MANCHESTER
1824
The University of Manchester

THE APPLICATIONS OF RECONFIGURABLE INTELLIGENT SURFACE IN NON-ORTHOGONAL MULTIPLE ACCESS NETWORKS

A thesis submitted to the University of Manchester for the degree of

Doctor of Philosophy

in the Faculty of Science and Engineering

2024

Tianqi Wang

Department of Electrical and Electronic Engineering

Contents

Word count: 61440

List of figures

List of tables

List of publications

T. Wang, F. Fang and Z. Ding, ”An SCA and Relaxation Based Energy Efficiency Optimization for Multi-User RIS assisted NOMA Networks” published in *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6843 - 6847, Mar. 2022. DOI: 10.1109/TVT.2022.3162197.

T. Wang, F. Fang and Z. Ding, ”Joint Phase Shift and Beamforming Design in a Multi-User MISO STAR-RIS Assisted Downlink NOMA Network” published in *IEEE Transactions on Vehicular Technology*, vol. 72, no. 7, pp. 9031 - 9043, Feb. 2023. DOI: 10.1109/TVT.2023.3248789.

T. Wang, F. Fang, Y. Sun and Z. Ding, ”Joint Beamforming and Phase Shift Design in RIS Assisted CoMP-NOMA Systems” submitted in *IEEE Transactions on Vehicular Technology*.

T. Wang, F. Fang and Z. Ding, ”Beamforming and Phase Shift Design in STAR-RIS Assisted ISAC NOMA Networks” in preparation.

Terms and abbreviations

5G	The Fifth Generation Wireless Communications
6G	The Sixth Generation Wireless Communications
AP	Access Point
AR	Augmented Reality
AWGN	Additive White Gaussian Noise
B5G	Beyond 5G Wireless Communications
BackCom	Backscatter Communications
BS	Base Station
CDMA	Code Domain Multiple Access
CSI	Channel State Information
D2D	Device to Device
DDPG	Deep Deterministic Policy Gradient
DL	Deep Learning
DRL	Deep Reinforcement Learning
EE	Energy Efficiency
EVD	Eigenvalue Decomposition
FDMA	Frequency Division Multiple Access

FP	Fractional Programming
IoT	Internet of Things
ISAC	Integrated Sensing and Communications
KKT	Karush-Kuhn-Tucker
LoS	Line of Sight
LP	Linear Programming
LTE	Long Term Evolution
MA	Multiple Access
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Machine Learning
MR	Mixed Reality
NOMA	Non-orthogonal Multiple Access
nLoS	non Line of Sight
OFDMA	Orthogonal Frequency Division Multiple Access
OMA	Orthogonal Multiple Access
PSD	Positive Semidefinite
QCQP	Quadratically Constrained Quadratic Programming
QoS	Quality of Service
QP	Quadratic Programming
RIS	Reconfigurable Intelligent Surface

RL	Reinforcement Learning
SCA	Successive Convex Approximation
SDMA	Space Division Multiple Access
SDP	Semidefinite Programming
SDR	Semidefinite Relaxation
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal Noise Ratio
SROCR	Sequential Rank-one Constraint Relaxation
STAR-RIS	Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surface
SVM	Support Vector Machine
TDMA	Time Division Multiple Access
VR	Virtual Reality
XR	Extended Reality

Abstract

The development of six generation (6G) wireless communication technology is becoming increasingly imminent to achieve internet of everything (IoE) in today's digital era. Non-orthogonal multiple access (NOMA) is the pivotal technology in the upcoming 6G networks due to its high spectrum efficiency. The reconfiguable intelligent surface (RIS) has been widely acknowledged in both industry and academia since it is able to improve the quality of channels. Based on RIS, simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS) can both transmit and reflect the signals and achieve full space communications. Meanwhile, the integrated sensing and communications (ISAC) is also widely considered as a paradigm-shifting approach to combine the sensing and communications into an unified framework in 6G.

This thesis proposes four different communication frameworks based on NOMA scheme including a multi-user RIS assisted downlink NOMA network, multi-cluster STAR-RIS assisted downlink NOMA network, RIS assisted coordinated multi-point (CoMP) NOMA network and STAR-RIS assisted ISAC NOMA network. In the first network, an energy efficiency (EE) maximization problem is formulated by jointly designing the beamforming of the users and phase shifts of the RIS elements. Alternating optimization and relaxation are utilized to find the suboptimal solutions of original problem. In the second network, an EE maximzation problem is also formulated by jointly optimizing the phase shifts and beamforming. In particular, sequential rank-one constraint relaxation is proposed to tackle the rank-one constraints in phase shift semidefinite relaxation (SDR) problem. In the third network, the sum rate maximization problem of NOMA users is formulated. Specifically, manifold optimization is applied to tackle the phase shift problem since the norm constraints of RIS elements can be considered as a manifold. In the fourth network, a sum rate maximization problem of multiple primary users is formulated in an ISAC case. A penalty-based successive convex cancellation (SCA) is leveraged to solve the rank constraints in both beamforming and phase shift SDR problems.

Declaration of originality

I hereby confirm that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

i The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made *only* in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see `http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420`), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see `http://www.library.manchester.ac.uk/about/regulations/`) and in The University's policy on Presentation of Theses.

List of thesis revisions

Put list of revisions here. Only required for resubmitted theses. Delete if not needed

Acknowledgements

First of all, please allow me to deliver my great gratitude to my supervisor, Prof. Zhiguo Ding for his professional guidance, unique views and selfless support throughout my Ph.D. journey. His expertise, patience, and encouragement always motivate me to make progress on research step by step. Meanwhile, his enthusiasm and rigor on research also have a profound impact on me and give me the courage to never give up at any time. It is a great fortune that I could have such a meaningful and unforgettable experience under the supervision of Prof. Ding.

Besides, I would like to express my gratitude to all the current and past members in the group at the University of Manchester. They include Dr. Kaidi Wang, Dr. Haodong Li, Dr. Shiyu Jiao, Dr. Yunus Dursun, Dr. Ximing Xie, Mr. Burak Goktas, Mr. Dingjia Lin, Mr. Yushen Lin and Ms. Wenqi Huang. Especially, I sincerely appreciate Dr. Fang Fang and Dr. Yanshi Sun for offering me deep and precious insights and suggestions on my research throughout my Ph.D. studies. Their kind efforts enable me to rapidly overcome the technical difficulties and master the research skills.

Last but not least, I would like to express my deep gratitude to all my family members. They include my paternal grandparents, Mr. Zuofu Wang and Mrs. Chengyun Wang and my maternal grandparents, Mr. Zhongyu Fan and Mrs. Ruolan Fan. Especially, I would like express my deepest appreciation to my parents, Mr. Xin Wang and Mrs. Wei Fan for their eternal love. Their unconditional support, accompaniment and dedication have driven me to bravely face all the difficulties in my life.

The author

If desired, a brief statement for External Examiners giving the candidate's degree(s) and research experience, even if the latter consists only of the work done for this thesis.

Chapter 1

Introduction

1.1 Overview

1.1.1 The Era of Future 6G Technique

The mobile communication system evolves on the average of every ten years. Nowadays, The application value of fifth generation (5G) has widely exploited in academia, industry and even medicine [1]. Besides, 5G wireless technology has successfully led humans to completely move from the era of 'connected people' to 'connected everything' [2]. The development of the sixth generation (6G) wireless techniques is predicted to be a significant step forward in the area of telecommunications. Based on the existing progress on the 5G technique, it is promising that 6G can radically lead our interaction with technology to a more digital one and thus facilitate the transformation from 'connected everything' to 'connected intelligence' [2]. The motivations of promoting 6G are formulated in two folds as followed:

Firstly, 6G empowers plenty of emerging applications. Based on the development of the emerging applications such as augmented reality (AR) and virtual reality (VR), mixed reality (MR) integrates the characteristics of AR and VR techniques and enables the interactions between the virtual and real worlds. For instance, Apple Vision Pro, as an MR headset device developed by Apple Inc. in 2023 has gained great success as a 'spatial computer' in the field of MR and it can provide users with immersive experience from work to entertainment [3]. However, the performance of these applications is severely constrained by the limitations of 5G technology. Differently, 6G can support the connections of the exponential increasing number of high traffic consumption devices and guarantee their high quality performance

due to its faster data speeds, lower latency, higher reliability and capacity than 5G [4]. Furthermore, the cutting-edge concepts such as the metaverse, blockchain and digital twins, which are difficult to be implemented in 5G era, will become possible [5]–[7].

Secondly, 6G empowers intelligence [8]. With the advent of the internet of things era, big data aided AI and machine learning (ML) are bounded to be widely applied and provide solutions across data privacy, reaction speed and stability of AI. In fact, 6G communication systems have been tailor-made for AI and ML. In the architecture of 6G, edging computing and distributed AI will be the main trend to protect the data privacy [9]. It is because that each device's data is employed to train the local model and will not be sent to the server. It is the parameters of local models shared to the central servers, thereby avoiding the direct data breach [2]. Besides, different from centralized AI, distributed AI can process the data in its creation point, thereby decreasing latency and increasing response speed. Furthermore, even if a particular node unfortunately suffers from failure, the overall system can still keep operating, enhancing the robustness and reliability of AI networks [10], [11].

In summary, with the further studies about next generation wireless communication networks, 6G technology is expected to move beyond the realm of wireless communications. In the future, it is promising that 6G will continuously lead humans to a more digital and intelligent world and provide solid supports for intelligence of everything.

1.1.2 Multiple Access Techniques

Multiple access (MA) is the key technique of wireless communications. It allows multiple users to transmit signals in the several communication resources including frequency, time and code etc.

Reviewing the history of multiple access, frequency division multiple access (FDMA) is one of the earliest MA techniques and adopted in the first generation (1G) wireless networks [12], [13]. By splitting the overall bandwidth resource, each user can communicate in its independent frequency band.

With the development of the digital communications, time division multiple access (TDMA) is the key technique of the second generation (2G). TDMA splits time into multiple time slots and each user can communicate within its allocated time slot. TDMA allows users to communicate alternatively in the same bandwidth, improving the spectrum efficiency.

Code division multiple access (CDMA) adopts different techniques from FDMA and TDMA [14]. It allows all the users to communicate in the same bandwidth. users' signals are distinguished by using different code sequences. Similar to TDMA, CDMA can still improve the spectrum efficiency and it is widely used in the late 2G and the third generation (3G) wireless communications.

Orthogonal frequency division multiple access (OFDMA) is the MA which the fifth generation communication technique applies. In fact, OFDMA is the extension of FDMA [15]. Different from FDMA, the entire bandwidth is divided into multiple subcarriers, which are orthogonal with each other. Therefore, there is no interference between subcarriers even though there are closely spaced. However, in FDMA, there is a guard band between two adjacent bandwidths and its spectrum efficiency is far less than that of OFDMA [15], [16]. Nowadays, OFDMA is widely used in the long term evolution (LTE), fourth generation (4G) and 5G networks.

Non-orthogonal multiple access (NOMA), as an emerging MA technique, is widely considered as the key of next generation wireless communication systems [17], [18]. NOMA allows each user to share the same bandwidth. To facilitate the best performance of NOMA, successive interference cancellation (SIC) is applied to remove the interference from other users and extract the target signals according to users' decoding capability [19], [20]. NOMA can have a better performance than FDMA and OFDMA in improving the spectrum efficiency and enable more devices to be connected to the networks due to its high capacity [21], [22]. Although several technical challenges remain including power allocation and complexity of SIC, as the development of application of deep learning and other advanced techniques, these issues are expected to be solved in the future.

1.1.3 Reconfigurable Intelligent Surface

Reconfigurable intelligent surface (RIS) is developed to overcome several tough challenges such as poor environment of high frequency channels, high power consumption of 5G/6G hardware and passive adaption to the environment [23]. The electromagnetic characteristics of the elements on RIS can be smartly tuned by programmable controllers. Thus, the smart radio environment (SRE) can be constructed. Besides, RIS is naturally a low cost and power consumption device, significantly practicing the philosophy of green communications.

Furthermore, RIS also has the low complexity for installation and deployment [24].

The development of RIS is also not confined to the academic studies. Since 2020, the project implementation of RIS has been carried out by both academia and industry. In September 2020, ZTE Corp. and China Unicom Ltd. joined forces with more than twenty universities and established the 'RIS Research Project' in China Communications Standards Association (CCSA). In September, 2020, ZTE Corp., China Unicom Ltd. and Southwest University, China jointly held the conference '1st Reconfigurable Intelligent Surface Technology Forum' [25]. In September 2023, China Telecom Ltd. and ZTE Corp. jointly employed dynamic RIS to provide a seamless 5G connections for all attendees in the 19^{th} Hangzhou Asian Games [26]. This showcases the great potential of RIS in promoting the smart communications in major events.

Nowadays, there has been significant progress on the materials, theories and applications of RIS. As a potential direction of future 6G, it is convincing that RIS has the opportunity to empower 6G and the large-scale implementation and commercialization of RIS can be realized in the near future.

1.1.4 Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surface

Although RIS has gained great reputations in the field of wireless communications, its inadequacy is becoming increasingly obvious. It is noted that RIS can only serve the users at the front of RIS since it can only reflect the incident signals. However, if users are located at the back side of RIS, it will be difficult for RIS to cover this blind zone. Thanks to the properties of metasurfaces [27], simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS) is developed to effectively compensate for the deficiencies of the limited coverage. Specifically, STAR-RIS can both reflect and transmit the incident signals to achieve the full space coverage[28]. As the advancing manufacturing process of metasurfaces, it is convincing for STAR-RIS to be a reliable alternative of RIS in some cases and promote the progress of 6G.

1.1.5 Coordinated Multi-Point Technique

Coordinated Multi-Point (CoMP) technique is crucial in cell-edge wireless communications and it can improve the communication quality and stability. especially in LTE and 5G [29]. The original idea of CoMP is to apply multiple BSs to serve the users which are located at the edge of cells. The advantages can be included in two aspects. Firstly, the edge users are easily affected by the interference from multiple cells. When BSs jointly serve the edge users, the quality of their signals can be greatly enhanced. This enables edge users to have better services by the joint effort of BSs [30], [31]. Meanwhile, by dynamic power allocation or beamforming designs, the interference are likely to be mitigated to improve the total performance of the cell-edge networks. Based on the above points, CoMP transmission can not only improve the rate of the cell users but enable more edge users to be connected, thereby enhancing the capacity of the networks. It is convincing that the applications of CoMP technique will become the mainstream in cell-edge systems.

1.1.6 Integrated Sensing and Communications

It is noted that larger scale of antenna arrays can both enable massive connections of the users and improve the sensing accuracy of radars. Therefore, in some aspects such as hardware and signal processing, sensing systems are gradually resembling the communication systems with a large scale of antenna arrays. Therefore, as massive MIMO becomes widespread, the boundaries between sensing and communications will be smaller and smaller. Actually, current technology enables radar systems and communication systems to coexist in the same bandwidth. For instance, the L-band (1-2 GHz) is used for the functionality long-range air-surveillance radars, i.e., air traffic control (ATC) radars [32]. Meanwhile, this band is also used for 5G new radio (NR) and frequency division duplexing - long term evolution (FDD-LTE) cellular systems [32]. Although coexistence makes sense in practical scenarios, it does not realize the sharing of bandwidth between radars and communications.
Thanks to the technical progress such as MIMO, advanced signal process and software defined radio, integrated sensing and communications (ISAC) is considered as a feasible technology to make radars and communications share the same bandwidth, effectively improving the spectrum efficiency [33]. Meanwhile, ISAC can become an economical

strategy. On the one hand, since the bandwidth is shared by both sensing and communication systems, the cost for the usage of bandwidth can be effectively reduced. On the other hand, the sensing and communication functionalities can be realized by the same set of hardware, meaning that the investments on hardware can be saved and the cost of energy consumption can be greatly decreased. In the future 6G era, ISAC will be more likely to be embedded into the NOMA networks and jointly improve the performance of sensing and communications.

1.1.7 Convex Optimization Theory

Convex optimization is an important branch in the field of optimization. Its development spans several centuries in history [34]. But the formation of this theory mainly occurred in late 20 century [35].

In 1917, the first book on optimization 'Maxima and Minima' was published by several mathematicians. In 1939, the theory of linear programming (LP) was proposed and it was used to determine how to obtain the optimal outcome. In the same year, William Karush first proposed the concept, which later became known as Karush-Kuhn-Tucker (KKT) conditions in his master dissertation. However, it did not attract enough attentions in academia at that time.

In 1950s, the concepts such as convex sets and convex functions were proposed, which provided mathematical foundations for solving the optimization problems. Meanwhile, in 1951, Harold W. Kuhn and Albert W. Tucker jointly published a paper, which comprehensively introduced several conditions and applied them to nonlinear programming. These conditions were later named as Karush-Kuhn-Tucker conditions for the contributions of these three researchers.

In 1960s, researchers started to systematically study convex optimization, including the properties of convex functions, dual theory and conditions of optimization. At this stage, convex optimization officially became the branch of mathematics.

1970s-1980s witnessed the rapid development of convex optimization, especially the development of the algorithms such as interior point method [36], marking great progress on solving large-scale convex optimization problems.

In 21 century, not only convex optimization can promote the progress on plenty of subjects such as operations research, economics and electrical engineering [37], [38]. As the

development of computer science and data science, convex optimization theory also plays a significant role in the development of machine learning and artificial intelligence. For instance, the classic machine learning techniques such as support vector machine (SVM) and principal component analysis [39] are derived from the convex optimization theory. Besides, convex optimization theory is also the basis of deep learning. Backpropagation algorithm as well as other deep learning algorithms [40] based on gradient descents have been wide-acknowledged tools in training neural networks and even large models.

1.2 Literature Review

Firstly, inspired by the great application values of the RIS and NOMA, RIS assisted NOMA networks have been intensively studied by plenty of works. For example, in [41], the energy efficiency maximization problem was formulated based on the RIS assisted downlink system with two-user case, where the beamforming optimization problem and phase shift optimization problems were alternatively optimized. The successive convex approximation (SCA) and semidefinite relaxation (SDR) were applied as the mathematical tools to tackle the optimization problems. In [42], the authors proposed an RIS UAV-NOMA downlink network to achieve the sum rate maximization in two-user case. SCA and SDR were also applied to tackle the optimization of beamforming vectors and phase shifts. Besides, [43] and [44] also studied the power consumption minimization problem in an RIS assisted multi-cluster system, where the concept of central and edge users were proposed. Particularly, [44] applied a novel algorithm based on second-order cone programming (SOCP) and alternating direction method of multipliers (ADMM) to optimize the beamforming of the users. Furthermore, machine learning, especially deep deterministic policy gradient (DDPG) in reinforcement learning was also an useful mathematical tool in solving the optimization problems related to RIS assisted NOMA systems [45]–[47].

Secondly, as the studies of RIS have been comprehensive, the potentials of STAR-RIS assisted NOMA systems have also gradually been studied by researchers. In [48], the coverage characterization of two-user STAR-RIS networks was discussed and compared between the NOMA and OMA cases. The key idea of this work was to apply KKT conditions to transform the non-convex decoding order constraint to a linear one. Besides,

the authors in [49] proposed a joint power allocation and beamforming design by applying various mathematical tools, i.e., SCA, SDR and alternating optimization (AO). Moreover, the authors in [50] also proposed the STAR-RIS aided NOMA and OMA networks. In this work, SDR, convex upper bound approximation, and geometry programming were used to optimize the decoding order, beamforming coefficient vectors and power allocations to achieve the sum rate maximization of the network. Similarly, DDPG was also applied to design the efficiency in a STAR-RIS NOMA network [47].

Thirdly, reviewing the previous works related to CoMP technique, it can be found that conventional CoMP technique is originally developed based on the orthogonal multiple access, which could result in the low spectrum efficiency as mentioned above. To solve this shortage, plenty of works have been completed to exploit the networks by combining CoMP and NOMA together. In [51], the authors developed the closed-form solutions for the beamforming vectors of the power consumption minimization problem in a CoMP-NOMA system. In [52], the authors exploited two power allocation schemes, namely joint power optimization (JPO) and distributed power optimization (DPO), respectively. The spectrum efficiency and energy efficiency of these two schemes were also both compared and evaluated in the simulation results. In [53], the relay selection and power allocation scheme were studied to improve the energy efficiency in an UAV-assisted CoMP NOMA system, which two schemes are based on Lagrange multiplier method and machine learning, respectively. The authors in [54] also designed a novel CoMP-empowered C-NOMA network to maximize the sum rate of the cell-edge users by power allocation. Since great progress has been made in the fields of RIS assisted NOMA and CoMP-NOMA systems, the RIS assisted CoMP-NOMA systems are becoming increasingly potential in further improving the performance of the system. In [55], the power minimization problem was studied in an RIS assisted CoMP-NOMA network. In this work, A novel framework with group-level SIC (GSIC) was constructed and it could effectively remove the decoded group's interference. In [56], the maximum sum rate was achieved by jointly optimizing power allocation, user clustering and phase shifts. In this paper, the closed-formed solutions of the power allocation were obtained and user clustering was solved by using iterative Hungarian method. Meanwhile, the SCA-based semidefinite relaxation was leveraged to tackle the phase shifts of the elements.

Fourth, RIS assisted ISAC NOMA, as an emerging research direction, is attracting increasing attentions from researchers. So far, many studies on RIS assisted ISAC NOMA have been

made. In [57], the authors built an RIS assisted ISAC NOMA network with multiple users and one sensing target and maximized the sum secrecy rate by using SDR, SCA-based iterative scheme and Gaussian randomization. Besides, the authors in both [58] and [59] studied the minimum beampattern gain (MBPG) maximization problems in RIS assisted ISAC NOMA networks. The difference is that the authors [58] applied SDR based fixed point iterative method to jointly optimize the power allocation and passive beamforming and introduced SROCR to solve the rank one issue while authors in [59] only applied SCA-based fixed point iterative method to jointly optimize the beamforming and phase shifts.

1.3 Motivations of the Thesis

This thesis is motivated by a promising technique NOMA which can significantly improve the spectrum efficiency and further enhance the performance metrics of networks. Besides, it is also promising that NOMA can become one of the scenarios of next generation multiple access (NGMA). Therefore, it is necessary and essential to accumulate sufficient research findings and lay a solid foundation on NOMA before the widespread deployment of 6G in around 2030. However, the potentials of NOMA cannot be fully realized without the assistance of other techniques. For example, 6G is likely to apply higher bandwidth with the shorter wavelength than 5G. While further improving the spectrum efficiency, rate and capacity of systems, the defects of its wavelength are more evident. It is because propagation distance of signals is shorter and the capability of wave of 6G is even worse to penetrate the obstacles than 5G. Following the conventional schemes, more BSs, access points or relays should be deployed to guarantee quality of communications, which is high cost both in hardware and power consumption. RIS is an innovative technique to effectively solve these issues. Firstly, RIS is a low-cost device and it is easy to be equipped. Secondly, RIS can effectively enhance the status of channels and improve the quality of communications by tuning the phase shifts of incident signals instead of adopting amplification of the signals to compensate for the path loss during the propagation. Besides, due to the simple installation of RIS, it can be deployed in more flexible positions, i.e., building facades and lifts to serve the mobile users in service dead zone. The above statements strongly demonstrate the research values of RIS on applications.

However, RIS can only reflect the signals and achieve the half space communications, which greatly restricts the application scenarios of RIS. STAR-RIS, as an extension of RIS, can effectively solve the coverage issues. It is because that STAR-RIS is equipped with both transmitting and reflecting elements. These elements allow the incident signals to transmit and reflect from the surface and achieve the full-space coverage. For RIS, which has been progressively matured in academia and applied to a certain extent in industry, studies related to STAR-RIS is also imminent to provide more insights for future development of STAR-RIS.

Meanwhile, some previous techniques, which have been adopted in the old multiple access is still useful in NOMA. For example, CoMP technique. CoMP was originally designed for the edge users in the cell-edge systems. In CoMP networks, the edge users are jointly served by multiple users. In this way, the power of the edge users' signals can be effectively increased, guaranteeing the communication quality of edge users. Besides, by dynamic control at the side of BSs, the bandwidth resource and power can be allocated to users in a more flexible way, improving the spectrum efficiency. In fact, CoMP is also useful in NOMA networks. By applying CoMP in NOMA, the communication quality can be enhanced to a higher level than CoMP OMA networks. Besides, the spectrum efficiency can be further improved since the entire bandwidth resource is shared by all the cell and edge users. Nowadays, the concept of RIS assisted CoMP NOMA has gradually attracted interests in academia and the number of research findings is still less in quantity. Therefore, the research on RIS assisted CoMP NOMA is worth to be carried out to fulfill this blank area.

