Efficient Design of Reconfigurable Intelligent Surface Assisted Multi-group Multicast Beamforming

by

Mohammad Ebrahimi

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 Electrical and Computer Engineering

The Faculty of Engineering and Applied Science

Department of Electrical and Computer Engineering

University of Ontario Institute of Technology(Ontario Tech University) Oshawa, Ontario, Canada

September 2023

THESIS EXAMINATION INFORMATION

Submitted by: **Mohammad Ebrahimi**

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

Thesis title: Efficient Design of Multi-group Multicast Beamforming via Reconfigurable Intelligent Surface

An oral defense of this thesis took place on September 25, 2023 in front of the following examining committee:

Examining Committee:

Chair of Examining Committee :	Dr. Akramul Azim
Research Supervisor :	Dr. Min Dong
Examining Committee Member :	Dr. Shahram ShahbazPanahi
Thesis Examiner:	Dr. Ying Wang, Faculty of Engineering and Applied Science, Ontario Tech University

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

This thesis considers a multi-group multicasting scenario facilitated by a reconfigurable intelligent surface (RIS). We propose low-complexity scalable algorithms for the joint design of the base station (BS) multicast beamformer and the RIS passive beamformer to minimize the transmit power subject to the quality-of-service (QoS) constraints and maximize the minimum quality-of-service among all users mainting a power budget constraint which is also known as max-min fair (MMF) problem. By exploring the interaction of the BS and RIS beamforming in the QoS problem, we formulate two subproblems, a BS multicast QoS problem and a RIS max-min-fair multicast problem, to be solved alternatingly. Our alternating multicast beamforming approach (AMBA) not only enables us to exploit the optimal multicast beamforming structure at the BS, but also allows us to employ a fast first-order projected subgradient algorithm (PSA) to solve the RIS MMF subproblem. Furthermore, we show that QoS problem and MMF problem in our RIS-aided scenario are inverse problems, and by using the optimal beamforming structure for BS beamformer, we propose a low complexity and efficient algorithm to solve MMF problem as well. The mentioned algorithm first reformulate the MMF problem by turning the two objective parameters in MMF problem (BS and RIS beamforming vectors) into a single one and then

employs fast first-order projected subgradient algorithm here, and at the same time avoids usual alternating process which is commonly used for two-objective parameters problems. After analysing the multicast scenario, we will show that our algorithms is applicable for general case of unicast as well. Simulation results show the effectiveness of our proposed alternating approach for QoS problem and its advantage in terms of both performance and computational complexity over other alternative methods. Further simulations show the behavior of our PSA-based algorithm for MMF problem for different set-up parameters, and its effectiveness over similar proposed methods in special cases.

Keywords: multicast; reconfigurable intelligent surface (RIS); beamforming, optimal solution structure; projected subgradient algorithm (PSA)

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.

_____ Mohammad Ebrahimi

Statement of Contributions

Part of this thesis has been accepted to :

M.Ebrahimi, and M.Dong, "Efficient Design of Multi-group Multicast Beamforming via Reconfigurable Intelligent Surface," in *the 57th Asilomar Conference on Signals, Syst., Compt., Asilomar, California, USA, October 29 to November 1, 2023.*

I developed the research design and performed all the simulations and also the majority of witting the manuscript under the supervision of Prof. Min Dong. I have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others.

Acknowledgements

I wish to convey my heartfelt appreciation to Prof. Min Dong, my supervisor, for granting me the chance to engage in research and for providing invaluable encouragement, support, enthusiasm, and vast expertise.

Furthermore, I extend special gratitude to my family for their unwavering love, support, and the sacrifices they made to afford me a top-notch education and prepare me for a promising future.

Contents

List of Figures

List of Tables

List of Abbreviations

3G	3rd generation
4G	4th generation
5G	5th generation
6G	6th generation
ADMM	alternating direction method of multipliers
AO	alternating optimization
BS	base station
CSI	channel state information
DC	difference of convex
EE	energy efficiency
IoT	internet of things
IoV	internet of vehicles nulling and alignment
M2M	machine to machine
MMF	max-min fair

NP-hard	non-deterministic polynomial-time hard
OFDM	orthogonal frequency division multiplexing
OLP	optimal linear precoder
PSA	projected subgradient algorithm
QCQP	quadratically constrained quadratic program
QoS	quality of service
RIS	reconfigurable intelligent surface
SCA	successive convex approximation
SDR	semi-definite relaxation
SINR	signal-to-interference-plus-noise ratio

Chapter 1

Introduction

1.1 Overview

Over the past few decades, mobile communication technology has undergone remarkable evolution. The introduction of the fourth generation (4G) wireless technology was a pivotal moment in mobile communication. Rolled out in the late 2000s, 4G wireless networks offered significant advancements over its predecessor, 3G networks. It provided faster data speed, improved network reliability, and the capability to support data-intensive activities such as video streaming and mobile gaming. This innovation reshaped the mobile landscape, fostering the growth of smartphones and app ecosystems, fundamentally altering how people consume multimedia and interact with technology on the move. Furthermore, it made possible the invention of new technologies such as Internet of Things (IoT), Internet of Vehicles (IoV), and machine-to-machine (M2M) communications.

The increasing demand for higher data rate, lower latency, and greater network capacity to support a growing array of data-intensive applications and devices, posed significant challenges for 4G networks, pushing the systems to their limits in handling the soaring demand placed upon them. To address these challenges and meet the re-

quirements of the modern era, the 5th generation (5G) networks have been developed. The 5G networks offer ultra-high data rates, extensive coverage, and seamless connectivity while ensuring ultra-low latency and energy consumption [1–10]. In particular, massive Multiple-Input Multiple-Output (MIMO) is the underpinning technology for 5G networks to cope with the ever-increasing growth in wireless traffic and high data rate service demands. Massive MIMO utilizes a large number of antennas at the base station (BS) to form transmission beams for multi-user data transmission to improve data throughput and increase network capacity. This technology can significantly enhance spectral efficiency and enable better spatial multiplexing of data to multiple users simultaneously. Massive MIMO-based transmission design holds both promise and challenges. Optimizing arrays of numerous antennas requires balancing power allocation, interference management, and efficient communication channels. Yet, the computational complexity of processing abundant data from these arrays can lead to latency and resource concerns. In response, numerous research efforts have been devoted to developing advanced algorithms that navigate these complexities, aiming to harness the potential of massive MIMO systems effectively [11–13].

1.1.1 Multicast Beamforming

With the rise of data-heavy tasks such as social medias with millions of users, and the need of sending the same information to groups of users at once, multicasting has become necessary. Instead of sending the same data separately to each user, multicasting transmits it to a group of users all at once, saving radio resources and reducing network congestion. This is crucial for tasks like live video streaming, distributing software

updates, and efficiently sharing data with multiple devices. Multi-antenna multicast beamforming is a physical layer transmission technique that can efficiently deliver common messages to multiple users simultaneously, enabling rapid content distribution in wireless systems. Multicast beamforming design has been investigated for a single-group scenario [14], multiple-group scenario [15–17], multi-cell networks [18,19], relay networks [20–23], cognitive radio networks [24–27] and cloud-radio networks [28]. Unlike unicast beamforming that uses dedicated beam for transmitting private data to each user [29], tackling multi-group multicast beamforming problems is challenging since these problems fall into the category of non-deterministic polynomial-time hard (NP-hard) problems. Different optimization algorithms or signal processing techniques have been developed to provide design solutions for various transmission scenarios or systems in the literature [14, 30–33]. Most of these works have utilized numerical algorithms to find approximate solutions with good performance.

1.1.2 Reconfigurable Intelligent Surface

With new applications such as real-time holographic communication or Augmented Reality (AR), the quest for ultra-high data rates, remarkably low latency, and the efficient management of a sprawling network of interconnected devices has pushed the boundaries of the 5G technologies. One important technology that has emerged recently to aid wireless networks to evolve them into future beyond 5G (B5G) such as the sixth-generation (6G) is reconfigurable intelligent surface (RIS). RIS uses a planar surface consisting of passive reflective elements to control the phase shifts of the reflected wireless signal towards the desired direction. Furthermore, regarding to its

passive reflective nature, it consumes a neglecting amount of power [34]. By locating RIS between the BS and users and forming passive beamforming via the RIS path, RIS can actively control and improve the wireless propagation channel conditions, and thus creates a smart reconfigurable wireless environment to enhance communications performance [35,36]. RIS-assisted designs have been explored in a wide range of wireless applications for performance enhancement in the recent literature, including coverage, spectral efficiency, and power consumption [37–40]. In densely populated cities where sky-scraper and other city structures block the direct line-of-sight signals from the transmitter that results in weak channel conditions, RIS provides promising opportunities to create a RIS path to strengthen wireless channels as well as mitigate interference among users. As wireless multicasting becomes increasingly popular to support content distribution and delivery in future wireless services and applications, it is important to investigate how to effectively utilize RIS to enhance the wireless prorogation environment to improved multicast performance. Although adding RIS to wireless networks add a significant challenge to solve beamforming problems, it will come with numerous benefits. The integration of reconfigurable intelligent surface technology into both multicast and unicast communication scenarios offers a multitude of benefits that significantly enhance the efficiency and performance of wireless networks. In multicasting, RIS aids in directing signals towards multiple recipient groups simultaneously, effectively mitigating interference and enabling tailored signal shaping for each group. This not only improves the overall spectral efficiency but also enhances coverage and signal strength, resulting in more reliable multicast transmissions. Similarly, in unicast scenarios, RIS technology optimizes signal paths by

dynamically adjusting phase and amplitude, resulting in precise beamforming that enhances signal quality, reduces interference, and extends coverage. These benefits collectively translate into increased network capacity, improved quality of service, and seamless connectivity, making RIS-aided multicasting and unicast communication pivotal in addressing the growing demands of modern wireless networks.

1.2 Motivation and Objective

In practical downlink multicast scenarios, multicast beamforming performance typically is limited by the weakest user in a group with the worst channel condition. This limitation by the wireless prorogation environment cannot be improved by multicast beamforming design optimization. To break this barrier, RIS offers a promising solution by intelligently manipulating the wireless propagation environment. In particular, RIS can form passive beamforming by controlling the phase shifts of the reflective elements, which effectively controls the signal paths and and improves the channel conditions among users to enhance the multicast beamforming performance. With this potential, it is important to investigate RIS-assisted multicast beamforming design to enhance the wireless multicast performance. This enables the optimization of resource allocation to achieve both fairness and enhanced system performance simultaneously. The ability of RIS to adaptively modify the wireless environment has the potential to revolutionize wireless communication systems, enabling equitable and efficient utilization of network resources. The promising enhancement of the RIS come with numerous challenges. As we discussed earlier, transmit multicast beamforming is already an NP-hard optimization problem. Adding RIS, which typically contains

a large number of reflective elements, into the system will increase the overall design difficulty in term of joint beamforming design and finding a computationally efficient solution.

Motivated by the above, in this thesis, we consider the joint design of active beamforming at the BS and passive beamforming at the RIS in a general multi-group multicasting scenario. We consider two commonly considered problem formulations for the joint beamforming design: the quality of service (QoS) problem, which is to minimize transmission power while meeting the QoS target requirements at users; the max-min fair (MMF) problem, which is to maximize the minimum QoS measure among users subject to the BS transmit power constraint. Our goal is to design high-performing algoirthms that are also computationally efficient.

1.3 Thesis Contribution

This thesis aims to provide an efficient mulitcast beamforming design for an RIS-assisted downlink multi-group multicast scenario. In this thesis, both the QoS and MMF problems are considered for the joint beamforming design. That is to minimize the transmit power subject to the signal-to-interference-and-noise ratio (SINR) requirements at users, and to maximize the minimum SINR among all users subject to the BS transmit power budget constraint. The primary contributions of this thesis can be summarized as follows:

- We propose a fast alternating multicast beamforming approach for the QoS problem. We explore the structure of the QoS problem in terms of the interaction of the BS multicast and RIS passive beamforming to transform it into

a favorable format. The resulting joint optimization problem can be naturally broken into two subproblems to be solved using alternating optimization technique. The two subproblems are essentially two types of multicast beamforming problems: the BS multicast beamforming QoS subproblem, and the RIS passive multicast beamforming MMF subproblem. We can judiciously exploit the optimal multicast beamforming structure to efficiently solve the BS multicast subproblem. Furthermore, for the RIS beamforming subproblem, we propose to adopt a first-order projected subgradient algorithm (PSA) [41] to solve it, which has simple closed-form updates and convergence guarantee.

- For the MMF problem, we first show the inverse relationship between the QoS and MMF problems for the RIS-aided multicast beamforming. Then, we use the optimal BS beamforming structure again to transform the MMF problem into an equivalent form, for which employs PSA to solve it with low computational complexity. In particular, we propose a PSA-based iterative algorithm to acquire a sub-optimal solution for both BS and RIS beamforming vectors using closed-form expression updates and convergence guarantee.

- We show that our overall approach for multi-group multicasting scenario is applicable to the general downlink multiuser unicast beamforming design as well.

- Finally, simulation results demonstrate the effectiveness of our proposed algorithms for both QoS and MMF problems in both performance and computational complexity over other alternative methods.

The research work from this thesis has resulted in the following accepted or in preparation publications:

- M.Ebrahimi,and M.Dong, "Efficient Design of Multi-group Multicast Beamforming via Reconfigurable Intelligent Surface," in *the 57th Asilomar Conference on Signals, Syst., Compt., Asilomar, California, USA, October 29 to November 1, 2023.*

- M.Ebrahimi, and M.Dong, "Efficient Design of Reconfigurable Intelligent Surface Assisted Downlink Multiuser Beamforming," in preparation to be submitted to IEEE journals.

1.4 Thesis Organization

The subsequent sections of this thesis are structured in the following manner. In Chapter 2, literature review on RIS-aided beamforming techniques, and benefits and challenges of this newly developed technology is presented. In Chapter 3, we propose our alternating multicast beamforming algorithm for the RIS-aided multicast QoS problem, and reformulate the QoS problem into the desired shape. Furthermore, we propose our fast algorithm to solve QoS beamforming problem, and generalize it to unicast scenario. In Chapter 4, we disscuss the connection between the QoS and MMF problems and propose a first-order fast algorithm to solve it. Chapter 5 provides simulation study for our proposed algorithms for the QoS and MMF problems, and Chapter 6 presents the conclusion and future work.

1.5 Notations

The main notations used in this thesis are summarized below. The symbols used to represent Hermitian, transpose, and conjugate are $(.)^H$, $(.)^T$, and $(.)^*$, respectively. The Euclidean norm of a vector is symbolized by $\|.\|$. The symbol $\mathbf{a} \succcurlyeq 0$ represents element-wise nonnegative. When referring to a matrix, $\mathbf{A} \succcurlyeq 0$ indicates that matrix $\mathbf{A}$ is positive semidefinite. The trace of matrix $\mathbf{A}$ is denoted as $\mathrm{tr}(\mathbf{A})$. The real part of $\mathbf{x}$ is represented as $\mathfrak{Re}\{\mathbf{x}\}$, and $E(x)$ denotes the expectation of x. The abbreviation i.i.d. stands for independent and identically distributed, and signifies that $\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I})$ is a complex Gaussian random vector with zero mean and covariance $\mathbf{I}$.

Chapter 2

Literature Review

2.1 Downlink Beamforming

Downlink beamforming, a critical technology in wireless communication systems, plays a pivotal role in optimizing the efficiency and quality of data transmission. By directing the transmission of signals from a BS towards individual users or devices with precision, downlink beamforming enhances signal strength, mitigates interference, and increases overall network capacity. This technology enables the customization of signal delivery, focusing on specific areas or users, thereby improving network coverage and user experience in high-density environments. Downlink beamforming also contributes to energy efficiency (EE) by minimizing wasteful broadcasting and reducing power consumption. In essence, downlink beamforming revolutionizes wireless communication by maximizing the utilization of available resources, improving spectral efficiency, and ensuring seamless connectivity, making it an indispensable tool for modern communication networks and their ability to meet the escalating demands of data-driven applications and services. In following subsections, we explain briefly about the literature works in this area.

2.1.1 BS Active Beamforming

Beamforming represents a signal processing method to transmit signals directionally, aiming to enhance signal strength at the desired recipient while minimizing signal interference with neighboring users in the vicinity. In conventional downlink communication systems, beamforming was only done at the BS, and because it consumes power for transmission, we will refer to it as active beamforming. As we mentioned earlier, beamforming can be used for both unicast and multicast wireless systems.

2.1.2 Unicast Beamforming

In unicast beamforming the BS will send a distinct message to each indivdual user. This technique is one of the most common one in the field of wireless data transmission. As we mentioned in the previous chapter, unicast beamforming has been studied extensively in the past [29, 42–46]. In the context of downlink multi-user unicast beamforming, an uncomplicated closed-form solution for the problem of minimizing power has been derived in [44]. This solution exhibits a straightforward and easily understandable configuration, requiring just a single user-specific design parameter.

2.1.3 Multicast Beamforming

In multicasting, the BS sends a common message to each group of users. Multicasting brings about various advantages and challenges. One of the primary benefits is its ability to efficiently deliver data from a single source to multiple receivers, reducing the need for redundant transmissions and conserving valuable bandwidth. Multicasting can also enhance network scalability by reducing the overall load on the network

infrastructure. However, wireless multicasting design faces many challenges. The issue of ensuring reliable and synchronized delivery of data to all intended recipients in a wireless environment can be complex, as interference, fading, and varying signal strengths can impact reception quality. Numerous works has been done to find the best approach for multicast beamforming for both the QoS and MMF problems. The first work on multicast beamforming design considered a system with just a single group [14]. The design approach was later extended to multi-group multicasting [15–17]. The earlier literature works adopted semi-definite relaxation (SDR) approach [14–17, 30], but SDR has a high computational complexity, especially when used in large-scale problems where BS is equipped with a large number of antenna, and as an approximate method, its performance deteriorates as the problem size becomes large. and inaccurate. With the number of antennas increasing, the successive convex approximation (SCA) technique has become more attractive for its lower computational complexity and better performance [32, 33]. In order to face the challenges posed by computational complexity, new multicast beamforming technique that reduced complexity has been introduced for large-scale MIMO systems which designed to serve multiple groups [47], and the mentioned problem has been further studied in [48] using alternating direction method of multipliers (ADMM). Later in [49, 50] the multi-group secnario has extended to multi-cell, and low-complexity beamforming techniniqes have been introduced for large-scale problems.

