}$ is $\mathbf{y}^o \triangleq [0, \ldots, 0, 1, 0, \ldots, 0]^T$, the index of 1 shows the worst-SINR-user position among the total JK users Note that $f(\mathbf{a}, \mathbf{y})$ is concave in $\mathbf{y}$ and nonconvex in $\mathbf{a}$. Thus, $\mathcal{S}_3$ is a nonconvex-concave min-max problem. Let $g(\mathbf{a}) \triangleq \max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{a}, \mathbf{y})$. Then, the problem $\mathcal{S}_3$ can be further expressed as:

$$\mathcal{S}_4 : \min_{\mathbf{a}\in\mathcal{A}} g(\mathbf{a}) \tag{4.22}$$

Since $g(\mathbf{a})$ may not be differentiable, and its gradient $\nabla_{\mathbf{a}} g(\mathbf{a})$ may not exits, we apply the PSA method to find a solution at the vicinity of a stationary point for $\mathcal{S}_4$

4.2.4 The Projected Subgradient Algorithm

At iteration l:

$$\mathbf{y}^{(l)} \in \arg\max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{a}^{(l)}, \mathbf{y}) \tag{4.23}$$

$$\mathbf{a}^{(l+1)} = \Pi_{\mathcal{A}}\left(\mathbf{a}^{(l)} - \alpha \nabla_{\mathbf{a}} f(\mathbf{a}^{(l)}, \mathbf{y}^{(l)})\right) \tag{4.24}$$

where $\alpha > 0$ is the step size, and $\Pi_{\mathcal{A}}(\mathbf{a})$ denotes the projection of point $\mathbf{a}$ onto set $\mathcal{A}$. Denotes the ratio of the transmit power over its power budget at BS i as $\rho_{\mathbf{a}_i} \triangleq \frac{1}{p_i}\|\mathbf{G}_i\mathbf{a}_i\|^2$, then the projection is given by

$$\Pi_{\mathcal{A}}(\mathbf{a}_i) = \begin{cases} \sqrt{\frac{1}{\rho_{\mathbf{a}_i}}}\mathbf{a}_i, & \mathbf{a}_i \notin \mathcal{A}, i \in \mathcal{J} \\ \mathbf{a}_i, & \mathbf{a}_i \in \mathcal{A}, i \in \mathcal{J} \end{cases} \tag{4.25}$$

For update in (4.24), we need to compute the gradient $\nabla_{\mathbf{a}} f(\mathbf{a}, \mathbf{y})$ w.r.t. $\mathbf{a}$, which can be obtained in closed-form. In particular, the gradient of $f(\mathbf{a}, \mathbf{y})$ is given by

$$\nabla_{\mathbf{a}} f(\mathbf{a}, \mathbf{y}) = [\nabla_{\mathbf{a}_1} f(\mathbf{a}, \mathbf{y}), \ldots, \nabla_{\mathbf{a}_J} f(\mathbf{a}, \mathbf{y})]^H \tag{4.26}$$

where

$$\nabla_{\mathbf{a}_{\hat{i}}} f(\mathbf{a}, \mathbf{y}) = \nabla_{\mathbf{a}_{\hat{i}}} \boldsymbol{\psi}^H(\mathbf{a})\mathbf{y} = \nabla_{\mathbf{a}_{\hat{i}}} \psi_{ik}(\mathbf{a}),\ i, \hat{i} \in \mathcal{N}, k \in \mathcal{K}.$$

The gradient $\nabla_{\mathbf{a}} \psi_{ik}(\mathbf{a})$, for $k \in \mathcal{K}, i, \hat{i} \in \mathcal{J}$, is given by

$$\nabla_{\mathbf{a}_{\hat{i}}} \psi_{ik}(\mathbf{a}) = \begin{cases} -\dfrac{2\mathbf{B}_{i,ik}\mathbf{a}_i}{\sum_{j\neq i, j=1}^{J} \mathbf{a}_j^H \mathbf{B}_{j,ik}\mathbf{a}_j + \sigma^2}, & i = \hat{i}, \\ \dfrac{2\mathbf{a}_i^H \mathbf{B}_{i,ik}\mathbf{a}_i \mathbf{B}_{\tilde{i}ik}\mathbf{a}_{\hat{i}}}{(\sum_{j\neq i, j=1}^{J} \mathbf{a}_j^H \mathbf{B}_{j,ik}\mathbf{a}_j + \sigma^2)^2}, & i \neq \hat{i}. \end{cases} \tag{4.27}$$

The proposed PSA-based first-order algorithm for the MMF problem is summarized in Algorithm 3.

Remark: Algorithm 3 requires an initial point $\mathbf{a}^{(0)}$. To find a good initial point, we proposed solving $\mathcal{P}_{2\text{SDR}}$ along with one iteration bi-section over t, and use the Gaussian randomization method to generate the weight vector $\mathbf{a}$ from $\mathbf{X}$. Note that the initial point $\mathbf{a}^{(0)}$ does not need to be feasible, the projection operation $\Pi_{\mathcal{A}}(\cdot)$ in (4.25) will project $\mathbf{a}$ into $\mathcal{A}$ in (4.19) after one iteration.

Algorithm 3 MMF-PSA for $\mathcal{S}_1$

1: **Initialization**: Set initial $\mathbf{a}^{(0)}$.
2: **repeat**
3: $\quad \mathbf{y}^{(l)} \in \arg\max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{a}^{(l)}, \mathbf{y})$.
4: $\quad \mathbf{a}^{(l+1)} = \Pi_{\mathcal{A}}(\mathbf{a}^{(l)} - \alpha\nabla_{\mathbf{a}} f(\mathbf{a}^{(l)}, \mathbf{y}^{(l)}))$.
5: $\quad l \leftarrow l + 1$.
6: **until** convergence.
7: Obtain $\mathbf{w}$ by using (4.16).