ISAC is also an important framework in 6G. Its ultimate goal is to achieve the unification of sensing and communications. In the past, sensing and communication systems operated in two independent frequency bands or alternatively in the same bandwidth. However, thanks to the further development of massive MIMO in 6G, sensing and communication systems are able to share the same bandwidth and hardware and both these two functionalities can be realized simultaneously. The framework of ISAC can save additional cost on hardware and power consumption since only one set of hardware for both sensing and communications is needed. In the future, ISAC will become the crucial technique in 6G as the basis of some emerging technologies such as autonomous vehicle and high accurate localization. RIS/STAR-RIS assisted NOMA networks can also contribute to the performance of ISAC. The large-scale RIS/STAR-RIS performs the similar functionalities of massive MIMO

and it can be used to detect the sensing targets and communicate with the mobile users in the service dead zone and enhance the performance of ISAC by tuning the phase shifts of RIS/STAR-RIS elements. Particularly, it is noted that the studies related to STAR-RIS assisted ISAC-NOMA networks are also less in quantity. Therefore, it is potential for researchers to obtain more novel insights by deeply exploiting this field.
In communication systems, sum rate is an important measure since it reflects the communication quality of users. Especially, some emerging techniques such as AR and VR devices should be driven with the support of higher transmission rates. However, recently, a new measure, energy efficiency (EE) becomes valued by plenty of organizations especially telecommunication companies. It is because blindly pursuing the rate might lead to high cost not only in power consumption but also hardware investments, which is not economical. EE, defined as the ratio between sum rate and power consumption can be a measure to evaluate the balance between sum rate and power consumption. Therefore, not only sum rate, but EE should also be sufficiently studied in the communication systems.

In summary, motivated by the illustrations above, the thesis conducts the research on RIS/STAR-RIS assisted NOMA and is aimed to find more research values of the applications of RIS in NOMA networks. They include EE maximization in multi-user RIS and multi-cluster downlink STAR-RIS NOMA networks and sum rate maximization in RIS CoMP NOMA and STAR-RIS ISAC NOMA networks. EE and sum rate are the optimization objectives among these networks. To achieve the best performance of the networks, the high efficiency methods and algorithms are developed to design the beamforming and phase shifts of the RIS/STAR-RIS elements. Besides, the performance of proposed schemes is also compared to several benchmark schemes, i.e., OMA to demonstrate the crucial effect of RIS/STAR-RIS and NOMA on further improving the performance of networks.

1.4 Contributions of the Thesis

First of all, the thesis successfully develops the optimal design scenarios of beamforming and phase shifts in four networks mentioned above. Among these four networks, several high efficiency mathematical tools, including relaxation, SCA, SDR, fractional programming (FP) and manifold optimization are utilized in the designs of proposed methods and the proposed

methods effectively solve the formulated EE/sum rate original problems. Meanwhile, the complexity of the algorithms are also analyzed and compared to some conventional ones i.e., SDR by Gaussian randomization. By simulation, the conclusion that NOMA can perform better than OMA can be observed in all the four networks not only for EE but also for sum rate. Besides, the importance of phase shift optimization in refining the quality of channels and improving sum rate/EE in the networks is illustrated. Meanwhile, the influence of hardware's configurations on the performance of networks can also be observed. For example, higher number of antennas and elements can lead to higher EE/sum rate. The simulation results can also demonstrate the STAR-RIS can have a higher degree of freedom (DoF) and perform better in EE/sum rate than two independent RISs. Last but not least, the proposed schemes in this thesis are also universal to other RIS/STAR-RIS networks. In the final part of the thesis, the future directions based on RIS assisted NOMA networks are given to provide the audiences with more inspirations on research.

1.5 Organization of the Thesis

The organization of the thesis can be summarized as follows:

Chapter 2: In this chapter, some background information related to the thesis is presented. Firstly, several multiple access techniques including FDMA, TDMA, CDMA, OFDMA and NOMA are introduced. Their characteristics and working principles are described in graphical ways. Then, RIS and STAR-RIS are introduced including their structure, working mechanism and advantages. Besides, CoMP technique is also introduced as the basis of chapter 5. Furthermore, ISAC is discussed as an emerging framework in 6G. Finally, the background concepts, theory of convex optimization and some classic programming problems and algorithms are presented.

Chapter 3: This chapter designs a low-complexity scheme to achieve the balanced tradeoff between the sum rate and power consumption in an RIS assisted NOMA system, which can be measured by energy efficiency. To solve the formulated problem effectively, the original non-convex problem is first decomposed into two sub-problems, i.e., beamforming optimization and phase shift optimization. Alternating optimization is proposed to solve these two sub-problems iteratively. In particular, SCA is utilized to convert the non-convex

constraints to convex ones. The provided simulation results demonstrate that the proposed scheme can achieve superior performance on energy efficiency compared to the random phase shifts and orthogonal multiple access (OMA) schemes.

Chapter 4: This chapter studies a multiple input single output (MISO) STAR-RIS assisted NOMA downlink network and investigate the energy efficiency maximization to achieve the tradeoff between the sum rate and the power consumption. The original formulated problem is non-convex due to the coupled beamforming vectors of the users and phase shifts of the STAR-RIS. To efficiently solve the problem, the original non-convex problem is decoupled to the phase shift and beamforming optimization problems and they are solved alternatively. In the phase shift optimization, fractional programming is applied to transform the sum rate maximization problem to convex SDR one with the rank-one constraints. After this, a novel sequential rank-one constraint relaxation (SROCR) is proposed to convert the rank-one constraints into convex ones, which can effectively compensate for the inadequacy of Gaussian randomization, i.e., quality of the solutions and computational complexity. Similarly, FP is applied to solve the beamforming problem by transforming it to an SDR problem. It turns out that the optimal solutions of the SDR beamforming optimization problem can be guaranteed to be rank-one by the mathematical proof and experiments. The simulation results demonstrate the STAR-RIS NOMA system can achieve the superior performance in EE.

Chapter 5: This chapter studies a joint beamforming and phase shift design for an RIS assisted coordinated multi-point non-orthogonal multiple access (CoMP-NOMA) system. With the assistance of two RISs, two base stations (BSs) cooperatively serve an edge user (termed CoMP user), while each BS serves a near user (termed NOMA user) simultaneously by applying NOMA. The optimization problem is formulated to maximize the sum rate achieved by the NOMA users. To effectively resolve this, an efficient alternative algorithm is developed, by decoupling the original problem into two suboptimal problems, i.e., beamforming optimization and phase shift optimization problems. A penalty-based semidefinite relaxation is used to tackle the beamforming optimization problem. In the phase shift optimization problem, a penalty-based manifold method is proposed. Extensive simulation results are provided to demonstrate the superior performance of RIS assisted CoMP-NOMA. The advantage of the proposed algorithm compared to conventional SDR with Gaussian randomization is also demonstrated.

Chapter 6: This chapter studies a STAR-RIS assisted integrated sensing and communications non-orthogonal multiple access (ISAC-NOMA) framework. A dual function base station (BS) is employed to serve multiple users and targets with the assistance of STAR-RIS in NOMA networks. A joint beamforming and phase shift optimization problem is proposed to maximize the sum rate of the primary users, ensuring that each user's signal can be successfully decoded and each target can be successfully detected. To effectively address the original non-convex problem, the alternating optimization is firstly proposed to decouple it into two suboptimal problems, namely beamforming and phase shift problems. After this, SDR is both leveraged to these sub-problems. Two optimization problems are first transformed to SDR ones, respectively and fractional programming is both used to transform several non-convex constraints to convex. The non-convex rank-one constraints are tackled by invoking an SCA based penalty method. The simulation results demonstrate the convergence of the proposed algorithms. They also demonstrate that the proposed scheme can outperform several benchmark schemes such as RIS assisted ISAC-NOMA, STAR-RIS assisted NOMA with random phase shifts and STAR-RIS assisted OMA in sum rate maximization.

Chapter 7: This chapter presents the conclusions of the thesis and also provides some new sights on future research directions.

1.6 Notations

In this thesis, the bold capital letters and bold lowercase letters refer the matrices and vectors, respectively. $\mathbf{A}^T$, $\mathbf{A}^H$ and $\mathrm{Tr}(\mathbf{A})$ represent the transpose, hermitian, and trace of a matrix $\mathbf{A}$. The notation $\mathbf{A} \succeq \mathbf{0}$ represents a constraint which signifies that $\mathbf{A}$ is a positive semidefinite matrix. j refers to the imaginary unit of complex numbers. Besides, $\|\cdot\|_2$ is the 2-norm of a matrix or vector. $\|\cdot\|_F$ is the Frobenius norm of a matrix or vector. $a \sim \mathcal{CN}(0, \sigma^2)$ denotes a follows a Gaussian distribution with mean zero and variance σ^2 . $a^\dagger$ represents the conjugate of the complex number a. diag($\mathbf{a}$) represents the diagonal matrix of the vector $\mathbf{a}$.

Chapter 2

Background Information

2.1 Orthogonal Multiple Access Techniques

OMA enables the base station to transmit signals to the multiple users within the finite communication resources such as bandwidth, time and code. In fact, general OMA includes FDMA, TDMA, CDMA and OFDMA. These techniques are discussed as follows:

2.1.1 Frequency Division Multiple Access

FDMA divides the total bandwidth to multiple channels. Assume the total bandwidth is B and it is used to serve K users. Therefore, each user can be allocated with $\frac{B}{K}$ frequency band. At the receiver of each user, the frequency response of bandpass filter $H(f)$ is designed to allow the signals with specific frequency to pass. If frequencies of channels are far enough, The overlapping parts of the frequency responses will be extremely small and even zero, thereby removing the interference.

2.1.2 Time Division Multiple Access

TDMA divides a period T into several time slots. Different from FDMA, in TDMA, all the users can share the total bandwidth B, improving the spectrum efficiency.

2.1.3 Code Division Multiple Access

In CDMA, each user's signal is allocated with a particular pseudorandom sequence code. This code multiples user's original signal to spread it across a wider frequency band. At the

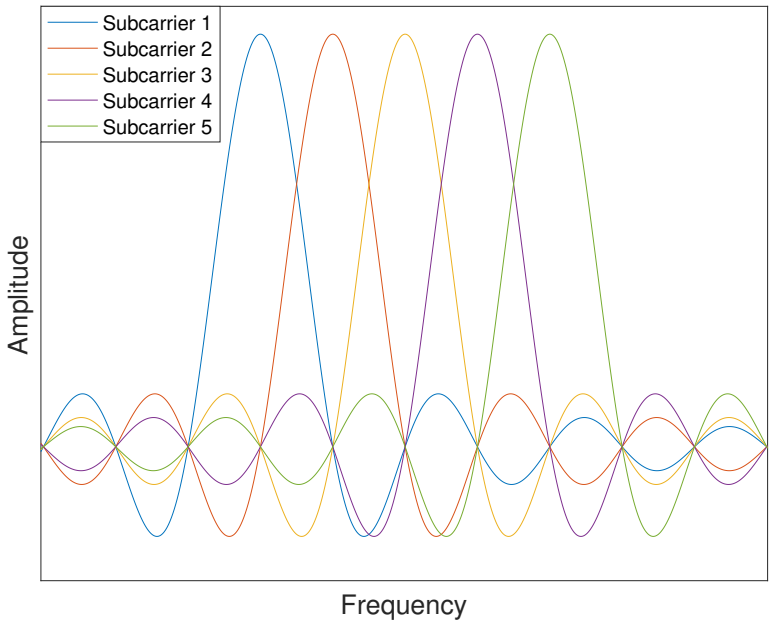

Fig. 2.1. Spectrum of OFDMA

receiver side, each user's particular pseudorandom sequence is also used to demodulate the signal. Specifically, one user's signal can be demodulated by integrating its pseudorandom sequence over an entire code period. Due to the orthogonality of pseudorandom sequence codes, other users' signals can be eliminated.

2.1.4 Orthogonal Frequency Division Multiple Access

OFDMA is the evolution of FDMA. compared to FDMA, OFDMA divides users to multiple subcarriers. These subcarriers are closely spaced but orthogonal with each other, achieving higher spectrum efficiency. Fig. 2.1 presents the spectrum of subcarriers in OFDMA networks. From the figure, it can be observed that the adjacent subcarriers are overlapped with each other. However, due to the orthogonality, the interference can be effectively removed though they are closely spaced.

2.2 Non-orthogonal Multiple Access Technique

NOMA is a promising technology in 6G. The basic idea of NOMA is that signals can be transmitted simultaneously by sharing the common bandwidth [60], [61]. This can lead to a great enhancement of spectrum efficiency. Not like OMA including FDMA, TDMA, CDMA

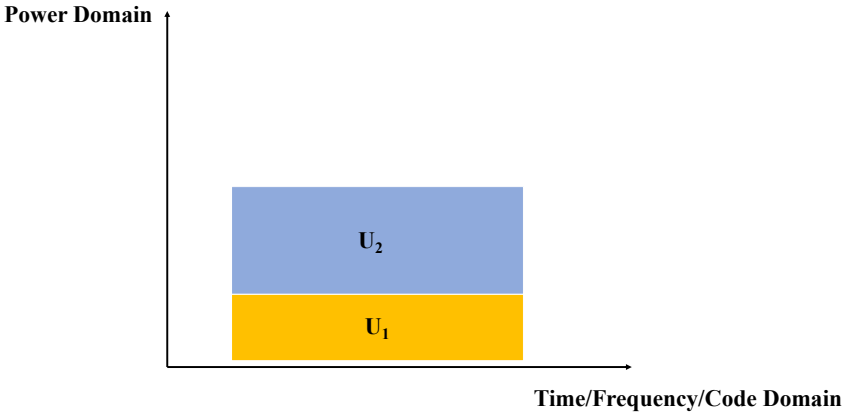

Fig. 2.2. Illustration of NOMA

and OFDMA, it should be noticed that the non-orthogonality of the channels inevitably introduces the interference among users. To overcome this technical difficulty, SIC is the key technique in NOMA. Specifically, the users in a NOMA network are given by a pre-defined decoding order. The decoding order is defined according to users' decoding capability. At the receiver, the user's signal with a lower decoding order is decoded before the one with higher decoding order. Then, SIC is employed to remove the interference of the signals which have been decoded before. For a more clear explanation, a two-user downlink NOMA network is shown to present the implementation of NOMA and SIC:

Fig. 2.3 demonstrates a two-user downlink NOMA network. The BS, equipped with M antennas transmits the superimposed signal of s_1 and s_2 to two single-antenna users with the beamforming of $\mathbf{w}_1 \in \mathbb{C}^{M\times 1}$ and $\mathbf{w}_2 \in \mathbb{C}^{M\times 1}$, respectively. The channel gains between two users are denoted as $\mathbf{h}_1^H \in \mathbb{C}^{1\times M}$ and $\mathbf{h}_2^H \in \mathbb{C}^{1\times M}$, respectively. Assume that U_2 has a higher decoding capability than U_1. The received signal of the k^{th} user can be first expressed as

$$y_k = \mathbf{h}_k^H(\mathbf{w}_1 s_1 + \mathbf{w}_2 s_2) + n_k, k = 1, 2, \tag{2.1}$$

where $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN). Since U_2 has a higher decoding capability than U_1, U_1's signal should be decoded before U_2's signal. According to SIC, at U_2, U_1's signal is decoded first. It is noted that U_2's signal has not been decoded at this moment, the interference of U_2's signal cannot be eliminated by SIC. Therefore, the

SINR of U_1's signal decoded at U_2, denoted as $\gamma_{2\to1}$ can be expressed as

$$\gamma_{2\to1} = \frac{\|\mathbf{h}_2^H \mathbf{w}_1\|^2}{\|\mathbf{h}_2^H \mathbf{w}_2\|^2 + \sigma^2}, \tag{2.2}$$

where σ^2 is the noise power. since U_1's signal has been decoded first, SIC can be employed to remove the interference of U_1. Hence, the SINR of U_2, denoted as γ_2 can be expressed as

$$\gamma_2 = \frac{\|\mathbf{h}_2^H \mathbf{w}_2\|^2}{\sigma^2}. \tag{2.3}$$

Similarly, at U_1, when decoding its signal first, SIC cannot remove the interference of U_2. The SINR of U_1's signal decoded at itself, denoted as $\gamma_{1\to1}$ can be expressed as:

$$\gamma_{1\to1} = \frac{\|\mathbf{h}_1^H \mathbf{w}_1\|^2}{\|\mathbf{h}_1^H \mathbf{w}_2\|^2 + \sigma^2}. \tag{2.4}$$

However, U_1 cannot decode U_2's signal because its limited decoding capability. According to (2.2) and (2.4), the achievable SINR of U_1's signal is given by

$$\gamma_1 = \min\{\gamma_{1\to1}, \gamma_{2\to1}\}. \tag{2.5}$$

It is noted that (2.5) is to guarantee each user's signal can be successfully decoded. Otherwise, SIC will fail to be carried out. This system model and NOMA scheme are also the basis of those of **Chapter 3**. Actually, NOMA can be applied in plenty of communication cases including multi-user, multi-cluster and cell-edge networks. In the following chapters, more diverse decoding and SIC schemes will be applied to demonstrate the superior performance of NOMA.

2.3 Reconfigurable Intelligent Surface

Benefiting from the development of surface electromagnetics, RIS is widely considered as an innovative technique and trending topic throughout both academia and industry in recent years [62]. As is shown in Fig. 2.4, there are three layers in RIS, in which the control circuit board constitutes the basic circuit. A microcontroller is connected to the BS and RIS [63]. Besides, multiple passive reflecting elements are deployed on RIS and their phase

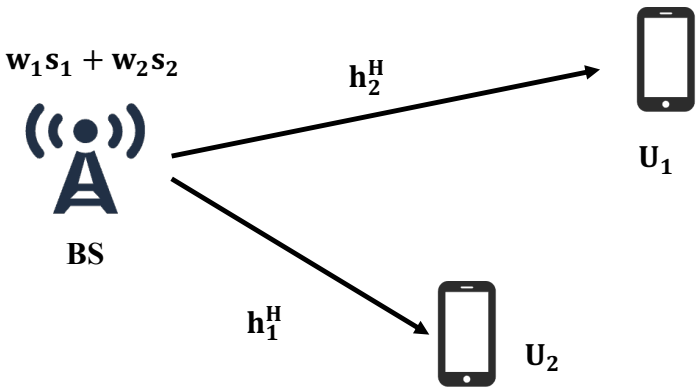

Fig. 2.3. A NOMA network

shifts can be smartly tuned by microcontroller to modify the incident electromagnetic waves [64]. With the assistance of RIS, some performance metrics of the networks can be greatly improved. Besides, as mentioned in **Chapter 1**, RIS can serve the users in the service dead zones. More importantly, until now, more BSs, access points (AP) and relays are deployed to tackle the issue of the path loss due to long distance propagation of signals, which requires more hardware and energy. However, RIS is a lower hardware/energy cost device and it is easy to be equipped. In Fig. 2.4, consider BS is equipped with M antennas and RIS has N elements. Denote the channel gain from BS to RIS is $\mathbf{G} \in \mathbb{C}^{N \times M}$, the channel gain from RIS to the mobile device is $\mathbf{h}_r^H \in \mathbb{C}^{1 \times N}$. The vector for RIS elements can be denoted as $\boldsymbol{\theta} = [\sqrt{\beta_1} e^{j\theta_1}, \cdots, \sqrt{\beta_N} e^{j\theta_N}]^T$, where θ_n is the phase shift angle for the n^{th}element and β_n is the amplitude coefficient. For passive RIS, phase shift angles can be tuned but amplitude coefficients cannot be changed. Usually, β_n is assigned to 1. Direct links in this network are neglected. Therefore, the total channel gain from BS to the mobile device can be expressed as follow:

$$\mathbf{h}^H = \mathbf{h}_r^H \text{diag}(\boldsymbol{\theta}) \mathbf{G}. \tag{2.6}$$

Let $\mathbf{w}$ denote the transmitting beamforming for mobile device. The rate (bps/Hz) of the device can be expressed as:

$$R = \log_2(1 + \frac{\|\mathbf{h}^H \mathbf{w}\|^2}{\sigma^2}). \tag{2.7}$$

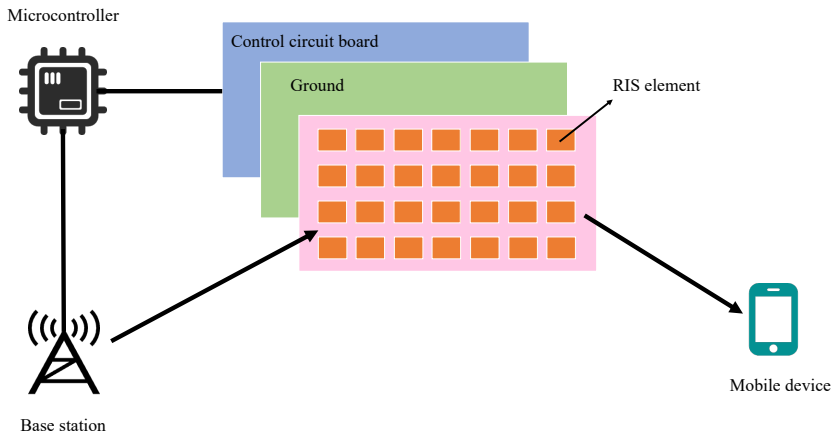

Fig. 2.4. An RIS network

Since the $\boldsymbol{\theta}$ is a variable in (2.7), the performance metrics can be improved by optimizing $\boldsymbol{\theta}$. For example, the rate maximization problem can be formulated as follow:

$$\max_{\boldsymbol{\theta}} \quad R, \tag{2.8a}$$

$$\text{s.t.} \quad \|\boldsymbol{\theta}_n\| = 1, n = 1, \cdots, N. \tag{2.8b}$$

RIS can be employed to assist the NOMA empowered networks and phase shift optimization can play the crucial role on enhancing the performance of networks.