The optimal multi-group multicast beamforming structure was recently obtained in [51]. This structure has been exploited to develop fast and scalable algorithms for large-scale multicast beamforming problems in massive MIMO systems [41, 52].

Furthermore, by using the inverse relationship between QoS and MMF problems [53], it is also shown in [51] that a solution to MMF problem can also be obtained using the optimal beamforming structure. By utilizing the optimal structure, a fast first order algorithm based on the projected subgradient algorithm (PSA) was proposed in [41] to find a sub-optimal solution to MMF problem. Finally, by utilizing the optimal beamforming structure, in [54] a combination of unicast and multicast scenario has been analyzed and a fast closed-form based iterative algorithm introduced using ADMM.

2.2 RIS-Aided Downlink Beamforming

By introducing RIS into wireless communication systems, It is shown that a joint beamforming operation can be done at both the BS and RIS. This opens a path to new opportunities to enhance quality of data transmission in unicast and multicast wireless systems. RIS-aided downlink beamforming stands as a transformative advancement in wireless communication systems to increase the capacity of wireless network for data transmission. By intelligently manipulating the phase and amplitude of incoming signal, RIS technology enables precise control over the direction of the transmitted message. This capability holds paramount importance in enhancing the performance of downlink communication. Like the conventional downlink systems, here we can categorize the beamforming problems into two unicast, and multicast as well.

2.2.1 RIS-Aided Unicast Beamforming

Unicast beamforming in RIS aided downlink systems has been analyzed in several works. The idea of joint active and passive beamforming was first introduced in [35, 55]. In this terminology, beamforming at RIS is referred as passive since in most RIS aided systems, RIS elements only change the phase of the received signal, and considered as passive elements, but in [56], an special reflective surface is used for RIS, which has the ability of damping the power of received signal as well. Analysis in [35, 55] coveres unicast beamforming for single user and multi-users system by considering QoS problem. The algorithm presented in [35] is based on alternating optimization (AO) between power minimization at base station and aligning the phase of direct path and RIS-created indirect path between BS and users, at RIS. This idea comes from this fact that in alternating optimization algorithm, by considering a constant value for BS active beamformer, the QoS problem will reduce to feasibility problem and stops searching after the first round. In [55] the beamforming at RIS only consider the feasibility problem. Both [35, 55] consider SDR as the main numerical method to solve QoS problem in their algorithms. In [57] the MMF problem for a multi-user unicast problem has been analysed using Optimal Linear Precoder (OLP) method, and [58] uses alternating descend round to maximize the users' worst rate which is equivalent to the MMF problem. [58] also study the mentiond problem in both cases of proper and improper Gaussian signalling. Furthermore, in [37, 38], solutions to sum-rate maximization and energy efficiency maximization problems have been presented using different numerical methods in different situation where beamformer

is aware and not aware of perfect channel state information (CSI).

2.2.2 RIS-Aided Multicast Beamforming

Adding RIS to a multicasting wireless communication system, will add more challenges to the mentioned problems, because usually RIS is considered as a passive elements which only changes the phase of received signal. Although this feature makes the design of physical layer easier, it adds a unit modulus constraint for RIS elements to both the QoS and MMF problems which is highly non-linear and non-convex.

RIS-assisted multicast beamforming design has recently been investigated for the single-group [59,60] and multi-group settings [61–63]. For the first group of problems, it is shown that using RIS for the purpose of phase-alignment between the direct path from BS to users and indirect path through the RIS is a favorable approach, and for the latter that contains inter-group interference, joint BS and RIS beamforming optimization for maximizing sum group rate was considered in [61, 62]. Specifically, with alternating optimization (AO), [61] proposed bounding and smoothing techniques, along with a majorization minimization method. In [62], the smoothing technique was applied directly to the rate objective, along with a low-complexity alternating projected gradient algorithm to improve performance and reduce computational complexity. In [63], for an RIS-assisted symbiotic radio multicast system that includes the primary receivers and an internet-of-things receiver, transmit power minimization subject to the quality-of-service constraints, *i.e.,* the QoS problem, was considered, where using AO with SDR was proposed. However, since SDR incurs high computational complexity for a large number of RIS elements and BS antennas, the method is

impractical for large-scale systems. Furthermore, in [64], problem QoS is approached using AO and SDR with the help of two layers of iterative algorithms and difference-of-convex (DC) function to force SDR method to converge to a rank 1 solution and reduce the inaccuracy of this method, the efficiency of this algorithm is not certain for one of the iterative layers of it reduces the QoS problem into a feasibility problem, and algorithm will stop searching after the first round. With these limited results, there is a need to develop effective RIS-aided multicast beamforming algorithms.

Multicast beamforming optimization is inherently an NP-hard problem, which makes finding the optimal solution to this problem almost impossible. As we mentioned earlier, both the QoS and MMF problems are non-convex NP-hard problems, and the existing numerical algorithms mainly try to find a suboptimal solutions to these problems [59–64]. One main challenge is the computational efficiency of proposed algorithms for RIS with large number of elements and in large-scale MIMO systems. Most of the proposed most of numerical methods such as SDR are extremely slow and also inaccurate for large-scale systems. Later works which have used SCA to tackle the non-convexity of QoS and MMF problems, still suffers relatively high computational complexity.

A main challenge of solving the QoS problem for a RIS-aided system is how to effectively use the alternating optimization, the problem will reduce to feasibility problem by considering a constant value for BS active beamforming vector. Due to this, most current works [cite] in this area to focus on the MMF problem (or maximizing the sum-rate problem), instead of the QoS problem, in both unicasting and multicasting.

Chapter 3

RIS-aided Multicast Beamforming Design

In this chapter, we consider the QoS problem for the RIS-aided multigroup multicast beamforming scenario. Utilizing the optimal beamforming structure at BS and using fast first-order projected subgradient algorithm (PSA) to solve the RIS MMF subproblem, we propose a low-complexity scalable algorithm for the joint design of the BS multicast beamformer and the RIS passive beamformer to minimize the transmit power subject to the quality-of-service. Furhtermore, different aspects of the proposed algorithm will be disscussed such as complexity analysis, and we will also discuss that our proposed algorithm is applicable for the general case of unicasting scenario as well.

3.1 System Model

We consider a RIS-assisted downlink multi-group multicast scenario, as shown in Fig.3.1. a base station (BS) equipped with N antennas multicasting messages to G user groups, and a RIS consisting of M passive reflective elements is deployed to assist the data transmission between the BS and the users. We assume that each

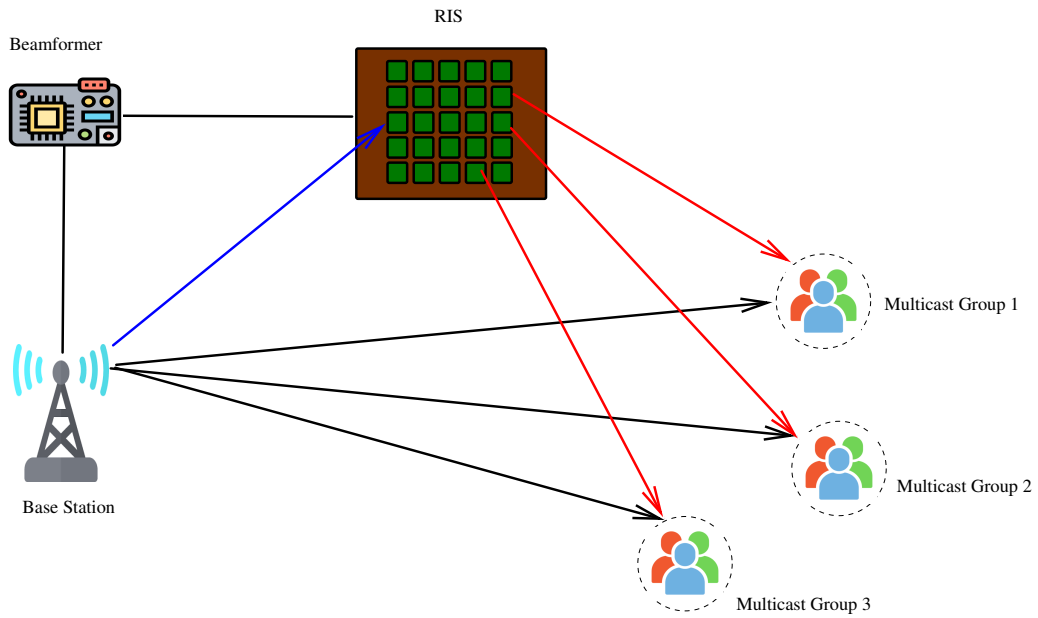

Figure 3.1: An example of RIS-assisted downlink multi-group multicast scenario.

group consists of K single-antenna users[1], and the users from the same group receive a common message that is independent to other groups. The BS controls the phases of the RIS array elements via a RIS controller to adjust the incident signals to the desired directions. We denote $\mathcal{G} \triangleq \{1, \ldots, G\}$ and $\mathcal{K} \triangleq \{1, \ldots, K\}$ as the set of all groups and the set of users in each group, respectively, and denote $\mathcal{M} \triangleq \{1, \ldots, M\}$ as the set of RIS elements.

Let $\mathbf{H}_{\mathrm{r}}$ denote the $M \times N$ channel matrix from the BS to the RIS, $\mathbf{h}_{ik}^{\mathrm{d}}$ the $N \times 1$ channel vector from the BS to user k in group i, and $\mathbf{h}_{ik}^{\mathrm{r}}$ the $M \times 1$ channel vector from the RIS to user k in group i, for $k \in \mathcal{K}, i \in \mathcal{G}$. We use $\mathbf{w}_i$ to represent the $N \times 1$ multicast beamforming vector at the BS for group $i \in \mathcal{G}$. Also, let $\mathbf{e} \triangleq [e_1, \ldots, e_M]^T$ denote the vector containing the M reflection coefficients at the RIS, where $e_m = e^{j\theta_m}$ with $\theta_m \in (-\pi, \pi]$ being the phase shift of element m. We refer to $\mathbf{e}$ as the RIS passive

[1]We consider the same size for all groups to simplify the notation. Generalization to individual group size K_i for group i is straightforward.

beamforming vector. It is also considered that the channel state information (CSI) is perfectly known at BS. The signal received at user k in group i is given by

$$y_{ik} = \sum_{j=1}^{G} \mathbf{w}_j^H \Big(\mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{H}_{\mathrm{r}}^H \mathrm{diag}(\mathbf{e}) \mathbf{h}_{ik}^{\mathrm{r}}\Big) s_j + n_{ik} \tag{3.1}$$

$$= \mathbf{w}_i^H \Big(\mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{G}_{ik}\mathbf{e}\Big) s_i + \sum_{j \neq i} \mathbf{w}_j^H \Big(\mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{G}_{ik}\mathbf{e}\Big) s_j + n_{ik} \tag{3.2}$$

where s_i is the symbol intended for group i with $\mathbb{E}[|s_i|^2] = 1$, $n_{ik} \sim \mathcal{CN}(0, \sigma^2)$ is the receiver additive white Gaussian noise with variance σ^2, and $\mathbf{G}_{ik} \triangleq \mathbf{H}_{\mathrm{r}}^H \mathrm{diag}(\mathbf{h}_{ik}^{\mathrm{r}})$ is the $N \times M$ cascaded channel from the BS to user k in group i via the RIS.

The received SINR at user k in group i is then given by

$$\mathrm{SINR}_{ik} = \frac{|\mathbf{w}_i^H \Big(\mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{G}_{ik}\mathbf{e}\Big)|^2}{\sum_{j \neq i} |\mathbf{w}_j^H \Big(\mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{G}_{ik}\mathbf{e}\Big)|^2 + \sigma^2}. \tag{3.3}$$

The total transmit power at the BS is $P_{\mathrm{tot}} = \sum_{i=1}^{G} \|\mathbf{w}_i\|^2$.

3.2 Problem Formulation

In this chapter, we focus on the QoS problem for the multicast beamforming design in this RIS-assisted transmission scenario. Specifically, our objective is to jointly design the multicast beamforming vectors $\mathbf{w}_i$ and RIS reflection coefficient vector $\mathbf{e}$ to minimize the BS transmit power while ensuring that the SINR target at each user is met. This joint beamforming optimization problem is formulated as

$$\mathcal{P}_o: \min_{\mathbf{w},\mathbf{e}} \quad \sum_{i=1}^{G} \|\mathbf{w}_i\|^2 \tag{3.4a}$$

$$\text{s.t.} \quad \mathrm{SINR}_{ik} \geq \gamma_{ik}, \ k \in \mathcal{K}, \ i \in \mathcal{G} \tag{3.4b}$$

$$|e_m|^2 = 1, \ m \in \mathcal{M} \tag{3.4c}$$

where $\mathbf{w} \triangleq [\mathbf{w}_1^H, \ldots, \mathbf{w}_G^H]^H$, and γ_{ik} is the SINR target for user k in group i.

Note that even in the conventional downlink multicast scenario without RIS, the QoS problem for multi-group multicast beamforming is well-known to be a non-convex quadratically constrained quadratic programming (QCQP) problem which is NP-hard. The joint optimization problem $\mathcal{P}_o$ is even more challenging with the additional passive beamformer $\mathbf{e}$ from RIS involved in the SINR_{ik} expression in (3.3) and the non-convex constraints in (3.4c). As a result, obtaining the optimal solution for $\mathcal{P}_o$ is difficult, and we focus on providing a good suboptimal solution. Furthermore, since M is expected to be large, $\mathcal{P}_o$ is a large-scale problem. Thus, our goal is to develop an effective computational method with high performance but low computational complexity.

3.3 Beamforming Design

To solve $\mathcal{P}_o$, we consider the alternating optimization (AO) technique, *i.e.,* to optimize $\mathcal{P}_o$ with respect to (w.r.t.) the BS multicast beamformer $\mathbf{w}$ and RIS passive beamformer $\mathbf{e}$ alternatingly. However, directly applying AO to $\mathcal{P}_o$ is not an effective method. This is because that the objective function in $\mathcal{P}_o$ is only a function of $\mathbf{w}$. For AO, with given $\mathbf{w}$, $\mathcal{P}_o$ w.r.t. $\mathbf{e}$ will reduce to a feasibility problem, for which the value of $\mathbf{e}$ from the previous iteration update is already a feasible solution. Thus, the algorithm will stuck at the value of $(\mathbf{w}, \mathbf{e})$ after the first round of updates, and the quality of this solution merely depends on the quality of the initial point, without any quality guarantee.

To address the above issue, we first analyze the relation between $\mathbf{w}$ and $\mathbf{e}$ to understand how they interact in $\mathcal{P}_o$ and then utilize it to devise an effective algorithm to solve the QoS problem.

3.3.1 QoS Problem Reformulation

To explore the structural relation between $\mathbf{w}$ and $\mathbf{e}$ in $\mathcal{P}_o$, we can equivalent the set of SINR constraints in (3.4b) as the follow constraint:

$$\begin{aligned} &\mathrm{SINR}_{ik} \geq \gamma_{ik},\ \forall k \in \mathcal{K}, i \in \mathcal{G} \\ \Leftrightarrow\ &\frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \geq 1,\ \forall k \in \mathcal{K}, i \in \mathcal{G} \\ \Leftrightarrow\ &\min_{i,k} \frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \geq 1 \end{aligned} \tag{3.5}$$

Following the above, we can replace (3.4b) with (3.5) and equivalently rewrite $\mathcal{P}_o$ as

$$\mathcal{P}_1:\ \min_{\mathbf{w},\mathbf{e}}\ \sum_{i=1}^{G} \|\mathbf{w}_i\|^2 \tag{3.6a}$$

$$\text{s.t.}\ \ \min_{i,k} \frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \geq 1 \tag{3.6b}$$

$$|e_m|^2 = 1,\ m \in \mathcal{M}. \tag{3.6c}$$

Noting that the RIS reflection vector $\mathbf{e}$ only affects SINR_{ik}'s in the SINR constraints in (3.6b), we can further transform $\mathcal{P}_1$ into the following equivalent problem:

$$\mathcal{P}_2:\ \min_{\mathbf{w}}\ \sum_{i=1}^{G} \|\mathbf{w}_i\|^2 \tag{3.7a}$$

$$\text{s.t.}\ \ \max_{\mathbf{e}:|e_m|^2=1, m\in\mathcal{M}} \min_{i,k} \frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \geq 1. \tag{3.7b}$$

Comparing $\mathcal{P}_2$ with $\mathcal{P}_o$, we note that the original set of SINR constraints in (3.4b) is replaced by the constraint in (3.7b), whose left hand side is a max-min weighted SINRs optimization problem w.r.t. $\mathbf{e}$, and the unit-modulus constraints for the elements in $\mathbf{e}$ in (3.4c) are now the constraints for this max-min optimization problem.

3.3.2 Alternating Multicast Beamforming Approach

Once transforming $\mathcal{P}_o$ into $\mathcal{P}_2$, we see that $\mathcal{P}_2$ inherently contains two optimization problems w.r.t. $\mathbf{w}$ and $\mathbf{e}$, respectively. Based on this, we naturally break down $\mathcal{P}_2$ into two subproblems to solve alternatingly. In particular, as we will show below, these two subproblems are essentially two types of multicast beamforming problems: 1) BS multicast beamforming QoS problem w.r.t. $\mathbf{w}$; 2) RIS passive multicast beamforming MMF problem w.r.t. $\mathbf{e}$. Our overall proposed alternating multicast beamforming (AMBF) approach includes the alternating optimization step and a final processing step to ensure a feasible solution to $\mathcal{P}_o$. We first describe the two subproblems and the final processing step below, leaving the proposed fast algorithms for solving the subproblems to Sections 3.4.1 and 3.4.2.