4.3 Simulation Results

Similar as the study on QoS problem, default system for MMF problem setup is $J = 3$ multicasting cells, with $K = 5$ users each cell. For MMF problem $\mathcal{S}_o$, we evaluate the following three proposed algorithms in this chapter:

- QoS2MMF-SDR: Solving $\mathcal{S}_1$ by iteratively solving the QoS problem $\mathcal{P}_1$ using OptSDR, along with bi-section search over t.
- QoS2MMF-SCA: Solving $\mathcal{S}_1$ by iteratively solving the QoS problem $\mathcal{P}_1$ using OptSCA-CVX, along with bi-section search over t.
- MMF-PSA: The proposed PSA-based fast algorithm in Algorithm 3.

For comparison, we also present the following the methods:

- Upper Bound for $\mathcal{S}_o$: Solving the relaxed problem of $\mathcal{S}_1$ via SDR along with bi-section search over t.
- DirectSDR: Solve the relaxed problem of $\mathcal{S}_1$ via the SDR method, combining with the Gaussian randomization approach to obtain $\{\mathbf{w}_i\}$.

- WeightedMRT [38]: A low-complexity suboptimal approach proposed in [38] that using the MRT beamforming structure.

4.3.1 Convergence Behavior

For the MMF problem $\mathcal{S}_o$, we study the convergence behavior of Algorithm 3. We set $M = 100, K = 5$ as the default system.

In MMF-PSA algorithm, we need to determine step size α. Fig. 4.1 shows the convergence behavior of the absolute difference of the minimum SINR under a series of step size α settings. From this figure, we can observe that if $\alpha \leqslant 10^{-4}$, there is an obvious trend shown for the objective value converge. Also, as the step size α decreases, the less iterations needed for the absolute difference of minimum SINR.

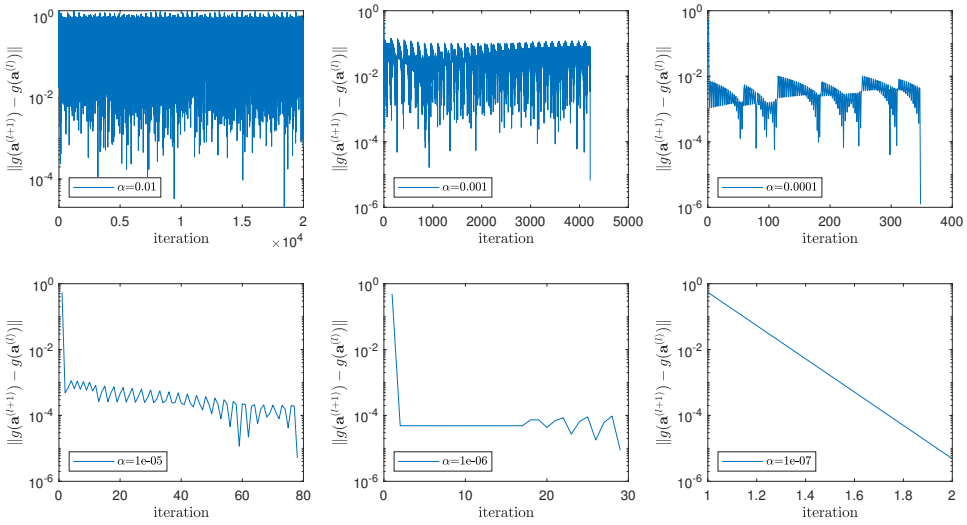

Figure 4.1: Convergence behavior of absolute difference of minimum SINR under different step size α ($M = 50, K = 5$).

Fig. 4.2 shows the minimum SINR that can achieved under different step size α settings when we set threshold for $|g(\mathbf{a}^{(l+1)} - g(\mathbf{a}^{(l)})|$ as $\epsilon = 10^{-5}$. From Fig. 4.2,

we can see that the minimum SINR stays relatively unchangeable if $\alpha \in (10^{-7}, 10^{-4})$. From these observations, overall for the rest simulation results, we set as $\alpha = 10^{-4}$ and $\epsilon = 10^{-3}$ to obtain a faster convergence.

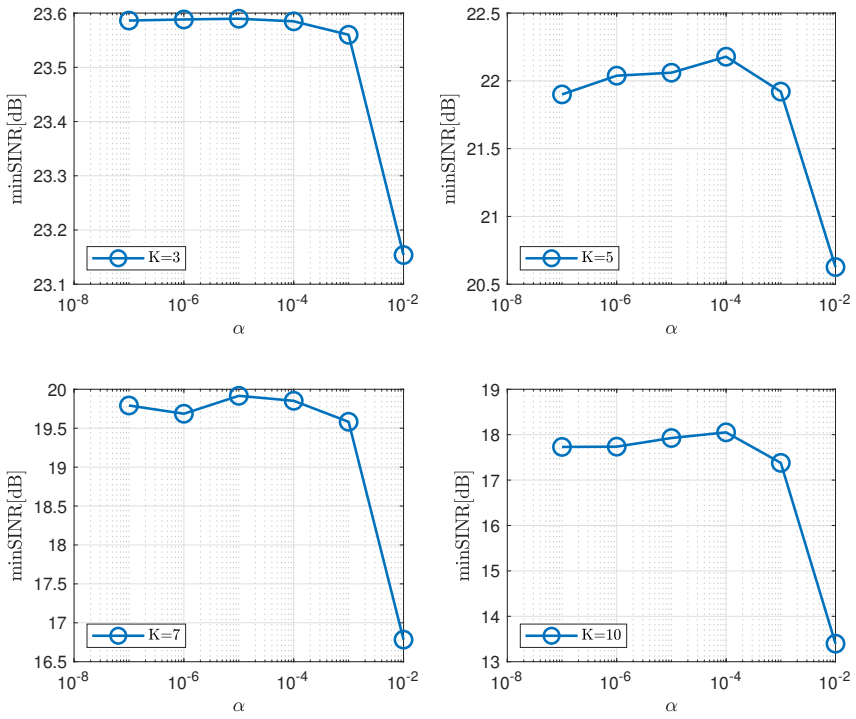

Figure 4.2: Influence on the minimum SINR affected by the step size α under different number of users K $(M = 50, J = 3, \epsilon = 10^{-5})$.