2.4 Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surface

From Fig. 2.5, it can be seen that the structure of STAR-RIS is similar to that of RIS. The difference is that there are two types of elements, namely transmitting elements (T elements) and reflecting elements (R elements) [65]. The users which are located at the side of BS are named as R user and the users which are located at the opposite side of BS are named as T user. In STAR-RIS, consider BS is equipped with M antennas and RIS has $2N$ elements, meaning that the number of each type of elements is N. The vector of $l, l \in \{T, R\}$ elements is denoted as $\boldsymbol{\theta}_l = [\sqrt{\beta_{1,l}}e^{j\theta_{1,l}}, \cdots, \sqrt{\beta_{N,l}}e^{j\theta_{N,l}}]^T$, where $\theta_{n,l}$ is the phase shift angle for the n^{th} l element and $\beta_{n,l}$ is the amplitude coefficient. It is noted that the particular characteristic of STAR-RIS is that it satisfy $\beta_{n,T} + \beta_{n,R} = 1$. Denote the channel gain from BS to RIS is

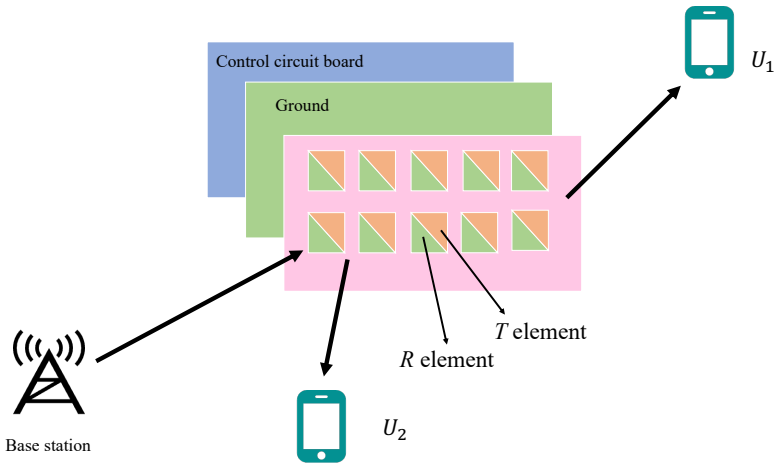

Fig. 2.5. A STAR-RIS network

$\mathbf{G} \in \mathbb{C}^{N\times M}$, the channel gain from STAR-RIS to mobile device $U_k, k = 1, 2$, is $\mathbf{h}_{k,r}^H \in \mathbb{C}^{1\times N}$. Therefore, similar to RIS, the total channel gains of two users can be expressed as

$$\mathbf{h}_1^H = \mathbf{h}_{1,r}^H \text{diag}(\boldsymbol{\theta}_T)\mathbf{G}, \tag{2.9}$$

$$\mathbf{h}_2^H = \mathbf{h}_{2,r}^H \text{diag}(\boldsymbol{\theta}_R)\mathbf{G}, \tag{2.10}$$

respectively. Assume NOMA is applied in this network and U_1's signal is decoded before U_2's signal. Then, the SINR of each user can be expressed as follows:

$$\gamma_2 = \frac{\|\mathbf{h}_2^H \mathbf{w}_2\|^2}{\sigma^2}, \tag{2.11}$$

$$\gamma_{1\to 1} = \frac{\|\mathbf{h}_1^H \mathbf{w}_1\|^2}{\|\mathbf{h}_1^H \mathbf{w}_2\|^2 + \sigma^2}, \tag{2.12}$$

$$\gamma_{2\to 1} = \frac{\|\mathbf{h}_2^H \mathbf{w}_1\|^2}{\|\mathbf{h}_2^H \mathbf{w}_2\|^2 + \sigma^2}, \tag{2.13}$$

and achievable SINR of U_1 can be expressed as

$$\gamma_1 = \min\{\gamma_{1\to 1}, \gamma_{2\to 1}\}. \tag{2.14}$$

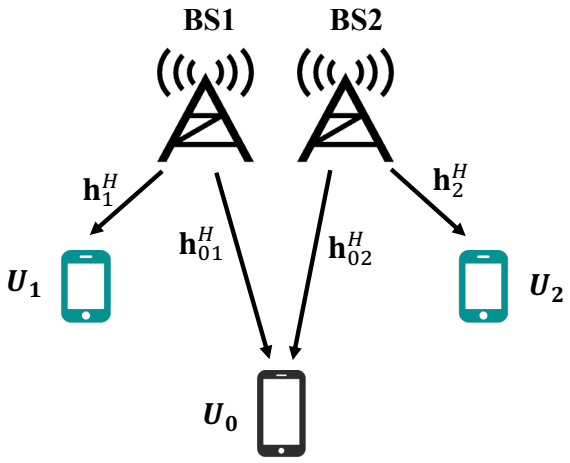

Fig. 2.6. A CoMP network

2.5 Coordinated Multi-Point Technique

The basic concept of CoMP is to allow multiple BSs to serve the edge users simultaneously [66]. By this means, the quality of communications for the edge users can be increased and their signals are likely to be successfully decoded, guaranteeing the stability and reliability of the networks. Consider the following system: In Fig. 2.6, there are two base stations, denoted by BS1 and BS2 in a CoMP network. Each BS is equipped with M antennas. two cell users are located as shown in the figure and each BS serves one cell user. Meanwhile, an edge user, also known as CoMP user is located at the edge of two cells and it is served by two BSs jointly. Therefore, the transmitting signals from two BSs can be respectively expressed as

$$\mathbf{x}_1 = \mathbf{w}_1 s_1 + \mathbf{w}_{01} s_0, \tag{2.15}$$

$$\mathbf{x}_2 = \mathbf{w}_2 s_2 + \mathbf{w}_{02} s_0, \tag{2.16}$$

where $\mathbf{w}_k$ is the beamforming of the k^{th} cell user and $\mathbf{w}_{0k}$ is the beamforming of the edge user from BSk. Denote $\mathbf{h}_k^H, k = 1, 2$ as the channel gain from the BS to the k^{th} cell user and $\mathbf{h}_{0k}^H, k = 1, 2$ denotes the channel gain from the BSk to edge user. Neglecting the inter-interference of the network, the received signals of the cell and edge users can be expressed as

$$y_k = \mathbf{h}_k^H (\mathbf{w}_k s_k + \mathbf{w}_{0k} s_0), k = 1, 2, \tag{2.17}$$

$$y_0 = \mathbf{h}_{01}^H (\mathbf{w}_1 s_1 + \mathbf{w}_{01} s_0) + \mathbf{h}_{02}^H (\mathbf{w}_2 s_2 + \mathbf{w}_{02} s_0). \tag{2.18}$$

If applying NOMA in this network, assume the cell users have the higher decoding capability than edge user. Therefore, edge user's signal should be decoded first. By applying SIC, the interference of U_0 can be removed when decoding U_1's or U_2's signal at itself. Therefore, the SINR for each users can be expressed as follows:

$$\gamma_{0\to 0} = \frac{\|\mathbf{h}_{01}^H \mathbf{w}_{01} + \mathbf{h}_{02}^H \mathbf{w}_{02}\|^2}{\|\mathbf{h}_{01}^H \mathbf{w}_1\|^2 + \|\mathbf{h}_{02}^H \mathbf{w}_2\|^2 + \sigma^2}, \tag{2.19}$$

$$\gamma_{k\to 0} = \frac{\|\mathbf{h}_k^H \mathbf{w}_{0k}\|^2}{\|\mathbf{h}_k^H \mathbf{w}_k\|^2 + \sigma^2}, k = 1, 2, \tag{2.20}$$

$$\gamma_k = \frac{\|\mathbf{h}_k^H \mathbf{w}_k\|^2}{\sigma^2}, k = 1, 2, \tag{2.21}$$

where $\gamma_{0\to 0}$ is the SINR when decoding edge user's signal at itself. $\gamma_{k\to 0}$ is the SINR when decoding edge user's signal at the cell user U_k. γ_k is the SINR when decoding the cell user U_k's signal. This system model and NOMA scheme are also the basis of those of **Chapter 5**.

2.6 Integrated Sensing and Communications

ISAC is an innovative framework in 6G networks [67]. In the future, ISAC can provide diverse services such as accurate position and tracking and mapping construction. An ISAC network consists of one dual function BS, multiple sensing targets and communication users. Both sensing and communications can be carried out within the same bandwidth. In Fig. 2.7, dual function BS, equipped with M antennas serves K communication users and R sensing targets. Denote $\mathbf{h}_k^H \in \mathbb{C}^{M\times 1}, k = 1, \cdots, K$ as the channel gain of the k^{th} user and $\mathbf{a}_r(\psi_r) = [1, e^{j\frac{2\pi d}{\lambda}\sin(\psi_r)}, \cdots, e^{j\frac{2\pi d}{\lambda}(N-1)\sin(\psi_r)}]^T, r = 1 \cdots, R$ as the steering vector of r^{th} target, where $\psi_{r,l}$ denotes the angle of departure. λ and d denote the wavelength and the space between the BS antennas, respectively. In conventional sensing system, radars transmit the sensing signals and obtain the locations of targets by receiving the reflected signals from targets. However, in ISAC systems, this scenario will cause the interference at the receivers of communication users since two functionalities should be operated in the same bandwidth. To prevent the unnecessary interference, the transmitting signals with beamforming $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ can be utilized for the purpose of both sensing and communications, which can be

expressed as

$$\mathbf{x} = \sum_{k=1}^{H} \mathbf{w}_k s_k. \tag{2.22}$$

If applying NOMA in this network, assume the decoding order is following that U_k's signal is decoded before U_j's signal, $k < j$. By this decoding scheme, SIC can be implemented to remove the interference of the signals which are decoded before U_k's signal. Therefore, the SINR of U_k's signal decoded at U_i, $k < i$ can be expressed as:

$$\gamma_{i \to k} = \frac{\|\mathbf{h}_i^H \mathbf{w}_k\|^2}{\sum\limits_{l=k+1}^{K} \|\mathbf{h}_i^H \mathbf{w}_l\|^2 + \sigma^2}, \tag{2.23}$$

and the achievable SINR for the k^{th} user is:

$$\gamma_k = \min\{\gamma_{i \to k}\}. \tag{2.24}$$

In sensing system, the power of transmitting signals is used to perform sensing. To evaluate the sensing performance, the covariance matrix is first designed as follow:

$$\mathbf{W} = \sum_{k=1}^{K} \mathbf{w}_k \mathbf{w}_k^H. \tag{2.25}$$

Then, the beampattern gain, also named as the effective sensing power of r^{th} target can be expressed as follow:

$$\mathcal{P}_r = \mathbf{a}_r(\psi_r)^H \mathbf{W} \mathbf{a}_r(\psi_r). \tag{2.26}$$

In this ISAC system, the beamforming can be designed to optimize various performance metrics. For example, the authors in [59] study a max/min optimization problem in an RIS assisted ISAC NOMA network. Similarly, a max/min optimization problem for beampattern can be formulated as follow to improve the accuracy of sensing:

$$\max_{\mathbf{w}} \quad \min\{\mathcal{P}_1, \cdots, \mathcal{P}_R\}, \tag{2.27a}$$

$$\text{s. t.} \quad \gamma_k \geq \gamma_{k,min}, k = 1 \cdots, K, \tag{2.27b}$$

$$\mathcal{P}_r \geq \mathcal{P}_{r,min}, r = 1 \cdots R, \tag{2.27c}$$

$$\sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \leq P_{max}, \tag{2.27d}$$

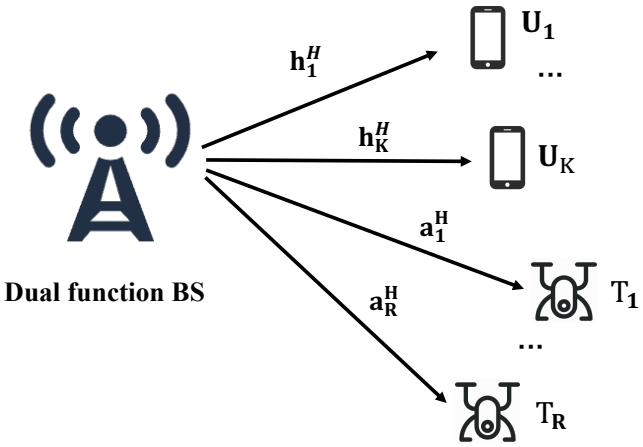

Fig. 2.7. An ISAC network

where (2.27b) is to guarantee the quality of service (QoS) constraints for communications. (2.27c) means that the beampattern gain for each target should be higher than a minimum requirement of gain to satisfy a certain accuracy of sensing. This system model and NOMA scheme are also the basis of those of **Chapter 6**.

2.7 Convex Optimization

2.7.1 Convex Sets

A set $\mathcal{C}$ is convex if any point in the line segment between two different points of $\mathcal{C}$ still lies in $\mathcal{C}$. Mathematically, for any two random points $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathcal{C}$, if

$$\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{C} \tag{2.28}$$

always holds for $0 \leq \theta \leq 1$. As is shown in Fig. 2.8, the left set is convex since any line segment is inside the hexagon. However, the right set is not convex since there exist two points, making that some points on the line segment of them are outside the set.

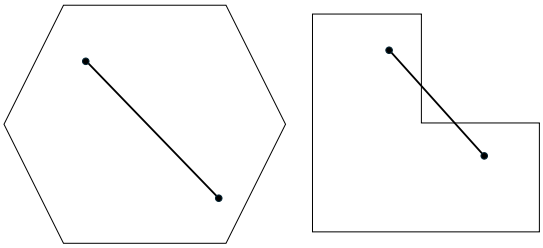

Fig. 2.8. Geometrical illustration of convex/nonconvex sets

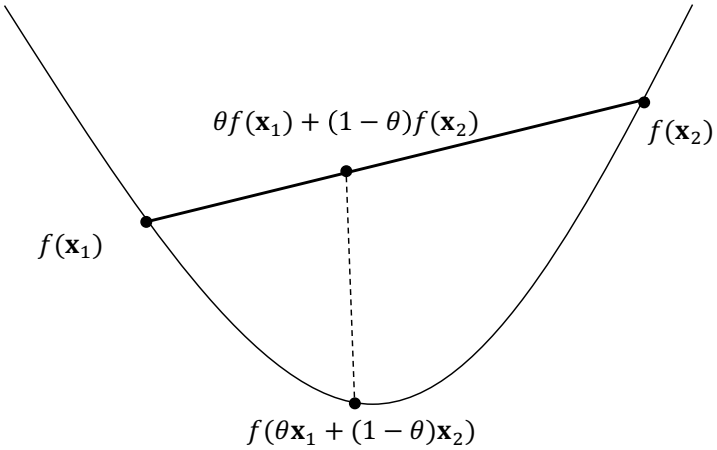

Fig. 2.9. Geometrical illustration of convex functions

2.7.2 Convex Functions

A function $f : R^N \rightarrow R$ is convex if **dom**f is convex and for any two points $\mathbf{x}_1, \mathbf{x}_2 \in$ **dom**f and $\theta \in [0, 1]$, there exists

$$f(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2). \tag{2.29}$$

Fig. 2.9 vividly describes the characteristic of convex functions. The inequality (2.29) means that the line segment of any $\mathbf{x}_1$ and $\mathbf{x}_2$ is always higher than the graph of f. If equality does not hold in (2.29), the function is strictly convex. In contrast, if '$\geq$' holds for (2.29), f is a concave function. Some classic convex and concave functions are given as follows:

- Affine. $ax + b$ is both convex and concave on R, for any real values a and b.
- Exponential. e^{ax} is convex on R, for any real value a.
- Powers. x^a is convex on if $a \geq 1$ or $a \leq 0$ and concave if $0 \leq a \leq 1$ on R_{++}.

- Powers of absolute value. $\|x\|^a$ is convex for $a \geq 1$ on R.
- Logarithm. $\log x$ is concave on R_{++}.
- Negative entropy. $x \log x$ is convex for $x > 0$.

Conditions of Convex Functions

- First order condition: Suppose f is differentiable in $\mathbf{dom}f$, f is convex if and only if $\mathbf{dom}f$ is convex and

$$f(\mathbf{x}_1) \geq f(\mathbf{x}_2) + \nabla f(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2) \tag{2.30}$$

  holds for any $\mathbf{x}_1$ and $\mathbf{x}_2$.

- Second order condition: Suppose f is twice differentiable in $\mathbf{dom}f$, f is convex if and only if $\mathbf{dom}f$ is convex and its Hessian matrix is positive semidefinite

$$\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}. \tag{2.31}$$

Convexity of Composite Functions

Consider there is a composite function

$$f(\mathbf{x}) = h(g(\mathbf{x})), \tag{2.32}$$

where $\mathbf{dom}g \in R^N$ and $\mathbf{dom}h \in R$. Without assuming differentiability of h and g, four properties are satisfied as follows:

- f is convex if h is convex, $\tilde{h}$ is nondecreasing and g is convex,
- f is convex if h is convex, $\tilde{h}$ is nonincreasing and g is concave,
- f is concave if h is concave, $\tilde{h}$ is nondecreasing and g is concave,
- f is concave if h is concave, $\tilde{h}$ is nonincreasing and g is convex,

where $\tilde{h}$ is the extended-value extension of the function, it means that if h is convex, $+\infty$ is assigned to h for the point outside $\mathbf{dom}h$. In contrast, if h is concave, $-\infty$ is assigned to h for the point outside $\mathbf{dom}h$.

2.7.3 Convex Optimization Problems

Generally, an optimization problem can be written as

$$\min_{\mathbf{x}} \quad f_0(\mathbf{x}), \tag{2.33a}$$

$$\text{s. t.} \quad f_i(\mathbf{x}) \leq 0, i = 1, \cdots m, \tag{2.33b}$$

$$h_i(\mathbf{x}) = 0, i = 1, \cdots p, \tag{2.33c}$$

where f_0 is the objective function. $f_i, i = 1 \cdots, n$ and $h_i, i = 1 \cdots, p$ are inequality constraints and equality constraints, respectively. The domain of $\mathbf{x}$ is defined as

$$\mathcal{D} = \bigcap_{i=0}^{m} f_i \cap \bigcap_{i=1}^{p} h_i. \tag{2.34}$$

(2.33) is a standard convex optimization problem if $\mathcal{D}$ is not empty and all $f_i, i = 0 \cdots m$ are convex and $h_i, i = 1 \cdots, p$ are affine. If a problem is a standard convex one, any locally optimal solution is also the global optimal one. If the objective function is strictly convex, the optimal solution of the problem is unique and global. Some classic convex optimization problems are presented as follows:

Linear Programming

A general linear programming problem can be written as

$$\min_{\mathbf{x}} \quad \mathbf{c}^T\mathbf{x} + d, \tag{2.35a}$$

$$\text{s. t.} \quad \mathbf{G}\mathbf{x} \preceq \mathbf{h}, \tag{2.35b}$$

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{2.35c}$$

where $\mathbf{c}, \mathbf{x} \in R^{N\times 1}$, $\mathbf{G} \in R^{M\times N}$, $\mathbf{h} \in R^{M\times 1}$, $\mathbf{A} \in R^{P\times N}$, $\mathbf{b} \in R^{P\times 1}$, $d \in R$. From the problem formulation, it is a standard convex optimization problem.

The linear-fractional programming is written as follow:

$$\min_{\mathbf{x}} \quad \frac{\mathbf{c}^T\mathbf{x} + d}{\mathbf{e}^T\mathbf{x} + f}, \tag{2.36a}$$

$$\text{s.t.} \quad \mathbf{Gx} \preceq \mathbf{h}, \tag{2.36b}$$

$$\mathbf{Ax} = \mathbf{b}, \tag{2.36c}$$

where $\mathbf{c} \in R^{N\times 1}$ and $f \in R$. From the problem formulation, (2.36) is not a standard convex optimization. However, it can be equivalently transformed to a linear one. Specifically, to show the equivalence, let $\mathbf{y} = \frac{\mathbf{x}}{\mathbf{e}^T\mathbf{x}+f}$ and $z = \frac{1}{\mathbf{e}^T\mathbf{x}+f}$, the problem can be written as

$$\min_{\mathbf{x},\mathbf{y},z} \quad \mathbf{c}^T\mathbf{y} + dz, \tag{2.37a}$$

$$\text{s.t.} \quad \frac{\mathbf{Gx}}{\mathbf{e}^T\mathbf{x}+f} \preceq \frac{\mathbf{h}}{\mathbf{e}^T\mathbf{x}+f}, \tag{2.37b}$$

$$\frac{\mathbf{Ax}}{\mathbf{e}^T\mathbf{x}+f} = \frac{\mathbf{b}}{\mathbf{e}^T\mathbf{x}+f}, \tag{2.37c}$$

$$\mathbf{e}^T\mathbf{y} + fz = 1. \tag{2.37d}$$

Thus, the problem can be transformed by replacing $\mathbf{x}$ with $\mathbf{y}$ and z as follow:

$$\min_{\mathbf{y},z} \quad \mathbf{c}^T\mathbf{y} + dz, \tag{2.38a}$$

$$\text{s.t.} \quad \mathbf{Gy} \preceq \mathbf{h}z, \tag{2.38b}$$

$$\mathbf{Ay} = \mathbf{b}z, \tag{2.38c}$$

$$\mathbf{e}^T\mathbf{y} + fz = 1. \tag{2.38d}$$

From the problem formulation, (2.38) is a LP problem and it is a standard convex one.

Quadratic Programming

The standard quadratic programming (QP) problem can be formulated as

$$\min_{\mathbf{x}} \quad \frac{1}{2}\mathbf{x}^T\mathbf{Px} + \mathbf{q}^T\mathbf{x} + r, \tag{2.39a}$$

$$\text{s.t.} \quad \mathbf{Gx} \preceq \mathbf{h}, \tag{2.39b}$$

$$\mathbf{Ax} = \mathbf{b}, \tag{2.39c}$$

where $\mathbf{P} \in R^{N\times N}$ and $\mathbf{P} \succeq \mathbf{0}$, $\mathbf{q} \in R^{N\times 1}$, $r \in R$. In a QP problem, a convex quadratic functions can be minimized over a polyhedron.

Quadratically Constrained Quadratic Programming

If the inequality in (2.39) are quadratic, the problem will become a quadratically constrained quadratic programming (QCQP) one as follow:

$$\min_{\mathbf{x}} \quad \frac{1}{2}\mathbf{x}^T\mathbf{P}_0\mathbf{x} + \mathbf{q}_0^T\mathbf{x} + r_0, \tag{2.40a}$$

$$\text{s.t.} \quad \frac{1}{2}\mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i^T\mathbf{x} + r_i \leq 0, i = 1 \cdots m, \tag{2.40b}$$

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{2.40c}$$

where $\mathbf{P}_i \in R^{N\times N}$, $\mathbf{P}_i \succeq \mathbf{0}$, $\mathbf{q}_i \in R^{N\times 1}, i = 1, \cdots m$, $r_i \in R, i = 1, \cdots m$. In a QCQP problem, the convex quadratic function can be minimized over the intersection of ellipsoids if $\mathbf{P}_i \succ \mathbf{0}$. In **Chapter 3**, **Chapter 4**, **Chapter 5** and **Chapter 6** of this thesis, QCQP-like problems are proposed in RIS/STAR-RIS assisted NOMA networks, indicating its critical status on the designs of beamforming and phase shifts.

2.7.4 Duality

Consider the following optimization problem:

$$\min_{\mathbf{x}} \quad f_0(\mathbf{x}), \tag{2.41a}$$

$$\text{s.t.} \quad f_i(\mathbf{x}) \leq 0, i = 1, \cdots m, \tag{2.41b}$$

$$h_i(\mathbf{x}) = 0, i = 1, \cdots u, \tag{2.41c}$$

where $\mathbf{x} \in R^N$. Its domain $\mathcal{D}$ is not empty. Denote the optimal value of (2.41) is p^*. The duality theory is to find the conditions of a convex optimization problem to obtain the optimal solutions.

First, the Lagrangian function is defined as followed:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{u} \nu_i h_i(\mathbf{x}), \tag{2.42}$$

where $\lambda_i \geq 0$ and ν_i are named as dual variables or Lagrangian multipliers. Then, the dual

function is defined as followed:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}). \tag{2.43}$$

The dual function is always concave regardless of whether the original problem is concave or convex or not. Since the dual function is the lower bound of the original problem,

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^*. \tag{2.44}$$

To further study the characteristics of dual functions, a dual problem is formulated as followed:

$$\max_{\boldsymbol{\lambda}, \boldsymbol{\nu}} \quad g(\boldsymbol{\lambda}, \boldsymbol{\nu}), \tag{2.45a}$$

$$\text{s. t.} \quad \lambda_i \geq 0, i = 1, \cdots m, \tag{2.45b}$$

where $\mathbf{dom} g \in \{(\boldsymbol{\lambda}, \boldsymbol{\nu}) | g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \geq -\infty\}$. For example, consider a LP problem

$$\min_{\boldsymbol{x}} \quad \mathbf{c}^T \mathbf{x}, \tag{2.46a}$$

$$\text{s. t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, i = 1, \cdots m, \tag{2.46b}$$

$$\mathbf{x} \succeq 0, i = 1, \cdots m. \tag{2.46c}$$

Its dual function is given by

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \begin{cases} -\mathbf{b}^T \boldsymbol{\nu} & \text{if } \mathbf{A}^T \boldsymbol{\nu} - \boldsymbol{\lambda} + \mathbf{c} = \mathbf{0} \\ -\infty & \text{Otherwise} \end{cases}. \tag{2.47}$$

Therefore, the dual problem can be formulated as:

$$\max_{\boldsymbol{\lambda}, \boldsymbol{\nu}} \quad g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \begin{cases} -\mathbf{b}^T \boldsymbol{\nu} & \text{if } \mathbf{A}^T \boldsymbol{\nu} - \boldsymbol{\lambda} + \mathbf{c} = \mathbf{0} \\ -\infty & \text{Otherwise} \end{cases}, \tag{2.48a}$$

$$\text{s. t.} \quad \lambda_i \geq 0. \tag{2.48b}$$

The purpose of studying is to find the relationship between the original problems and dual

problems. Denote d^* as the optimal value of dual problem, there are two types of duality as followed:

- Weak duality: $d^* \leq p^*$ if the original problem is not convex. It is noted that if the original problem is unbounded, the dual problem must be unbounded as well and $d^* = p^* = \pm\infty$.
- Strong duality: $d^* = p^*$ if the original problem is convex and satisfy the Slater's condition, which means that there exists at least one point $\mathbf{x}$, such that $f_i(\mathbf{x}) < 0, i = 1, \cdots m$ and $h_i(\mathbf{x}) = \mathbf{0}$ in (2.41).

2.7.5 Optimality Conditions

As for a convex optimization problem, in which the strong duality holds, there exists the following relationship

$$f_0(\mathbf{x}^*) = g(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*). \tag{2.49}$$

According to the definition of dual function (2.43), (2.49) can be further developed as

$$f_0(\mathbf{x}^*) = f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{u} \nu_i^* h_i(\mathbf{x}^*). \tag{2.50}$$

From (2.50), it is found that $\mathbf{x}^*$ is obtained by finding the derivative of $\mathcal{L}$ with respect to $\mathbf{x}^*$. Therefore, the following relationship can be obtained as followed:

$$\nabla f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(\mathbf{x}^*) + \sum_{i=1}^{u} \nu_i^* \nabla h_i(\mathbf{x}^*) = \mathbf{0}. \tag{2.51}$$

It is noted that $f_0(\mathbf{x}^*)$ appears in both left and right side of the equation of (2.50). Therefore,

$$\lambda_i^* f_i(\mathbf{x}^*) = 0 \tag{2.52}$$

and

$$\nu_i^* h_i(\mathbf{x}^*) = 0 \tag{2.53}$$

should be satisfied to hold the equality. Until now, several conditions of the optimality can

be concluded as

$$\begin{aligned}
&\nabla f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(\mathbf{x}^*) + \sum_{i=1}^{u} \nu_i^* \nabla h_i(\mathbf{x}^*) = \mathbf{0}, && \text{(Stationarity)} \\
&\lambda_i^* f_i(\mathbf{x}^*) = 0, && \text{(Complementary Slackness)} \\
&f_i(\mathbf{x}^*) \leq 0, && \text{(Primal Feasibility)} \\
&h_i(\mathbf{x}^*) = 0, && \text{(Primal Feasibility)} \\
&\lambda_i \geq 0. && \text{(Dual Feasibility)}
\end{aligned} \tag{2.54}$$

(2.54) is named as KKT conditions. KKT conditions are the milestone of the development of optimization theory and plenty of classic optimization problems such as water-filling problem can be solved by applying KKT conditions. In this thesis, KKT conditions are applied to demonstrate the relaxed SDR optimal solutions of beamforming matrices are strictly rank one in **Chapter 4** and **Chapter 5**.