3.3.3 BS multicast beamforming QoS problem for w

For given RIS passive beamforming vector $\mathbf{e}$, $\mathcal{P}_o$ is reduced to the following BS multicast beamforming problem for $\mathbf{w}$:

$$\mathcal{P}_{\mathbf{e}}: \min_{\mathbf{w}} \sum_{i=1}^{G} \|\mathbf{w}_i\|^2 \tag{3.8a}$$

$$\text{s.t. } \text{SINR}_{ik} \geq \gamma_{ik},\ k \in \mathcal{K},\ i \in \mathcal{G}. \tag{3.8b}$$

Note that $\mathcal{P}_{\mathbf{e}}$ is in the form of a typical QoS problem for downlink multi-group multicast beamforming, with the effective channel between BS and user k in group i consisted of the direct path and the RIS-path.

3.3.4 RIS multicast beamforming MMF problem for e

Optimizing the RIS passive beamformer $\mathbf{e}$ in $\mathcal{P}_o$ with given $\mathbf{w}$ is the main challenge. As discussed at the beginning, directly applying AO to the original problem $\mathcal{P}_o$ is not effective. However, once we convert $\mathcal{P}_o$ into $\mathcal{P}_2$, for given $\mathbf{w}$, we naturally have the following optimization problem w.r.t. $\mathbf{e}$, which is to maximize the minimum gap between the received SINR and the SINR target, *i.e.,* the ratio $\frac{\text{SINR}_{ik}}{\gamma_{ik}}$:

$$\mathcal{S}_{\mathbf{w}}: \max_{\mathbf{e}} \min_{i,k} \frac{\text{SINR}_{ik}}{\gamma_{ik}} \tag{3.9a}$$

$$\text{s.t.} \quad |e_m|^2 = 1, \ m \in \mathcal{M}. \tag{3.9b}$$

Remark 1. Note that $\mathcal{S}_{\mathbf{w}}$ can be viewed as a weighted MMF multicast beamforming problem, where the beamformer $\mathbf{e}$ is designed to transmitting messages to all the users in the system. However, it has three key differences from the conventional MMF multicast problem: **i)** Although we can equivalent this case to a single-group multicast beamforming, it contains intra-group self-interference. **ii)** The beamformer $\mathbf{e}$ appears in both the numerator and the denominator of the SINR expression in (3.3), where the denominator term can be viewed as the interference to all other users. Thus, from the perspective of $\mathbf{e}$, SINR_{ik} in (3.3) can in fact be interpreted as the signal-to-leakage-and-noise ratio (SLNR) for user k in group i. Thus, the MMF problem $\mathcal{S}_{\mathbf{w}}$ w.r.t. $\mathbf{e}$ is based on the SLNR metric. **iii)** Different from the conventional transmit power constraint, the constraints in (3.9b) can be viewed as a non-convex per-element power constraint, where it requires each element consumes exactly one unit power.

Due to the difference discussed above, $\mathcal{S}_{\mathbf{w}}$ is more challenging to solve than the

conventional MMF problem, which is already NP-hard. To tackle this problem, we note that the constraints in (3.9b) are equivalent to

$$|e_m|^2 \leq 1, m \in \mathcal{M}; \quad \mathbf{e}^H\mathbf{e} = M. \tag{3.10}$$

Thus, we can replace the constraints in (3.9b) with (3.10) and transform $\mathcal{S}_{\mathbf{w}}$ into the following equivalent problem:

$$\mathcal{S}_{\mathbf{w}}^{\text{eq}}: \ \max_{\mathbf{e}} \min_{i,k} \ \frac{\text{SINR}_{ik}}{\gamma_{ik}} \tag{3.11a}$$

$$\text{s.t.} \ \ |e_m|^2 \leq 1, m \in \mathcal{M}, \tag{3.11b}$$

$$\mathbf{e}^H\mathbf{e} = M. \tag{3.11c}$$

To make the problem more tractable, we relax $\mathcal{S}_{\mathbf{w}}^{\text{eq}}$ by transferring the constraint in (3.11c) into the objective function in (3.11a) as a penalty term with a penalty weight $\zeta > 0$. The relaxed problem is given by

$$\widetilde{\mathcal{S}}_{\mathbf{w}}: \ \max_{\mathbf{e}} \ \min_{i,k} \frac{\text{SINR}_{ik}}{\gamma_{ik}} + \zeta\frac{\mathbf{e}^H\mathbf{e}}{M} \tag{3.12a}$$

$$\text{s.t.} \ \ |e_m|^2 \leq 1, \ m \in \mathcal{M}. \tag{3.12b}$$

Remark 2. The penalty term in (3.12a) is to ensure that in solving $\widetilde{\mathcal{S}}_{\mathbf{w}}$, $\mathbf{e}^H\mathbf{e}$ is as close to M as possible. We choose the penalty term in this form such that the two terms in the objective function are both normalized: the SINR term $\min_{i,k} \frac{\text{SINR}_{ik}}{\gamma_{ik}} \geq 1$, and the RIS passive beamformer term $\frac{\mathbf{e}^H\mathbf{e}}{M} \leq 1$. In this way, the values of the two terms are both around 1, which are comparable to each other and are not affected by the values of SINR or $\|\mathbf{e}\|^2$. This simplifies the choice of penalty parameter ζ, which can remain the same for different values of SINR or $\|\mathbf{e}\|^2$.

Although $\widetilde{\mathcal{S}}_{\mathbf{w}}$ is still a non-convex and NP-hard problem, compared with $\mathcal{S}_{\mathbf{w}}$, it is more amenable to efficient algorithm design to compute a solution, which will be discussed later. Following the above, our AMBF approach is to solve the two multicast beamforming problems, *i.e.,* $\mathcal{P}_{\mathbf{e}}$ for the BS and $\widetilde{\mathcal{S}}_{\mathbf{w}}$ for the RIS, alternatingly.

3.3.5 Final processing

Let $(\mathbf{w}^\star, \mathbf{e}^\star)$ denote the solution after the above alternating optimization procedure. Since $\widetilde{\mathcal{S}}_{\mathbf{w}}$ is a relaxed problem of $\mathcal{S}_{\mathbf{w}}$, $(\mathbf{w}^\star, \mathbf{e}^\star)$ may not be feasible to $\mathcal{P}_o$. Thus, we have the final processing step to obtain the feasible solution $(\mathbf{w}^{\text{final}}, \mathbf{e}^{\text{final}})$ as follows:

- Project $\mathbf{e}^\star$ onto its feasible set: $\mathbf{e}^{\text{final}} = \exp\left(j\angle \mathbf{e}^\star\right)$, *i.e.,* take the phase of each element in $\mathbf{e}^\star$;
- Solve $\mathcal{P}_{\mathbf{e}^{\text{final}}}$ with $\mathbf{e}^{\text{final}}$ and obtain $\mathbf{w}^{\text{final}}$.

3.4 Fast Algorithms for the AMBF Approach

In the following sections, we present fast algorithms to solve subproblems $\mathcal{P}_{\mathbf{e}}$ and $\widetilde{\mathcal{S}}_{\mathbf{w}}$.

3.4.1 Fast Algorithm for BS Multicast Beamforming $\mathcal{P}_{\mathbf{e}}$

As mentioned earlier, $\mathcal{P}_{\mathbf{e}}$ is a typical downlink multicast beamforming QoS problem. Define $\tilde{\mathbf{h}}_{ik} \triangleq \mathbf{h}_{ik}^{\text{d}} + \mathbf{G}_{ik}\mathbf{e}$ as the $N \times 1$ equivalent channel from BS to user k in group i, which includes both the direct path and the RIS path. The received SINR at user k in group i in (3.3) can be rewritten as

$$\text{SINR}_{ik} = \frac{|\mathbf{w}_i^H \tilde{\mathbf{h}}_{ik}|^2}{\sum_{j\neq i}^{G} |\mathbf{w}_j^H \tilde{\mathbf{h}}_{ik}|^2 + \sigma^2},\ k \in \mathcal{K}, i \in \mathcal{G}. \tag{3.13}$$

Although the QoS problem $\mathcal{P}_{\mathbf{e}}$ is known to be NP-hard, the recent work in [51] has obtained the structure of the optimal multicast beamforming solution, which can be utilized to compute the beamforming solution $\mathbf{w}$ with high computational efficiency. Based on this structure, an ultra-low-complexity first-order fast algorithm has been recently developed for the QoS problem [52]. We directly employ this optimal structure to compute $\mathbf{w}$. Specifically, the optimal $\mathbf{w}_i$ is a weighted MMSE filter given by [51]

$$\mathbf{w}_i \;=\; \mathbf{R}^{-1}(\boldsymbol{\lambda})\widetilde{\mathbf{H}}_i\mathbf{a}_i\,,\; i \in \mathcal{G} \tag{3.14}$$

where $\widetilde{\mathbf{H}}_i \triangleq [\tilde{\mathbf{h}}_{i1}, \ldots, \tilde{\mathbf{h}}_{iK}]$ is the equivalent channel matrix for group i, $\mathbf{a}_i$ is the $K \times 1$ optimal weight vector containing the (complex) weight for each user channel in group i, and $\mathbf{R}(\boldsymbol{\lambda}) \triangleq \mathbf{I} + \sum_{i=1}^{G}\sum_{k=1}^{K} \lambda_{ik}^{o}\gamma_{ik}\tilde{\mathbf{h}}_{ik}\tilde{\mathbf{h}}_{ik}$ is the noise-plus-weighted-channel covariance matrix with λ_{ik}^{o} being the optimal Lagrange multipliers associated with the SINR constraints and $\boldsymbol{\lambda}$ being the vector containing all λ_{ik}^{o}'s.

The solution $\mathbf{w}_i$ in (3.14) is given in a semi-closed form with $\boldsymbol{\lambda}$ and $\{\mathbf{a}_i\}$ to be numerically determined. The value of $\boldsymbol{\lambda}$ can be efficiently computed by the fixed-point method proposed in [51]. With the optimal structure of $\mathbf{w}_i$ in (3.14), $\mathcal{P}_{\mathbf{e}}$ can be transformed into a weight optimization problem w.r.t. weight vectors $\{\mathbf{a}_i\}$, which has a smaller problem size with GK variables as compared to the original problem $\mathcal{P}_{\mathbf{e}}$ with GN variables for $N \gg 1$. We can compute $\{\mathbf{a}_i\}$ by using the SCA approach discussed in [51], and further adopt the fast ADMM algorithm proposed in [52] for computing each SCA update using closed-form expressions with significantly low complexity.

3.4.2 Fast Algorithm for RIS Multicast Beamforming $\widetilde{\mathcal{S}}_{\mathbf{w}}$

As discussed in Remark 1, we note that the RIS passive beamforming optimization problem $\widetilde{\mathcal{S}}_{\mathbf{w}}$ is a variant of the MMF problem, which is non-convex NP-hard. Since the number of RIS elements $M \gg 1$, the size of $\widetilde{\mathcal{S}}_{\mathbf{w}}$ is large. Existing methods in the literature often adopt SDR or SCA [61, 63], but have high computational complexity as M grows. It is important to develop a low-complexity effective algorithm to find a solution to $\widetilde{\mathcal{S}}_{\mathbf{w}}$, especially because it needs to be solved in each AO iteration.

PSA is a fast first-order algorithm recently proposed in [41] to solve the multi-group multicast beamforming MMF problem. Although the forms of the objective function and the constraints of $\widetilde{\mathcal{S}}_{\mathbf{w}}$ are different from those of the conventional MMF problem considered in [41], the method can still be adapted to tackle $\widetilde{\mathcal{S}}_{\mathbf{w}}$. Thus, we propose to apply PSA to $\widetilde{\mathcal{S}}_{\mathbf{w}}$ to efficiently compute a near-stationary solution of $\widetilde{\mathcal{S}}_{\mathbf{w}}$.

3.4.3 Problem Reformulation:

Denote the feasible set of $\widetilde{\mathcal{S}}_{\mathbf{w}}$ as $\mathcal{E} \triangleq \{\mathbf{e} : |e_m|^2 \leq 1,\ m \in \mathcal{M}\}$. From the objective function in (3.12a) and using the SINR expression in (3.3), we define

$$\phi_{ik}(\mathbf{e}) \triangleq -\frac{1}{\gamma_{ik}} \frac{|\mathbf{w}_i^H(\mathbf{h}_{ik}^{\mathrm{d}}+\mathbf{G}_{ik}\mathbf{e})|^2}{\sum_{j\neq i}|\mathbf{w}_j^H(\mathbf{h}_{ik}^{\mathrm{d}}+\mathbf{G}_{ik}\mathbf{e})|^2+\sigma^2} - \zeta\frac{\mathbf{e}^H\mathbf{e}}{M}, \tag{3.15}$$

for $k \in \mathcal{K}$, $i \in \mathcal{G}$. Then, $\widetilde{\mathcal{S}}_{\mathbf{w}}$ can be rewritten as

$$\widetilde{\mathcal{S}}_{\mathbf{w}} : \min_{\mathbf{e}\in\mathcal{E}} \max_{i,k} \phi_{ik}(\mathbf{e}). \tag{3.16}$$

We can further transform it into the following equivalent problem

$$\widetilde{\mathcal{S}}' : \min_{\mathbf{e}\in\mathcal{E}} \max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{e}, \mathbf{y}) \tag{3.17}$$

where $f(\mathbf{e}, \mathbf{y}) \triangleq \boldsymbol{\phi}^T(\mathbf{e})\mathbf{y}$, with $\boldsymbol{\phi}(\mathbf{e})$ being a $GK \times 1$ vector containing all $\phi_{ik}(\mathbf{e})$'s, and $\mathcal{Y} \triangleq \{\mathbf{y} : \mathbf{y} \succcurlyeq 0, \mathbf{1}^T\mathbf{y} = 1\}$ is a probability simplex. Since $\mathcal{Y}$ is a probability simplex, an optimal solution to the inner maximization problem in $\widetilde{\mathcal{S}}'$ is $\mathbf{y} = [0, \cdots, 1, \cdots, 0]^T$, with 1 at some jth position, which equivalent to the inner maximization in (3.16). Note that both $\mathcal{E}$ and $\mathcal{Y}$ are compact convex sets.

3.4.4 The Projected Subgradient Algorithm

Since $f(\mathbf{e}, \mathbf{y})$ is concave in $\mathbf{y}$ and non-convex in $\mathbf{e}$, problem $\widetilde{\mathcal{S}}'$ is a non-convex-concave min-max problem, which is NP-hard. Since $\widetilde{\mathcal{S}}'$ has the same structure as the problem considered in [41] (*i.e.,* min-max optimization with $f(\mathbf{e}, \mathbf{y}) \triangleq \boldsymbol{\phi}^T(\mathbf{e})\mathbf{y}$ and a convex feasible set), we employ PSA in [41] to find a near-stationary point of $\widetilde{\mathcal{S}}'$. PSA is an iterative algorithm. We use notation $\tilde{\mathbf{e}}^{(j)}$ to indicate the jth PSA update for computing $\mathbf{e}$ for $\widetilde{\mathcal{S}}_{\mathbf{w}}$ (in order to differentiate this iterative update on $\mathbf{e}$ from that of the alternating procedure between $\mathcal{P}_{\mathbf{e}}$ and $\widetilde{\mathcal{S}}_{\mathbf{w}}$. Specifically, the updating procedure at iteration j is given by[2]

$$\mathbf{y}^{(j)} \in \arg\max_{\mathbf{y} \in \mathcal{Y}} f(\tilde{\mathbf{e}}^{(j)}, \mathbf{y}); \tag{3.18}$$

$$\tilde{\mathbf{e}}^{(j+1)} = \Pi_{\mathcal{E}}(\tilde{\mathbf{e}}^{(j)} - \alpha \nabla_{\tilde{\mathbf{e}}} f(\tilde{\mathbf{e}}^{(j)}, \mathbf{y}^{(j)})) \tag{3.19}$$

where $\alpha > 0$ is the step size, and $\Pi_{\mathcal{E}}(\mathbf{e})$ is the projection of $\mathbf{e}$ onto the feasible set $\mathcal{E}$. Since we have the per-element constraint on each e_m, $\Pi_{\mathcal{E}}(\mathbf{e})$ performs per-element projection, given by

$$\Pi_{\mathcal{E}}(\mathbf{e}) = \begin{cases} e_m & \text{if } |e_m| \leq 1 \\ \frac{e_m}{|e_m|} & \text{o.w} \end{cases}, \text{ for } m \in \mathcal{M}.$$

[2]Note that PSA in [41] is provided using all real-valued variables. We provide the complex version of it, which we can show to be equivalent to the real versio.

From (3.18), as mentioned earlier, once $\mathbf{y}^{(j)}$ is obtained, we know that $f(\mathbf{e}, \mathbf{y}^{(j)}) = \phi_{i'k'}(\mathbf{e}) = \max_{i,k} \phi_{ik}(\mathbf{e})$, for some $i'k'$. With a slight abuse of notation to ease our presentation, we use $\phi_{ik}(\mathbf{e})$ to represent $\phi_{i'k'}(\mathbf{e})$. Then, the gradient $\nabla_{\mathbf{e}} f(\mathbf{e}, \mathbf{y}^{(j)})$ has the following expression:

$$\begin{aligned}\nabla_{\mathbf{e}} f(\mathbf{e}, \mathbf{y}^{(j)}) &= \nabla_{\mathbf{e}} \phi_{ik}(\mathbf{e}) \\ &= -\frac{(\mathbf{Q}_{iik}\mathbf{e} + \mathbf{q}_{iik}) I_{ik}(\mathbf{e}) - \sum_{j \neq i} (\mathbf{Q}_{jik}\mathbf{e} + \mathbf{q}_{jik}) S_{ik}(\mathbf{e})}{\gamma_{ik} I_{ik}^2(\mathbf{e})} \\ &\quad - \frac{\zeta}{M}\mathbf{e} \end{aligned} \tag{3.20}$$

where $I_{ik}(\mathbf{e}) \triangleq \sum_{j \neq i} |\mathbf{w}_j^H (\mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{G}_{ik}\mathbf{e})|^2 + \sigma^2$, $S_{ik}(\mathbf{e}) \triangleq |\mathbf{w}_i^H (\mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{G}_{ik}\mathbf{e})|^2$, $\mathbf{Q}_{jik} \triangleq \mathbf{G}_{ik}^H \mathbf{w}_j \mathbf{w}_j^H \mathbf{G}_{ik}$, and $\mathbf{q}_{jik} = \mathbf{G}_{ik}^H \mathbf{w}_j \mathbf{w}_j^H \mathbf{h}_{ik}^{\mathrm{d}}$.