Fig 4.3 shows the trajectory of the absolute difference $|g(\mathbf{a}^{(l+1)}) - g(\mathbf{a}^{(l)})|$ over iterations, where $g(\mathbf{a})$ is the objective function of $\mathcal{S}_4$ in (4.22). We see this absolute difference reaches 10^{-3} in less than 800 iterations.

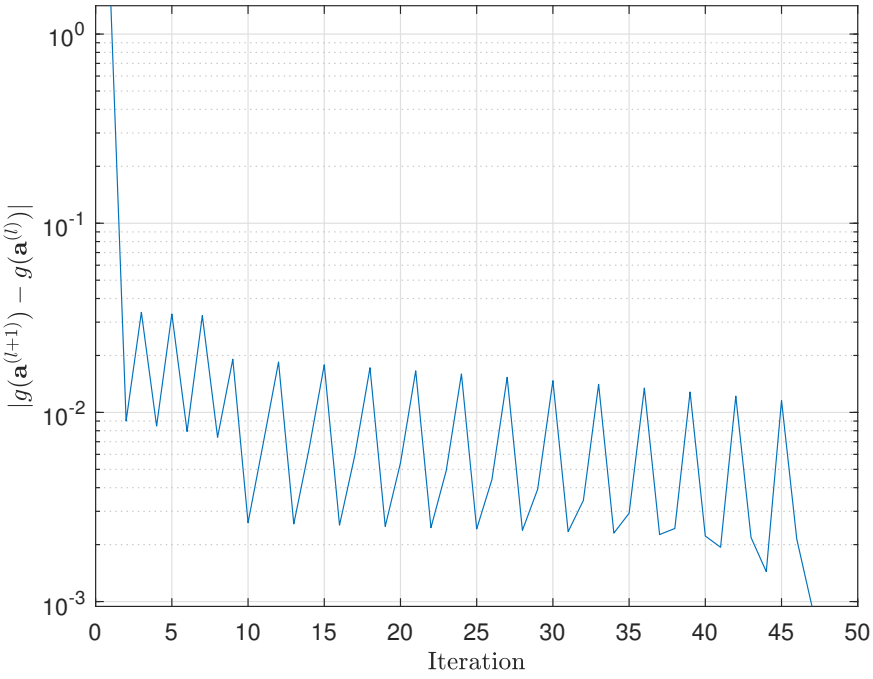

Figure 4.3: Convergence behavior for $\mathcal{S}_4$: Absolute difference of minimum SINR between two consecutive iterations among all users. ($M = 100$, $K = 5$).

4.3.2 Performance Analysis for MMF Problem

We now study the performance of our proposed MMF-PSA in Algorithm 3 for the MMF problem $\mathcal{S}_o$ over the number of antennas M and the number of users K per cell. Fig 4.4 shows the average mininum SINR among users vs. M under all the methods listed at the beginning of this section. We see that our proposed algorithm nearly attain the upper bound, indicating its near-optimal performance. WeightedMRT [38] uses a suboptimal beamforming structure to compute the weight vectors $\mathbf{a}$, and thus it has a increasing gap to our algorithm as M increases, and thus the computation time remains roughly constant over M.

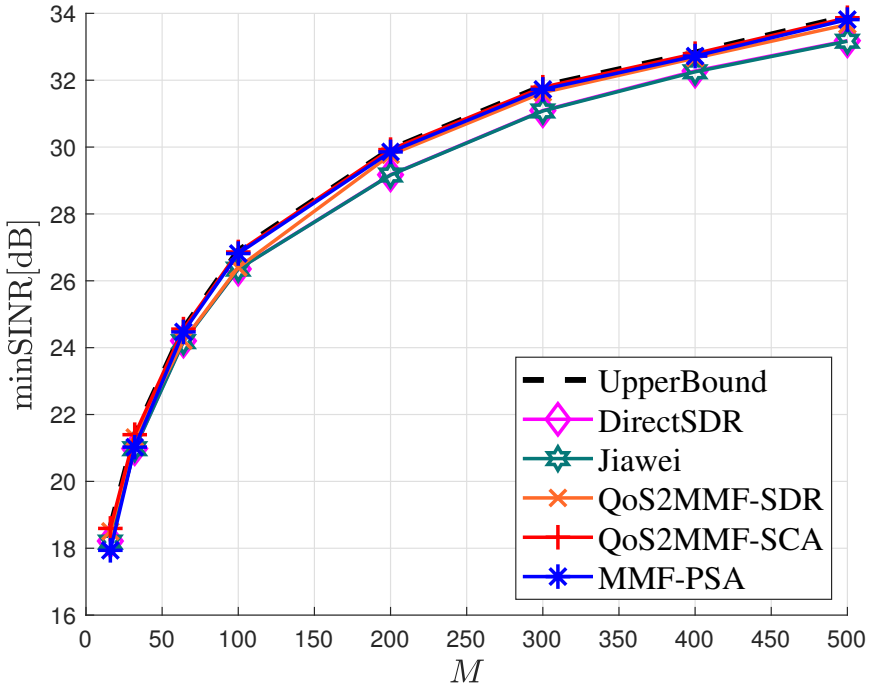

Figure 4.4: MMF problem: The average minimum SINR vs. the number of antennas M. $(K = 5)$.

The corresponding computation times of these methods in Fig. 4.4 are shown in Table 4.1. The computation time of QoS2MMF-SDR and QoSMMF-SCA is higher among these algorithms and increases slightly with M. As discussed in Section 4.2.1, this is because the two methods use bi-section search and requires to iteratively solve the QoS problem, which is not computationally efficient. In contrast, our proposed MMF-PSA is substantially faster than the rest methods,to compute the solution, and the computation time is roughly constant as M increases. The computation time for WeightedMRT [38] is comparable to QoSMMF-SDR and stays steadily as M increases. It uses the suboptimal structure applied and solve the relaxed MMF problem $\mathcal{S}_o$ directly.

We also show the average minimum SINR over K by different methods in Fig 4.5,

Table 4.1: Comparison of average computation time (sec) for MMF problem $\mathcal{S}_o$ $(J = 3, K = 5)$.