2.7.6 Algorithms

In fact, if an optimization problem is convex and feasible, the globally optimal solutions can be obtained by applying several existing techniques, i.e., Newton method. However, most proposed optimization problems are not convex in practical communication scenarios. To tackle this issue, several algorithms are employed to transform the non-convex problems to convex in this thesis as follows:

Successive Convex Approximation

The basic idea of SCA is to apply a series of convex functions to approximate the objective functions. Consider the following optimization problem:

$$\min_{\mathbf{x}} \quad f_0(\mathbf{x}), \tag{2.55a}$$

$$\text{s.t.} \quad \mathbf{x} \in \boldsymbol{\psi}, \tag{2.55b}$$

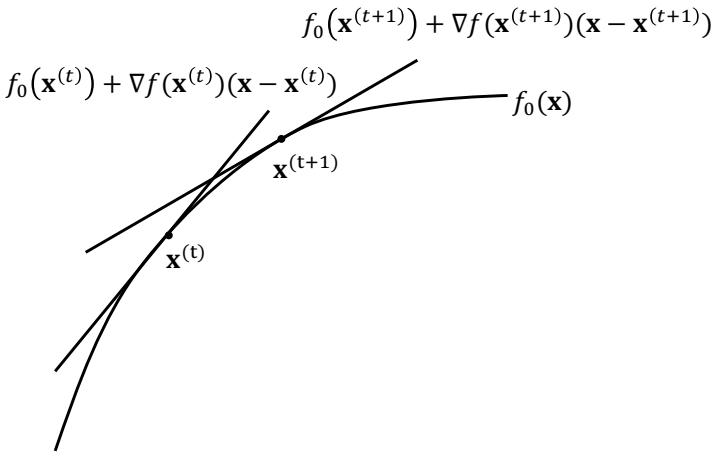

Fig. 2.10. Illustration of SCA

where f_0 may not be convex. To make (2.55a) convex, the first order Taylor series can be implemented as

$$f_0(\mathbf{x}) \approx f_0(\mathbf{x}^{(t)}) + \nabla f_0(\mathbf{x}^{(t)})(\mathbf{x} - \mathbf{x}^{(t)}), \tag{2.56}$$

where $\mathbf{x}^{(t)}$ is $\mathbf{x}$ at in the t^{th} iteration. From the formulation of (2.56), the approximated function is affine with respect to $\mathbf{x}$. By such the way of iterations, a locally optimal solution can be obtained. The graph illustration of SCA is shown in Fig. 2.10. In this thesis, SCA is employed in **Chapter 3**, **Chapter 5** and **Chapter 6**.

Semidefinite Relaxation

Consider the following QCQP optimization problem:

$$\min_{\mathbf{x}} \quad \mathbf{x}^T\mathbf{C}\mathbf{x}, \tag{2.57a}$$

$$\text{s.t.} \quad \mathbf{x}^T\mathbf{A}_i\mathbf{x} \trianglerighteq_i b_i, i = 1 \cdots m, \tag{2.57b}$$

where $\trianglerighteq_i$ can refer to $\geq$, $\leq$ or $=$ and $\mathbf{C}$ refers to a semidefinite matrix. $\mathbf{A}_k, k = 0, \cdots, m$ are $n \times n$ semidefinite matrices. From the problem formulation, it can be seen that the convexity of (2.57) is uncertain. To tackle this issue, SDR can be employed to transform (2.57) to linear. Specifically, let $\mathbf{X} = \mathbf{x}\mathbf{x}^T$, (2.57) can be equivalent as follows:

$$\min_{\mathbf{X}} \quad \text{Tr}(\mathbf{C}\mathbf{X}), \tag{2.58a}$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{A}_i\mathbf{X}) \trianglerighteq_i b_i, i = 1 \cdots m, \tag{2.58b}$$

$$\mathbf{X} \succeq \mathbf{0}, \tag{2.58c}$$

$$\text{Rank}(\mathbf{X}) = 1. \tag{2.58d}$$

It should be noted that (2.58d) is not convex. Several techniques are applied to solve the rank one constraints of SDR problems including Gaussian randomization, SROCR and penalty based SCA method.

Gaussian randomization is a basic method to solve SDR problems. Specifically, the relaxed solutions $\mathbf{X}^*$ can be obtained by solving (2.58) dropping (2.58d). Then, plenty of feasible and random points $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{X}^*)$ are generated to approximate the suboptimal solutions of (2.58). This method is utilized as the benchmark scheme in **Chapter 4** and **Chapter 5**.

SROCR is a relaxation scheme for the rank one constraints. Specifically, if a matrix is rank one, the following relationship holds that

$$e_{max}(\mathbf{X}) = \text{Tr}(\mathbf{X}). \tag{2.59}$$

This relationship can be further written as

$$\boldsymbol{\varphi}_{max}^T \mathbf{X} \boldsymbol{\varphi}_{max} = \text{Tr}(\mathbf{X}), \tag{2.60}$$

where $\boldsymbol{\varphi}_{max}$ is the eigenvector corresponding to the maximum eigenvalue of $\mathbf{X}$. Inspired by (2.60), a variable ω can be introduced to relax (2.60) to

$$\boldsymbol{\varphi}_{max}^T \mathbf{X} \boldsymbol{\varphi}_{max} = \omega \, \text{Tr}(\mathbf{X}), \tag{2.61}$$

where $0 \leq \omega \leq 1$. In the algorithm designs, we can first set a stepsize $\Delta \in [0, 1]$, during the iterations of the algorithm, ω is updated by $\omega^{(t+1)} = \min(1, \frac{e_{max}(\mathbf{X}^{(t)})}{\text{Tr}(\mathbf{X}^{(t)})} + \Delta)$. The suboptimal solutions can be obtained when both ω converges to 1 and objective function converges.

SROCR is applied as the proposed scheme in **Chapter 4**.

Penalty based SCA method apply one of the characteristics of rank one semidefinite matrix as

$$\text{Tr}(\mathbf{X}) = \|\mathbf{X}\|_F. \tag{2.62}$$

Penalty function method can be applied by considering (2.62) as the penalty term as follow:

$$\rho(\text{Tr}(\mathbf{X}) - \|\mathbf{X}\|_F), \tag{2.63}$$

where $\rho \geq 1$ is the penalty factor and it is updated in each iteration. However, (2.63) is not convex with respect to $\mathbf{X}$. To tackle the non-convexity issue, SCA is implemented to (2.63) as

$$\rho^{(t)}(\text{Tr}(\mathbf{X}) - \|\mathbf{X}^{(t)}\|_F - \text{Tr}(\nabla D(\mathbf{X}^{(t)})(\mathbf{X} - \mathbf{X}^{(t)}))), \tag{2.64}$$

where

$$\nabla D(\mathbf{X}) = \frac{\mathbf{X}}{\sqrt{\text{Tr}(\mathbf{X}^T\mathbf{X})}} \tag{2.65}$$

is the first order partial derivative of $\|\mathbf{X}\|_F$. As ρ is updated larger and larger, the penalty term is likely to tend to zero and locally optimal solutions can be obtained. Penalty based SCA method is applied as the proposed scheme in **Chapter 5** and **Chapter 6**.

Fractional Programming

Consider the following optimization problem:

$$\max_{\mathbf{x}} \quad \frac{A(\mathbf{x})}{B(\mathbf{x})}, \tag{2.66a}$$

$$\text{s.t.} \quad \mathbf{x} \in \psi, \tag{2.66b}$$

where $A(\mathbf{x})$ and $B(\mathbf{x})$ are quadratic function with respect to $\mathbf{x}$ and (2.66b) is convex. Although the formulated problem is not convex, an auxiliary variable y can be introduced to transform (2.66) to

$$\max_{\mathbf{x}} \quad 2y\sqrt{A(\mathbf{x}^{(t+1)})} - y^2 B(\mathbf{x}^{(t+1)}), \tag{2.67a}$$

$$\text{s.t.} \quad \mathbf{x} \in \psi, \tag{2.67b}$$

where y is updated by $y^{(t+1)} = \frac{\sqrt{A(\mathbf{x}^{(t+1)})}}{B(\mathbf{x}^{(t+1)})}$. According to [68], by iteratively solving (2.67), the globally optimal solutions of (2.66) can be obtained. Fractional programming is applied as the proposed scheme in **Chapter 4** and **Chapter 6**.

Chapter 3

An SCA and Relaxation Based Energy Efficiency Optimization for Multi-User RIS Assisted NOMA Networks

3.1 Introduction

3.1.1 Motivations of the Chapter

It is indicated in **Chapter 1** that NOMA and RIS can jointly improve the performance of the communication networks. Besides, **Chapter 1** has also pointed that solely pursuing the increase of the sum rate in a network might not be an economical and efficient solution in the future 6G networks. In order to achieve the tradeoff between the sum rate and power consumption [69], energy efficiency has been introduced and defined as a ratio between the sum rate and power consumption [70], [71]. Recently, some research works have demonstrated the feasibility of energy efficiency optimization for the two-user case in RIS assisted NOMA systems [41], [72]. Especially, it can be seen that the authors in [41] constructed a two-user NOMA network and maximize EE in this network by applying several mathematical techniques. However, due to the specificity of two-user network, the methodology proposed in [41] cannot apply to multi-user case. Different from the two-user case in [41], this chapter considers the multi-user case and designs an algorithm to maximize energy efficiency in a downlink MISO RIS assisted NOMA network. Meanwhile, the direct links from the BS to the users are considered to make the system model more generalized.

From the aspect of the proposed optimization methodology, a more efficient and universal method is expected to be developed. To achieve it, we first decouple the original problem to the beamforming and phase shift suboptimal problems and then solved alternatively [43]. By applying SCA, the energy efficiency will converge to a locally optimal value [73], [74]. Besides, different from the SDR method for the phase shift optimization in [41], in this chapter, we apply SCA to optimize phase shifts of the RIS. In simulation, we compare the energy efficiency of the proposed scheme with the benchmark schemes using random phase shifts and OMA and investigate the impacts of the number of reflecting elements of RIS, internal circuit power, power amplifier coefficient and transmitting power budget on the performance of RIS-NOMA networks.

3.1.2 Contributions of the Chapter

The contributions of this chapter are summarized as followed:

- The multi-user NOMA network is considered and a mathematical method and algorithm are designed to maximize energy efficiency in a downlink MISO RIS assisted NOMA network. Meanwhile, the direct links from the BS to the users are considered.
- The original problem is decoupled to the beamforming and phase shift suboptimal problems and then solved alternatively [43]. By applying SCA, the energy efficiency will converge to a locally optimal value. Besides, different from the SDR method for the phase shift optimization, in this chapter, SCA is applied to optimize phase shifts of the RIS.
- In simulation, the energy efficiency of the proposed scheme with the benchmark schemes using random phase shifts and OMA are compared and the impacts of the number of reflecting elements of RIS, internal circuit power, power amplifier coefficient and transmitting power budget are investigated on the performance of RIS-NOMA.

3.2 System Model

In Fig. 3.1, we consider a quasi-static downlink MISO RIS-NOMA network with one BS and multiple single-antenna users. The BS is equipped with M antennas and the number of users

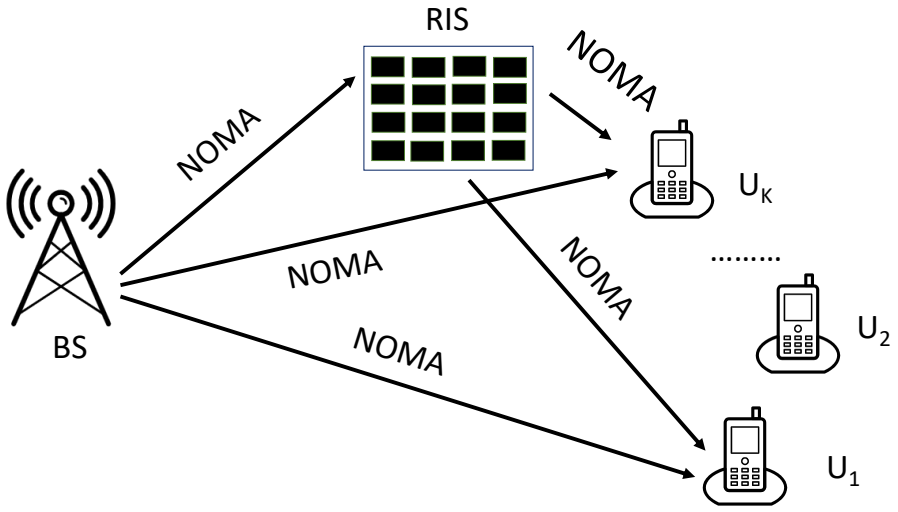

Fig. 3.1. A multi-user RIS assisted NOMA downlink network

is K, and the RIS consists of N passive reflecting elements to assist the transmission from the BS to the users. Besides, assume that the BS can communicate with the users through both the RIS and direct links from the BS to the users in this model. Define $\mathbf{G} \in \mathbb{C}^{N\times M}$ and $\mathbf{h}_{k,r}^H \in \mathbb{C}^{1\times N}$ as the channel gains from the BS to the RIS and the RIS to the users, respectively. The direct channel gains from the BS to the users are denoted by $\mathbf{h}_{k,d}^H \in \mathbb{C}^{1\times M}$, where k denotes the index of the k^{th} user, $k = 1, \cdots, K$. We define $\mathbf{\Theta}$=diag$(e^{j\theta_1}, \cdots, e^{j\theta_N})$ as the phase reflection matrix, where $\theta_n \in [0, 2\pi]$ is the reflection phase shift. In this system, consider the total channel gains of the K users are $\mathbf{h}_1^H, \cdots, \mathbf{h}_K^H$, where $\mathbf{h}_k^H = \mathbf{h}_{k,r}^H \mathbf{\Theta} \mathbf{G} + \mathbf{h}_{k,d}^H$. Therefore, the received signal at the k^{th} user can be written as follow:

$$y_k = \mathbf{h}_k^H (\sum_{i=1}^{K} \mathbf{w}_i s_i) + n_k, \tag{3.1}$$

where $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ is the beamforming vector from the BS to the k^{th} user, and $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN), where σ^2 is the variance of the noise. In the downlink RIS assisted NOMA system, the performance of SIC is highly affected by the decoding order, which indicates that when decoding U_k's signal, the interference from the users whose signals have been decoded before could be cancelled. In fact, the decoding order depends on the users's capability. In this chapter, assume U_k's signal is decoded before U_j's signal, for $k < j$. The signal-interference-plus-noise ratio (SINR) of U_k's signal decoded at

U_i ($k = 1, \cdots, K$ and $i = 1, \cdots, K, k \leq i$) can be expressed as followed:

$$SINR_{k,i} = \frac{|\mathbf{h}_i^H \mathbf{w}_k|^2}{\sum_{l=k+1}^{K} |\mathbf{h}_i^H \mathbf{w}_l|^2 + \sigma^2}. \tag{3.2}$$

In this network, the condition that SIC can be carried out is that the k^{th} user's signal can be decoded by the i^{th} user successfully ($k \leq i \leq K$). We first define $SINR_k$ as the achievable SINR for U_k's signal and $SINR_{k,i}$ is the SINR for U_k's signal to be decoded by the i^{th} user. The $SINR_k$ is defined as followed:

$$SINR_k = \min\{SINR_{k,i}\}. \tag{3.3}$$

It is noted that (3.3) can guarantee that the signal of U_k can be successfully decoded at U_i because $SINR_k$ is the minimal of $SINR_{k,i}$ as shown in (3.3). Otherwise, SIC cannot be implemented successfully.

By perfect SIC, all the interference from the other users can be removed when decoding U_K's signal because the K^{th} user's signal is decoded at the last stage of SIC. Therefore, the SINR of U_K is

$$SINR_K = \frac{|\mathbf{h}_K^H \mathbf{w}_K|^2}{\sigma^2}. \tag{3.4}$$

Thus, the achievable rate of the k^{th} user is

$$R_k = \log_2(1 + SINR_k). \tag{3.5}$$

3.3 Problem Formulation

This work is aimed to maximize the system energy efficiency which is the ratio between sum rate and total power consumption in an RIS-NOMA system. In particular, define the internal circuit power consumption of the BS as P_c and set a parameter $\eta \in [0, 1]$ to denote the power amplifier coefficient at the BS. Thus, the energy efficiency problem addressed in this chapter

can be formulated as followed:

$$\max_{\boldsymbol{\Theta},\mathbf{w}_1,\cdots,\mathbf{w}_K} \quad \frac{\sum_{k=1}^{K} R_k}{\frac{1}{\eta}\sum_{k=1}^{K}\|\mathbf{w}_k\|^2 + P_c}, \tag{3.6a}$$

$$\text{s. t.} \quad R_k \geq R_{k,min}, k = 1, \cdots, K, \tag{3.6b}$$

$$\sum_{k=1}^{K}\|\mathbf{w}_k\|^2 \leq P_{\max}, \tag{3.6c}$$

$$\theta_n \in [0, 2\pi], n = 1, \cdots, N. \tag{3.6d}$$

Firstly, (3.6b) indicates that each user's rate has to be no less than a minimal rate $R_{k,min}, k = 1, \cdots, K$ to satisfy the quality of service requirement. Constraint (3.6c) shows that the total power budget of beamforming at the BS for users is not supposed to be larger than a maximum threshold P_{max}. (3.6d) demonstrates that the angle of each reflecting element should be in the range of $[0, 2\pi]$. However, (3.6) is a non-convex problem due to the fact that R_k is not convex with respect to $\boldsymbol{\Theta}$ and $\mathbf{w}$. QoS requirement (3.6b) is also not convex, making this problem more complicated.

3.4 Proposed Optimization Methodology

In this section, we mainly design a low-complexity algorithm which can effectively obtain a suboptimal solution of (3.6). From the problem formulation, it can be seen that there are two types of variables, namely beamforming vectors and the phase shifts in this optimization problem. Due to the fact that the vector $\mathbf{w}_k$ is a complex-valued vector, the complexity of the problem could be greatly increased if solving the optimization problem with $\mathbf{w}$ and $\boldsymbol{\Theta}$ directly. Therefore, in this chapter, an optimization scheme is proposed by alternatively solving beamforming and phase shift optimization problems. In the process of addressing the non-convex functions and constraints in the beamforming and phase shift optimization problems, some slack variables are introduced to relax these non-convex functions. Moreover, SCA is utilized in the algorithm design to obtain the suboptimal solutions of the energy efficiency.

3.4.1 Beamforming Optimization

From beamforming optimization, to make the objective function convex, a slack variable α is introduced to (3.6) and (3.6) can be transformed to the following equivalent form:

$$\max_{\mathbf{w}_1,\cdots,\mathbf{w}_K,\alpha} \quad \alpha, \tag{3.7a}$$

$$\text{s. t.} \quad \frac{\sum_{k=1}^{K} R_k}{\frac{1}{\eta}\sum_{k=1}^{K} \|\mathbf{w}_k\|^2 + P_c} \geq \alpha, \tag{3.7b}$$

$$(3.6b), (3.6c). \tag{3.7c}$$

Although the objective function of (3.6a) has been transformed to an affine function, the constraint (3.7b) is not convex. In order to solve it effectively, one slack variable ρ is introduced to make the total power expression of (3.7b) convex. Thus, (3.7b) can be further transformed to the following form

$$\sum_{k=1}^{K} R_k \geq \alpha\rho, \tag{3.8}$$

and the constraint of total power expression can be written as followed:

$$\frac{1}{\eta}\sum_{k=1}^{K} \|\mathbf{w}_k\|^2 + P_c \leq \rho. \tag{3.9}$$

Based on the constraints transformation, constraint (3.8) is still not convex. Then, a new variable vector $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_K]^T$ is used to relax (3.8), which can be rewritten as

$$\sum_{k=1}^{K} \log_2(\gamma_k) \geq \alpha\rho, \tag{3.10a}$$

$$1 + SINR_k \geq \gamma_k. \tag{3.10b}$$

Subsequently, use another variable vector $\boldsymbol{\delta} = [\delta_1, \cdots, \delta_K]^T$ to make constraint (3.10a) convex, which can be rewritten as followed:

$$\sum_{k=1}^{K} \delta_k \geq \alpha\rho, \tag{3.11a}$$

$$\gamma_k \geq 2^{\delta_k}. \tag{3.11b}$$

The relationship between these slack variables introduced based on (3.8), (3.10a), (3.11a) and (3.11b) can be summarized as followed:

$$\sum_{k=1}^{K} \log_2(1 + SINR_k) \geq \sum_{k=1}^{K} \log_2(\gamma_k) \geq \sum_{k=1}^{K} \delta_k \geq \alpha\rho. \tag{3.12}$$

Considering the fractional form of SINR, one can introduce a series of variables with the form of $\Omega_{k,i}$, where $k \leq i, k, i = 1, \cdots, K$. Therefore, (3.10b) is equivalent as followed:

$$\|\mathbf{h}_i^H \mathbf{w}_k\|^2 \geq (\gamma_k - 1)\Omega_{k,i}, \tag{3.13a}$$

$$\sum_{l=k+1}^{K} \|\mathbf{h}_i^H \mathbf{w}_l\|^2 + \sigma^2 \leq \Omega_{k,i}. \tag{3.13b}$$

However, (3.13a) and (3.13b) are not applicable for the situation that i and k are equal to K simultaneously. When both of i and k are equal to K, (3.13a) and (3.13b) can be written as followed:

$$\|\mathbf{h}_K^H \mathbf{w}_K\|^2 \geq (\gamma_K - 1)\Omega_{K,K}, \tag{3.14a}$$

$$\sigma^2 \leq \Omega_{K,K}. \tag{3.14b}$$

In the following, those steps in [41] are used to convert (3.13a) to a convex form. Firstly, an arbitrary phase rotation of the beamforming vector is introduced to make the imaginary part of $\mathbf{h}_i^H \mathbf{w}_k$ equal to zero for the two-user case in this model. However, this method cannot be straightforwardly applied to the multi-user case considered in this chapter because only one given arbitrary phase rotation for $\mathbf{w}_k$ cannot guarantee that all $\Im\{\mathbf{h}_i^H \mathbf{w}_k\}$=0. Therefore, at this stage, one can apply additional variables in the form of $w_{k,i}^{Re}$ and $w_{k,i}^{Im}$, which

$$w_{k,i}^{Re} = \Re\{\mathbf{h}_i^H \mathbf{w}_k\}, \tag{3.15a}$$

$$w_{k,i}^{Im} = \Im\{\mathbf{h}_i^H \mathbf{w}_k\}, \tag{3.15b}$$

(3.15a) and (3.15b) are convex constraints. Although (3.11a) and (3.14a) are still non-convex constraints, instead of introducing new slack variables, SCA is applied to deal with these non-convex constraints. The main idea of SCA is to use the First-order Taylor approximation to convert the non-convex functions to affine [75]. In this chapter, the constraints (3.11a),

(3.13a) can be approximated to affine functions by applying SCA, which can be shown as followed:

$$\sum_{k=1}^{K} \delta_k \geq \alpha^{(t)}\rho^{(t)} + \rho^{(t)}(\alpha - \alpha^{(t)}) + \alpha^{(t)}(\rho - \rho^{(t)}), \tag{3.16}$$

$$\begin{aligned}((w_{k,i}^{Re})^{(t)})^2 + ((w_{k,i}^{Im})^{(t)})^2 + 2(w_{k,i}^{Re})^{(t)}(w_{k,i}^{Re} - (w_{k,i}^{Re})^{(t)}) + 2(w_{k,i}^{Im})^{(t)}(w_{k,i}^{Im} - (w_{k,i}^{Im})^{(t)}) \\ \geq ((\gamma_k)^{(t)} - 1)(\Omega_{k,i})^{(t)} + (\Omega_{k,i})^{(t)}(\gamma_k - (\gamma_k)^{(t)}) + ((\gamma_k)^{(t)} - 1)(\Omega_{k,i} - (\Omega_{k,i})^{(t)}),\end{aligned} \tag{3.17}$$

where t denotes the t^{th} iteration. (3.6b) is a non-convex constraint and it can be rewritten as followed:

$$\frac{\|\mathbf{h}_i^H \mathbf{w}_k\|^2}{\sum_{l=k+1}^{K} \|\mathbf{h}_i^H \mathbf{w}_l\|^2 + \sigma^2} \geq 2^{R_{k,min}} - 1, \tag{3.18}$$

where $R_{k,min}$ is a constant. After that, this constraint can be reformulated with SCA as

$$\begin{aligned}((w_{k,i}^{Re})^{(t)})^2 + ((w_{k,i}^{Im})^{(t)})^2 + 2(w_{k,i}^{Re})^{(t)}(w_{k,i}^{Re} - (w_{k,i}^{Re})^{(t)}) + 2(w_{k,i}^{Im})^{(t)}(w_{k,i}^{Im} - (w_{k,i}^{Im})^{(t)}) \\ \geq (2^{R_{k,min}} - 1)(\sum_{l=k+1}^{K} \|\mathbf{h}_{,i}^H \mathbf{w}_l\|^2 + \sigma^2),\end{aligned} \tag{3.19}$$

where $k+1 \leq l \leq K$. Constraint (3.19) is convex. In summary, the convex problem can be formulated as

$$\max_{\mathbf{w}_1,\cdots,\mathbf{w}_K,\alpha,\rho,\gamma,\delta,\Omega,w^{Re},w^{Im}} \quad \alpha, \tag{3.20a}$$

$$\text{s. t.} \quad \text{(3.6c), (3.9), (3.11b), (3.15a), (3.15b), (3.16), (3.17), (3.19),} \tag{3.20b}$$

where t is the t^{th} iteration of the First-order Taylor expansion. (3.20) is a convex optimization problem.