Note that the updates in each PSA iteration are all in closed-form as in (3.18)–(3.20), with only matrix and vector multiplications. Thus, PSA has a low computational complexity and is particularly suitable for solving $\widetilde{\mathcal{S}}_{\mathbf{w}}$ with large value of M. The convergence analysis of PSA in [41] is applicable to our problem. It shows that the above PSA procedure is guaranteed to converge in finite time to a point at the vicinity of a stationary point for $\widetilde{\mathcal{S}}_{\mathbf{w}}$.

3.5 Discussion on the AMBF Algorithm

We summarize our proposed fast AMBF algorithm for the RIS-assisted multicast QoS problem $\mathcal{P}_o$ in Algorithm 1. It combines the proposed AMBF approach in Section 3.4 and the fast algorithms proposed above for computing the updates in AMBF. A few aspects of the algorithm are discussed below:

3.5.1 Initialization

The AMBF approach requires an initial feasible point of the RIS passive beamformer $\mathbf{e}^{(0)}$ for the BS multicast beamforming subproblem $\mathcal{P}_{\mathbf{e}}$. The feasible point need to satisfy the unit-modulus constraints in (3.4c), which can be easily obtained by generating a random phase for each element in $\mathbf{e}^{(0)}$. In each iteration n of the alternating optimization, solving $\tilde{\mathcal{S}}_{\mathbf{w}}$ by PSA also requires an initial point $\tilde{\mathbf{e}}^{(0)}$ for the update in (3.18). It should directly take $\mathbf{e}^{(n)}$ used in $\mathcal{P}_{\mathbf{e}}$, which is updated from the previous iteration, *i.e.,* $\tilde{\mathbf{e}}^{(0)} = \mathbf{e}^{(n)}$.

3.5.2 Computational Complexity

Algorithm 1 consists of solving two optimization problems $\mathcal{P}_{\mathbf{e}}$ and $\widetilde{\mathcal{S}}_{\mathbf{w}}$ at each iteration. We discuss the computational complexity for solving each subproblem below:

i) As mentioned in Section 3.4.1, using the optimal multicast beamforming structure $\mathbf{w}$ in (3.14), $\mathcal{P}_{\mathbf{e}}$ is converted into the weight optimization problem w.r.t. $\{\mathbf{a}_i\}$ with total GK variables, which can be solved by SCA [51]. If each SCA subproblem is solved using a typical interior point method [65] by the standard convex solver, the computational complexity is $O((GK)^3)$ per SCA iteration which does not grow with neither M nor N. Note that this computational complexity can be further reduced by using the first-order fast ADMM algorithm proposed recently in [66], which only involves closed-form updates and has the computational complexity in the order of $O((GK)^2)$.

ii) Problem $\mathcal{S}_{\mathbf{w}}$ is solved using PSA with updates in (3.18) and (3.19) in each iteration. At each iteration, obtaining $\mathbf{y}^{(j)}$ in (3.18) involves computing SINR for all

Algorithm 1 Alternating Multicast Beamforming Algorithm for solving RIS-assisted QoS problem $\mathcal{P}_o$.

Initialization: Set feasible initial point $\mathbf{e}^{(0)}$; Set $n = 0$.
repeat
    With $\mathbf{e}^{(n)}$, solve $\mathcal{P}_{\mathbf{e}}$ to obtain $\mathbf{w}^{(n+1)}$ by using (3.14).
    With $\mathbf{w}^{(n)}$, set $\mathbf{e}^{(0)} = \mathbf{e}^{(n)}$; Solve $\widetilde{\mathcal{S}}_{\mathbf{w}}$ using PSA updates (3.18) and (3.19) to obtain $\mathbf{e}^{(n+1)}$.
    Set $n \leftarrow n + 1$.
until convergence
Set $\mathbf{e}^{\text{final}} = \exp\left(j\angle\mathbf{e}^{(n)}\right)$.
Obtain $\mathbf{w}^{\text{final}}$ by solving $\mathcal{P}_{\mathbf{e}^{\text{final}}}$.
return $(\mathbf{w}^{\text{final}}, \mathbf{e}^{\text{final}})$.

GK users to find the maximum SINR, which requires $G^2K(NM + M + N)$ flops. The computation in (3.19) is mainly at computing the gradient $\nabla_{\mathbf{e}} f(\mathbf{e}, \mathbf{y}^{(j)})$ as in (3.20). We note that all the terms in (3.20), *i.e.,* $\mathbf{Q}_{jik}\mathbf{e}$, $\mathbf{q}_{jik}$, $I_{ik}(\mathbf{e})$, and $S_{ik}(\mathbf{e})$ are part of the SINR expression that has already been computed. Thus, calculating $\nabla_{\mathbf{e}}\phi_{ik}(\mathbf{e})$ only requires $2GM + 2M$ flops. Thus, the entire updates per iteration requires $G^2K(NM + M + N) + 2GM + 2M$ flops, with the leading complexity being G^2KMN flops. This shows the complexity of PSA grows linearly over M, N, and K.

The final per-element projection complexity is $O(M)$. Thus, the overall computational complexity for each AO iteration of Algorithm 1 is as low as $O((GK)^2 + G^2KMN)$, which is linear in both M and N.

3.5.3 Convergence

For solving $\mathcal{P}_{\mathbf{e}}$ and $\widetilde{\mathcal{S}}_{\mathbf{w}}$ alternatingly, note that the objective of the minimization problem $\mathcal{P}_{\mathbf{e}}$ is lower bounded and that of the maximization problem $\widetilde{\mathcal{S}}_{\mathbf{w}}$ is upper bounded. Thus, the alternating procedure is guaranteed to converge.

3.6 Generalization to RIS-Aided Multi-user Unicast Beamforming

In the special case of multicast beamforming, for $K = 1$ for each group, $\mathcal{P}_{\mathbf{e}}$ reduces to a downlink multi-user unicast beamforming problem w.r.t. $\mathbf{w}$, which yields a closed-form solution. In particular, (3.14) still holds with $\mathbf{a}_i$ being a scalar now that can be derived in closed-form. At the same time, $\mathcal{S}_{\mathbf{w}}$ is still a multicast beamforming MMF problem as discussed in Remark 1 of Section 3.3.4. Algorithm 1 is directly applicable to this case. As we mentioned in Section 2.2.1, in [55] the QoS problem for RIS-aided multi-user unicast has been discussed using AO as well, but unlike our proposed Algorithm 1, they only consider solving problem $\mathcal{P}_o$ alternatively which is reduced to feasibility problem considering a constant $\mathbf{w}$. Thus, the algorithm will stop after first round of iteration, and wont show any significant performance comparing to initial starting point. Analysis in [35] which consider a different form of AO is more potent to find a better solution, for instead of solving the feasibility problem, they also introduce a new optimization problem which try to align the phase of direct Chanel between BS and users with the phase of indirect Chanel through RIS.

Chapter 4

RIS-aided Multicast MMF Problem

In this chapter we consider the weighted max-min fair (MMF) problem for our RIS-assisted multicast scenario we defined in Section 3.1. The objective is to maximize the minimum weighted SINR among users, subject to the BS transmit power budget and the unit-modulus constraint on each RIS reflection coefficient. The weighted SINR essentially is the ratio between SINR and a pre-specified SINR target, where this target sets certain fairness among users. This joint optimization problem over $(\mathbf{w}, \mathbf{e})$ can be formulated as

$$\mathcal{S}_o : \max_{\mathbf{w},\mathbf{e}} \min_{i,k} \quad \frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \tag{4.1a}$$

$$\text{s.t.} \quad \sum_{i=1}^{G} \|\mathbf{w}_i\|^2 \leq P, \tag{4.1b}$$

$$|e_m|^2 = 1, \ m \in \mathcal{M}. \tag{4.1c}$$

where P is the transmit power budget. Similar to the QoS problem $\mathcal{P}_o$, problem $\mathcal{S}_o$ is also nonconvex and NP-hard. Below, we first examine the relation of the MMF problem $\mathcal{S}_o$ to the QoS problem $\mathcal{P}_o$ we have considered previously. Then, we propose a fast algorithm to solve $\mathcal{S}_o$ efficiently.

4.1 Inverse Relation between $\mathcal{P}_o$ and $\mathcal{S}_o$

For the conventional downlink multi-group multicast beamforming design at the BS, it is known that the QoS and MMF problems are inverse problems [53]: the role of the objective function and the constraints are switched in the two problems, and we can solve the MMF problem by iteratively solving the QoS problem along with a bi-section search over the minimum SINR target for the QoS problem. Thus, a natural question arises that whether for such relation also holds for the RIS-assisted multicast beamforming design, which now also includes the RIS passive beamformer optimization, in addition to the BS transmit beamformers. We show below that this is indeed the case.

Proposition 1. For RIS-assisted multicast beamforming, the QoS problem $\mathcal{P}_o$ and the MMF problem $\mathcal{S}_o$ are inverse problems. In particular, explicitly parameterize $\mathcal{S}_o$ as $\mathcal{S}_o(\boldsymbol{\gamma}, P)$ for given $\boldsymbol{\gamma}$ and P with the maximum weighted SINR obtained as $t^o = \mathcal{S}_o(\boldsymbol{\gamma}, P)$, and parameterize $\mathcal{P}_o$ as $\mathcal{P}_o(\boldsymbol{\gamma})$ with the minimum power obtained as $P = \mathcal{P}_o(\boldsymbol{\gamma})$. Then, $\mathcal{P}_o(\boldsymbol{\gamma})$ and $\mathcal{S}_o(\boldsymbol{\gamma}, P)$ have the following relations:

$$t^o = \mathcal{S}_o(\boldsymbol{\gamma}, \mathcal{P}_o(t^o\boldsymbol{\gamma})), \quad P = \mathcal{P}_o(\mathcal{S}_o(\boldsymbol{\gamma}, P)\boldsymbol{\gamma}). \tag{4.2}$$

proof: We can equivalently rewrite $\mathcal{S}_o$ as

$$\begin{aligned} \mathcal{S}_o' : \ & \max_{\mathbf{w}} \left(\max_{\mathbf{e}:|e_m|^2=1, m\in\mathcal{M}} \min_{i,k} \frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \right) \\ & \text{s.t.} \ \sum_{i=1}^{G} \|\mathbf{w}_i\|^2 \le P. \end{aligned}$$

For the QoS problem $\mathcal{P}_o$, consider its equivalent formulation $\mathcal{P}_2$ in (3.7) derived in Section 3.3.1. Comparing $\mathcal{S}_o'$ and $\mathcal{P}_2$, we see that the objective functions and the constraints in these two problems are the same but switched, clearly indicating that they (equivalently $\mathcal{S}_o$ and $\mathcal{P}_o$) are the inverse problems. It is straightforward to see that, at the optimality, the respective constraints in $\mathcal{S}_o'$ and $\mathcal{P}_2$ are attained with equality. As the result, we have the relations in (4.2).

Based on the above relation, we can solve the MMF problem $\mathcal{S}_o$ using Algorithm 1, where by the relation in (4.2), $\mathcal{S}_o$ can be solved via solving its inverse QoS problem $\mathcal{P}_o$ iteratively along with a bi-section search for $t_{\mathbf{e}}^o$. However, since the size of the RIS reflection elements is typically large $M \gg 1$, iteratively solving $\mathcal{P}_o$ will incur relatively high computational complexity. We need a more computationally efficient algorithm to solve the large-scale joint optimization problem $\mathcal{S}_o$. Next, we present a fast algorithm to solve $\mathcal{S}_o$ directly.

4.2 Proposed Fast Algorithm for Solving $\mathcal{S}_o$

As mentioned in Section 3.4.2, for the multicast beamforming MMF problem in the conventional downlink scenario without RIS, PSA is proposed in [41] to obtain a near-stationary solution. Below, we show that we can adopt PSA to directly solve the MMF problem efficiently in the RIS-assisted multicast scenario as well.

Note from Proposition 1 that problems $\mathcal{P}_o$ and $\mathcal{S}_o$ are the inverse problems. This means the optimal structure of the BS multicast beamformer $\mathbf{w}$ still has the form shown in (3.14). Indeed, treating the RIS-assisted channel between the BS and user k in group i as the equivalent channel $\tilde{\mathbf{h}}_{ik}$, the problem reduces to the conventional

multicast MMF problem, for which the optimal beamforming structure is shown as in (3.14) [51]. However, for the MMF problem, computing $\boldsymbol{\lambda}$ in $\mathbf{R}(\boldsymbol{\lambda})$ is not straightforward. Fortunately, the asymptotic expression of $\mathbf{R}(\boldsymbol{\lambda})$ for large N is obtained in closed-form in [51], which can be used as an approximate expression of $\mathbf{R}(\boldsymbol{\lambda})$ to further simplify the computation. Specifically, let $\tilde{\mathbf{h}}_{iK} = \sqrt{\tilde{\beta}_{ik}}\mathbf{g}_{ik}$ where $\tilde{\beta}_{ik}$ is the large-scale channel variation and $\tilde{\mathbf{g}}_{ik} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Then, the asymptotic expression of $\mathbf{R}(\boldsymbol{\lambda})$ has the following simple closed form

$$\mathbf{R}(\boldsymbol{\lambda}) \approx \mathbf{I} + \frac{P\bar{\beta}_{\mathbf{h}}}{\sigma^2 K_{\text{tot}}} \sum_{i=1}^{G} \sum_{k=1}^{K_i} \mathbf{g}_{ik}\mathbf{g}_{ik}^H \triangleq \widetilde{\mathbf{R}} \tag{4.3}$$

where $\bar{\beta}_{\mathbf{h}} \triangleq 1/\Big(\frac{1}{K_{\text{tot}}} \sum_{i=1}^{G} \sum_{k=1}^{K_i} \frac{1}{\tilde{\beta}_{ik}}\Big)$ is the harmonic mean of the large-scale channel variations of all users. Using $\widetilde{\mathbf{R}}$ in (4.3), we can replace $\mathbf{w}_i$ in (3.14) with the following expression:

$$\mathbf{w}_i \;=\; \widetilde{\mathbf{R}}^{-1}\widetilde{\mathbf{H}}_i\mathbf{a}_i \,,\; i \in \mathcal{G}. \tag{4.4}$$

Below, we develop our fast algorithm to solve $\mathcal{S}_o$ based on $\mathbf{w}_i$ in (4.4) to reduce the computational complexity.

Using $\mathbf{w}_i$ in (4.4), we first convert $\mathcal{S}_o$ into a problem of smaller size, which is a joint optimization problem w.r.t. $\{\mathbf{a}_i\}$ and $\mathbf{e}$, given by

$$\mathcal{S}_o^{\text{eq}} : \max_{\mathbf{a},\mathbf{e}} \min_{i,k} \quad \frac{\text{SINR}_{ik}}{\gamma_{ik}} \tag{4.5a}$$

$$\text{s.t.} \quad \sum_{i=1}^{G} \|\widetilde{\mathbf{C}}_i\mathbf{a}_i\|^2 \le P, \tag{4.5b}$$

$$|e_m|^2 \le 1,\; m \in \mathcal{M}, \tag{4.5c}$$

$$\mathbf{e}^H\mathbf{e} = M. \tag{4.5d}$$

where $\mathbf{a} \triangleq [\mathbf{a}_1^T, \cdots, \mathbf{a}_G^T]^T$, $\widetilde{\mathbf{C}}_i \triangleq \widetilde{\mathbf{R}}^{-1}\widetilde{\mathbf{H}}_i$, and the constraints in (4.1c) are replaced by the equivalent set of constraints in (4.5c) and (4.5d).

Using the technique similar to that in Section 3.3.4, we relax S_o^{eq} into the following problem by moving the equality constraint in (4.5d) into the objective function as a penalty term with a penalty weight $\delta > 0$:

$$\widetilde{\mathcal{S}}_o : \max_{\mathbf{a},\mathbf{e}} \min_{i,k} \quad \frac{\text{SINR}_{ik}}{\gamma_{ik}} + \delta \frac{\mathbf{e}^H\mathbf{e}}{M} \tag{4.6a}$$

$$\text{s.t.} \quad \sum_{i=1}^{G} \|\widetilde{\mathbf{C}}_i\mathbf{a}_i\|^2 \le P, \tag{4.6b}$$

$$|e_m|^2 \le 1, \ m \in \mathcal{M}. \tag{4.6c}$$

Problem $\widetilde{\mathcal{S}}_o$ has a similar structure as $\widetilde{\mathcal{S}}_{\mathbf{w}}$ in (3.12). The difference is that $\widetilde{\mathcal{S}}_o$ is a joint optimization problem for $(\mathbf{a}, \mathbf{e})$. Following Section 3.4.2, we can again apply PSA to efficiently compute a near-stationary joint solution $(\mathbf{a}, \mathbf{e})$ for $\widetilde{\mathcal{S}}_o$. The details are described below.