M	100	200	300	400	500
MMF-PSA	1.3054	1.2672	1.3967	1.3985	1.1924
QoS2MMF-SDR	13.28	14.01	16.46	20.24	25.01
QoS2MMF-SCA	96.91	104.71	91.73	115.58	119.48
WeightedMRT [38]	15.4778	15.0249	15.2444	15.1870	15.3118
DirectSDR	2258	18334	-	-	-

for $M = 50, 100, 200$. Overall,our proposed MMF-PSA and QoSMMF-SCA nearly attain the upper bound for all values of M. They significantly outperform WeightedMRT [38] , which shows $1 \sim 2$ dB gap to ours as K increases. Also, the SDR-based algorithms, DirectSDR and QoS2MMF-SDR, have deteriorating performance as K increases, which is due to its inaccurate approximation as the problem size grows.

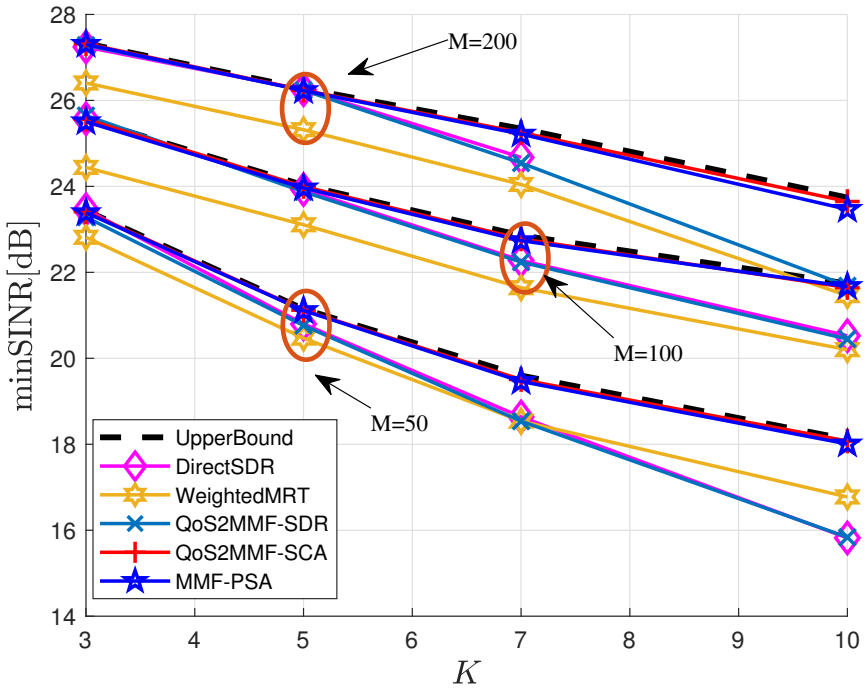

Figure 4.5: MMF problem: The average minimum SINR vs. the number of users K. $(M = 100)$.

Table 4.2: Comparison of average computation time (sec) for MMF problem $\mathcal{P}_o$ ($J = 3, M = 100$).

K	3	5	7	10
MMF-PSA	0.7828	1.0558	1.3884	2.4346
QoS2MMF-SDR	5.2624	6.7534	8.6851	12.9281
QoS2MMF-SCA	26.96	56.60	77.55	137.57
WeightedMRT [38]	6.1583	7.8044	10.3540	18.4640
DirectSDR	1240.5	2474.8	4083.8	7517.6

In Table 4.2, we summarize the average computation time of the methods shown in Fig 4.5. The computation time of all the methods increase with K, Our proposed MMF-PSA is the fastest algorithm. Compared with other methods, its computation time is only 1.7% $\sim$ 2.8% of that for QoS2MMF-SCA and 13% $\sim$ 18% of that for QoS2MMF-SDR or WeightedMRT. This shows MMF-PSA is highly effective in both performance and computational efficiency.

4.4 Summary

This chapter derives the optimal beamforming structure for the MMF problem under a multi-cell multicast scenario. However, the structure is almost impossible to be utilized directly for the MMF problem since the parameters needed are challenging to determine. To address this issue, we proposed iteratively solving the QoS problem to obtain the solution to the MMF problem.

Indeed, we apply the PSA method [42] and combine the asymptotic beamforming structure with the first-order approach. Compared to the iterative QoS2MMF method, the MMF-PSA shows low complexity. Besides the QoS2MMF method, the other proposed methods can all achieve a stationary upper bound even if the number

of users is large.

Chapter 5

Conclusions and Future Work

In this thesis, we considered both the QoS and the MMF problem for multi-cell multicast beamforming. We applied the derivation of the optimal beamforming solution structure in [40] to develop the optimal solution for a multiple-cell case. The derivation by the Lagrangian duality shows the beamforming solution has a similar structure as in [40]: the weighted user channels in each cell can be viewed as a group-channel direction for the cells. By the derived structure, we proposed three algorithms, OptSDR, OptSCA-CVX, and OptSCA-ADMM, to solve the QoS problem numerically. Theoretically, the two last algorithms are both based on the SCA method. The difference is that the former utilizes a CVX solver, and the latter focuses on the ADMM techniques to solve the SCA problem iteratively. The simulation result presented that our proposed algorithms can obtain near-optimal solutions compared with directly using the SDR method. Also, our proposed methods have less computational complexity.

For the MMF problem, we developed the optimal beamforming structure and applied the inverse relation between it and the QoS problem. Since determining the parameters in MMF optimal beamforming structure directly is almost impossible, we mainly focus on obtaining the MMF beamformer by iteratively solving the QoS

problem. From the simulation result, our QoS2MMF-SCA can achieve the upper bound closely. However, these QoS2MMF methods have relatively high computation complexity. To overcome this, we apply another approach, PSA [42], to solve the MMF problem. Compared with other methods, *i.e.,* Weighted MRT [**?**] or DirectSDR, this PSA method provided a closed-form gradient and utilized a projection to get closer to the near-optimal solution. Simulation results showed that our proposed algorithms can perform very closely to the upper bound, besides the QoS2MMFSDR method.