3.4.2 Phase Shift Optimization

In this section, we focus on the phase shift optimization for RIS-NOMA. In the above section, the beamforming optimization has been solved with the fixed phase shifts. Given the optimized beamforming vectors $\mathbf{w}^f$, the denominator of the objective function (3.6a) is a constant and the energy-efficiency problem can be transformed to a sum rate optimization

problem as followed:

$$\max_{\mathbf{\Theta}} \quad \sum_{k=1}^{K} R_k, \tag{3.21a}$$

$$\text{s. t.} \quad R_k \geq R_{k,min}, k = 1, \cdots, K, \tag{3.21b}$$

$$\theta_n \in [0, 2\pi], n = 1, \cdots, N. \tag{3.21c}$$

In [42], the phase shift optimization is dealt with SDR. In this work, (3.21) is optimized via SCA. Similar to (3.6), (3.21) is still a non-convex problem because R_k is not convex with respect to $\mathbf{\Theta}$. To deal with it, Define a variable vector $\mathbf{v} = [v_1, \cdots, v_N]^H = [e^{j\theta_1}, \cdots, e^{j\theta_N}]^H$ and vectors $\mathbf{a}_{k,i}$=diag$(\mathbf{h}_{i,r}^H)\mathbf{G}\mathbf{w}_k^f$ which $k \leq i \leq K$ and $\mathbf{w}_k^f$ denotes the optimized beamforming vector, which is fixed. Thus, the SINR of the k^{th} user decoded at the i^{th} user ($k \leq i$) can be expressed as

$$SINR_{k,i} = \frac{\|\mathbf{v}^H\mathbf{a}_{k,i} + \mathbf{h}_{i,d}^H\mathbf{w}_k^f\|^2}{\sum\limits_{l=k+1}^{K} \|\mathbf{v}^H\mathbf{a}_{l,i} + \mathbf{h}_{i,d}^H\mathbf{w}_l^f\|^2 + \sigma^2}. \tag{3.22}$$

For the case of $k = i = K$, the SINR expression can be written as

$$SINR_K = \frac{\|\mathbf{v}^H\mathbf{a}_{K,K} + \mathbf{h}_{K,d}^H\mathbf{w}_K^f\|^2}{\sigma^2}. \tag{3.23}$$

Similar to process of beamforming optimization, define a slack variable q such that

$$\sum_{k=1}^{K} R_k \geq q. \tag{3.24}$$

Then, a new vector $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_K]$ is further introduced to rewrite (3.24)

$$\sum_{k=1}^{K} \log_2(\lambda_k) \geq q, \tag{3.25a}$$

$$1 + SINR_k \geq \lambda_k. \tag{3.25b}$$

Then, another variable vector named $\boldsymbol{\mu} = [\mu_1, \cdots, \mu_K]$ is applied such that,

$$\sum_{k=1}^{K} \mu_k \geq q, \tag{3.26a}$$

$$\lambda_k \geq 2^{\mu_k}. \tag{3.26b}$$

By applying the similar method to deal with the SINR part, additional variables $\varphi_{k,i}$ which have the same purpose with $\Omega_{k,i}$ are applied as

$$\|\mathbf{v}^H \mathbf{a}_{k,i} + \mathbf{h}_{i,d}^H \mathbf{w}_k^f\|^2 \geq (\lambda_k - 1)\varphi_{k,i}, \tag{3.27a}$$

$$\sum_{l=k+1}^{K} \|\mathbf{v}^H \mathbf{a}_{l,i} + \mathbf{h}_{i,d}^H \mathbf{w}_l^f\|^2 + \sigma^2 \leq \varphi_{k,i}. \tag{3.27b}$$

For the situation of $i = k = K$, the expressions should be written as

$$\|\mathbf{v}^H \mathbf{a}_{K,K} + \mathbf{h}_{K,d}^H \mathbf{w}_K^f\|^2 \geq (\lambda_K - 1)\varphi_{K,K}, \tag{3.28a}$$

$$\sigma^2 \leq \varphi_{K,K}. \tag{3.28b}$$

Then, the First-order Taylor approximation is used to make (3.27a) convex. Introduce a series of variables $v_{k,i}^{Re}$ and $v_{k,i}^{Im}$, such that

$$v_{k,i}^{Re} = \Re\{(\mathbf{v})^H \mathbf{a}_{k,i} + \mathbf{h}_{i,d}^H \mathbf{w}_k^f\}, \tag{3.29a}$$

$$v_{k,i}^{Im} = \Im\{(\mathbf{v})^H \mathbf{a}_{k,i} + \mathbf{h}_{i,d}^H \mathbf{w}_k^f\}. \tag{3.29b}$$

Then, extend (3.27a) via SCA,

$$\begin{aligned}((v_{k,i}^{Re})^{(t)})^2 + ((v_{k,i}^{Im})^{(t)})^2 + 2(v_{k,i}^{Re})^{(t)}(v_{k,i}^{Re} - (v_{k,i}^{Re})^{(t)}) + 2(v_{k,i}^{Im})^{(t)}(v_{k,i}^{Im} - (v_{k,i}^{Im})^{(t)}) \\ \geq ((\lambda_k)^{(t)} - 1)(\varphi_{k,i})^{(t)} + (\varphi_{k,i})^{(t)}(\lambda_k - (\lambda_k)^{(t)}) + ((\lambda_k)^{(t)} - 1)(\varphi_{k,i} - (\varphi_{k,i})^{(t)}).\end{aligned} \tag{3.30}$$

To deal with the QoS constraints, the same technique can still be applied to make it become a convex expression,

$$\begin{aligned}((v_{k,i}^{Re})^{(t)})^2 + ((v_{k,i}^{Im})^{(t)})^2 + 2(v_{k,i}^{Re})^{(t)}(v_{k,i}^{Re} - (v_{k,i}^{Re})^{(t)}) + 2(v_{k,i}^{Im})^{(t)}(v_{k,i}^{Im} - (v_{k,i}^{Im})^{(t)}) \\ \geq (2^{R_{k,min}} - 1)(\sum_{l=k+1}^{K} \|(\mathbf{v})^H \mathbf{a}_{l,i} + \mathbf{h}_{i,d}^H \mathbf{w}_l^f\|^2 + \sigma^2).\end{aligned} \tag{3.31}$$

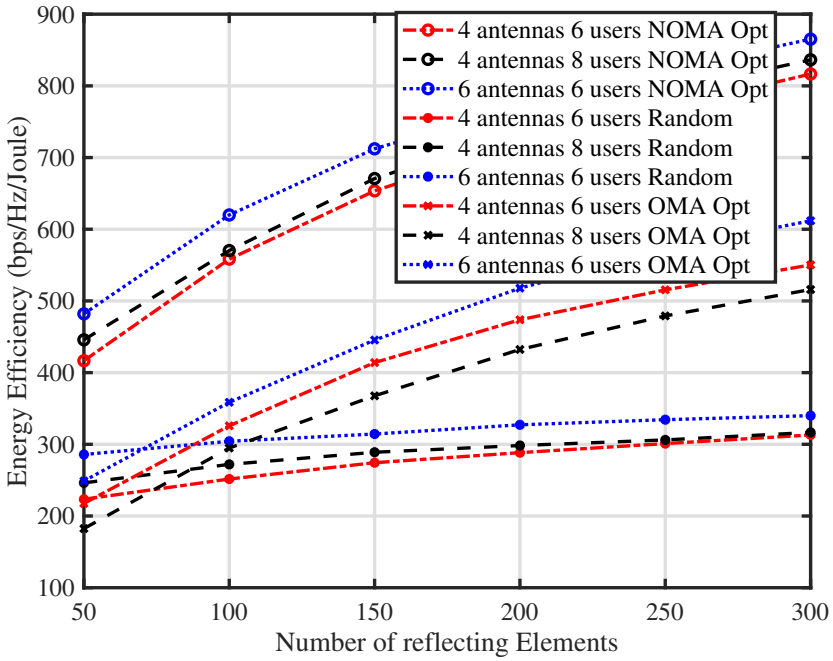

Fig. 3.2. Energy efficiency versus the number of reflecting elements

After the transformation of (3.21), the phase shift optimization can be expressed as

$$\max_{\mathbf{v},q,\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\varphi},v^{Re},v^{Im}} \quad q, \tag{3.32a}$$

$$\text{s. t.} \quad \text{(3.21c), (3.26b), (3.29a), (3.29b), (3.30), (3.31).} \tag{3.32b}$$

(3.32) is a convex optimization problem.

3.5 Simulation Results

In this section, we present the simulation results of the proposed algorithm to illustrate the performance of the optimization scheme in this chapter. We consider that the distance between the BS and the RIS is 10 m [43]. The distance between the BS and every user is within the range between 16 m and 20 m and each user is also 16-20 m away from the RIS [43]. The channels from the BS to the RIS and from the RIS to the users are Rician fading channels. The Rician factor κ of the channel is set to 2. Besides, the channels from the BS to the users obey Rayleigh distribution. The path loss exponents from the BS to the RIS, the RIS to the users and the BS to the users are all set to 2.5. Without the loss of generality, the minimal SINR requirement for all users is -10 dB and the additive white Gaussian noise (AWGN) σ^2 is -60 dB [43].

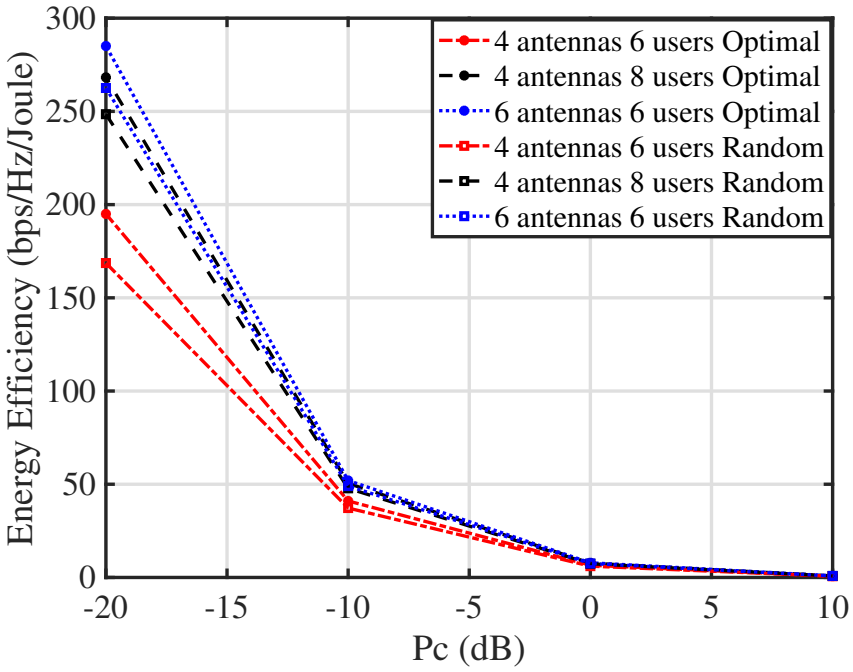

Fig. 3.3. Energy efficiency versus internal circuit power

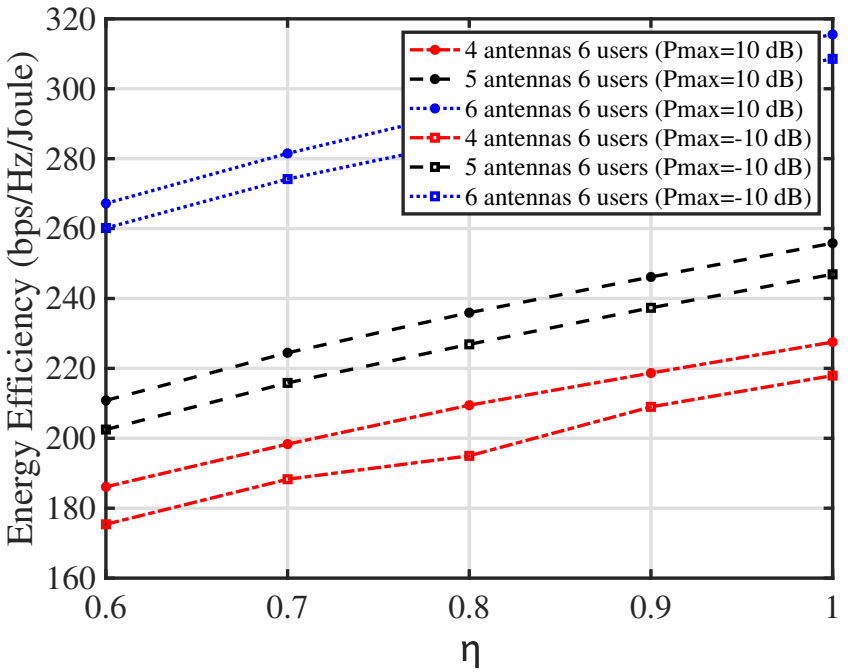

Fig. 3.4. Energy efficiency versus the power amplifier coefficient

Fig. 3.2 shows the relationship between energy efficiency and the number of the reflecting elements. Under this circumstance, P_c and P_{max} are set to -20 dB and -10 dB, respectively. η is set to 0.8. For a comprehensive analysis, three different combinations of the numbers of the antennas and the users are selected to compare their energy efficiency values with those with random phase shifts and OMA cases. Initially, we use '4 antennas, 6 users' as an example. From this graph, it can be seen that the energy efficiency values optimized by the proposed algorithm are from 416.63 bps/Hz/Joule to 816.57 bps/Hz/Joule as the number of

elements increases. Meanwhile, the growth of energy efficiency values with random phase shifts is only from 223.4 bps/Hz/Joule to 313.05 bps/Hz/Joule and the energy efficiency for OMA case is within the range between 217.58 bps/Hz/Joule and 550.22 bps/Hz/Joule. It can be observed that the energy efficiency values optimized by the proposed algorithm are much higher than those with random phase shifts and OMA cases. Besides, it can be also found that the value of energy efficiency increases as the number of elements increases. At the beginning, energy efficiency increases very rapidly. However, the curve for each case is saturated gradually as the number of elements grows. This is because the effect of the phase shift optimization for EE becomes weaker as the number of the elements grows. When the number of elements reaches to a certain large value, the growth rate of the energy efficiency cannot be as high as those with less elements and the curve of the EE is likely to be saturated. Therefore, a further increase of the elements cannot improve the energy efficiency as significantly as the cases with a relatively small quantity of the elements.

Fig. 3.3 shows the effect of the internal circuit consumption on energy efficiency value. the number of reflecting elements is set to 10 and P_{max} is set to -10 dB. The circuit power is set within the range from -20 dB to 10 dB and amplifier coefficient is 0.8, respectively. From the curves under both optimized and random phase shifts circumstances, it can be found that the energy efficiency value decreases as the internal circuit power increases. Besides, when P_c rises from -20 dB to -10 dB, there is a sharp decrease in the energy efficiency value. The curve approaches to zero when P_c further rises. It is because the sum power consumption of beamforming is no more than P_{max} . Energy efficiency could decline as P_c increases. As P_c approaches and is far more than P_{max}, the rate of decline of energy efficiency is lower and lower because energy efficiency should not be less than zero in a network.

Fig. 3.4 shows the relationship between energy efficiency and the power amplifier coefficient. The number of elements is set to 10 and P_c is set to -20 dB. Meanwhile, two schemes with different transmitting power P_{max} are considered. From the curves, it can be observed that the energy efficiency can increase as the value of η increases. More importantly, when the number of antennas and users is fixed, the energy efficiency with a higher transmitting power budget is greater than that with a lower P_{max}. It is because that the increase of P_{max} expands the domain of the beamforming vectors. Hence, the energy efficiency is more likely to be optimized to a higher value during the optimization process.

3.6 Conclusions of the Chapter

In this chapter, we have designed an algorithm to maximize the system in an RIS assisted NOMA system with the multi-user case. Firstly, we decoupled the primal problem into two suboptimal problems. By alternatively optimizing the beamforming vectors and phase shifts, the optimal energy efficiency could be obtained through the proposed iterative algorithm. To deal with the non-convex constraints in these two sub-problems, the problem relaxation method was firstly applied to simplify it by introducing slack variables. SCA was also applied as a mathematical tool to convert some non-convex constraints to convex. It is noted that this work can be extended to a MIMO RIS assisted NOMA system. Machine learning is a promising method to be applied to address this new model.

Chapter 4

Joint Phase Shift and Beamforming Design in a Multi-User MISO STAR-RIS Assisted Downlink NOMA Network

4.1 Introduction

4.1.1 Motivations of the Chapter

In **Chapter 3**, a multi-user MISO RIS assisted NOMA network has been studied, in which a significant improvement of RIS and NOMA on EE is demonstrated. Motivated by the great success of RIS and NOMA, more research potentials of STAR-RIS assisted NOMA networks are expected to be exploited. In fact, there have been plenty of progress on the research of STAR-RIS assisted NOMA networks. For example, the authors in [43] and [44] proposed the mathematical scenarios to minimize the power consumption in the multi-cluster NOMA networks. However, as mentioned in **Chapter 1** and **Chapter 3**, EE is also a crucial measure to evaluate the performance of the networks. In addition, there has not been adequate studies related to EE maximization in STAR-RIS NOMA networks. Therefore, following the networks in [43] and [44], one of the motivations of this chapter is to achieve the energy efficiency maximization in a multi-cluster MISO STAR-RIS assisted downlink system, where each cluster has one near user and one far user and NOMA is applied on each cluster to reduce the interference. Besides, the number of transmitting clusters is the same as the reflecting counterparts. Without the direct links between the BS and the users,

when the STAR-RIS receives the incident signals, the elements can transmit and reflect them to the users in the clusters. By tuning the electromagnetic properties of the elements in STAR-RIS, the transmitting and reflected signals can be modified. From the mathematics perspective, the RIS or STAR-RIS phase shift and beamforming optimization problems are usually related to the method of SDR. The rank-one constraints can be introduced after the SDR, which makes the problems NP-hard. In practice, this type of problems is challenging to solve. To tackle this type of problems, most existing works usually first solved the convex SDR problem by dropping rank-one constraints and then check whether the optimal solutions were rank-one. If not, Gaussian randomization was applied to find the suboptimal solutions of the original problem [76]–[78]. Besides, in some cases such as [78]–[81], the solutions to the SDR problems by dropping rank constraints could happen to be rank-one mathematically. However, in this system, plenty of the constraints in the proposed optimization problems indicate that it might be difficult to generate the feasible points by Gaussian randomization. To obtain the desired suboptimal solutions, more random points have to be sampled, which dramatically increases the complexity of the algorithms. Therefore, in this work, one novel mathematical tool and mathematical proof are adopted to replace the Gaussian randomization to tackle the rank-one constraints in the phase shift and beamforming optimization problems, respectively.

4.1.2 Contributions of the Chapter

The contributions of this chapter can be summarized below:

- A multi-cluster MISO STAR-RIS assisted downlink NOMA network is proposed to find the maximum energy efficiency by optimizing phase shifts of the STAR-RIS and beamforming vectors of the users. However, these two types of variables are coupled with each other and it is difficult to formulate the closed form of each variable. To address this dilemma, we decouple the original problem to two suboptimal problems and solve them alternatively.
- To solve the phase shift optimization problem, fractional programming [68], [82] is first exploited to address the non-convexity issue of objective function and constraints. Then, we transform the problem to an SDR one with rank-one constraints. Instead of applying

the Gaussian randomization, we propose another method called SROCR to address the rank-one issue and obtain a suboptimal solution [83].

- The challenge to solve the beamforming optimization problem is due to the fact that the objective function is non-convex fractional form. The key idea is to first apply the novel fractional programming combined with SDR and convert the QoS constraints to concave/convex forms. After this, instead of applying Dinkelbach's method to transform the objective function which can bring extra computational complexity, we apply fractional programming directly to transform the objective function to a concave form and it will greatly increase the efficiency of the algorithm. In this chapter, different from the phase shift optimization, the optimal solution of the SDR beamforming problem dropping rank constraints can be proven to be rank-one, which avoids the extra steps to tackle the rank issue.

- By comprehensively considering the quality of the solutions and the complexity of the algorithms, the proposed scheme can achieve better energy efficiency performance than the benchmark schemes, such as SDR by Gaussian randomization. Besides, this work also demonstrates the great value of the combination between the NOMA technology and STAR-RIS. For instance, the simulation results show that the energy efficiency performance of the STAR-RIS NOMA system can outperform that of the other systems, such as RIS NOMA, STAR-RIS OMA, RIS OMA. Furthermore, the proposed scheme also adapts to other types of the optimization problems, which involve non-convex quadratic fractional objective functions and constraints or have the rank-one constraints. Meanwhile, the proposed method can also be extended in the networks with more than two users in a cluster by a given pre-defined decoding order.

4.1.3 Particular Notations of the Chapter

Except the common mathematical notations in **Chapter 1**, there are also some particular variables and notations presenting in Table 4.1 for the convenience of the readers.