4.2.1 Problem Reformulation

Let $\mathbf{x} \triangleq [\mathbf{a}^H, \mathbf{e}^H]^H$. Denote the feasible set of $\widetilde{\mathcal{S}}_o$ as $\mathcal{U} \triangleq \{\mathbf{x} : |e_m|^2 \le 1, \ m \in \mathcal{M}; \ \sum_{i=1}^{G} \|\widetilde{\mathbf{C}}_i\mathbf{a}_i\|^2 \le P\}$. By substituting the SINR expression in (3.3) into the objective function in (4.6a), we define

$$\varphi_{ik}(\mathbf{x}) \triangleq -\frac{1}{\gamma_{ik}} \frac{|\mathbf{a}_i^H\widetilde{\mathbf{C}}_i^H(\mathbf{h}_{ik}^{\text{d}}+\mathbf{G}_{ik}\mathbf{e})|^2}{\sum_{j\neq i}|\mathbf{a}_j^H\widetilde{\mathbf{C}}_j^H(\mathbf{h}_{ik}^{\text{d}}+\mathbf{G}_{ik}\mathbf{e})|^2+\sigma^2} - \delta\frac{\mathbf{e}^H\mathbf{e}}{M}, \tag{4.7}$$

for $k \in \mathcal{K}$, $i \in \mathcal{G}$. Then, $\widetilde{\mathcal{S}}_o$ can be equivalently rewritten as

$$\widetilde{\mathcal{S}}_o' : \min_{\mathbf{x}\in\mathcal{U}} \max_{i,k} \varphi_{ik}(\mathbf{x}). \tag{4.8}$$

Let $\boldsymbol{\varphi}(\mathbf{x})$ be a $GK \times 1$ vector containing all $\varphi_{ik}(\mathbf{x})$'s, and let $\mathbf{y}$ denote a $GK \times 1$ vector with nonnegative elements. Then, $\widetilde{\mathcal{S}}_o'$ can be further transform into the following

equivalent problem

$$\widetilde{\mathcal{S}_{\mathbf{x}}} : \min_{\mathbf{x}\in\mathcal{U}} \max_{\mathbf{y}\in\mathcal{Y}} g(\mathbf{x},\mathbf{y}) \tag{4.9}$$

where $g(\mathbf{x},\mathbf{y}) \triangleq \boldsymbol{\varphi}^T(\mathbf{x})\mathbf{y}$, and $\mathcal{Y} \triangleq \{\mathbf{y} : \mathbf{y} \succcurlyeq 0, \mathbf{1}^T\mathbf{y} = 1\}$. Both $\mathcal{U}$ and $\mathcal{Y}$ are compact convex sets. Note that $g(\mathbf{x},\mathbf{y})$ is non-convex in $\mathbf{x}$ and concave in $\mathbf{y}$, and problem $\widetilde{\mathcal{S}_{\mathbf{x}}}$ is a non-convex-concave min-max problem, which is NP-hard.

4.2.2 The Projected Subgradient Algorithm

Since $\widetilde{\mathcal{S}_{\mathbf{x}}}$ has the same structure as $\widetilde{\mathcal{S}}'$ in (3.17), we can again apply PSA with updating steps similar to (3.18) and (3.19). In particular, the updating procedure at iteration j is given by

$$\mathbf{y}^{(j)} \in \arg\max_{\mathbf{y}\in\mathcal{Y}} g(\mathbf{x}^{(j)},\mathbf{y}); \tag{4.10}$$

$$\mathbf{x}^{(j+1)} = \Pi_{\mathcal{U}}\Big(\mathbf{x}^{(j)} - \eta\nabla_{\mathbf{x}} g(\mathbf{x}^{(j)},\mathbf{y}^{(j)})\Big) \tag{4.11}$$

where $\eta > 0$ is the step size, and $\Pi_{\mathcal{U}}(\mathbf{x})$ is the projection of $\mathbf{x}$ onto the feasible set $\mathcal{U}$. Since the constraints on $\mathbf{a}$ and $\mathbf{e}$ in $\mathcal{U}$ are separate, $\Pi_{\mathcal{U}}(\mathbf{x})$ is performed for $\mathbf{a}$ and $\mathbf{e}$ separately:

$$\Pi_{\mathcal{U}}(\mathbf{x}) = \begin{cases} \Pi_{\mathcal{U}}(\mathbf{a};\cdot) = \begin{cases} \mathbf{a} & \text{if } \sum_{i=1}^{G} \|\tilde{\mathbf{C}}_i\mathbf{a}_i\|^2 \le P \\ \sqrt{\frac{P}{P_{\text{tot}}}}\mathbf{a} & \text{o.w.} \end{cases} \\ \Pi_{\mathcal{U}}(\cdot;\mathbf{e}) = \begin{cases} e_m & \text{if } |e_m| \le 1 \\ \frac{e_m}{|e_m|} & \text{o.w.} \end{cases}, \text{ for } m \in \mathcal{M}. \end{cases} \tag{4.12}$$

Since $\mathcal{Y}$ is a probability simplex, an optimal solution to the maximization problem in (4.10) is $\mathbf{y} = [0,\cdots,1,\cdots,0]^T$, with 1 at some jth position. Thus, we have

$$g(\mathbf{x},\mathbf{y}^{(j)}) = \varphi_{\hat{i}\hat{k}}(\mathbf{x}) = \max_{i,k} \varphi_{ik}(\mathbf{x}), \tag{4.13}$$

for some $(\hat{i}, \hat{k})$. Then, the gradient $\nabla_{\mathbf{x}} g(\mathbf{x}, \mathbf{y}^{(j)})$ is given by

$$\begin{aligned}\nabla_{\mathbf{x}} g(\mathbf{x}, \mathbf{y}^{(j)}) &= \nabla_{\mathbf{x}} \varphi_{\hat{i}\hat{k}}(\mathbf{x}) \\ &= \left[\nabla_{\mathbf{a}} \varphi_{\hat{i}\hat{k}}^{H}(\mathbf{x}), \nabla_{\mathbf{e}} \varphi_{\hat{i}\hat{k}}^{H}(\mathbf{x})\right]^{H}. \end{aligned} \tag{4.14}$$

For $\nabla_{\mathbf{e}} \varphi_{\hat{i}\hat{k}}(\mathbf{x})$, note that $\varphi_{ik}(\mathbf{x})$ is the same as $\phi_{ik}(\mathbf{e})$ by using $\mathbf{w}_i$ in (4.4) and change δ to ζ. Thus, the gradient $\nabla_{\mathbf{e}} \varphi_{\hat{i}\hat{k}}(\mathbf{x})$ can be computed using (3.20), except that $\mathbf{w}_i$ is given in (4.4) and ζ is replaced by δ.

To compute $\nabla_{\mathbf{a}} \varphi_{\hat{i}\hat{k}}(\mathbf{x})$, we use the equivalent channel $\tilde{\mathbf{h}}_{ik} = \mathbf{h}_{ik}^{\mathrm{d}} + \mathbf{G}_{ik}\mathbf{e}$ in (4.7) and rewrite $\varphi_{ik}(\mathbf{x})$ in (4.7) as

$$\varphi_{ik}(\mathbf{x}) = -\frac{1}{\gamma_{ik}} \frac{\mathbf{a}_i^H \widetilde{\mathbf{A}}_{iik} \mathbf{a}_i}{\sum_{j \neq i} \mathbf{a}_j^H \widetilde{\mathbf{A}}_{jik} \mathbf{a}_j + \sigma^2} - \delta \frac{\mathbf{e}^H \mathbf{e}}{M},$$

where $\widetilde{\mathbf{A}}_{jik} \triangleq \widetilde{\mathbf{C}}_j^H \tilde{\mathbf{h}}_{ik} \tilde{\mathbf{h}}_{ik}^H \widetilde{\mathbf{C}}_j$. Then, we have

$$\nabla_{\mathbf{a}} \varphi_{\hat{i}\hat{k}}(\mathbf{x}) = [\nabla_{\mathbf{a}_1} \varphi_{\hat{i}\hat{k}}^{H}(\mathbf{x}), \cdots, \nabla_{\mathbf{a}_G} \varphi_{\hat{i}\hat{k}}^{H}(\mathbf{x})]^{H} \tag{4.15}$$

where

$$\nabla_{\mathbf{a}_i} \varphi_{\hat{i}\hat{k}}(\mathbf{x}) = \begin{cases} -\dfrac{1}{\gamma_{\hat{i}\hat{k}}} \dfrac{\widetilde{\mathbf{A}}_{\hat{i}\hat{i}\hat{k}} \mathbf{a}_{\hat{i}}}{\sum_{j \neq \hat{i}} \mathbf{a}_j^H \widetilde{\mathbf{A}}_{j\hat{i}\hat{k}} \mathbf{a}_j + \sigma^2} & \text{if } i = \hat{i} \\ \dfrac{1}{\gamma_{\hat{i}\hat{k}}} \dfrac{(\mathbf{a}_{\hat{i}}^H \widetilde{\mathbf{A}}_{\hat{i}\hat{i}\hat{k}} \mathbf{a}_{\hat{i}}) \widetilde{\mathbf{A}}_{i\hat{i}\hat{k}} \mathbf{a}_i}{\left(\sum_{j \neq \hat{i}} \mathbf{a}_j^H \widetilde{\mathbf{A}}_{j\hat{i}\hat{k}} \mathbf{a}_j + \sigma^2\right)^2} & \text{o.w.} \end{cases} \tag{4.16}$$

4.2.3 Final Processing

Since the PSA is to solve $\widetilde{\mathcal{S}}_o$, which is a relaxed problem of $\mathcal{S}_o$, its solution may not be feasible to $\mathcal{S}_o$, *i.e.,* $\mathbf{e}^\star$ may not satisfy (4.1c). Thus, we have this final processing step to obtain a feasible solution to $\mathcal{S}_o$. Let $\mathbf{x}^\star = [\mathbf{a}^{\star H}, \mathbf{e}^{\star H}]^H$ be the solution produced by

the updating procedure in (4.10) and (4.11). If $\mathbf{e}^\star$ does not satisfy (4.1c), we have the following step to obtain a feasible solution based on $(\mathbf{a}^\star, \mathbf{e}^\star)$:

- Project $\mathbf{e}^\star$ onto its feasible set: $\mathbf{e}^{\text{final}} = \exp(j\angle\mathbf{e}^\star)$, which takes the phase of each element in $\mathbf{e}^\star$;

- Given $\mathbf{e}^{\text{final}}$, we solve $\mathcal{S}_o^{\text{eq}}$ in (4.5) w.r.t. $\mathbf{a}$ to obtain the final beamforming solution $\mathbf{w}^{\text{final}}$. Note that in this case, $\mathcal{S}_o^{\text{eq}}$ is reduced to the following conventional MMF problem:

$$\mathcal{S}^{\text{eq}}(\mathbf{e}^{\text{final}}) : \max_{\mathbf{a}} \min_{i,k} \ \frac{\text{SINR}_{ik}}{\gamma_{ik}} \tag{4.17a}$$

$$\text{s.t.} \ \sum_{i=1}^{G} \|\widetilde{\mathbf{C}}_i \mathbf{a}_i\|^2 \leq P. \tag{4.17b}$$

  where SINR_{ik} is as shown in (3.13) with the equivalent channel $\tilde{\mathbf{h}}_{ik}$ computed based on $\mathbf{e}^{\text{final}}$. As mentioned earlier, the PSA-based fast algorithm has been proposed in [41] for this MMF problem, and we use directly apply it to compute the solution $\mathbf{a}^{\text{final}}$ to $\mathcal{S}^{\text{eq}}(\mathbf{e}^{\text{final}})$ efficiently. In particular, we use $\mathbf{a}^\star$ as a good initial point to this algorithm to achieve fast convergence and good performance.

.

We summarize our proposed first-order fast algorithm for the RIS-assisted multicast MMF problem $\mathcal{S}_o$ in Algorithm 2. Again, based on the existing convergence analysis [41], the PSA updating procedure in (4.10) and (4.11) is guaranteed to converge in finite time to a point at the vicinity of a stationary point for $\widetilde{\mathcal{S}}_o$.

4.3 Discussions

Like the previous chapter, in the following lines, we will discuss the different aspects of our peoposed algorithm to solve our RIS-aided multicast MMF problem.

4.3.1 Initialization

The proposed updating procedure in (4.10) and (4.11) requires an initial point $\mathbf{x}^{(0)} = [\mathbf{a}^{(0)^H}, \mathbf{e}^{(0)^H}]^H$. This initial point can be generated randomly, as the projection (4.11) will ensure that the subsequent points $\{\mathbf{x}^{(n)}\}$ are feasible. However, a good initial point improves the rate of convergence. Thus, instead of a random initial point, we consider the following approach.

We first generate $\mathbf{e}^{(0)}$ using a random phase for each element. Given $\mathbf{e}^{(0)}$, $\mathcal{S}_o^{\text{eq}}$ in (4.5) is reduced to the MMF problem w.r.t. $\mathbf{a}$, which is similar to $\mathcal{S}^{\text{eq}}(\mathbf{e}^{\text{final}})$ in (4.17). In particular, the MMF problem w.r.t. $\mathbf{a}$ can be further equivalently written as

$$\mathcal{S}^{\text{eq}}(\mathbf{e}^{(0)}) : \max_{\mathbf{a},t} \quad t \tag{4.18a}$$

$$\text{s.t.} \quad \frac{\text{SINR}_{ik}}{\gamma_{ik}} \geq t, \ k \in \mathcal{K}, i \in \mathcal{G} \tag{4.18b}$$

$$\sum_{i=1}^{G} \|\widetilde{\mathbf{C}}_i \mathbf{a}_i\|^2 \leq P. \tag{4.18c}$$

We generate $\mathbf{a}^{(0)}$ by solving this MMF problem. Note that in this way, $(\mathbf{a}^{(0)}, \mathbf{e}^{(0)})$ is feasible to $\widetilde{\mathcal{S}}_o$. In the literature, a common approach for the above MMF problem is to solve its inverse problem, *i.e.,* the QoS problem similar to $\mathcal{P}_{\mathbf{e}}$ in (3.8) for given t, along with bi-section search over t [51, 53]. Since our goal is only to generate a feasible point for $\widetilde{\mathcal{S}}_o$, to reduce the computational complexity, we only need to solve

the QoS problem[1] along with one single bi-section search over t, to obtain the initial point $\mathbf{a}^{(0)}$.

4.3.2 Computational Complexity

For the updates in (4.10) and (4.11) at each iteration, the update $\mathbf{y}^{(j)}$ is obtained in (4.13) by finding the maximum SINR among GK users, which requires $G^2K(NM + M + N)$ flops. The update in (4.11) requires to calculate $\nabla_{\mathbf{a}}\varphi_{\hat{i}\hat{k}}(\mathbf{x})$ and $\nabla_{\mathbf{e}}\varphi_{\hat{i}\hat{k}}(\mathbf{x})$. Since $\nabla_{\mathbf{e}}\varphi_{\hat{i}\hat{k}}(\mathbf{x})$ is computed using (3.20), as discussed in Section 3.5.2.ii), it requires $2GM + 2M$ flops. To compute $\nabla_{\mathbf{a}_i}\varphi_{\hat{i}\hat{k}}(\mathbf{x})$ in (4.16), we need to calculate $\mathbf{a}_j^H\widetilde{\mathbf{C}}_j^H\tilde{\mathbf{h}}_{\hat{i}\hat{k}}$ for all j's. Note that $\widetilde{\mathbf{C}}_j$'s only need to computed once before the updating procedure. Computing $\tilde{\mathbf{h}}_{\hat{i}\hat{k}}$ based on $\mathbf{e}^{(j)}$ requires $N(M+1)$ flops. Then, obtaining $\mathbf{a}_j^H\widetilde{\mathbf{C}}_j^H\tilde{\mathbf{h}}_{\hat{i}\hat{k}}$ for all $j \in \mathcal{G}$ needs $GK(N+1)$ flops. We also need to calculate $\widetilde{\mathbf{A}}_{\hat{i}\hat{i}\hat{k}}\mathbf{a}_{\hat{i}} = \widetilde{\mathbf{C}}_{\hat{i}}^H\tilde{\mathbf{h}}_{\hat{i}\hat{k}}\tilde{\mathbf{h}}_{\hat{i}\hat{k}}^H\widetilde{\mathbf{C}}_{\hat{i}}\mathbf{a}_{\hat{i}}$. Note that since $\widetilde{\mathbf{C}}_{\hat{i}}^H\tilde{\mathbf{h}}_{\hat{i}\hat{k}}$ and $\tilde{\mathbf{h}}_{\hat{i}\hat{k}}^H\widetilde{\mathbf{C}}_{\hat{i}}\mathbf{a}_{\hat{i}}$ are already computed when calculating $\mathbf{a}_j^H\widetilde{\mathbf{C}}_j^H\tilde{\mathbf{h}}_{\hat{i}\hat{k}}$ for all j, we have $\widetilde{\mathbf{A}}_{\hat{i}\hat{i}\hat{k}}\mathbf{a}_{\hat{i}}$ readily available without extra computation. Finally, we need $K(2G-1)$ flops to calculate the final value of $\nabla_{\mathbf{a}_i}\varphi_{\hat{i}\hat{k}}(\mathbf{x})$ for all $\mathbf{a}_i$'s. Thus, the overall leading complexity of the updating procedure in each iteration is $G^2KMN + GKN + GM + MN$, which grows linearly with M and N.

4.4 Generalization to RIS-Assisted Multi-user Downlink Beamforming

Similar to the generalization to the downlink RIS-assisted multi-user beamforming for power minimization discussed in Section 3.6, our proposed fast algorithm for the mul-

[1]To solve the QoS problem, we use the classical approach of SDR with the Gaussian randomization procedure.

Algorithm 2 First-order fast algorithm for RIS-assisted MMF problem $\mathcal{S}_o$.