In the future, the ADMM techniques and QoS2MMF-SCA can be combined to improve the performance of the MMF method further. Although this approach may still need a bi-section search, the complexity at each iteration can be reduced compared to utilizing CVX solver. Moreover, the extra gradient-based SCA algorithm [64] can also be utilized to solve the QoS problem, which is also a first-order algorithm. Finally, our proposed algorithms can be extended to study a common scenario in reality, such as a jointly unicast with a multi-cell multicast system.

Appendix A

Proof of Proposition 3.1

Proof. Based on (3.7), the minimization of $\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z})$ over $\mathbf{w}$ and t in (3.10) can be separated. Assuming the optimal Lagrange multipliers $(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ for the dual problem $\mathcal{D}_{1,\text{SCA}}(\mathbf{z})$, solving the the minimization of $\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star; \mathbf{z})$ by setting the derivative of $\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star; \mathbf{z})$ w.r.t. $\mathbf{w}_i^H$, $i \in \mathcal{J}$, to 0 , we have

$$\frac{\partial}{\partial \mathbf{w}_i^H} \mathcal{L}(\mathbf{z}, \mathbf{w}, t, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) = \mathbf{R}_{i^-}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)\mathbf{w}_i(\mathbf{z}) - \boldsymbol{\nu}_i = 0. \tag{A.1}$$

Also, by setting the derivative of $\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star; \mathbf{z})$ w.r.t. t, we obtain $\mathbf{1}^T \boldsymbol{\mu}^\star = 1$. We now show that $\mu_i^\star > 0$ and thus $\mathbf{R}_{i^-}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ is invertible. This can be proved by contradiction. Assume $\mu_i^\star = 0$ for some i. Then, $\mathbf{R}_{i^-}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ in (3.8) is rank deficient. Notice that it is spanned by channels from BS i to all users in other cells $\{\mathbf{h}_{i,jk}, k \in \mathcal{K}, j \in \mathcal{J}, j \neq i\}$, while $\boldsymbol{\nu}_i$ from (3.9) is a linear combination of channels from BS i to its own users in cell i $\{\mathbf{h}_{i,ik}, k \in \mathcal{K}\}$. Since channels of different users are independent, this means that there is no solution to the linear equation in (A.1), regardless of the values of γ_{ik} or $\mathbf{z}$. In other words, the gradient in (A.1) will not be 0 at optimality. This contradict with the condition of the optimal solution to $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$.

Thus, the optimal $\mu_i^\star > 0$, for $i \in \mathcal{J}$. Following this, from (A.1) we have

$$\mathbf{w}_i^\star(\mathbf{z}) = \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)\left(\sum_{k=1}^{K} \lambda_{ik}^\star \mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H\right)\mathbf{z}_i$$

which leads to (3.11). □

Appendix B

Proof of Theorem 1

Proof. The proof follows the proof of [40, Theorem 1]. Specifically, by the optimal $\mathbf{w}_i^\star(\mathbf{z})$ for $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ in (3.11), we have the following

$$\mathbf{R}_{i^-}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)\mathbf{w}_i^\star(\mathbf{z}) = \sum_{k=1}^{K} \lambda_{ik}\mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H\mathbf{z}_i. \tag{B.1}$$

Since $\mathbf{R}_i(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{R}_{i^-}(\boldsymbol{\lambda}, \boldsymbol{\mu}) + \sum_{k=1}^{K} \lambda_{ik}\gamma_{ik}\mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H$, we can rewrite (B.1) as

$$\begin{aligned}
&\mathbf{R}_i(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)\mathbf{w}_i^\star(\mathbf{z}) \\
&= \left(\mathbf{R}_{i^-}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) + \sum_{k=1}^{K} \lambda_{ik}\mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H\right)\mathbf{w}_i^\star(\mathbf{z}) \\
&= \sum_{k=1}^{K} \lambda_{ik}^\star\mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H\mathbf{z}_i + \sum_{k=1}^{K} \lambda_{ik}^\star\gamma_{ik}\mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H\mathbf{w}_i^\star(\mathbf{z}) \\
&= \sum_{k=1}^{K} \lambda_{ik}^\star(1+\gamma_{ik})\left(\mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H\mathbf{z}_i + \mathbf{h}_{i,ik}\mathbf{h}_{i,ik}^H\mathbf{w}_i^\star(\mathbf{z})\right).
\end{aligned} \tag{B.2}$$

As $\mathbf{z} \to \mathbf{w}^o$, we have $\mathbf{w}_i^\star(\mathbf{z}) \to \mathbf{w}_i^o$. At the same time, the optimal $(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ for the dual problem $\mathcal{D}_{1\text{SCA}}(\mathbf{z})$ also converges to the optimal $(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ of $\mathcal{D}_{1,\text{SCA}}(\mathbf{w}^o)$, which is the dual problem of $\mathcal{P}_1$. Thus, at the limit as $\mathbf{z} \to \mathbf{w}^o$, (B.2) becomes

$$\begin{aligned}
\mathbf{R}_i(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)\mathbf{w}_i^o &= \sum_{k=1}^{K} \underbrace{\lambda_{ik}^o(1+\gamma_{ik})(\mathbf{h}_{i,ik}^H\mathbf{w}^o)}_{\triangleq a_{ik}^o}\mathbf{h}_{i,ik} \\
&= \mathbf{H}_i\mathbf{a}_i^o
\end{aligned} \tag{B.3}$$

Following the argument in Appendix A, we can similarly show that $\mu_i^o > 0$, for $i \in \mathcal{J}$, and thus $\mathbf{R}_i(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ is full rank and invertible. Thus, we obtain the optimal solution in (3.13) for $i \in \mathcal{J}$.