Variables/Notations	Descriptions
M, N, K	Number of antennas, elements and clusters
$\boldsymbol{\theta}_t$, $\boldsymbol{\theta}_r$	Phase shift angle vectors of transmitting and reflecting elements
β_t, β_r	amplitude coefficients of transmitting and reflecting vectors
k, $l \in \{t, r\}$, $i \in \{\varrho, \varsigma\}$	the index of the clusters, transmission or reflection region, near or far user
$U_{k,l,i}$	the i user in the k^{th} cluster of the region l
$\gamma_{k,l,i}$, $\gamma_{k,l,\varrho\to\varsigma}$	The SINR when decoding $U_{k,l,i}$'s signal at itself, the SINR when decoding $U_{k,l,\varsigma}$'s signal at $U_{k,l,\varrho}$
$\mathbf{z}$, $\boldsymbol{\alpha}$	the collections of slack variables for the SINRs in the logarithms
$\boldsymbol{\xi}$, $\boldsymbol{\zeta}$, $\boldsymbol{\delta}$	The collections of the auxiliary variables for $\gamma_{k,l,\varrho}$, $\gamma_{k,l,\varsigma\to\varsigma}$ and $\gamma_{k,l,\varrho\to\varsigma}$ to satisfy the QoS constraints in the phase shift optimization problem
$\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\lambda}$	The collections of the auxiliary variables for $\gamma_{k,l,\varrho}$, $\gamma_{k,l,\varsigma\to\varsigma}$ and $\gamma_{k,l,\varrho\to\varsigma}$ to satisfy the QoS constraints in the beamforming optimization problem
ω	The relaxation coefficient for rank constraint in the phase shift optimization problem
$\boldsymbol{\Lambda}$, $\hat{\boldsymbol{\Lambda}}$	The collections of the inter-cluster interference in the phase shift optimization problem
$\boldsymbol{\Omega}$, $\hat{\boldsymbol{\Omega}}$	The collections of the inter-cluster interference in the beamforming optimization problem

Table 4.1. Particular Variables/Notations

4.2 System Model

As is shown in Fig. 4.1, we consider a multi-cluster MISO STAR-RIS system. BS is equipped with M transmit antennas. Firstly, BS, STAR-RIS and users are located in the Cartesian coordinate system. the STAR-RIS elements of this system are N. There are totally $4K$ users and $2K$ clusters in this system, where $2K$ clusters are evenly distributed in transmission and reflection regions. Each cluster contains one near user and one far user. $l \in \{t, r\}$ refers to the transmission or reflection region, $i \in \{\varrho, \varsigma\}$ refers to the near user or far user. As is shown in Fig. 4.1, each region has totally K clusters and they are notated from 1 to K. $U_{k,l,i}$ refers to the i user in the k^{th} cluster of the region l, $k \in 1, \cdots, K$. The phase shift vector is $\boldsymbol{\theta}_l = \beta_l(e^{\mathrm{j}\theta_{1,l}}, \cdots, e^{\mathrm{j}\theta_{N,l}})$, where θ_n is to adjust the phase shift of the incident signals and $\sqrt{\beta_l}$ is to control the power allocation of the l elements. It is noted that $\beta_t + \beta_r = 1$. Because of the existence of the block, there is no direct links between the BS and users. $\mathbf{G} \in \mathbb{C}^{N\times M}$ is the channel gain from the BS to the STAR-RIS. $\widetilde{\mathbf{h}}_{k,l,i}^H \in \mathbb{C}^{1\times N}$ refers to the channel gain from the STAR-RIS to the i user of the k^{th} cluster in the region l. For example, $\widetilde{\mathbf{h}}_{k,t,\varrho}$ refers to the channel gain from the STAR-RIS to the near user of the k^{th} cluster in the transmission region. Therefore the total channel gain for the i user of the k^{th} cluster in region l, $\mathbf{h}_{k,l,i}$ can be expressed as

$$\mathbf{h}_{k,l,i}^H = \widetilde{\mathbf{h}}_{k,l,i}^H \mathrm{diag}(\boldsymbol{\theta}_l)\mathbf{G}, \tag{4.1}$$

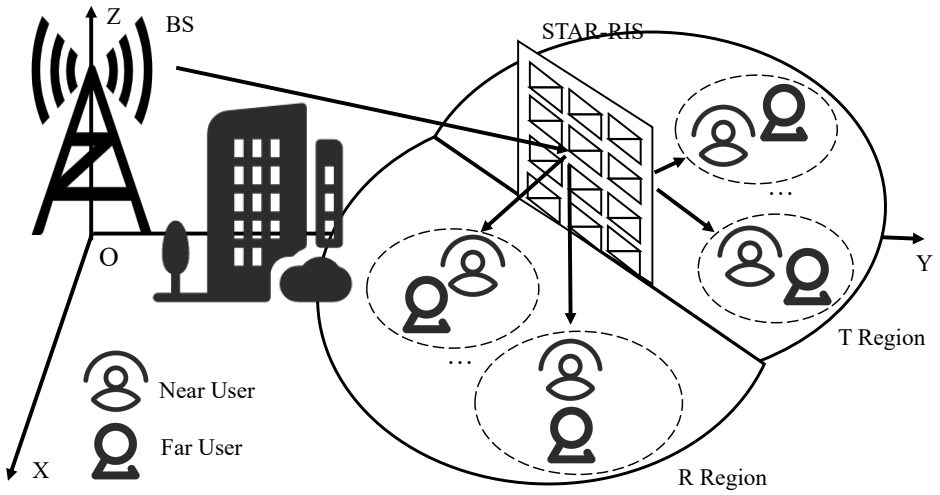

Fig. 4.1. MISO STAR-RIS NOMA downlink system

where $\mathbf{w}_{k,l,i} \in \mathbb{C}^{M\times 1}$ is the beamforming vector of the i user of the k^{th} cluster in the l region. $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN). Recall in **Chapter 3** that SIC is implemented with a given decoding order of all K users $\{U_1, \cdots, U_K\}$. Specifically, given two of these K users, denoted by U_k and U_j , $k < j$, the U_k's signal is decoded before U_j's one. By applying the SIC, the interference of the users whose indexes are less than k can be removed when decoding U_k's signal. However, if using such a scheme, it requires that each user's signal has to be decoded at itself and all the users who have a higher decoding priority than it. Thus, the decoding complexity of this system can be very high with a larger number of the users. To decrease the complexity during the decoding process, we group the users into clusters and apply NOMA only within each cluster. Specifically, the decoding order of this network is decided according to the user's decoding capability. In this chapter, we assume that the near user can have a higher decoding capability than the far user within a cluster. Therefore, by applying NOMA in each cluster, the far user's signal is decoded before the near user. For the near user's signal, SIC can be employed to cancel the intra-cluster interference from the far user in its cluster. In contrast, for the far user, it cannot decode the near user's signal. So the intra-cluster interference from the near user when decoding the far user's signal cannot be removed. This scheme enables each near user and far user's signals to be decoded only once and twice, respectively and greatly decreases the decoding complexity of the network. $\gamma_{k,l,\varrho}$ is the SINR for the k^{th} near user in region l. $\gamma_{k,l,\varrho\to\varsigma}$ is the SINR for signal of the k^{th} far user decoded at the near user. $\gamma_{k,l,\varsigma\to\varsigma}$ is the SINR for signal of the far user of the k^{th} cluster decoded by itself. After SIC, the SINR for each user with different

decoding situations can be expressed as followed:

$$\gamma_{k,l,\varrho} = \frac{|\mathbf{h}_{k,l,\varrho}^H \mathbf{w}_{k,l,\varrho}|^2}{\sigma^2 + \Omega_{k,l,\varrho}(\mathbf{w}) + \hat{\Omega}_{k,l,\varrho}(\mathbf{w})}, \tag{4.2a}$$

$$\gamma_{k,l,\varsigma\to\varsigma} = \frac{|\mathbf{h}_{k,l,\varsigma}^H \mathbf{w}_{k,l,\varsigma}|^2}{\sigma^2 + |\mathbf{h}_{k,l,\varsigma}^H \mathbf{w}_{k,l,\varrho}|^2 + \Omega_{k,l,\varsigma}(\mathbf{w}) + \hat{\Omega}_{k,l,\varsigma}(\mathbf{w})}, \tag{4.2b}$$

$$\gamma_{k,l,\varrho\to\varsigma} = \frac{|\mathbf{h}_{k,l,\varrho}^H \mathbf{w}_{k,l,\varsigma}|^2}{\sigma^2 + |\mathbf{h}_{k,l,\varrho}^H \mathbf{w}_{k,l,\varrho}|^2 + \Omega_{k,l,\varrho}(\mathbf{w}) + \hat{\Omega}_{k,l,\varrho}(\mathbf{w})}, \tag{4.2c}$$

where

$$\Omega_{k,l,i}(\mathbf{w}) = \sum_{j=1,j\neq k}^{K} \sum_{i'\in\{\varrho,\varsigma\}} (|\mathbf{h}_{k,l,i}^H \mathbf{w}_{j,l,i'}|^2), \tag{4.3a}$$

$$\hat{\Omega}_{k,l,i}(\mathbf{w}) = \sum_{j=1}^{K} \sum_{i'\in\{\varrho,\varsigma\}} (|\mathbf{h}_{k,l,i}^H \mathbf{w}_{j,l',i'}|^2), \tag{4.3b}$$

where $\mathbf{w}$ is the collection of all the beamforming vectors, $i, i' \in \{\varrho, \varsigma\}$ and $l' \in \{t, r\}, l' \neq l$. In this network, the SIC can be carried out by satisfying the conditions that the near user's signal of the k^{th} cluster can be decoded by itself and the far user's signal can be decoded by itself and the near user in the same cluster successfully. Therefore, the achievable SINR of the far user in the k^{th} cluster at l region should be $\gamma_{k,l,\varsigma} = min(\gamma_{k,l,\varsigma\to\varsigma}, \gamma_{k,l,\varrho\to\varsigma})$ and the achievable rate for the near and far users can be expressed as followed:

$$R_{k,l,\varrho} = \log_2(1 + \gamma_{k,l,\varrho}), \tag{4.4a}$$

$$R_{k,l,\varsigma} = \log_2(1 + \gamma_{k,l,\varsigma}). \tag{4.4b}$$

Define the power amplifier coefficient and power consumption for internal circuit are η and P_c, respectively. The energy efficiency optimization problem for the entire users can be formulated as:

$$\max_{\mathbf{w},\boldsymbol{\theta}} \quad \frac{f(\mathbf{w},\boldsymbol{\theta})}{\ln(2)(\frac{1}{\eta}P + P_c)}, \tag{4.5a}$$

$$\text{s. t.} \quad \gamma_{k,l,\varrho} \geq \gamma_{k,l,\varrho,min}, \tag{4.5b}$$

$$\gamma_{k,l,\varsigma\to\varsigma} \geq \gamma_{k,l,\varsigma,min}, \tag{4.5c}$$

$$\gamma_{k,l,\varrho\to\varsigma} \geq \gamma_{k,l,\varsigma,min}, \tag{4.5d}$$

$$P \leq P_{max}, \tag{4.5e}$$

$$\theta_{n,l} \in [0, 2\pi], n = 1, \cdots, N, \tag{4.5f}$$

where $\boldsymbol{\theta}$ is denoted as a collection of the phase shift vectors, respectively and $f(\mathbf{w}, \boldsymbol{\theta})$ is

$$f(\mathbf{w}, \boldsymbol{\theta}) = \sum_{k=1}^{K} \sum_{l \in \{t,r\}} \sum_{i \in \{\varrho,\varsigma\}} \ln(1 + \gamma_{k,l,i}), \tag{4.6}$$

and

$$P = \sum_{k=1}^{K} \sum_{l \in \{t,r\}} \sum_{i \in \{\varrho,\varsigma\}} ||\mathbf{w}_{k,l,i}||^2. \tag{4.7}$$

(4.5b)-(4.5d) are to satisfy the quality of service requirements, which is to ensure that the rate of each user should be greater than a certain value. $\gamma_{k,l,\varrho,min}$, and $\gamma_{k,l,\varsigma,min}$ are constant values. (4.5e) is to guarantee the total transmitting power consumption of the beamforming P is limited by the maximum power budget P_{max}.

4.3 Proposed Optimization Method

In this section, we aim to design a scheme to effectively transform the non-convex optimization problem (4.5) to convex and propose the corresponding algorithms. From the formulation of the original problem, it can be observed that there are totally two types of variables to be optimized, one is the phase shifts of the STAR-RIS elements and another is the beamforming of the users. However, it is difficult to solve this problem directly because two types of variables are coupled with each other. To address the coupling issue, we propose the alternating optimization scheme to split the original problem to phase shift and beamforming optimization problems, respectively and alternatively solve them. Besides, the original non-convex objective function and constraints can also greatly increase the difficulty to tackle this problem. Therefore, it is necessary to transform the original problem to a solvable convex optimization problem by applying mathematical techniques.

4.3.1 Phase Shift Optimization

In the phase shift optimization, we first convert the original problem to a SDR one with rank-one constraints. After the transformation of the SDR of the original problem, we apply

the fractional programming method to transform the non-convex objective function and QoS constraints to convex by introducing the auxiliary variables. To solve the non-convex rank-one constraints, SROCR is proposed to partially relax the rank-one constraints.
From the proposed optimization problem, it can be observed that the phase shift optimization is actually a sum rate maximization problem related to the phase shifts of the elements because the denominator part of the objective function is only decided by the beamforming vectors. Thus, the constraint (4.5e) can be temporarily neglected. In the first step, given the fixed beamforming vectors $\mathbf{w}$, we can define a collection of constant vectors $\{\mathbf{a}_{k,l,i}^{j,l',i'}\}$, where $\mathbf{a}_{k,l,i}^{j,l',i'} = \text{diag}(\widetilde{\mathbf{h}}_{k,l,i}^H)\mathbf{G}\mathbf{w}_{j,l',i'}$ and the phase shift vector $\mathbf{v}_l = \beta_l[v_{1,l}, \cdots, v_{N,l}]^H = [e^{\mathrm{j}\theta_{1,l}}, \cdots, e^{\mathrm{j}\theta_{N,l}}]^H$, where $k, j = 1, \cdots, K$, $l, l^{'} \in \{t, r\}$ and $i, i^{'} \in \{\varrho, \varsigma\}$. Therefore, the SINR for each user with different decoding situations could be expressed as followed:

$$\gamma_{k,l,\varrho} = \frac{\|\mathbf{v}_l^H \mathbf{a}_{k,l,\varrho}^{k,l,\varrho}\|^2}{\sigma^2 + \Gamma_{k,l,\varrho}(\mathbf{v}_l) + \hat{\Gamma}_{k,l,\varrho}(\mathbf{v}_l)}, \tag{4.8a}$$

$$\gamma_{k,l,\varsigma\to\varsigma} = \frac{\|\mathbf{v}_l^H \mathbf{a}_{k,l,\varsigma}^{k,l,\varsigma}\|^2}{\sigma^2 + \|\mathbf{v}_l^H \mathbf{a}_{k,l,\varsigma}^{k,l,\varrho}\|^2 + \Gamma_{k,l,\varsigma}(\mathbf{v}_l) + \hat{\Gamma}_{k,l,\varsigma}(\mathbf{v}_l)}, \tag{4.8b}$$

$$\gamma_{k,l,\varrho\to\varsigma} = \frac{\|\mathbf{v}_l^H \mathbf{a}_{k,l,\varrho}^{k,l,\varsigma}\|^2}{\sigma^2 + \|\mathbf{v}_l^H \mathbf{a}_{k,l,\varrho}^{k,l,\varrho}\|^2 + \Gamma_{k,l,\varrho}(\mathbf{v}_l) + \hat{\Gamma}_{k,l,\varrho}(\mathbf{v}_l)}, \tag{4.8c}$$

where

$$\Gamma_{k,l,i}(\mathbf{v}_l) = \sum_{j=1, j\neq k}^{K} \sum_{i'\in\{\varrho,\varsigma\}} (\|\mathbf{v}_l^H \mathbf{a}_{k,l,i}^{k,l,i'}\|^2), \tag{4.9a}$$

$$\hat{\Gamma}_{k,l,i}(\mathbf{v}_l) = \sum_{j=1}^{K} \sum_{i'\in\{\varrho,\varsigma\}} (\|\mathbf{v}_l^H \mathbf{a}_{k,l,i}^{k,l',i'}\|^2). \tag{4.9b}$$

The sum rate function is not concave because the fractional terms are inside the logarithms. To solve this problem, we introduce a collection of auxiliary slack variables $\{z_{k,l,i}\}$ to make the SINR expression out of the logarithms. Therefore, (4.5b), (4.5c), (4.5d) can be transformed as followed in this network:

$$\gamma_{k,l,\varrho} \geq z_{k,l,\varrho}, \tag{4.10a}$$

$$\gamma_{k,l,\varsigma\to\varsigma} \geq z_{k,l,\varsigma}, \tag{4.10b}$$

$$\gamma_{k,l,\varrho\to\varsigma} \geq z_{k,l,\varsigma}, \tag{4.10c}$$

and

$$z_{k,l,\varrho} \geq \gamma_{k,l,\varrho,min}, \tag{4.11a}$$

$$z_{k,l,\varsigma} \geq \gamma_{k,l,\varsigma,min}. \tag{4.11b}$$

The original EE problem can be converted to following form

$$\max_{\mathbf{v},\boldsymbol{\beta},\boldsymbol{z}} \quad \sum_{k=1}^{K} \sum_{l\in\{t,r\}} \sum_{i\in\{\varrho,\varsigma\}} \ln(1+z_{k,l,i}), \tag{4.12a}$$

$$\text{s. t.} \quad (4.10a), (4.10b), (4.10c), (4.11a), (4.11b), \tag{4.12b}$$

$$\|\mathbf{v}_l(n)\| = \beta_l, n = 1, \cdots, N, \tag{4.12c}$$

$$\beta_t + \beta_r = 1. \tag{4.12d}$$

It is noted that (4.12) is still not a convex optimization problem. On the one hand, although the fractional forms have been removed out of the logarithms and the objective function become convex, the non-convex fractional forms still exist in the constraints. On the other hand, for each STAR-RIS elements, the amplitude of each phase shift has to be β_l but this type of equality constraint is not affine, which cannot be solved by some optimization toolboxes. Therefore, more mathematical techniques are needed to further transform it to convex.

Inspired by [41], by introducing a new variable matrix $\mathbf{V}_l = \mathbf{v}_l^H \mathbf{v}_l$, which is a positive semidefinite and rank-one matrix. After this, it can be found that $\|\mathbf{v}_l \mathbf{a}_{k,l,i}^{j,l',i'}\|^2$ is equivalent to $\text{Tr}(\mathbf{A}_{k,l,i}^{j,l',i'} \mathbf{V}_l)$, which $\mathbf{A}_{k,l,i}^{j,l',i'} = \mathbf{a}_{k,l,i}^{j,l',i'} (\mathbf{a}_{k,l,i}^{j,l',i'})^H$. The SINR for each user can be rewritten as:

$$\gamma_{k,l,\varrho} = \frac{\text{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varrho} \mathbf{V}_l)}{\sigma^2 + \Gamma_{k,l,\varrho}(\mathbf{V}_l) + \hat{\Gamma}_{k,l,\varrho}(\mathbf{V}_l)}, \tag{4.13a}$$

$$\gamma_{k,l,\varsigma\to\varsigma} = \frac{\text{Tr}(\mathbf{A}_{k,l,\varsigma}^{k,l,\varsigma} \mathbf{V}_l)}{\sigma^2 + \text{Tr}(\mathbf{A}_{k,l,\varsigma}^{k,l,\varrho} \mathbf{V}_l) + \Gamma_{k,l,\varsigma}(\mathbf{V}_l) + \hat{\Gamma}_{k,l,\varsigma}(\mathbf{V}_l)}, \tag{4.13b}$$

$$\gamma_{k,l,\varrho\to\varsigma} = \frac{\text{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varsigma} \mathbf{V}_l)}{\sigma^2 + \text{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varrho} \mathbf{V}_l) + \Gamma_{k,l,\varrho}(\mathbf{V}_l) + \hat{\Gamma}_{k,l,\varrho}(\mathbf{V}_l)}, \tag{4.13c}$$

and (4.9a) and (4.9b) can be reformulated as:

$$\Gamma_{k,l,i}(\mathbf{V}_l) = \sum_{j=1, j\neq k}^{K} \sum_{i'\in\{\varrho,\varsigma\}} (\text{Tr}(\mathbf{A}_{k,l,i}^{j,l,i'} \mathbf{V}_l)), \tag{4.14a}$$

$$\hat{\Gamma}_{k,l,i}(\mathbf{V}_l) = \sum_{j=1}^{K} \sum_{i' \in \{\varrho,\varsigma\}} (\mathrm{Tr}(\mathbf{A}_{k,l,i}^{j,l',i'} \mathbf{V}_l)). \tag{4.14b}$$

By transforming the quadratic expressions to affine, (4.12c) can be equivalent that each diagonal element of $\mathbf{V}_l$ is one and it should be the rank-one matrix. Meanwhile, to tackle the non-convex QoS constraints, the fractional programming method can be utilized to transform them to convex ones. Inspired by the *Corollary 3* of [68], by introducing several real auxiliary variables $\xi_{k,l}$, $\zeta_{k,l}$ and $\delta_{k,l}$, (4.13a), (4.13b), (4.13c) can be equivalent to

$$\max_{\boldsymbol{\xi}} \quad g_{k,l,\varrho}(\boldsymbol{\xi}) = 2\xi_{k,l}\sqrt{\mathrm{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varrho}\mathbf{V}_l)} - \xi_{k,l}^2(\sigma^2 + \Lambda_{k,l,\varrho}(\mathbf{V}_l)) \geq z_{k,l,\varrho}, \tag{4.15a}$$

$$\max_{\boldsymbol{\zeta}} \quad g_{k,l,\varsigma\to\varsigma}(\boldsymbol{\zeta}) = 2\zeta_{k,l}\sqrt{\mathrm{Tr}(\mathbf{A}_{k,l,\varsigma}^{k,l,\varsigma}\mathbf{V}_l)} - \zeta_{k,l}^2(\sigma^2 + \mathrm{Tr}(\mathbf{A}_{k,l,\varsigma}^{k,l,\varrho}\mathbf{V}_l) + \Lambda_{k,l,\varsigma}(\mathbf{V}_l)) \geq z_{k,l,\varsigma}, \tag{4.15b}$$

$$\max_{\boldsymbol{\delta}} \quad g_{k,l,\varrho\to\varsigma}(\boldsymbol{\delta}) = 2\delta_{k,l}\sqrt{\mathrm{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varsigma}\mathbf{V}_l)} - \delta_{k,l}^2(\sigma^2 + \mathrm{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varrho}\mathbf{V}_l) + \Lambda_{k,l,\varrho}(\mathbf{V}_l)) \geq z_{k,l,\varsigma}, \tag{4.15c}$$

where $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$, $\boldsymbol{\delta}$ are the collection of $\xi_{k,l}$, $\zeta_{k,l}$, $\delta_{k,l}$ and

$$\Lambda_{k,l,i}(\mathbf{V}_l) = \Gamma_{k,l,i}(\mathbf{V}_l) + \hat{\Gamma}_{k,l,i}(\mathbf{V}_l). \tag{4.16}$$

The optimal solutions for $\xi_{k,l}$, $\zeta_{k,l}$ and $\delta_{k,l}$ can be found by finding partial derivatives $\frac{\partial g_{k,l,\varrho}}{\partial \xi_{k,l}}$, $\frac{\partial g_{k,l,\varsigma\to\varsigma}}{\partial \zeta_{k,l}}$ and $\frac{\partial g_{k,l,\varrho\to\varsigma}}{\partial \delta_{k,l}}$ as followed:

$$\xi_{k,l}^* = \frac{\sqrt{\mathrm{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varrho}\mathbf{V}_l)}}{\sigma^2 + \Lambda_{k,l,\varrho}(\mathbf{V}_l)}, \tag{4.17a}$$

$$\zeta_{k,l}^* = \frac{\sqrt{\mathrm{Tr}(\mathbf{A}_{k,l,\varsigma}^{k,l,\varsigma}\mathbf{V}_l)}}{\sigma^2 + \mathrm{Tr}(\mathbf{A}_{k,l,\varsigma}^{k,l,\varrho}\mathbf{V}_l) + \Lambda_{k,l,\varsigma}(\mathbf{V}_l)}, \tag{4.17b}$$

$$\delta_{k,l}^* = \frac{\sqrt{\mathrm{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varsigma}\mathbf{V}_l)}}{\sigma^2 + \mathrm{Tr}(\mathbf{A}_{k,l,\varrho}^{k,l,\varrho}\mathbf{V}_l) + \Lambda_{k,l,\varrho}(\mathbf{V}_l)}. \tag{4.17c}$$

By substituting (4.17) into (4.15), we can obtain the same expressions of (4.13). Therefore, it can be observed that the problem can be alternatively solved between $\{\mathbf{V}_l, \beta_l\}$ and $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$, $\boldsymbol{\delta}$. However, since SDR is applied to transform the quadratic expressions to the affine trace form,

the rank constraints are introduced to the problem formulation:

$$\max_{\mathbf{v},\boldsymbol{\beta},\boldsymbol{z}} \quad \sum_{k=1}^{K} \sum_{l\in\{t,r\}} \sum_{i\in\{\varrho,\varsigma\}} \ln(1+z_{k,l,i}), \tag{4.18a}$$

$$\text{s. t.} \quad (4.11a), (4.11b), (4.12d), (4.15a), (4.15b), (4.15c), \tag{4.18b}$$

$$\mathbf{V}_l(n,n) = \beta_l, l \in \{t,r\}, n = 1, \cdots N, \tag{4.18c}$$

$$\mathbf{V}_l \succeq 0, l \in \{t,r\}, \tag{4.18d}$$

$$\text{Rank}(\mathbf{V}_l) = 1, l \in \{t,r\}. \tag{4.18e}$$

$$\beta_t + \beta_r = 1. \tag{4.18f}$$

Until now, all the objective function and constraints in (4.18) are convex except (4.18e). According to [78], Gaussian randomization is usually a very effective technique to address the rank-one constraints [84], which is to firstly solve the convex SDR problem and obtain an optimal solution by dropping the rank-one constraints. Then, by generating several approximated feasible vectors according to [23], we can obtain a suboptimal solution of the original problem. The advantage of this method is that it can be easily implemented in practice and a high quality suboptimal solution can be obtained in some problems. For example, the authors in [78] pointed that the Gaussian randomization can guarantee $\frac{\pi}{4}-$approximation of the optimal value of the standard semidefinite programming (SDP) problem dropping rank-one constraints if the number of the approximated samples is large enough. However, the inadequacy of this method is also obvious. If the solution of the SDR problem is a high-rank matrix, it will be difficult to generate an approximated solution to satisfy all the other constraints of the problem [83]. To overcome such a dilemma, the author in [83] also proposed a novel algorithm called SROCR. It is designed to find a suboptimal solution by partially relaxing the rank-one constraints. Firstly, we combine $\mathbf{V}_t$ and $\mathbf{V}_r$ to one variable matrix $\mathbf{V} \in \mathbb{C}^{2N\times 2N}$, which is

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_t & \mathbf{Q} \\ \mathbf{Q}^H & \mathbf{V}_r \end{bmatrix}, \tag{4.19}$$

where $\mathbf{Q} \in \mathbb{C}^{N\times N}$ and $\mathbf{Q}(\iota, o) = \mathbf{v}_t(\iota)\mathbf{v}_r(o)$ for $\iota, o = 1, \cdots, N$. Due to the fact that $\mathbf{V}$ is a

rank-one matrix, (4.18e) can be equivalent to the following form

$$e_{max}(\mathbf{V}) = \text{Tr}(\mathbf{V}), \tag{4.20}$$

which one e_{max} is the largest eigenvalue of the matrix $\mathbf{V}$. Further, we can write that

$$\max_{\boldsymbol{\varphi}, \|\boldsymbol{\varphi}\| \leq 1} \boldsymbol{\varphi}^H \mathbf{V} \boldsymbol{\varphi} = \text{Tr}(\mathbf{V}). \tag{4.21}$$

From the characteristic of the rank-one matrix, the optimal solution $\boldsymbol{\varphi}^*$ can be obtained by finding the unit eigenvector corresponding to the maximum eigenvalue of $\mathbf{V}$. Inspired by this idea, we can find the optimal $\mathbf{V}^*$ by alternatively optimizing $\{\{\mathbf{V}, \boldsymbol{\beta}\}, \boldsymbol{\varphi}\}$. Until now, we can relax the rank-one constraint (4.18e) by introducing $\omega \in [0, 1]$ in the following:

$$(\boldsymbol{\varphi}^*)^H \mathbf{V} \boldsymbol{\varphi}^* \geq \omega \, \text{Tr}(\mathbf{V}), \tag{4.22}$$

where $\boldsymbol{\varphi}^*$ can be obtained in last iteration. (4.18) has been transformed to a convex problem:

$$\max_{\mathbf{V}, \boldsymbol{\beta}, \boldsymbol{z}} \quad \sum_{k=1}^{K} \sum_{l \in \{t,r\}} \sum_{i \in \{\varrho, \varsigma\}} \ln(1 + z_{k,l,i}), \tag{4.23a}$$

$$\text{s. t.} \quad (4.12\text{d}), (4.15\text{a}), (4.15\text{b}), (4.15\text{c}), (4.22), \tag{4.23b}$$

$$\mathbf{V} \succeq 0, \tag{4.23c}$$

$$\mathbf{V}(n, n) = \beta_t, n = 1, \cdots, N, \tag{4.23d}$$

$$\mathbf{V}(n, n) = \beta_r, n = N + 1, \cdots, 2N. \tag{4.23e}$$

$$\beta_t + \beta_r = 1. \tag{4.23f}$$

In summary, by alternatively optimizing $\{\mathbf{V}, \boldsymbol{\beta}\}$ and $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$, $\boldsymbol{\delta}$, $\boldsymbol{\varphi}$, we can obtain a rank-one suboptimal solution $\mathbf{V}^*$ and $\mathbf{v}_l^*$ can be obtained by the eigenvalue decomposition.