1: **Initialization:** Set feasible point $\mathbf{x}^{(0)}$; Set $n = 0$.
2: **repeat**
3: Update $\mathbf{x}^{(n)}$ using (4.10) and (4.11) to obtain $\mathbf{x}^{(n+1)}$.
4: Set $n \leftarrow n + 1$.
5: **until** convergence
6: Set $\mathbf{e}^{\text{final}} = \exp(j\angle\mathbf{e}^{(n)})$.
7: Solving $\mathcal{S}(\mathbf{e}^{\text{final}})$ to obtain $\mathbf{w}^{\text{final}}$.
8: **return** $(\mathbf{w}^{\text{final}}, \mathbf{e}^{\text{final}})$.

ticast beamforming MMF problem is also directly applicable to the general downlink RIS-assisted multi-user beamforming for the max-min fair objective. As we mentioned in Section 2.2.1, [57, 58] have been studied the MMF problem and maximize the users' worst rate problem which is equivalent to MMF problem as well. In [57], OLP is used to find a suboptimal answer to MMF problem iterativly, and in [58] alternating descend round has been used to solve the users' worst rate problem.

Chapter 5

Simulation Results

In this chapter we discuss the results of our simulation for our RIS-aided multigroup-multicast wireless communication system, and compare it with the current state of the art.

5.1 Simulation Setup

We consider a downlink RIS-assisted multicast scenario with a BS and an RIS. Unless specified otherwise, our default system setup is $G = 2$ groups and $K = 2$ users per group, $N = 4$ antennas at the BS, and $M = 5 \times 5$ reflective elements for the RIS. The locations of the BS and the RIS using the (x, y, z)-coordinates in meters are $(0, 0, 0)$ and $(70, 70, 0)$, respectively. We consider the RIS as a rectangular surface with array elements placed on the (y, z)-plane. The users are located randomly within a circle centered at the RIS with the radius 20 m on the (x, y)-plane. We set the target SINR $\gamma_{ik} = \gamma$, $\forall i, k$. We assume a line-of-sight (LOS) channel between the BS and the RIS. Thus, $\mathbf{H}_{\mathrm{r}}$ is modeled as a Rician fading channel matrix given by $\mathbf{H}_{\mathrm{r}} = \beta_{\mathrm{BR}}\Big(\sqrt{\frac{K_{\mathrm{r}}}{1+K_{\mathrm{r}}}}\mathbf{H}_{\mathrm{r}}^{\mathrm{LOS}} + \sqrt{\frac{1}{1+K_{\mathrm{r}}}}\mathbf{H}_{\mathrm{r}}^{\mathrm{NLOS}}\Big)$, where the path gain β_{BR} is modeled as $\beta_{\mathrm{BR}}[\mathrm{dB}] = -30 - 22\log 10(d_{\text{B-R}})$ with $d_{\text{B-R}}$ being the BS-RIS distance in meters, the

Rician factor $K_{\mathrm{r}} = 10$, $\mathbf{H}_{\mathrm{r}}^{\mathrm{NLOS}}$ is the non-line-of-sight (NLOS) component modeled as $[\mathbf{H}_{\mathrm{r}}^{\mathrm{NLOS}}]_{ij} \sim \mathcal{CN}(0, \mathbf{I})$, and $\mathbf{H}_{\mathrm{r}}^{\mathrm{LOS}}$ is the LOS component. The LOS component $\mathbf{H}_{\mathrm{r}}^{\mathrm{LOS}}$ is a function of the BS and RIS locations and is modeled as

$$\mathbf{H}_{\mathrm{r}}^{\mathrm{LOS}} = \mathbf{b}_{\mathrm{RIS}}(\psi_2, \theta_2)\mathbf{b}_{\mathrm{BS}}(\psi_1, \theta_1)^H. \tag{5.1}$$

In (5.1), ψ_1 and θ_1 are the azimuth and elevation angles of departure (AoD) from the BS to the IRS, and $\mathbf{b}_{\mathrm{BS}}(\psi_1, \theta_1)$ is the BS steering vector given by

$$\mathbf{b}_{\mathrm{BS}}(\psi_1, \theta_1) = \left[1, \ldots, e^{\frac{2\pi(N-1)d_{\mathrm{BS}}}{\lambda_c}\cos(\psi_1)\cos(\theta_1)}\right] \tag{5.2}$$

where d_{BS} is the distance between two adjacent BS antennas, λ_c is the carrier wavelength, and $\cos(\psi_1)\cos(\theta_1) = \frac{x_{\mathrm{RIS}} - x_{\mathrm{BS}}}{d_{\mathrm{B\text{-}R}}}$. We assume $d_{\mathrm{BS}} = \frac{\lambda_c}{2}$. Similarly, ψ_2 and θ_2 are the azimuth and elevation angles of arrival (AoA) from the BS to the RIS, and the m^{th} entry of the RIS steering vector $\mathbf{b}_{\mathrm{RIS}}(\psi_2, \theta_2)$ is given by

$$[\mathbf{b}_{\mathrm{RIS}}(\psi_2, \theta_2)]_m = e^{\frac{2\pi d_{\mathrm{RIS}}}{\lambda_c}[y_m \sin(\psi_2)\cos(\theta_2) + z_m \sin(\theta_2)]} \tag{5.3}$$

where d_{RIS} is the distance between two adjacent RIS elements, $y_m = \mathrm{mod}(m-1, M_y)$ with M_y being the number of RIS elements along the y-axis, $z_m = \lfloor m-1, M_z \rfloor$ with M_z being the number of RIS elements along the z-axis, $\sin(\psi_2)\cos(\theta_2) = \frac{y_{\mathrm{BS}} - y_{\mathrm{RIS}}}{d_{\mathrm{B\text{-}R}}}$, and $\sin(\theta_2) = \frac{z_{\mathrm{BS}} - z_{\mathrm{RIS}}}{d_{\mathrm{B\text{-}R}}}$. We assume $d_{\mathrm{RIS}} = \frac{\lambda_c}{2}$. We consider two different scenarios for the channels between the RIS and users: channels with and without a LOS path. Thus, the channel $\mathbf{h}_{ik}^{\mathrm{r}}$ between the RIS and user k in group i is modeled as $\mathbf{h}_{ik}^{\mathrm{r}} = \beta_{ik}^{\mathrm{r}}\left(\sqrt{\frac{K_{ik}^{\mathrm{r}}}{1+K_{ik}^{\mathrm{r}}}}\mathbf{h}_{ik}^{\mathrm{r,LOS}} + \sqrt{\frac{1}{1+K_{ik}^{\mathrm{r}}}}\mathbf{h}_{ik}^{\mathrm{r,NLOS}}\right)$, where β_{ik}^{r} is the path gain modeled as $\beta_{ik}^{\mathrm{r}}[\mathrm{dB}] = -30 - 22\log 10(d_{ik}^{\mathrm{R\text{-}U}})$, with $d_{ik}^{\mathrm{R\text{-}U}}$ being the distance between the RIS and the user, K_{ik}^{r} is the Rician factor, and $\mathbf{h}_{ik}^{\mathrm{r,NLOS}} \sim \mathcal{CN}(0, \mathbf{I})$. We set $K_{ik}^{\mathrm{r}} = 10$ for the scenario with

a LOS path and $K_{ik}^{\mathrm{r}} = 0$ for the scenario without a LOS path. The LOS component $\mathbf{h}_{ik}^{\mathrm{r,LOS}}$ is modeled as $\mathbf{h}_{ik}^{\mathrm{r,LOS}} = \mathbf{b}_{\mathrm{RIS}}(\psi_{3,ik}, \theta_{3,ik})$, where $\psi_{3,ik}$ and $\theta_{3,ik}$ are the azimuth and elevation AoD, and the m^{th} entry of the RIS steering vector $\mathbf{b}_{\mathrm{RIS}}(\psi_{3,ik}, \theta_{3,ik})$ is expressed as

$$[\mathbf{b}_{\mathrm{RIS}}(\psi_{3,ik}, \theta_{3,ik})]_m = e^{\frac{2\pi d_{\mathrm{RIS}}}{\lambda_c}[y_m \sin(\psi_{3,ik}) \cos(\theta_{3,ik}) + z_m \sin(\theta_{3,ik})]} \tag{5.4}$$

where $\sin(\psi_{3,ik})\cos(\theta_{3,ik}) = \frac{y_{ik} - y_{\mathrm{RIS}}}{d_{ik}^{\text{R-U}}}$ and $\sin(\theta_{3,ik}) = \frac{z_{ik} - z_{\mathrm{RIS}}}{d_{ik}^{\text{R-U}}}$. We model the channel between BS and each user as Rayleigh fading and generated i.i.d as $\mathbf{h}_{ik}^{\mathrm{d}} \sim \mathcal{CN}(\mathbf{0}, \beta_{ik}^{\mathrm{d}}\mathbf{I})$, where β_{ik}^{d} is the respective path gain modeled as $\beta_{ik}^{\mathrm{d}}[\mathrm{dB}] = -32.6 - 36.7 \log 10(d_{ik}^{\text{B-U}})$, with $d_{ik}^{\text{B-U}}$ being the respective BS-user distances, $\forall k, i$. The receiver noise power is set to $\sigma^2 = -100$ dBm. For subproblems $\widetilde{\mathcal{S}}_{\mathbf{w}}$ and $\widetilde{\mathcal{S}}_o$, we set the penalty parameter $\zeta = 0.1$. For PSA updates in (3.19) for the QoS problem, we set the step size $\alpha = 10$, and for the updates in (4.11) for the MMF problem, we set $\eta = 1$.[1]

5.2 Convergence Behavior of the Proposed Algorithms

We first study the convergence behavior of our proposed AMBF algorithm in Algorithm 1, which alternatingly solves two multicast subproblems at the BS and the RIS using our proposed fast algorithms. Fig. 5.1 shows the convergence behavior of subproblem $\widetilde{\mathcal{S}}_{\mathbf{w}}$, which is solved using PSA, in one iteration round of AMBF algorithm for three random channel realizations. As we expected from the convergence analysis in [41], the PSA starts converging after $1000 \sim 3000$ iterations rounds. Furthermore,

[1]We have conducted extensive experiments using different values of α, η, and ζ and found that $\zeta = 0.1$, $\alpha = 10$, and $\eta = 1$ in general provide the best balance between the performance and the convergence speed.

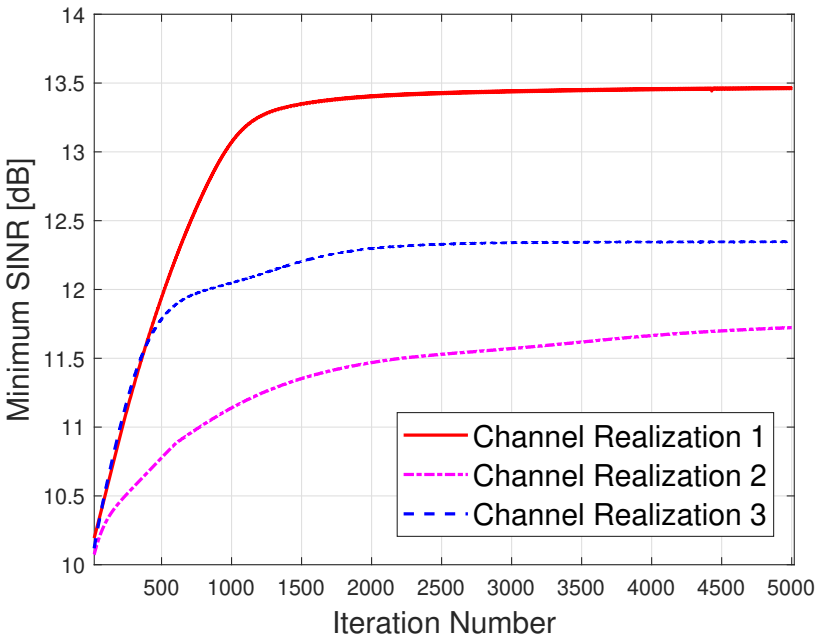

Figure 5.1: Convergence behavior of the MMF subproblem $\widetilde{\mathcal{S}}_{\mathbf{w}}$ ($N = 4, M = 25, G = 2, K = 2, \gamma = 10$ dB, $K_{ik}^{\mathrm{r}} = 0$).

denote the total transmit power at the BS as $P_{\mathrm{tot}} = \sum_{i=1}^{G} \|\mathbf{w}_i\|^2$. Fig. 5.2 shows the trajectory of P_{tot} over the number of iterations by Algorithm 1. We assume the channel between the RIS and each user follows Rayleigh fading with $K_{ik}^{\mathrm{r}} = 0$, $\forall k, i$. Fig. 5.2 shows the trajectory for three random channel realizations. We observe that our proposed AMBF approach converges within about $6 \sim 12$ iterations, which is relatively fast, especially given that we use fast algorithms for computing the solution to each subproblem.

Finally, Fig. 5.3 shows the convergence behavior of the MMF problem solved using Algorithm 2 for three random channel realizations. Here we also consider the channel between the RIS and each user follows Rayleigh fading with $K_{ik}^{\mathrm{r}} = 0$, $\forall k, i$. Similar to subproblem $\widetilde{\mathcal{S}}_{\mathbf{w}}$, in Algorithm 2, we also use PSA, which makes the convergence analysis in [41] applicable here as well. As Fig. 5.3 shows, even in worse scenarios,

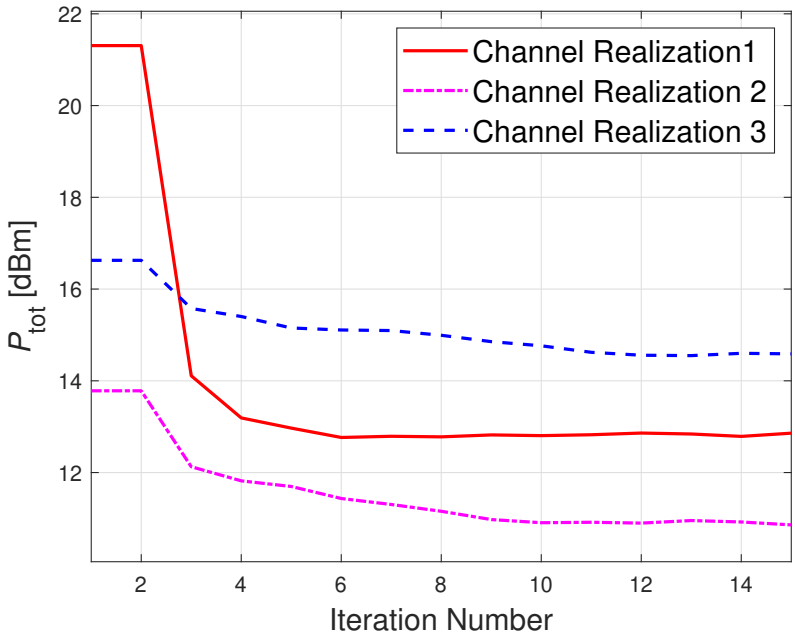

Figure 5.2: Convergence behavior of the AMBF Algorithm ($N = 4, M = 25, G = 2, K = 2, \gamma = 10$ dB, $K_{ik}^{\mathrm{r}} = 0$).

Algorithm 2 starts to converge after $700 \sim 1000$ iterations.

5.3 Performance Comparison for the QoS Problem

We now evaluate the performance of our proposed AMBF algorithm in Algorithm 1. For performance comparison, we consider the following methods:

1. No RIS: A conventional downlink multi-group multicast scenario without the RIS;

2. Random RIS: Apply random phase-shift for the RIS elements. For the QoS problem $\mathcal{P}_o$, with given $\mathbf{e}$, solve the BS multicast beamforming problem $\mathcal{P}_{\mathbf{e}}$ for $\mathbf{w}$;

3. Direct AO [64]: Apply the AO approach directly to $\mathcal{P}_o$, where the subproblem w.r.t. $\mathbf{e}$ is a feasibility problem with the SINR constraints in (3.4b) and the

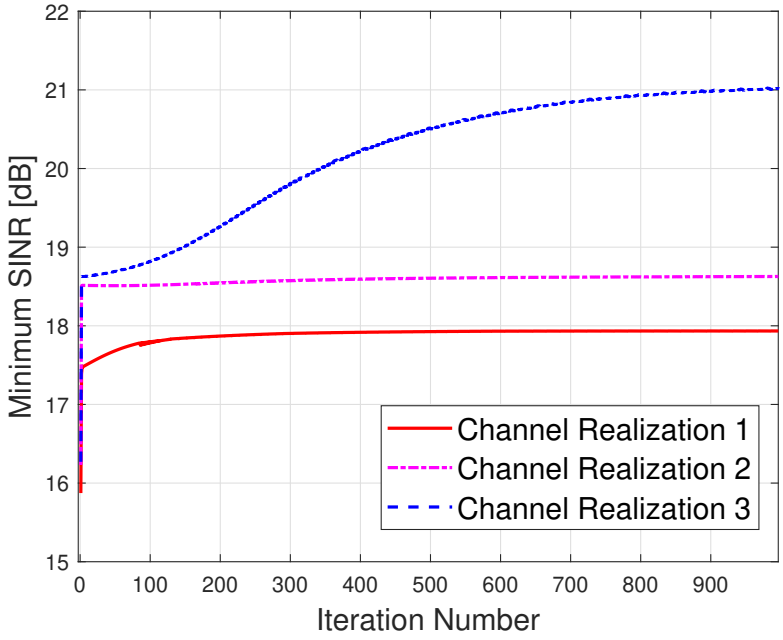

Figure 5.3: Convergence behavior of the Algorithm 2 ($N = 30, M = 100, G = 2, K = 2, \gamma = 10, P = 10$dB dB, $K_{ik}^{\mathrm{r}} = 0$).

unit-modulus constraints in (3.4c);

4. RIS-SDR: Apply Algorithm 1, except that instead of solving subproblem $\tilde{\mathcal{S}}_{\mathbf{w}}$ w.r.t. $\mathbf{e}$, we solve the following relaxed problem of $\mathcal{S}_{\mathbf{w}}^{\mathrm{eq}}$ by dropping the constraint in (3.11c):

$$\max_{\mathbf{e}} \min_{i,k} \quad \frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \tag{5.5a}$$

$$\text{s.t.} \quad |e_m|^2 \leq 1, m \in \mathcal{M}, \tag{5.5b}$$

which can be further expressed as

$$\mathcal{S}_{\mathbf{w}}^{\mathrm{Relaxed}}: \max_{\mathbf{e}} \quad t \tag{5.6a}$$

$$\text{s.t.} \quad \frac{\mathrm{SINR}_{ik}}{\gamma_{ik}} \geq t, \ k \in \mathcal{K}, \ i \in \mathcal{G} \tag{5.6b}$$

$$|e_m|^2 \leq 1, m \in \mathcal{M}. \tag{5.6c}$$

Problem $\mathcal{S}_{\mathbf{w}}^{\mathrm{Relaxed}}$ is solved by SDR with bi-section search over t;

5. RIS-SCA: Similar to RIS-SDR, apply Algorithm 1, except that we solve $\mathcal{S}_{\mathbf{w}}^{\text{Relaxed}}$ by SCA with bi-section search over t. In particular, $\mathcal{S}_{\mathbf{w}}^{\text{Relaxed}}$ is convexified into a convex approximation problem and solved iteratively using a standard convex solver.