The optimal value of $\mathcal{P}_o$ is also the optimal t^o in $\mathcal{P}_1$. Consider the optimal solution to $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ in each SCA iteration in (3.11). We can express $\boldsymbol{\nu}_i$ in (3.9) in a compact matrix form as

$$\boldsymbol{\nu}_i = \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{z}_i, \tag{B.4}$$

where $\mathbf{D}_{\boldsymbol{\lambda}_i} \triangleq \text{diag}(\boldsymbol{\lambda}_i)$. Thus, the optimal beamforming vector $\mathbf{w}_i^\star(\mathbf{z})$ can be rewritten as below

$$\mathbf{w}_i^\star(\mathbf{z}) = \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{z}_i. \tag{B.5}$$

Denote $\boldsymbol{\gamma}_i \triangleq [\gamma_{i1}, \ldots \gamma_{iK}]^T$, $i \in \mathcal{J}$, and $\boldsymbol{\gamma} \triangleq [\boldsymbol{\gamma}_1^T, \ldots, \boldsymbol{\gamma}_J^T]^T$. Substituting the above expression (B.5) into (3.10), the Lagrangian is equivalent to

$$\begin{aligned}
&g(\boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}) \triangleq \min_{\mathbf{w}, t} \mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}) \\
&= (1 - \sum_{i=1}^{J} \mu_i) t + \sum_{i=1}^{J} \sum_{k=1}^{K} \lambda_{ik} \left(\sigma^2 \gamma_{ik} + |\mathbf{z}_i^H \mathbf{h}_{i,ik}|^2 \right) \\
&\quad + \sum_{i=1}^{J} \mathbf{z}_i^H \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{z}_i \\
&\quad - 2 \sum_{i=1}^{J} \mathfrak{Re} \left\{ \mathbf{z}_i^H \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{z}_i \right\} \\
&= (1 - \sum_{i=1}^{J} \mu_i) t + \sigma^2 \sum_{i=1}^{J} \boldsymbol{\lambda}_i^T \boldsymbol{\gamma}_i + \sum_{i=1}^{J} \mathbf{z}_i^H \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{z}_i \\
&\quad - \sum_{i=1}^{J} \mathbf{z}_i^H \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \mathbf{z}_i \\
&= \sum_{i=1}^{J} \mathbf{z}_i^H \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \left(\mathbf{I} - \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) \mathbf{H}_i \mathbf{D}_{\boldsymbol{\lambda}_i} \mathbf{H}_i^H \right) \mathbf{z}_i \\
&\quad + (1 - \sum_{i=1}^{J} \mu_i) t + \sigma^2 \sum_{i=1}^{J} \boldsymbol{\lambda}_i^T \boldsymbol{\gamma}_i.
\end{aligned} \tag{B.6}$$

Based on the proof of Proposition 3.1, the solution for QoS problem at the optimality is represented as (3.12), substituting the optimal weight vector $\boldsymbol{\alpha}_i^o$ into (3.12),

$$\begin{aligned}\mathbf{w}_i^o &= \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)\mathbf{H}_i\boldsymbol{\alpha}_i^o \\ &= \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)\mathbf{H}_i\mathbf{D}_{\boldsymbol{\lambda}^o}\mathbf{H}_i^H\mathbf{w}_i^o. \end{aligned} \tag{B.7}$$

Thus,

$$\left(\mathbf{I} - \mathbf{R}_{i^-}^{-1}(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)\mathbf{H}_i\mathbf{D}_{\boldsymbol{\lambda}_i^o}\mathbf{H}_i^H\right)\mathbf{w}_i^o = \mathbf{0}. \tag{B.8}$$

Through the iterative SCA method and under optimal setting of $\boldsymbol{\mu}^o$, if the primal problem $\mathcal{P}_{1\text{SCA}}(\mathbf{z})$ converges to the optimum, the Lagrangian multiplier in dual problem $\mathcal{D}_{1\text{SCA}}(\boldsymbol{\lambda})$ would also converge, which is $\boldsymbol{\lambda}^\star(\mathbf{z}) \to \boldsymbol{\lambda}^o$ based the optimal convergence of $\mathbf{z}_i \to \mathbf{w}_i^o$ in (3.12). Thus, the first term in (B.6) would become zero.

Also, the KKT conditions for Lagrangian $\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star; \mathbf{z})$ (3.6) still hold. By applying the partial derivative of $\mathcal{L}(\mathbf{w}, t, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star; \mathbf{z})$ w.r.t. t set as zero, we obtain $\mathbf{1}^T\boldsymbol{\mu}^\star = 1$ at the optimality, which is equivalent to

$$\sum_{j=1}^{J} \mu_i^\star = 1. \tag{B.9}$$

Now, we conclude that the first and second term in (B.6) become zero as $\mathbf{z}_i$ converges to the optimum $\mathbf{w}_i^o$, which yields the optimum value is the term left

$$t^\star = \min\ t = \max_{\boldsymbol{\lambda},\boldsymbol{\mu}} g(\boldsymbol{\lambda}, \boldsymbol{\mu}; \mathbf{z}) = \sigma^2 \sum_{i=1}^{J} \boldsymbol{\lambda}_i^\star\boldsymbol{\gamma}_i = \sigma^2\boldsymbol{\lambda}^{\star T}\boldsymbol{\gamma}. \tag{B.10}$$

Based on the above proof, as $\mathbf{z} \to \mathbf{w}^o$ through iterative SCA methods, $\mathcal{D}_{1\text{SCA}}(\boldsymbol{\lambda}^\star) \to \mathcal{D}_{1\text{SCA}}(\boldsymbol{\lambda}^o)$, $\mathcal{P}_{\text{SCA}}(\mathbf{z}) \to \mathcal{P}_o$, thus $t^\star \to t^o$ and the minimum individual objective value is obtained as (B.10).

□

References

[1] E. Björnson, Y. C. Eldar, E. G. Larsson, A. Lozano, and H. V. Poor, "Twenty-five years of signal processing advances for multiantenna communications: From theory to mainstream technology," *IEEE Signal Processing Magazine*, vol. 40, no. 4, pp. 107–117, 2023.

[2] 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 Processing Mag.*, vol. 30, pp. 40–60, Jan. 2013.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] G. C. Alexandropoulos, P. Ferrand, J.-M. Gorce, and C. B. Papadias, "Advanced coordinated beamforming for the downlink of future LTE cellular networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 54–60, 2016.

[5] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE transactions on wireless communications*, vol. 9, no. 5, pp. 1748–1759, 2010.

[6] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, I. Chih-Lin *et al.*, "Millimeter wave communications for future mobile networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909–1935, 2017.

[7] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 104–110, 2014.