4.3.2 Beamforming Optimization

In the beamforming optimization, we also first convert the objective function and QoS constraints to convex by applying fractional programming and SDR and introducing the relevant variables. Different from the phase shift optimization, it can be proven that the

optimal solution of the SDR problem is rank one. Therefore, no more steps are needed to be implemented to tackle rank-one constraints.

Since the phase shifts of the STAR-RIS elements have been solved which can be considered fixed and beamforming vectors of the users have become the targeted optimization variables, the total channel gain for each i user of the k^{th} cluster in the region l can be expressed by (4.1). Similar to the phase shift optimization, we still use SDR to transform the quadratic beamforming expressions to affine. let $\mathbf{H}_{k,l,i} = \mathbf{h}_{k,l,i}\mathbf{h}_{k,l,i}^H$ and $\mathbf{W}_{k,l,i} = \mathbf{w}_{k,l,i}\mathbf{w}_{k,l,i}^H$, $\mathbf{H}_{k,l,i}$ and $\mathbf{W}_{k,l,i}$ is rank-one matrix . Therefore, the SINR for each user can be rewritten as

$$\gamma_{k,l,\varrho} = \frac{\text{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varrho})}{\sigma^2 + \Omega_{k,l,\varrho}(\mathbf{W}) + \hat{\Omega}_{k,l,\varrho}(\mathbf{W})}, \tag{4.24a}$$

$$\gamma_{k,l,\varsigma\to\varsigma} = \frac{\text{Tr}(\mathbf{H}_{k,l,\varsigma}\mathbf{W}_{k,l,\varsigma})}{\sigma^2 + \text{Tr}(\mathbf{H}_{k,l,\varsigma}^H\mathbf{W}_{k,l,\varrho}) + \Omega_{k,l,\varsigma}(\mathbf{W}) + \hat{\Omega}_{k,l,\varsigma}(\mathbf{W})}, \tag{4.24b}$$

$$\gamma_{k,l,\varrho\to\varsigma} = \frac{\text{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varsigma})}{\sigma^2 + \text{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varrho}) + \Omega_{k,l,\varrho}(\mathbf{W}) + \hat{\Omega}_{k,l,\varrho}(\mathbf{W})}, \tag{4.24c}$$

where

$$\Omega_{k,l,i}(\mathbf{W}) = \sum_{j=1,j\neq k}^{K} \sum_{i'\in\{\varrho,\varsigma\}} (\text{Tr}(\mathbf{H}_{k,l,i}\mathbf{W}_{j,l,i'})), \tag{4.25a}$$

$$\hat{\Omega}_{k,l,i}(\mathbf{W}) = \sum_{j=1}^{K} \sum_{i'\in\{\varrho,\varsigma\}} (\text{Tr}(\mathbf{H}_{k,l,i}\mathbf{W}_{j,l',i'})), \tag{4.25b}$$

where $\mathbf{W}$ is the collection of the beamforming matrices $\{\mathbf{W}_{k,l,i}\}$ $i, i' \in \{\varrho,\varsigma\}$ and $l' \in \{t, r\}, l' \neq l$. Different from the phase shift optimization, the objective function is fractional form which the variables exist both in the nominator and denominator. Firstly, the nominator part is still a sum rate expression. Therefore, the same method in the phase shift optimization problem can also be used to remove the fractional terms out of the logarithms by introducing several slack auxiliary variables $\{\alpha_{k,l,\varrho}, \alpha_{k,l,\varsigma}\}$ such that

$$\gamma_{k,l,\varrho} \geq \alpha_{k,l,\varrho}, \tag{4.26a}$$

$$\gamma_{k,l,\varsigma\to\varsigma} \geq \alpha_{k,l,\varsigma}, \tag{4.26b}$$

$$\gamma_{k,l,\varrho\to\varsigma} \geq \alpha_{k,l,\varsigma}, \tag{4.26c}$$

and

$$\alpha_{k,l,\varrho} \geq \gamma_{k,l,\varrho,min}, \tag{4.27a}$$

$$\alpha_{k,l,\varsigma} \geq \gamma_{k,l,\varsigma,min}. \tag{4.27b}$$

Therefore, the beamforming optimization problem can be reformulated as:

$$\max_{\mathbf{W},\boldsymbol{\alpha}} \quad \frac{\sum_{k=1}^{K} \sum_{l\in\{t,r\}} \sum_{i\in\{\varrho,\varsigma\}} \ln(1+\alpha_{k,l,i})}{\ln(2)(\frac{1}{\eta}P + P_c)}, \tag{4.28a}$$

$$\text{s. t.} \quad (4.26a), (4.26b), (4.26c), (4.27a), (4.27b), \tag{4.28b}$$

$$P \leq P_{max}, \tag{4.28c}$$

$$\mathbf{W}_{k,l,i} \succeq \mathbf{0}, \tag{4.28d}$$

$$\text{Rank}(\mathbf{W}_{k,l,i}) = 1, \tag{4.28e}$$

where

$$P = \sum_{k=1}^{K} \sum_{l\in\{t,r\}} \sum_{i\in\{\varrho,\varsigma\}} \text{Tr}(\mathbf{W}_{k,l,i}). \tag{4.29}$$

From the problem formulation, it can be observed that (4.28a) is a concave/convex objective function which can be solved by the Dinkelbach's method to obtain the optimal solution. However, the complexity of the algorithm can be greatly increased since the massive iterations will be inevitably introduced by the Dinkelbach's method. Therefore, in this work, the proposed optimization method for the beamforming aims to obtain the optimal solution with a relatively low complexity compared to the Dinkelbach's method. The first step is to apply the fractional programming method to transform the objective function to a concave form by introducing y such that

$$\max_{\mathbf{W},\boldsymbol{\alpha}} \quad 2y\sqrt{\sum_{k=1}^{K} \sum_{l\in\{t,r\}} \sum_{i\in\{\varrho,\varsigma\}} \ln(1+\alpha_{k,l,i})} - y^2 \ln(2)(\frac{1}{\eta}P + P_c), \tag{4.30}$$

where y is updated by each iteration in the algorithm. Next, to solve the QoS constraints (4.26a), (4.26b) and (4.26c), we can still apply fractional programming method to be

transformed to the following constraints:

$$\max_{\boldsymbol{\mu}} \quad 2\mu_{k,l}\sqrt{\mathrm{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varrho})} - \mu_{k,l}^2(\sigma^2 + I_{k,l,\varrho}(\mathbf{W})) \geq \alpha_{k,l,\varrho}, \tag{4.31a}$$

$$\max_{\boldsymbol{\tau}} \quad 2\tau_{k,l}\sqrt{\mathrm{Tr}(\mathbf{H}_{k,l,\varsigma}\mathbf{W}_{k,l,\varsigma})} - \tau_{k,l}^2(\sigma^2 + \mathrm{Tr}(\mathbf{H}_{k,l,\varsigma}\mathbf{W}_{k,l,\varrho}) + I_{k,l,\varsigma}(\mathbf{W})) \geq \alpha_{k,l,\varsigma}, \tag{4.31b}$$

$$\max_{\boldsymbol{\lambda}} \quad 2\lambda_{k,l}\sqrt{\mathrm{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varsigma})} - \lambda_{k,l}^2(\sigma^2 + \mathrm{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varrho}) + I_{k,l,\varrho}(\mathbf{W})) \geq \alpha_{k,l,\varsigma}, \tag{4.31c}$$

where

$$I_{k,l,i}(\mathbf{W}) = \Omega_{k,l,i}(\mathbf{W}) + \hat{\Omega}_{k,l,i}(\mathbf{W}). \tag{4.32}$$

Therefore, the beamforming optimization problem can be further written as:

$$\max_{\mathbf{W},\boldsymbol{\alpha}} \quad 2y\sqrt{\sum_{k=1}^{K}\sum_{l\in\{t,r\}}\sum_{i\in\{\varrho,\varsigma\}}\ln(1+\alpha_{k,l,i})} - y^2\ln(2)(\frac{1}{\eta}P + P_c), \tag{4.33a}$$

$$\text{s. t.} \quad \text{(4.28c), (4.28d), (4.28e), (4.31a), (4.31b), (4.31c).} \tag{4.33b}$$

It can be observed that the (4.28e) is still non-convex rank-one constraints.

Lemma 1 *: If solving* (4.33) *dropping* (4.28e)*, it can be guaranteed that any optimal beamforming solution* $\mathbf{W}_{k,l,i}$ *is rank-one. Proof: Please see the Appendix 4.6.*

Therefore, the optimal EE can be obtained by alternatively optimizing $\mathbf{W}_{k,l,i}$ and y, $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\lambda}$.

4.3.3 Algorithm Design

Before solving the phase shift and beamforming optimization problems alternatively, the feasible initial points for phase shift and beamforming vectors $\mathbf{w}^{(0)}$ and $\boldsymbol{\Theta}^{(0)}$ are supposed to be searched to guarantee the convergence of the algorithms we design to realize the EE maximization. Firstly, we are able to generate a series of random phase shift angles whose range is from 0 to 2π and β_l which satisfies $\beta_t + \beta_r = 1$ of the elements. For finding a feasible beamforming initial solution, we can apply the fractional programming for the complex case [68]. Back to the original problem formulation (4.5), consider there is a

feasible beamforming solution $\widetilde{\mathbf{w}}$, (4.5b), (4.5c), (4.5d) can be equivalent to the following form:

$$f_{k,l,\varrho}(\widetilde{\boldsymbol{\mu}}) = 2\Re\{\widetilde{\mu}_{k,l}^{\dagger}\mathbf{h}_{k,l,\varrho}^{H}\widetilde{\mathbf{w}}_{k,l,\varrho}\} - \widetilde{\mu}_{k,l}^{\dagger}(\sigma^2 + I_{k,l,\varrho}(\widetilde{\mathbf{w}}))\widetilde{\mu}_{k,l} \geq \gamma_{k,l,\varrho,min}, \tag{4.34a}$$

$$f_{k,l,\varsigma\to\varsigma}(\widetilde{\boldsymbol{\tau}}) = 2\Re\{\widetilde{\tau}_{k,l}^{\dagger}\mathbf{h}_{k,l,\varsigma}^{H}\widetilde{\mathbf{w}}_{k,l,\varsigma}\} - \widetilde{\tau}_{k,l}^{\dagger}(\sigma^2 + |\mathbf{h}_{k,l,\varsigma}^{H}\widetilde{\mathbf{w}}_{k,l,\varrho}|^2 + I_{k,l,\varsigma}(\widetilde{\mathbf{w}}))\widetilde{\tau}_{k,l} \geq \gamma_{k,l,\varsigma,min}, \tag{4.34b}$$

$$f_{k,l,\varrho\to\varsigma}(\widetilde{\boldsymbol{\lambda}}) = 2\Re\{\widetilde{\lambda}_{k,l}^{\dagger}\mathbf{h}_{k,l,\varrho}^{H}\widetilde{\mathbf{w}}_{k,l,\varsigma}\} - \widetilde{\lambda}_{k,l}^{\dagger}(\sigma^2 + |\mathbf{h}_{k,l,\varrho}^{H}\widetilde{\mathbf{w}}_{k,l,\varrho}|^2 + I_{k,l,\varrho}(\widetilde{\mathbf{w}}))\widetilde{\lambda}_{k,l} \geq \gamma_{k,l,\varsigma,min}, \tag{4.34c}$$

where

$$\widetilde{\mu}_{k,l} = \frac{\mathbf{h}_{k,l,\varrho}^{H}\widetilde{\mathbf{w}}_{k,l,\varrho}}{\sigma^2 + I_{k,l,\varrho}(\widetilde{\mathbf{w}})}, \tag{4.35a}$$

$$\widetilde{\tau}_{k,l} = \frac{\mathbf{h}_{k,l,\varsigma}^{H}\widetilde{\mathbf{w}}_{k,l,\varsigma}}{\sigma^2 + |\mathbf{h}_{k,l,\varsigma}^{H}\widetilde{\mathbf{w}}_{k,l,\varsigma}|^2 + I_{k,l,\varsigma}(\widetilde{\mathbf{w}})}, \tag{4.35b}$$

$$\widetilde{\lambda}_{k,l} = \frac{\mathbf{h}_{k,l,\varrho}^{H}\widetilde{\mathbf{w}}_{k,l,\varsigma}}{\sigma^2 + |\mathbf{h}_{k,l,\varsigma}^{H}\widetilde{\mathbf{w}}_{k,l,\varrho}|^2 + I_{k,l,\varrho}(\widetilde{\mathbf{w}})}, \tag{4.35c}$$

and $\widetilde{\boldsymbol{\mu}}$, $\widetilde{\boldsymbol{\tau}}$ and $\widetilde{\boldsymbol{\lambda}}$ are the collection of $\{\widetilde{\mu}_{k,l}\}$, $\{\widetilde{\tau}_{k,l}\}$, $\{\widetilde{\lambda}_{k,l}\}$ and only determined by $\widetilde{\mathbf{w}}$. If we generate random complex values $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\lambda}}$ to replace $\widetilde{\boldsymbol{\mu}}$, $\widetilde{\boldsymbol{\tau}}$ and $\widetilde{\boldsymbol{\lambda}}$, the following inequalities are satisfied:

$$f_{k,l,\varrho}(\widetilde{\boldsymbol{\mu}}) \geq f_{k,l,\varrho}(\hat{\boldsymbol{\mu}}), \tag{4.36a}$$

$$f_{k,l,\varsigma\to\varsigma}(\widetilde{\boldsymbol{\tau}}) \geq f_{k,l,\varsigma\to\varsigma}(\hat{\boldsymbol{\tau}}), \tag{4.36b}$$

$$f_{k,l,\varrho\to\varsigma}(\widetilde{\boldsymbol{\lambda}}) \geq f_{k,l,\varrho\to\varsigma}(\hat{\boldsymbol{\lambda}}). \tag{4.36c}$$

Therefore, as is shown in the **Algorithm 1**, we firstly generate the random complex values $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\lambda}}$. If there exists a $\mathbf{w}$ which satisfies (4.37a), (4.37b), (4.37c), it can be considered as a feasible and initial solution of the phase shift and beamforming algorithms.

In **Algorithm 2**, it can be found that the proposed alternating optimization algorithm consists of the phase shift and beamforming sub-algorithms. In the phase shift optimization with SROCR, we first need to define the initial ω from zero, and a positive step size $\Delta \in (0, 1)$. EE_1 and EE_2 refer to the energy efficiency in phase shift, beamforming sub-problems,

Algorithm 1 The Initial Point Algorithm

1: **Initialize:** Generate arbitrary phase shift angles in the range of $[0, 2\pi]$ and β_l for STAR-RIS elements.
2: **repeat:**
3: Generate arbitrary complex-valued $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\lambda}}$, which obey normal distribution.
4: Find a $\mathbf{w}$, which satisfies

$$2\Re\{\hat{\mu}_{k,l}^{\dagger}\mathbf{h}_{k,l,\varrho}^{H}\mathbf{w}_{k,l,\varrho}\} - \hat{\mu}_{k,l}^{\dagger}(\sigma^2 + I_{k,l,\varrho}(\mathbf{w}))\hat{\mu}_{k,l} \geq \gamma_{k,l,\varrho,min}, \tag{4.37a}$$

$$2\Re\{\hat{\tau}_{k,l}^{\dagger}\mathbf{h}_{k,l,\varsigma}^{H}\mathbf{w}_{k,l,\varsigma}\} - \hat{\tau}_{k,l}^{\dagger}(\sigma^2 + |\mathbf{h}_{k,l,\varsigma}^{H}\mathbf{w}_{k,l,\varrho}|^2 + I_{k,l,\varsigma}(\mathbf{w}))\hat{\tau}_{k,l} \geq \gamma_{k,l,\varsigma,min}, \tag{4.37b}$$

$$2\Re\{\hat{\lambda}_{k,l}^{\dagger}\mathbf{h}_{k,l,\varrho}^{H}\mathbf{w}_{k,l,\varsigma}\} - \hat{\lambda}_{k,l}^{\dagger}(\sigma^2 + |\mathbf{h}_{k,l,\varrho}^{H}\mathbf{w}_{k,l,\varrho}|^2 + I_{k,l,\varrho}(\mathbf{w}))\hat{\lambda}_{k,l} \geq \gamma_{k,l,\varsigma,min}. \tag{4.37c}$$

5: **until:** (4.37) is solvable.
6: **output:** The initial feasible solution $\boldsymbol{\theta}^{(0)}$ and $\mathbf{w}^{(0)}$.

Algorithm 2 Alternating Optimization Algorithm

1: **Initialize:** Given the initial $\mathbf{w}^{(0)}$ and $\mathbf{V}^{(0)} = \mathbf{v}^{(0)}\mathbf{v}^{(0),H}$, define $\omega^{(0)} = 0$, $\Delta \in (0, 1)$. Define $EE^{(0)} = 0$.
2: **repeat**
3: Phase Shift Optimization with SROCR: Given the fixed beamforming vectors $\mathbf{w}^{(l)}$ and $EE_1^{(0)}$.
4: **repeat**
5: update $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$, $\boldsymbol{\delta}$ and solve the Optimization Problem (4.23).
6: **if** solvable:
Obtain the Optimal $\mathbf{V}^{(n+1)}$ and let

$$\Delta^{(n+1)} = \Delta^{(n)}. \tag{4.38}$$

else
Let $\mathbf{V}^{(n+1)} = \mathbf{V}^{(n)}$ and

$$\Delta^{(n+1)} = \Delta^{(n)}/2. \tag{4.39}$$

7: **end**
8: Update $\omega^{(n+1)} = \min(1, \frac{e_{max}(\mathbf{V}^{(n+1)})}{\mathrm{Tr}(\mathbf{V}^{(n+1)})}) + \Delta^{(n+1)})$.
9: **until** EE_1, ω convergence.
10: Beamforming Optimization with SDR: Given the optimal phase shifts $\mathbf{v}^*$ by EVD of $\mathbf{V}^*$ and $EE_2^{(0)}$.
11: **repeat**
12: update y, $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\lambda}$ by solving the problem (4.33).
13: **until:** EE_2 convergence.
14: **until:** EE convergence.
15: **output:** The Optimal $\mathbf{V}^*$ and $\mathbf{W}^*$ and $\mathbf{v}^*$ and $\mathbf{w}^*$ by EVD.

respectively. EE refers to the energy efficiency in the outer loop. In the first sub-algorithm, we iteratively solve the Problem (4.23) and update the parameters as shown in **Algorithm 2**. Then, the optimal solution $\mathbf{V}^*$ can be obtained when the EE_1 and ω converge. In the second sub-algorithm, the values of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\lambda}$ are updated by each iteration and the problem can be iteratively solved. When the objective function reaches to convergence, we can obtain the optimal EE_2 and the corresponding optimal beamforming vectors. After alternatively solving these two sub-algorithms, we can obtain the suboptimal solution of beamforming vector and phase shifts if EE converges.

4.3.4 Complexity Analysis

Since **Algorithm 1** is to find a feasible initial point of the beamforming vectors. To discuss its computational complexity, the matrix calculation of (4.37a), (4.37b) and (4.37c) should be analyzed. According to [85], the computational complexity of the matrix multiplication AB is $\mathcal{O}(LRP)$, where A and B are the $L \times R$ and $R \times P$ matrices, respectively. Therefore, the complexity of executing (4.37a), (4.37b) or (4.37c) is $\mathcal{O}(KM)$. Assuming the **Algorithm 1** spends S_1 iterations finding a feasible point **w**, the total complexity of **Algorithm 1** is $\mathcal{O}(S_1KM)$.

To discuss the complexity of **Algorithm 2**, we can first analyze two sub-algorithms, respectively and then combine them together. To simplify the analysis, we assume that the phase shift optimization problem is solvable in each iteration. Therefore, it can take a complexity of $\mathcal{O}(KN^3)$ when solving (4.23) with the same method applied in **Algorithm 1**. Assuming two thresholds in this sub-algorithm are assumed to be ϵ_1 and ϵ_2 for EE and ω, respectively, the total complexity of phase shift sub-algorithm is $\mathcal{O}(\log \frac{1}{\epsilon_1} \log \frac{1}{\epsilon_2} KN^3)$ in the worst case. Similarly, it can take a complexity of $\mathcal{O}(\log \frac{1}{\epsilon_3} KM^3)$ for beamforming sub-algorithm in the worst case by assuming the convergence of EE in the beamforming optimization sub-algorithm is ϵ_3. After adding these two complexity together, we can get the total complexity of **Algorithm 2** is $\mathcal{O}(\log \frac{1}{\epsilon_4} \log \frac{1}{\epsilon_1} \log \frac{1}{\epsilon_2} KN^3 + \log \frac{1}{\epsilon_4} \log \frac{1}{\epsilon_3} KM^3)$ by assuming the convergence threshold of the outer iterations is ϵ_4.