5.3.1 Performance over target SINR γ

We show the total transmit power P_{tot} vs. target SINR γ under different methods in Fig. 5.4 for RIS-user channels being Rayleigh fading with $K_{ik}^{\text{r}} = 0, \forall k, i$, and in Fig. 5.5 for RIS-user channels being Rician fading with $K_{ik}^{\text{r}} = 10$. First, for both Figs. 5.4 and 5.5, we see that RIS with random passive beamforming or direct AO on $\mathcal{P}_o$ bring nearly no benefit as compared with the case of no RIS. Thus, RIS passive beamforming needs to be effectively designed to improve the channel conditions among users to enhance the overall system performance. The substantial gain can be seen when the RIS passive beamforming is effectively optimized. In particular, our proposed Algorithm 1 provides about 4 dB power reduction over the no RIS case for $\gamma \geq 4$ dB in Fig. 5.4 and $\gamma \geq 6$ dB in Fig. 5.5. The performance of RIS-SCA is nearly the identical to that of Algorithm 1, while RIS-SDR performs worse than Algorithm 1 or RIS-SCA with a noticeable gap, especially in Fig. 5.4. This is expected, as SDR is an approximation method with a degraded performance when the problem size is relatively large.

Since Algorithm 1 implements the proposed fast algorithms in Section 3.4 to solve subproblems, its computational advantage is demonstrated in Table 5.1, where we provide the average computation time of Algorithm 1, RIS-SCA, and RIS-SDR,

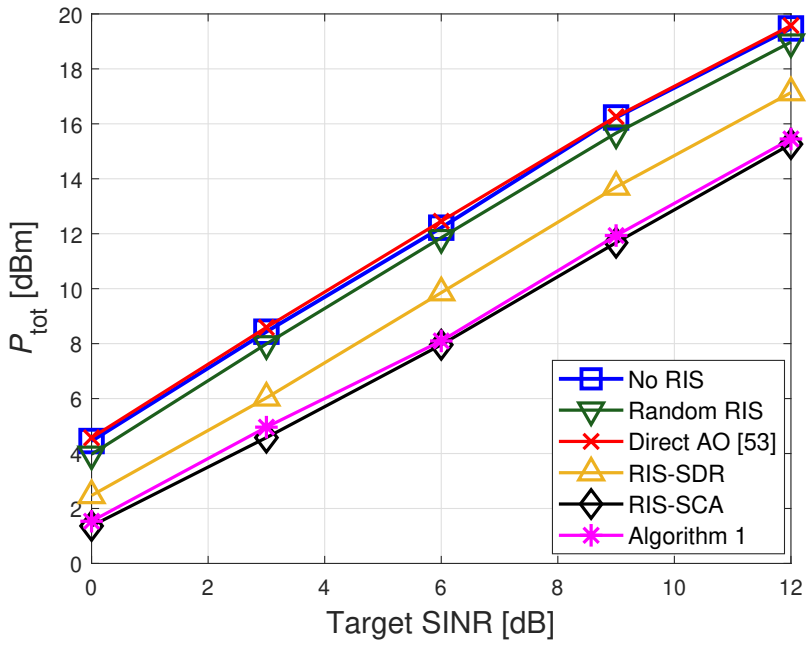

Figure 5.4: Total transmit power P_{tot} vs. target SINR γ ($N = 4, M = 25, G = 2, K = 2, K_{ik}^{\text{r}} = 0$).

as the number of RIS elements M increases from 20 to 100. We see that besides providing the best performance, Algorithm 1 is also the fastest algorithm with much lower computational complexity than the other two methods, especially as the value of M becomes large. We see that the computational complexity of Algorithm 1 only grows mildly with M.

5.3.2 Performance over the number of RIS elements M

In Figs. 5.6 and 5.7, we show the total transmit power P_{tot} as the number of RIS reflective elements M increases under different methods for $\gamma = 10$ dB, where we set $K_{ik}^{\text{r}} = 0$ in Fig. 5.6 and $K_{ik}^{\text{r}} = 10$ in Fig. 5.7. As expected, the transmit power for random RIS and direct AO remains roughly flat without change as M increases, as they are not effectively using RIS. In contrast, Algorithm 1 jointly optimizes the RIS and the BS beamforming vectors, and the resulting transmit power is decreased by

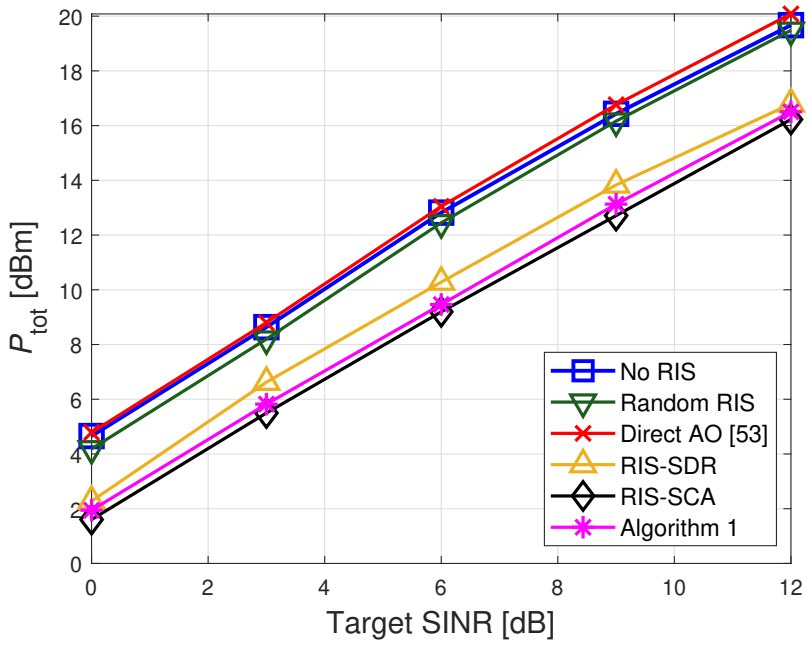

Figure 5.5: Total transmit power P_{tot} vs. target SINR γ ($N = 4, M = 25, G = 2, K = 2, K_{ik}^{\text{r}} = 10$).

4 dB as M increases from 20 to 100. Similar to Figs. 5.4 and 5.5, Algorithm 1 and RIS-SCA have nearly identical performance, while RIS-SDR performs worse with a noticeable gap as large as 2 dB. In particular, compared with Fig. 5.5 with a relatively smaller gap for $M = 25$, we see from Fig. 5.7 that the gap becomes much more noticeable as M increases.

5.4 Performance Comparison for the MMF Problem

We now evaluate the proposed RIS-assisted design for the MMF problem. For performance comparison, besides our proposed Algorithm 2, we consider the following methods:

1. No RIS: A conventional downlink multi-group multicast scenario without RIS;

Table 5.1: Average Computation Time for Different M (sec).

M	20	40	60	80	100
Algorithm 1	4.58	4.80	5.21	5.64	6.38
RIS-SDR	23.52	27.30	38.13	57.02	84.29
RIS-SCA	21.16	25.52	31.84	40.71	51.98

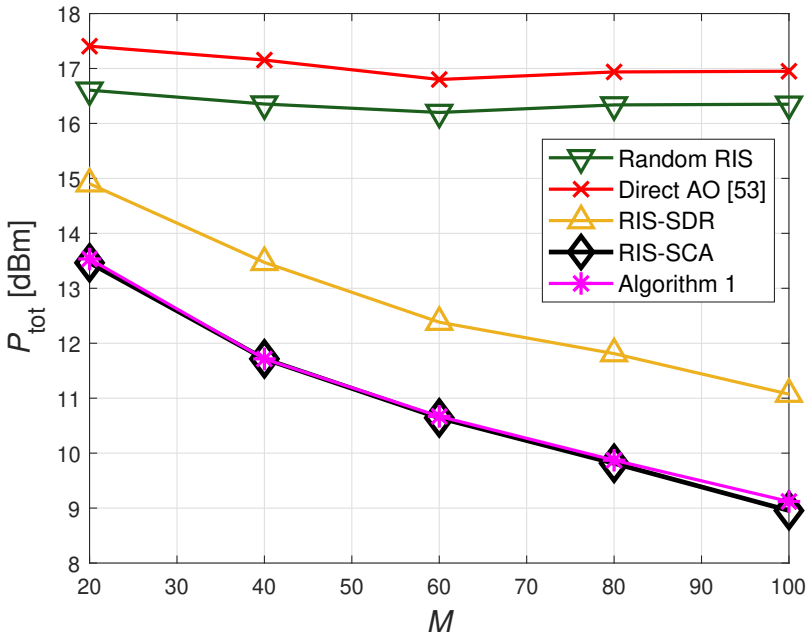

Figure 5.6: Total transmit power P_{tot} vs. M ($\gamma = 10$ dB, $G = 2, K = 2, K_{ik}^{\text{r}} = 0$).

2. Random RIS: Apply random phase-shift for the RIS elements, and with given $\mathbf{e}$, solve the BS multicast beamforming MMF problem for $\mathbf{w}$;

3. Relaxed $\widetilde{\mathcal{S}}_o$: Only solving our proposed relaxed problem $\widetilde{\mathcal{S}}_o$ in (4.6), which uses a penalty term and each e_m is not guaranteed to have a unit modulus. In other words, Algorithm 2 without the final processing step to obtain the feasible solution $\mathbf{e}$. The purpose is to quantify the loss in the final processing step in

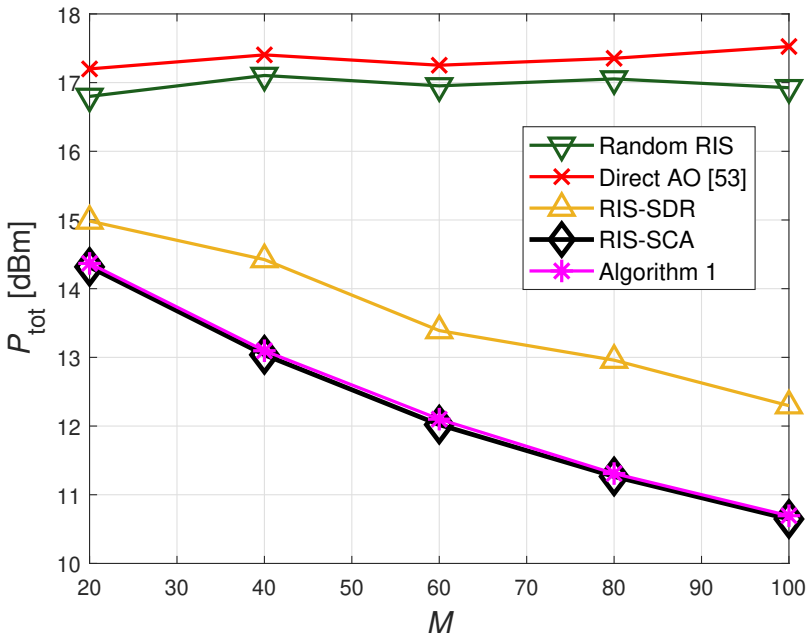

Figure 5.7: Total transmit power P_{tot} vs. M ($\gamma = 10$ dB, $G = 2, K = 2, K_{ik}^{\text{r}} = 10$).

Algorithm 2 and show how good is the relaxed problem as an approximation of the original problem $\mathcal{S}_o$.

4. Alternating projected gradient (APG) method [62]: An alternating projected gradient (APG) method recently proposed in [62] to maximize the achievable sum group-rate of all groups, where the group-rate is defined as the minimum rate among users in the group.

Remark 3. Note that the objective for the APG method in [62] is different from our MMF max-min SINR (or rate) objective. However, for the single group multicast scenario where $G = 1$, the two objectives becomes the same. Since there are no existing methods designed for the RIS-assisted multicast MMF problem, we use this special single-group case to compare Algorithm 2 with the APG method proposed in [62]. Also, we note that besides [62], [61]

also proposed the multicast beamforming algorithms for the sum-group-rate maximization. It is shown in [62] that the APG outperforms the method in [61] in both performance and the computational complexity. Thus, we choose APG in [62] as the current state-of-the-art for comparison.

In the following simulation, we set the BS transmit power $P = 10$ dBm.

5.4.1 Multiple Groups

Fig. 5.8 shows the average minimum SINR vs. M by different methods for $N = 30$. First, we see that Algorithm 2 and relaxed $\widetilde{\mathcal{S}}_o$ provide a nearly identical performance for all values of M. This demonstrates that the solution to $\widetilde{\mathcal{S}}_o$ is mostly feasible to $\mathcal{S}_o$ and thus the final processing step causes a negligible difference. This further indicates that our proposed relaxed problem $\widetilde{\mathcal{S}}_o$ is a good approximation of the original problem $\mathcal{S}_o$. Next, we see that the minimum SINR achieved by Algorithm 2 increases noticeably as M increases, as compared with random RIS, especially when $M > N$. This demonstrates the effectiveness of our proposed algorithm to utilize RIS to improve the channel conditions among users.

We also show the average minimum SINR vs. N under different methods in Fig. 5.9 for $M = 100$ and $K_{ik}^{\mathrm{r}} = 0$. Again, we see that random phase-shift provides little gain over the case of no RIS. Compared with no RIS and random RIS, our proposed Algorithm 2 provides about 2 dB SINR gain for different values of N. In Fig. 5.10, we show the average minimum SINR vs. the number of users K in each group for $N = 30$ and $M = 100$. The gain of Algorithm 2 over random RIS and no RIS is clearly seen for all values of K. The gain is smaller for a larger value of K,

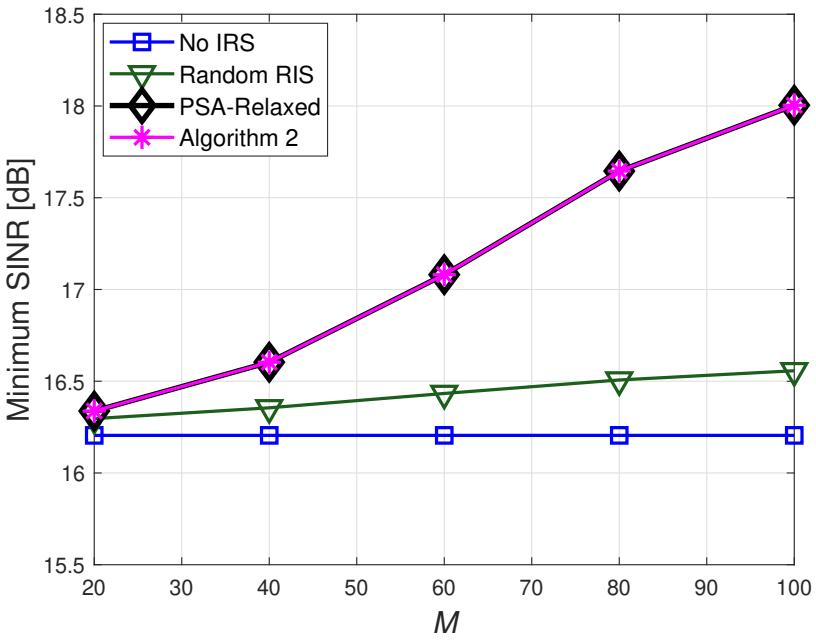

Figure 5.8: Average minimum SINR vs. M ($N = 30, G = 2, K = 2, P = 10\,\text{dB}, K_{ik}^{\text{r}} = 0$).

which is expected because in the multicasting scenario, the beamforming vector needs to cover more users in a larger group, which results in a decreased beamforming gain.

5.4.2 Single Group

We now compare our proposed algorithm with the APG method in [62] for the MMF problem in the single group scenario. Fig. 5.11 shows the average minimum SINR vs. the number o RIS elements M for $G = 1$, $K = 4$, and $N = 30$. We see that our proposed algorithm outperforms APG substantially with 1 dB $\sim$ 2.5 dB gain for M ranging from 40 $\sim$ 100.

Besides the performance, we also compare the average computation time of Algorithm 2 and the APG method Table 5.2 for the setting used in Fig. 5.11. We see that both methods have similar computation time and scale with M similarly. Although

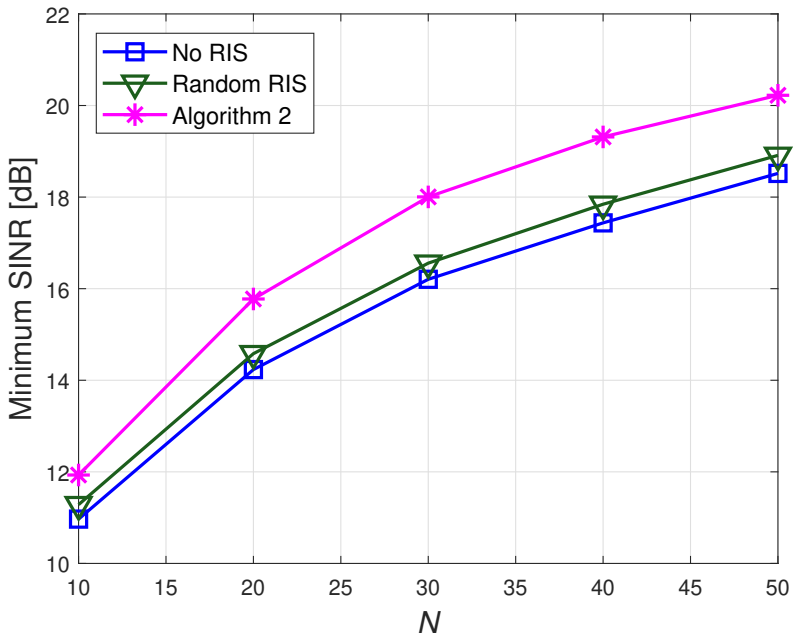

Figure 5.9: Average minimum SINR vs. N ($M = 100, G = 2, K = 2, P = 10\,\text{dB}, K_{ik}^{\text{r}} = 0$).