[8] M.-L. Ku, W. Li, Y. Chen, and K. R. Liu, "Advances in energy harvesting communications: Past, present, and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1384–1412, 2015.

[9] F. Amirnavaei and M. Dong, "Online power control optimization for wireless transmission with energy harvesting and storage," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4888–4901, 2016.

[10] M. Dong, W. Li, and F. Amirnavaei, "Online joint power control for two-hop wireless relay networks with energy harvesting," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 463–478, 2017.

[11] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 82–89, 2014.

[12] G. S. Paschos, G. Iosifidis, M. Tao, D. Towsley, and G. Caire, "The role of caching in future communication systems and networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 6, pp. 1111–1125, 2018.

[13] Y. Deng and M. Dong, "Fundamental structure of optimal cache placement for coded caching with nonuniform demands," *IEEE Transactions on Information Theory*, vol. 68, no. 10, pp. 6528–6547, 2022.

[14] ——, "Memory-rate tradeoff for caching with uncoded placement under nonuniform random demands," *IEEE Transactions on Information Theory*, vol. 68, no. 12, pp. 7850–7870, 2022.

[15] R. AliHemmati, B. Liang, M. Dong, G. Boudreau, and S. H. Seyedmehdi, "Power allocation for underlay device-to-device communication over multiple channels," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 3, pp. 467–480, 2017.

[16] "Cisco visual networking index: Forecast and trends, 2017-2022," [Online]. Available: https://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html.

[17] T. L. Marzetta, "Massive MIMO: an introduction," *Bell Labs Tech. J.*, vol. 20, pp. 11–22, 2015.

[18] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Processing*, vol. 54, pp. 2239–2251, Jun. 2006.

[19] S. Wu, W.-K. Ma, and A.-C. So, "Physical-layer multicasting by stochastic transmit beamforming and Alamouti space-time codings," *IEEE Trans. Signal Processing*, vol. 61, pp. 4230–4245, Sept 2013.

[20] A. Abdelkader, A. Gershman, and N. Sidiropoulos, "Multiple-antenna multicasting using channel orthogonalization and local refinement," *IEEE Trans. Signal Processing*, vol. 58, pp. 3922–3927, Jul. 2010.

[21] E. Karipidis, N. Sidiropoulos, and Z.-Q. Luo, "Quality of Service and Max-Min Fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Processing*, vol. 56, pp. 1268–1279, 2008.

[22] A. Abdelkader, A. B. Gershman, and N. D. Sidiropoulos, "Multiple-antenna multicasting using channel orthogonalization and local refinement," *IEEE Trans. Signal Processing*, vol. 58, pp. 3922–3927, Jul. 2010.

[23] N. Bornhorst, M. Pesavento, and A. Gershman, "Distributed beamforming for multi-group multicasting relay networks," *IEEE Trans. Signal Processing*, vol. 60, pp. 221–232, Jan. 2012.

[24] G. Venkatraman, A. Tölli, M. Juntti, and L. Tran, "Multigroup multicast beamformer design for MISO-OFDM with antenna selection," *IEEE Trans. Signal Processing*, vol. 65, no. 22, pp. 5832–5847, Nov 2017.

[25] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Weighted fair multicast multigroup beamforming under per-antenna power constraints," *IEEE Trans. Signal Processing*, vol. 62, no. 19, pp. 5132–5142, Oct. 2014.

[26] A. Schad and M. Pesavento, "Max-min fair transmit beamforming for multi-group multicasting," in *2012 International ITG Workshop on Smart Antennas (WSA)*. IEEE, 2012, pp. 115–118.

[27] M. Sadeghi, L. Sanguinetti, R. Couillet, and C. Yuen, "Reducing the computational complexity of multicasting in large-scale antenna systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2963–2975, May 2017.

[28] E. Karipidis, N. D. Sidiropoulos, and Z. Luo, "Far-field multicast beamforming for uniform linear antenna arrays," *IEEE Trans. Signal Processing*, vol. 55, no. 10, pp. 4916–4927, Oct 2007.

[29] O. Tervo, L.-N. Tran, H. Pennanen, S. Chatzinotas, B. Ottersten, and M. Juntti, "Energy-efficient multicell multigroup multicasting with joint beamforming and antenna selection," *IEEE Transactions on Signal Processing*, vol. 66, no. 18, pp. 4904–4919, 2018.

[30] M. A. Hanif and M. H. Lee, "An efficient shared relay beamforming scheme for multi-cell multi-user LTE-A cooperative networks," in *2013 13th International Symposium on Communications and Information Technologies (ISCIT)*, 2013, pp. 57–59.

[31] Z. Xiang, M. Tao, and X. Wang, "Coordinated multicast beamforming in multi-cell networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 12–21, Jan. 2013.

[32] Z. q. Luo, W. k. Ma, A. M. c. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Mag.*, vol. 27, pp. 20–34, May 2010.

[33] L. Tran, M. F. Hanif, and M. Juntti, "A conic quadratic programming approach to physical layer multicasting for large-scale antenna arrays," *IEEE Signal Processing Lett.*, vol. 21, no. 1, pp. 114–117, Jan 2014.

[34] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup beamforming for per-antenna power constrained large-scale arrays," in *Proc. IEEE Workshop on Signal Processing advances in Wireless Commun.(SPAWC)*, Jun. 2015, pp. 271–275.

[35] T. Biermann, L. Scalia, C. Choi, W. Kellerer, and H. Karl, "How backhaul networks influence the feasibility of coordinated multipoint in cellular networks [accepted from open call]," *IEEE Communications Magazine*, vol. 51, no. 8, pp. 168–176, 2013.

[36] J. Wang, B. Liang, M. Dong, and G. Boudreau, "Online multicell coordinated MIMO wireless network virtualization with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 10 455–10 471, 2022.

[37] J. Wang, M. Dong, B. Liang, G. Boudreau, and H. Abou-zeid, "Distributed coordinated precoding for MIMO cellular network virtualization," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 106–120, 2021.

[38] J. Yu and M. Dong, "Low-complexity weighted MRT multicast beamforming in massive MIMO cellular networks," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 3849–3853.

[39] ——, "Distributed low-complexity multi-cell coordinated multicast beamforming with large-scale antennas," in *Proc. IEEE Int. Workshop on Signal Processing advances in Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.

[40] M. Dong and Q. Wang, "Multi-group multicast beamforming: Optimal structure and efficient algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3738–3753, 2020.

[41] E. Chen and M. Tao, "ADMM-based fast algorithm for multi-group multicast beamforming in large-scale wireless systems," *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2685–2698, 2017.

[42] C. Zhang, M. Dong, and B. Liang, "Fast first-order algorithm for large-scale max-min fair multi-group multicast beamforming," *IEEE Wireless Communications Letters*, vol. 11, no. 8, pp. 1560–1564, 2022.

[43] F. Rashid-Farrokhi, K. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, 1998.

[44] M. Schubert and H. Boche, "An efficient algorithm for optimum joint downlink beamforming and power control," in *Vehicular Technology Conference. IEEE*

55th Vehicular Technology Conference. VTC Spring 2002 (Cat. No.02CH37367), vol. 4, 2002, pp. 1911–1915 vol.4.

[45] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," *Handbook of Antennas in Wireless Communications*, 01 2001.

[46] Z.-Q. Luo, N. Sidiropoulos, P. Tseng, and S. Zhang, "Approximation bounds for quadratic optimization with homogeneous quadratic constraints," *SIAM J. Optim.*, vol. 18, pp. 1–28, 2007.

[47] T.-H. Chang, Z.-Q. Luo, and C.-Y. Chi, "Approximation bounds for semidefinite relaxation of max-min-fair multicast transmit beamforming problem," *IEEE Trans. Signal Processing*, vol. 56, pp. 3932–3943, 2008.

[48] L.-N. Tran, M. F. Hanif, and M. Juntti, "A conic quadratic programming approach to physical layer multicasting for large-scale antenna arrays," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 114–117, 2014.

[49] Yan Gao and M. Schubert, "Group-oriented beamforming for multi-stream multicasting based on quality-of-service requirements," in *Proc. IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec 2005, pp. 193–196.

[50] M. Jordan, X. Gong, and G. Ascheid, "Multicell multicast beamforming with delayed SNR feedback," in *Proc. IEEE Global Telecommn. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–6.

[51] M. Tao, E. Chen, H. Zhou, and W. Yu, "Content-centric sparse multicast beamforming for cache-enabled cloud RAN," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6118–6131, Sep. 2016.

[52] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[53] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[54] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, March 1998.

[55] D. Gesbert, M. Kountouris, R. W. Heath, C. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Processing Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.

[56] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[57] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug 2003.

[58] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[59] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, Oct 2006, pp. 359–363.

[60] R. Jiang, H. Liu, and A. M. So, "LPA-SD: An efficient first-order method for single-group multicast beamforming," in *Proc. IEEE Int. Workshop on Signal Processing advances in Wireless Commun. (SPAWC)*, June 2018, pp. 1–5.

[61] A. Konar and N. D. Sidiropoulos, "A fast approximation algorithm for single-group multicast beamforming with large antenna arrays," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2016, pp. 1–5.

[62] M. Sadeghi, E. Björnson, E. G. Larsson, C. Yuen, and T. L. Marzetta, "Max-min fair transmit precoding for multi-group multicasting in massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1358–1373, Feb 2018.

[63] E. Chen and M. Tao, "ADMM-based fast algorithm for multi-group multicast beamforming in large-scale wireless systems," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2685–2698, Jun. 2017.

[64] C. Zhang, M. Dong, and B. Liang, "Ultra-low-complexity algorithms with structurally optimal multi-group multicast beamforming in large-scale systems," *IEEE Transactions on Signal Processing*, vol. 71, pp. 1626–1641, 2023.

[65] T. X. Tran and G. Yue, "Grab: Joint adaptive grouping and beamforming for multi-group multicast with massive MIMO," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[66] B.-X. Wu, K. C.-J. Lin, K.-C. Hsu, and H.-Y. Wei, "Hybridcast: Joint multicast-unicast design for multiuser MIMO networks," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 1724–1732.

[67] X. Zhang, S. Sun, F. Qi, R. Bo, R. Q. Hu, and Y. Qian, "Massive MIMO based hybrid unicast/multicast services for 5G," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6.

[68] S. Mohammadi, M. Dong, and S. ShahbazPanahi, "Fast algorithm for joint unicast and multicast beamforming for large-scale massive MIMO," *IEEE Transactions on Signal Processing*, vol. 70, pp. 5413–5428, 2022.

[69] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative multicell networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2014, pp. 4777–4782.

[70] M. A. Mosleh, S. M. Shahabi, M. Ghasimi, and M. Ardebilipour, "A low-complexity AoA-driven multi-cell constant envelope precoding for massive MIMO

systems," in *2020 28th Iranian Conference on Electrical Engineering (ICEE)*, 2020, pp. 1–4.

[71] O. Tervo, L.-N. Trant, S. Chatzinotas, B. Ottersten, and M. Juntti, "Multigroup multicast beamforming and antenna selection with rate-splitting in multicell systems," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018, pp. 1–5.

[72] H. Joudeh and B. Clerckx, "Rate-splitting for max-min fair multigroup multicast beamforming in overloaded systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7276–7289, 2017.

[73] A. Z. Yalcin and M. Yuksel, "Precoder design for multi-group multicasting with a common message," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7302–7315, 2019.

[74] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, 2011.

[75] N. Mohamadi, M. Dong, and S. ShahbazPanahi, "Low-complexity ADMM-based algorithm for robust multi-group multicast beamforming in large-scale systems," *IEEE Transactions on Signal Processing*, vol. 70, pp. 2046–2061, 2022.

[76] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, March 2004.

[77] S. X. Wu, W. Ma, and A. M. So, "Physical-layer multicasting by stochastic transmit beamforming and alamouti space-time coding," *IEEE Trans. Signal Processing*, vol. 61, no. 17, pp. 4230–4245, Sep. 2013.

[78] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.