APG is slightly faster than Algorithm 2, the difference is not significant.

5.4.3 Unicast Scenario

We also compare the performance of Algorithm 2 with the APG method in the unicast scenario for RIS-assisted downlink beamforming. We set $G = 4$ and $K = 1$, which presents a 4-user case. Note that in this special case, Algorithm 2 and APG have different design objectives. APG is essentially to maximize the sum rate of all users, while Algorithm 2 is to maximize the minimum SINR among users. For comparison, we plot the average minimum SINR and the sum rate performance over M in Figs. 5.12 and 5.13, respectively, for $N = 30$ and $K_{ik}^{\text{r}} = 0, \forall k, i$. As we see, for both the minimum SINR and the sum rate, our proposed algorithm outperforms APG.

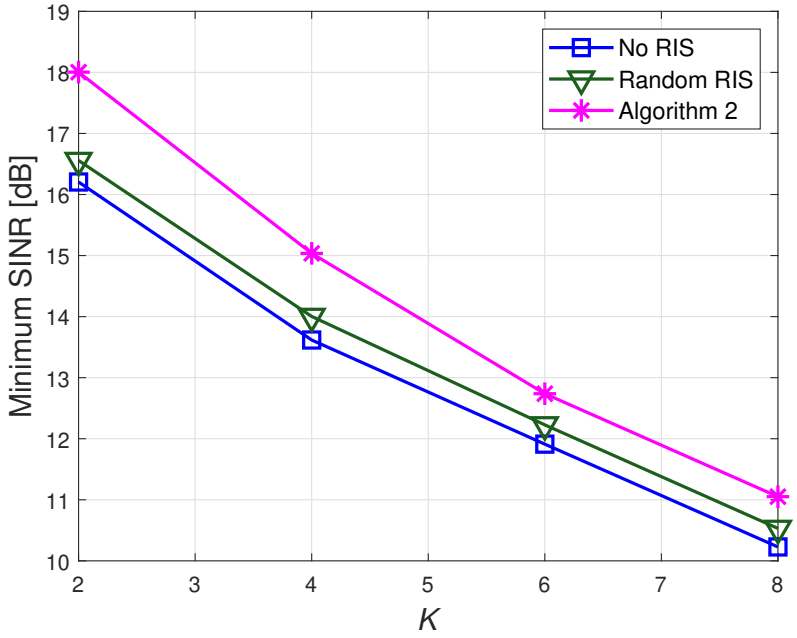

Figure 5.10: Average minimum SINR vs. K ($N = 30, M = 100, G = 2, P = 10\,\text{dB}, K_{ik}^{\text{r}} = 0$).

Table 5.2: MMF-Average Computation Time for Different M (sec).

M	20	40	60	80	100
Algorithm 2	0.43	0.65	0.98	1.41	2.1
APG [51]	0.29	0.47	0.71	1.03	1.61

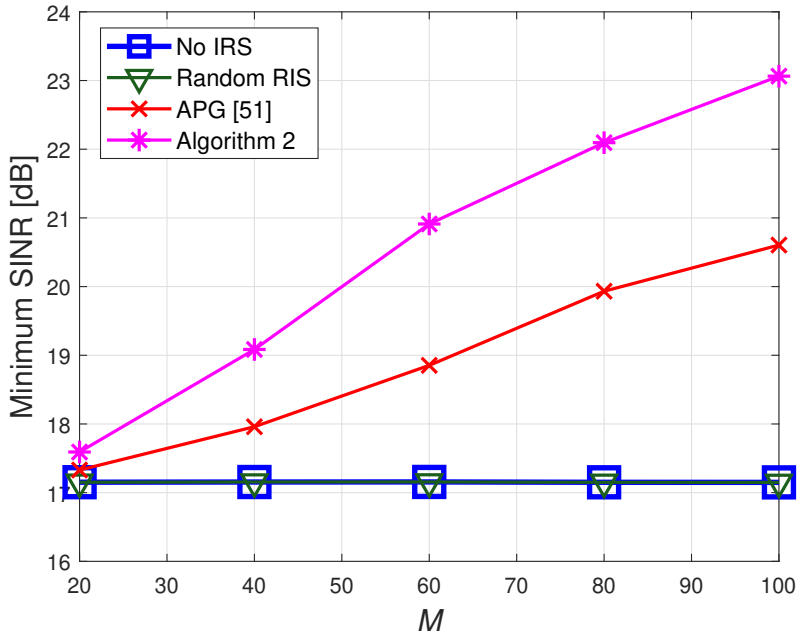

Figure 5.11: Average minimum SINR vs. M ($N = 30, G = 1, K = 4, P = 10\,\text{dB}, K_{ik}^{\text{r}} = 0$).

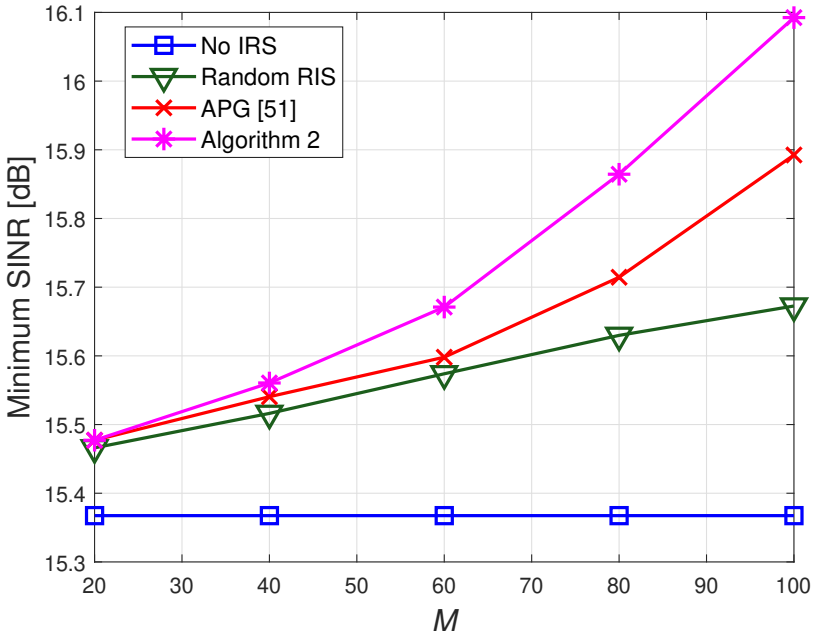

Figure 5.12: Average minimum SINR vs. M ($N = 30, G = 4, K = 1, P = 10\,\text{dB}, K_{ik}^{\text{r}} = 0$).

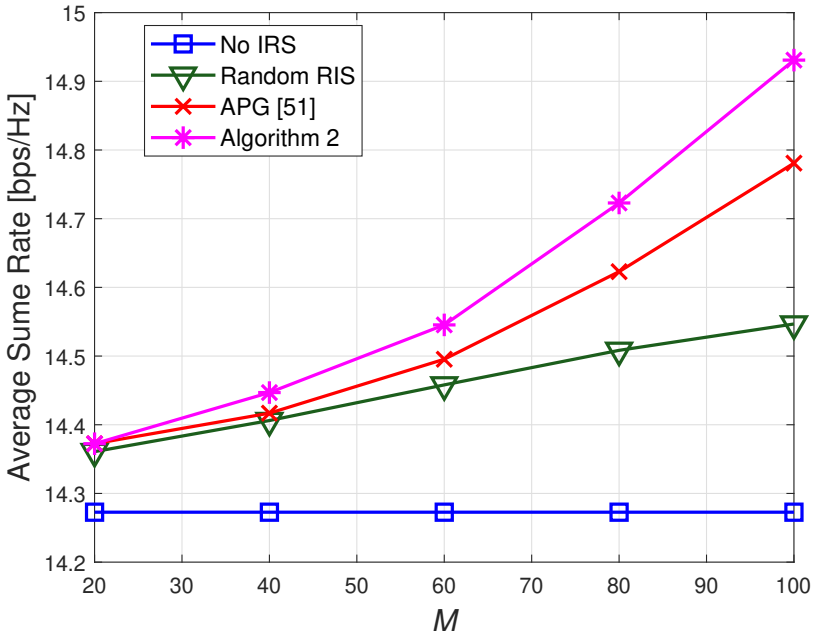

Figure 5.13: Average Sum Rate vs. M ($N = 30, G = 4, K = 1, P = 10\,\mathrm{dB}, K_{ik}^{\mathrm{r}} = 0$).

Chapter 6

Conclusions and Future Work

In this thesis, we consider joint BS and RIS beamforming design for a RIS-assisted multi-group multicast scenario for both QoS and MMF problems. Aiming to provide efficient and scalable algorithms to solve these challenging NP-hard problems: 1) We have proposed a fast alternating multicast beamforming algorithm for the QoS problem to minimize the BS transmit power. We have formulated the BS multicast and RIS passive beamforming subproblems by exploring their interaction in the QoS problem and solve them by the alternating optimization technique. Furthermore, we have proposed to use PSA to solve the challenging RIS beamforming problem with computationally-cheap closed-form updates. 2) We have discovered that like the traditional multigroup multicast problem without the assistance of RIS, here also there is a duality realtionship between QoS and MMF problem. 3) Using this knowledge, we have proposed to solve MMF problem by turning it into a single variable optimization problem and solve it by PSA which as we mentioned is a low-complexity updating procedure. Using the duality realtionship between QoS and MMF problems, we also use the optimal beamforming structure for BS beamforming vector to further reduce the computational complexity of PSA. 4) Simulation results have demonstrated

the advantage of our proposed algorithms in both performance and computational complexity over other numerical-based and alternative methods.

There are multiple potential extensions that can be explored in relation to this study. We proposed our algorithms with the assumption that BS is aware of perfect CSI, but in reality usually it is not the case, and a constant channel estimation process should be performed due to the random nature of the wireless channel and potential movements of users. It is critical to explore the potential of RIS in respect to channel estimation, and learn about the possible improvements that RIS can achieve. Furthermore, it is also important to investigate whether and how to extend our methods to optimize RIS in the multicarrier systems, such as orthogonal frequency division multiplexing (OFDM) systems. Another important missing factor which can be the subject of future studies is finding different methods of RIS phase-shift initialization, for in this thesis usually random initialization has been used for the phase of RIS elements. It is undeniable that in solving non-convex problems such as multicast beamforming using iterative algorithm, the quality of final solution is under the effect of starting point.

Bibliography

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[2] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, C.-L. I, and A. Ghosh, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, 2017.

[3] M.-L. Ku, W. Li, Y. Chen, and K. J. Ray Liu, "Advances in energy harvesting communications: Past, present, and future challenges," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 2, pp. 1384–1412, 2016.

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

[5] M. Dong, W. Li, and F. Amirnavaei, "Online joint power control for two-hop wireless relay networks with energy harvesting," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 463–478, 2018.

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

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

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

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

[10] R. AliHemmati, B. Liang, M. Dong, G. Boudreau, and S. H. Seyedmehdi, "Multi-channel power allocation for device-to-device communication underlaying cellular networks," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2016, pp. 3596–3600.

[11] 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, 2014.

[12] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, 2016.

[13] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[14] N. D. Sidiropoulos, T. N. Davidson, and Zhi-Quan Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, 2006.

[15] E. Karipidis, N. D. Sidiropoulos, and Z. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1268–1279, 2008.

[16] T. Chang, Z. Luo, and C. Chi, "Approximation bounds for semidefinite relaxation of max-min-fair multicast transmit beamforming problem," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3932–3943, 2008.

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

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

[19] 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, 2013.

[20] M. Dong and B. Liang, "Multicast relay beamforming through dual approach," in *2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2013, pp. 492–495.

[21] N. Bornhorst, M. Pesavento, and A. B. Gershman, "Distributed beamforming for multi-group multicasting relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 221–232, 2012.

[22] M. R. A. Khandaker and Y. Rong, "Precoding design for MIMO relay multicasting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3544–3555, 2013.

[23] Y. C. B. Silva and A. Klein, "Linear transmit beamforming techniques for the multigroup multicast scenario," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4353–4367, 2009.

[24] K. Cumanan, L. Musavian, S. Lambotharan, and A. B. Gershman, "SINR balancing technique for downlink beamforming in cognitive radio networks," *IEEE Signal Processing Lett.*, vol. 17, no. 2, pp. 133–136, 2010.

[25] A. H. Phan, H. D. Tuan, H. H. Kha, and D. T. Ngo, "Nonsmooth optimization for efficient beamforming in cognitive radio multicast transmission," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2941–2951, 2012.

[26] K. T. Phan, S. A. Vorobyov, N. D. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via qos-aware secondary multicast beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2323–2335, 2009.

[27] Y. Huang, Q. Li, W.-K. Ma, and S. Zhang, "Robust multicast beamforming for spectrum sharing-based cognitive radios," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 527–533, 2012.

[28] Q.-D. Vu, K.-G. Nguyen, and M. Juntti, "Weighted max–min fairness for c-ran multicasting under limited fronthaul constraints," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1534–1548, 2018.

[29] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Processing Mag.*, vol. 31, no. 4, pp. 142–148, 2014.

[30] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Weighted fair multicast multigroup beamforming under per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 62, pp. 5132–5142, Oct. 2014.

[31] E. Chen and M. Tao, "ADMM-based fast algorithm for multi-group multicast beamforming in large-scale wireless systems," *IEEE Trans. Commun.*, vol. 65, pp. 2685–2698, Jun. 2017.

[32] 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, pp. 114–117, Jan 2014.

[33] 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), 2015, pp. 271–275.

[34] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of data transmission with large intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2746–2758, 2018.

[35] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Aug. 2019.

[36] ——, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Nov. 2020.

[37] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, 2019.

[38] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, Feb. 2020.

[39] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, May 2021.

[40] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, Jan. 2021.

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

[42] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, 1998.

[43] C. Farsakh and J. Nossek, "Spatial covariance based downlink beamforming in an SDMA mobile radio system," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1497–1506, 1998.

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

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

[46] J. Wang, M. Dong, B. Liang, G. Boudreau, and H. Abou-zeid, "Distributed coordinated precoding for MIMO cellular network virtualization," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 106–120, 2022.

[47] 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, 2017.

[48] 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, 2017.

[49] J. Yu and M. Dong, "Low-complexity weighted MRT multicast beamforming in massive MIMO cellular networks," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, 2018, pp. 3849–3853.

[50] J. Yu and M. Dong, "Distributed low-complexity multi-cell coordinated multicast beamforming with large-scale antennas," in *Proc. IEEE Workshop on Signal Processing advances in Wireless Commun.(SPAWC)*, 2018, pp. 1–5.

[51] M. Dong and Q. Wang, "Multi-group multicast beamforming: Optimal structure and efficient algorithms," *IEEE Trans. Signal Process.*, vol. 68, pp. 3738–3753, May 2020.

[52] C. Zhang, M. Dong, and B. Liang, "Ultra-low-complexity algorithms with structurally optimal multi-group multicast beamforming in large-scale systems," *IEEE Trans. Signal Process.*, 2023.

[53] 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 Process.*, vol. 56, pp. 1268–1279, 2008.

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

[55] M. Fu, Y. Zhou, and Y. Shi, "Intelligent reflecting surface for downlink non-orthogonal multiple access networks," in *2019 IEEE Globecom Workshops (GC Wkshps)*, 2019, pp. 1–6.

[56] H. Xie, J. Xu, and Y.-F. Liu, "Max-min fairness in IRS-aided multi-cell MISO systems via joint transmit and reflective beamforming," in *Proc. IEEE Int. Conf. Communications (ICC)*, Jul. 2020, pp. 1–6.

[57] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Asymptotic max-min sinr analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7748–7764, 2020.

[58] H. Yu, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "Joint design of reconfigurable intelligent surfaces and transmit beamforming under proper and improper gaussian signaling," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2589–2603, 2020.

[59] L. Du, S. Shao, G. Yang, J. Ma, Q. Liang, and Y. Tang, "Capacity characterization for reconfigurable intelligent surfaces assisted multiple-antenna multicast," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6940–6953, May 2021.

[60] Q. Tao, S. Zhang, C. Zhong, and R. Zhang, "Intelligent reflecting surface aided multicasting with random passive beamforming," *IEEE Wireless Commun. Letters*, vol. 10, no. 1, pp. 92–96, Sep. 2021.

[61] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3236–3251, 2020.

[62] M. Farooq, V. Kumar, M. Juntti, and L.-N. Tran, "On the achievable rate of IRS-assisted multigroup multicast systems," in *Proc.IEEE Global Telecommn. Conf. (GLOBECOM)*, Dec. 2022, pp. 5844–5849.

[63] F. Shu, G. Yang, and Y.-C. Liang, "Reconfigurable intelligent surface enhanced symbiotic radio over multicasting signals," in *Proc. IEEE Vehicular Technology Conf. (VTC)*, 2021, pp. 1–6.

[64] D. Li, Q. An, Y. Shi, and Y. Zhou, "Multigroup multicast transmission via intelligent reflecting surface," in *IEEE Trans. Veh. Technol.*, Feb. 2020, pp. 1–6.

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

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