4.4 Simulation Results

In this section, we present the simulation results to demonstrate the superior performance of the proposed algorithms. Firstly, we set the coordinates of the transmitter in the BS and centre of the STAR-RIS are (0, 0, 20) and (0, 50, 10), respectively. Besides, all the users are deployed in the *T* region and *R* region. The radius of these two half-circle regions are both 20 m and the radius of each cluster is 5 m. Without loss of generality, we consider that the total bandwidth of the network is 100 MHz [44] and the noise spectrum density is -170 dBm/Hz [86]. Besides, the Rician fading is applied for all the channels involved in this system. For

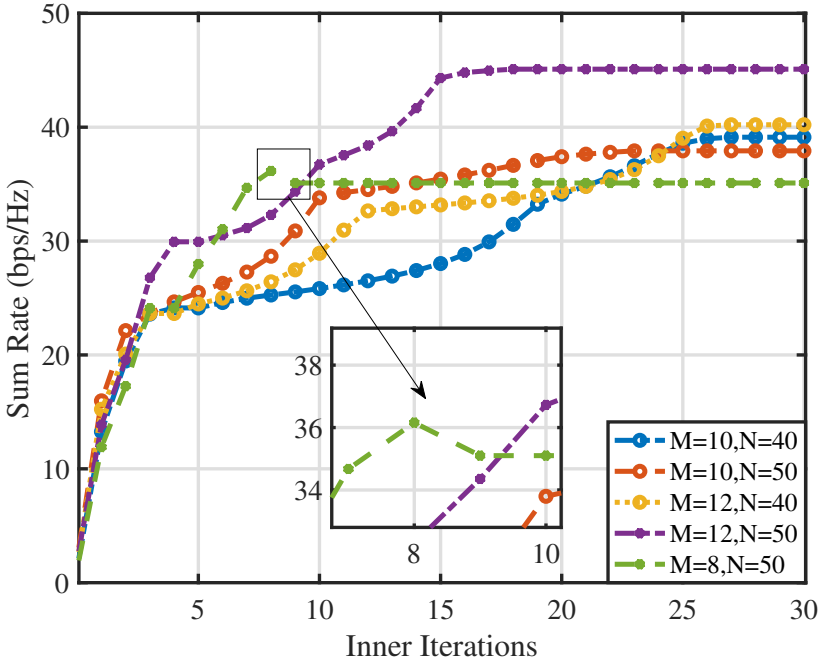

Fig. 4.2. Convergence of the phase shift algorithm, $K = 2$, $P_{max} = 30$ dBm

example, the Rician fading channel channel $\mathbf{F}$ can be expressed as

$$\mathbf{F} = \sqrt{\frac{\kappa}{1+\kappa}}\mathbf{F}_{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}}\mathbf{F}_{\text{nLoS}}, \tag{4.40}$$

where κ denotes the Rician factor and it is set to 2. The $\mathbf{F}_{LoS}$ and $\mathbf{F}_{nLoS}$ denote the line of sight (LoS) and non line of sight (nLoS) components, respectively. The path loss exponent for all the links is set to 2.5 [41]. The minimum SINR for each user is 10 dBm. The internal circuit power P_c and power amplifier coefficient η are set to 20 dBm and 0.8, respectively.

4.4.1 Convergence of the Algorithms

In Fig. 4.2, we demonstrate the convergence the proposed **Algorithm 2** by the numerical analysis. From the figure, we can find that the phase shift sub-algorithm can converge within 30 iterations for all the five cases. Except the convergence of the algorithm, it can also be observed that the **Algorithm 2** sometimes does not converge to the maximum value during the iterations. For example, in the case of $M = 8, N = 50$, the optimal value reaches 36.1547 bps/Hz at the eighth iteration. However, it drops to 35.1003 bps/Hz at the next iteration. It is because that the rank of the optimal $\mathbf{V}$ at the eighth iteration is not one and it does not satisfy one of the convergence conditions of the this sub-algorithm.

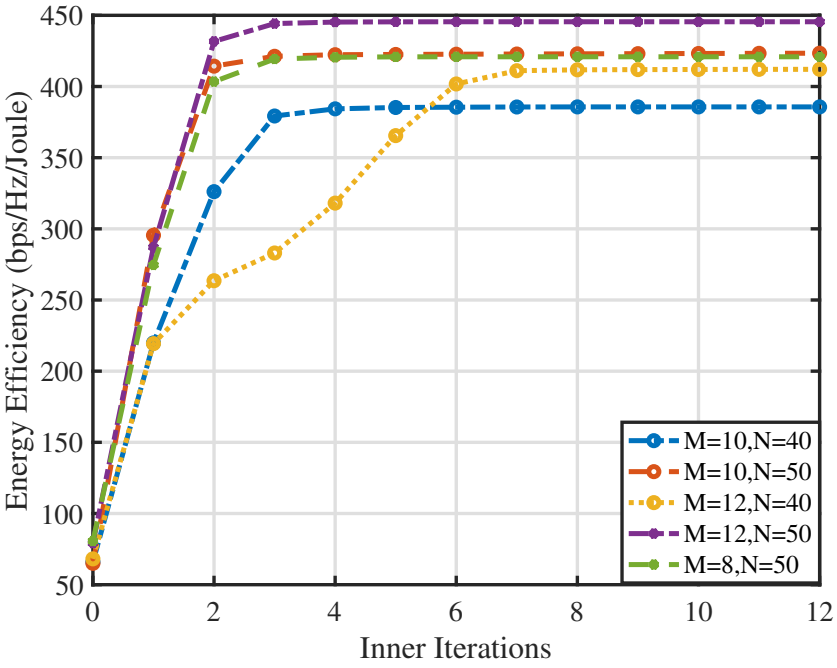

Fig. 4.3. Energy efficiency convergence of beamforming algorithm, $K = 2$, $P_{max} = 30$ dBm

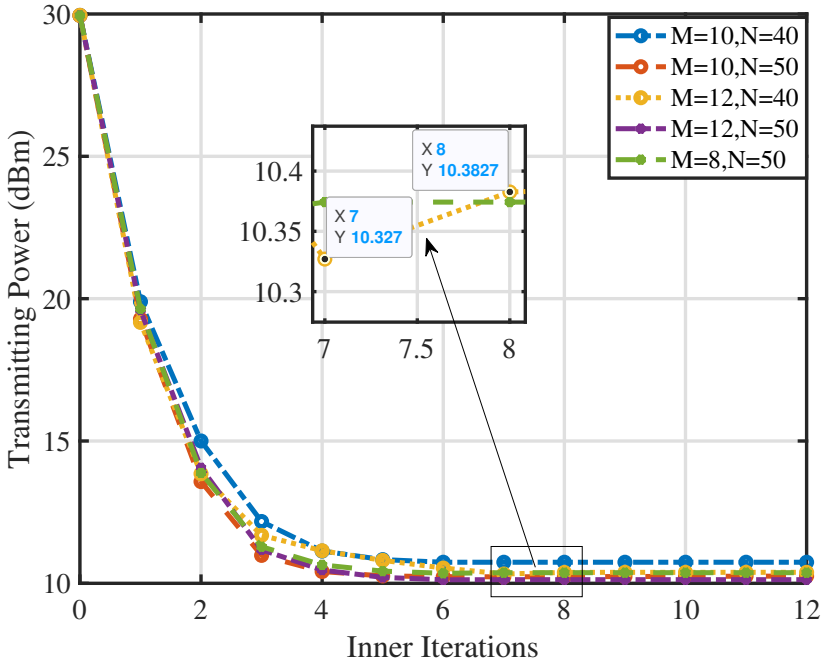

Fig. 4.4. Power consumption convergence of beamforming algorithm, $K = 2$, $P_{max} = 30$ dBm

Fig. 4.3 and Fig. 4.4 demonstrate the convergence of the **Algorithm 2** both in energy efficiency and power consumption. From the figure, it can be found that the beamforming sub-algorithm is able to converge within 12 iterations. Besides, interestingly, during each iteration, the ranks of all the beamforming variables can be verified to be always one after the optimization, which demonstrates the correctness of the *Lemma 1* from the perspective of experiments. In Fig. 4.4, the energy efficiency is always increasing as the number of iterations grows while the transmitting power consumption initially underpins

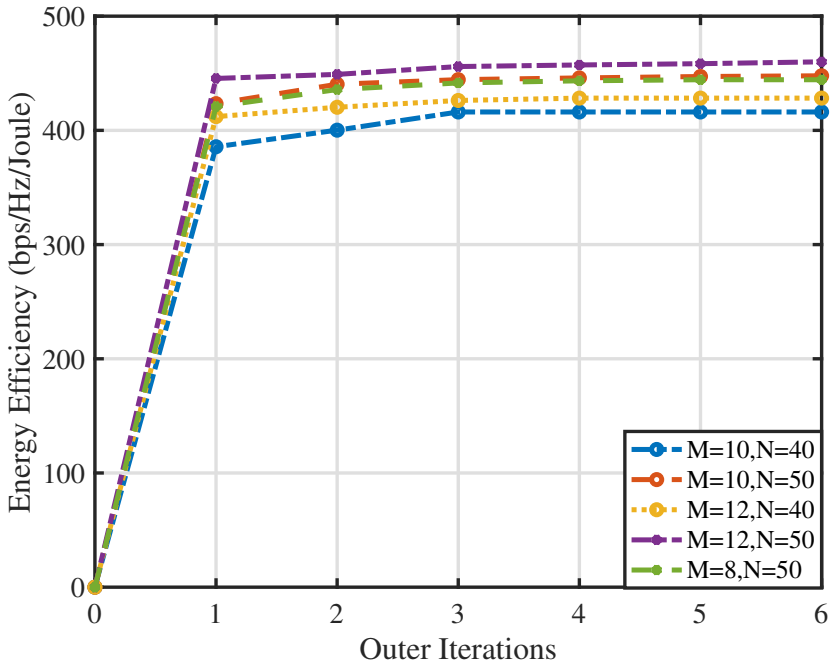

Fig. 4.5. Convergence of Algorithm 2, $K = 2$, $P_{max} = 30$ dBm

a sharp decrease. However, in some cases such as $M = 12, N = 40$, after reaching the minimum, which is 10.327 dBm at the seventh iteration, it slowly rises to 10.3827 dBm and eventually converges at the 10.4253 dBm. Therefore, it cannot be guaranteed that the proposed beamforming algorithm can achieve the energy efficiency maximization and power consumption minimization simultaneously in this network as when the energy efficiency reaches to the highest point and converges, the value of the total transmitting power might be higher than the minimum value during the iterations.

Fig. 4.5 demonstrates the convergence of the proposed alternating algorithm. It can be found that there is a sharp increase when the outer iteration is from zero to one. After several iterations and slow increasing of the EE values, the energy efficiency of these five cases finally converges within 6 iterations.

4.4.2 Impact of the Number of Elements

Fig. 4.6 shows the relationship between the energy efficiency and the number of elements. Firstly, four different schemes are proposed as benchmarks, namely SDR by Gaussian randomization [78], RIS NOMA, STAR-RIS random phase shift, STAR-RIS OMA. For RIS systems, one transmitting-only RIS and reflecting-only RIS are deployed adjacent to the original position of STAR-RIS to realize the full-space communications. It is noted

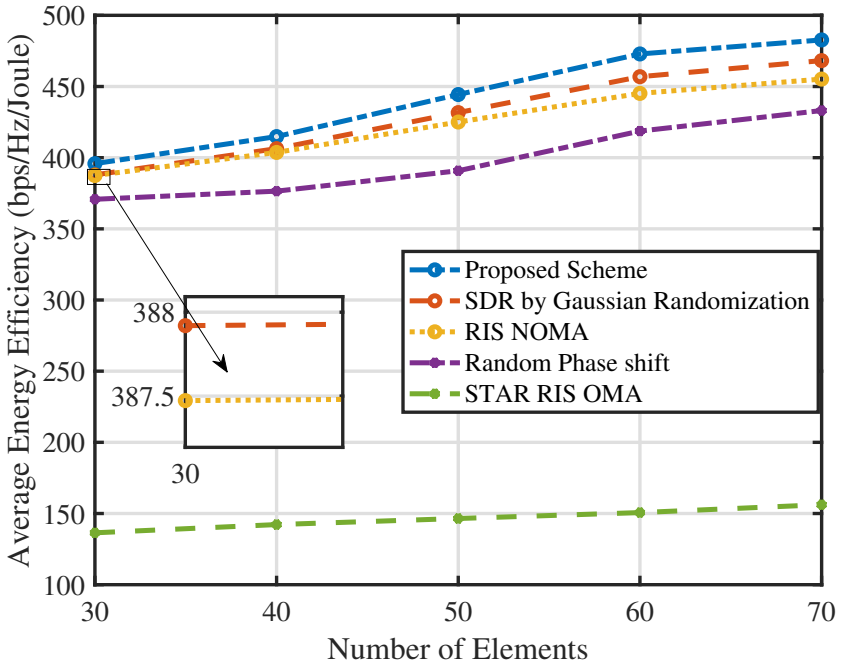

Fig. 4.6. Energy efficiency versus number of elements, $M = 12$, $K = 2$, P_{max}= 30 dBm

that since two RISs are independent, β_t and β_r are not coupled with each other. Therefore, β_t and β_r can be randomly generated with guaranteeing $\beta_t + \beta_r = 1$. Meanwhile, to guarantee the number of elements for RIS systems is same as that of STAR-RIS, each RIS is equipped $\frac{N}{2}$ elements [86]. In the figure, it can be observed the STAR-RIS NOMA scheme by either the SROCR or Gaussian randomization can outperform other schemes including RIS NOMA, random phase shift and STAR-RIS OMA in energy efficiency with the same number of elements. Especially, the STAR-RIS NOMA can outperform the RIS NOMA system. It is because the STAR-RIS NOMA system can have a higher degree-of-freedom (DoF). This DoF reflects on the amplitude coefficients, which means that STAR-RIS can optimize β_l and θ_l jointly to adjust the transmitting and reflected signals and mitigate inter-cluster and intra-cluster interference. This property allows a better performance of the STAR-RIS NOMA than RIS NOMA system in energy efficiency [86]. Besides, it can be also observed that the energy efficiency can increase as the number of elements increases. It is because a higher number of elements is able to produce higher transmitting and reflecting gains [86]. From the perspective of th mathematical tools, the performance of the proposed scheme can always be better than that of the SDR by Gaussian randomization. Besides, From the perspective of computational complexity, assuming the convergence threshold of the proposed algorithm is equal to that of the Gaussian randomization, different from the SROCR scheme, Gaussian randomization requires additional complexity of $\mathcal{O}(LN^3)$ to

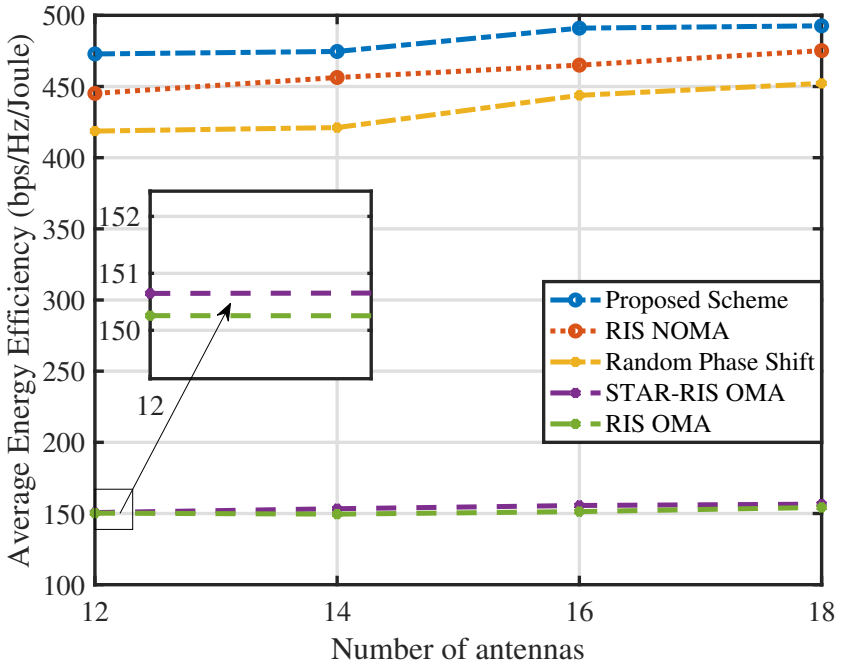

Fig. 4.7. Energy efficiency versus number of antennas, $N = 60$, $K = 2$, P_{max}= 30 dBm

generate L points since Gaussian distribution involves eigenvalue decomposition (EVD) of the SDR optimal solution of $\mathbf{V}$. The complexity of EVD is $\mathcal{O}(b^3)$ for a matrix $\mathbf{B} \in \mathbb{C}^{b\times b}$ [44]. Therefore, it is not worthwhile spending such an extra complexity solving the phase shift optimization problem by using Gaussian randomization in this network.

4.4.3 Impact of the Number of Antennas

Fig. 4.7 shows the relationship between the energy efficiency and the number of BS antennas. In this figure, the proposed scheme is compared to RIS NOMA, STAR-RIS NOMA with the random phase shifts, STAR-RIS OMA and RIS-OMA. From the figure, it can be seen that energy efficiency increases as the number of antennas grows for all the five schemes. Besides, we can also find that the STAR-RIS NOMA system performs better than the RIS NOMA case. Meanwhile, similar to the relationship between STAR-RIS NOMA and RIS NOMA, it can be also found that STAR-RIS OMA outperforms RIS OMA. It is because with the fixed summation of β_t and β_r, the STAR-RIS can have a better performance in EE by flexibly tuning $\boldsymbol{\beta}$ and phase shifts of the elements jointly while only phase shifts can be tuned in the RIS network.

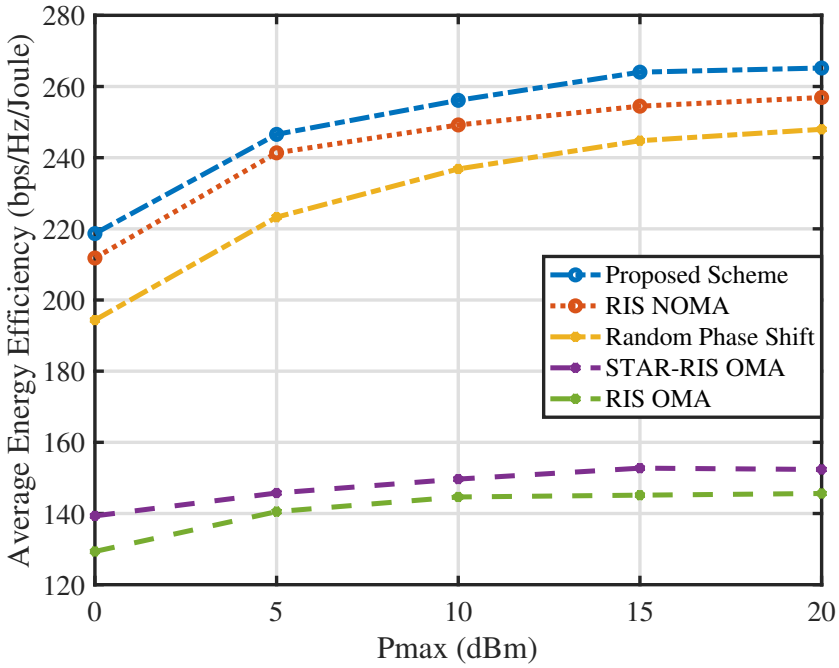

Fig. 4.8. Energy efficiency versus P_{max}, $M = 10$, $N = 40$, $K = 1$

4.4.4 Impact of the Maximum Power Budget

Fig. 4.8 shows the relationship between the energy efficiency and the maximum power budget. From the figure, we can find that the energy efficiency increases as the maximum power budget increases. However, as P_{max} further increases, there is no obvious increase in energy efficiency for all the four cases. Make STAR-RIS NOMA as an example, the energy efficiency rises from 218.6660 bps/Hz/Joule to 264.0331 bps/Hz/Joule when P_{max} grows from 0 dBm to 15 dBm. However, when P_{max} is 20 dBm, the energy efficiency for STAR-RIS NOMA is 265.1941 bps/Hz/Joule, which is almost the same as the value of P_{max} =15 dBm. It is because when the P_{max} is relatively small, the domain of the beamforming is limited. Therefore, the transmitting power of the beamforming is likely to be equal to P_{max} to satisfy the QoS constraints and realize the energy efficiency maximization. As the P_{max} increases, the total channel gains of the users can increase and energy efficiency is likely to improve because the domain of the beamforming expands. When P_{max} is large enough, the transmitting power does not have to be equal to P_{max} to meet QoS constraints and maximize the energy efficiency. Therefore, the energy efficiency can be kept the same in total no matter how the P_{max} increases.

4.5 Conclusions of the Chapter

In this chapter, we have investigated a novel scheme in a multi-cluster MISO STAR-RIS assisted downlink NOMA network, which is to achieve the energy efficiency maximization by alternatively optimizing the phase shifts and beamforming vectors. Firstly, to efficiently solve the non-convex original problem, we first split it into the phase shift and beamforming optimization problems. In these two problems, both SDR and fractional programming were introduced to convert the quadratic non-convex QoS constraints to convex. the difference was in the phase shift optimization, the rank constraints could not be neglected and they were solved by SROCR. However, in the beamforming optimization, the rank of the optimal SDR beamforming solutions could be proven to be equal to one by both theoretical analysis and experimental results. The simulation results have demonstrated the STAR-RIS NOMA network can outperform the RIS NOMA, STAR-RIS OMA, and RIS OMA systems. Besides, the simulation also has demonstrated that the proposed SROCR scheme can have a better performance than the Gaussian randomization method by comprehensively considering the difference between the optimal energy efficiency they can reach and the complexity of the algorithm.

4.6 Proof of Lemma 1

From the problem formulation, it can be seen that (4.33) dropping rank-one is jointly convex with respected to each beamforming matrix $\mathbf{W}_{k,l,i}$ and Slater's condition is satisfied for (4.33) dropping rank-one. Therefore, the strong duality is guaranteed, meaning that the dual gap is zero and the optimal value of the original problem is equal to the optimal value of its dual problem. Based on these, make $\mathbf{W}_{k,l,\varrho}$ as an example, the Lagrangian function with respect to

$\mathbf{W}_{k,l,\varrho}$ can be expressed as followed:

$$\begin{aligned}\mathcal{L} = \Upsilon + y^2 \ln(2)\frac{1}{\eta}\operatorname{Tr}(\mathbf{W}_{k,l,\varrho}) + u\operatorname{Tr}(\mathbf{W}_{k,l,\varrho}) - \operatorname{Tr}(\mathbf{X}_k\mathbf{W}_{k,l,\varrho}) - 2\varpi_{k,l,\varrho}\mu_{k,l}\sqrt{\operatorname{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varrho})} \\ + \sum_{j=1}^{K}\rho_{j,l}\tau_{j,l}^2\operatorname{Tr}(\mathbf{H}_{j,l,\varsigma}\mathbf{W}_{k,l,\varrho}) + \sum_{j=1}^{K}\vartheta_{j,l}\lambda_{j,l}^2\operatorname{Tr}(\mathbf{H}_{j,l,\varrho}\mathbf{W}_{k,l,\varrho}) + \sum_{j=1}^{K}\chi_{j,l}\tau_{j,l'}^2\operatorname{Tr}(\mathbf{H}_{j,l',\varrho}\mathbf{W}_{k,l,\varrho}) \\ + \sum_{j=1}^{K}\upsilon_{j,l}\lambda_{j,l'}^2\operatorname{Tr}(\mathbf{H}_{j,l',\varrho}\mathbf{W}_{k,l,\varrho}).\end{aligned} \tag{4.41}$$

Since we only study the Lagrangian function with respect to $W_{k,l,\varrho}$, the additional terms of the Lagrangian function which have no relation to the $W_{k,l,\varrho}$ can be included as the term Υ. Besides, $l' \in \{t, r\}$ and $l' \neq l$. $u, \varpi, \rho, \vartheta, \chi, \upsilon$ and $\mathbf{X}$ represent the Lagrangian multipliers. For $k = 1, \cdots, K$, the KKT conditions should hold and be formulated such that

$$u^*, \varpi_{k,l,\varrho}^*, \rho_{k,l,\varrho}^*, \vartheta_{k,l,\varrho}^*, \chi_{k,l,\varrho}^*, \upsilon_{k,l,\varrho}^* \geq 0, \tag{4.42}$$

$$\mathbf{X}_{k,l,\varrho}^*\mathbf{W}_{k,l,\varrho}^* = \mathbf{0}, \tag{4.43}$$

$$\mathbf{X}_{k,l,\varrho}^* \succeq \mathbf{0}, \tag{4.44}$$

$$\frac{\partial\mathcal{L}}{\partial\mathbf{W}_{k,l,\varrho}} = \mathbf{0}. \tag{4.45}$$

For the further proof, we reformulate constraint (4.45) as followed:

$$\mathbf{X}_{k,l,\varrho}^* = (u^* + y^2\ln(2)\frac{1}{\eta})\mathbf{E}_{M\times M} - \mathbf{R}_{M\times M}, \tag{4.46}$$

which $\mathbf{E}_{M\times M}$ is a unit matrix and $\mathbf{R}_{M\times M}$ is

$$\begin{aligned}\mathbf{R}_{M\times M} = \varpi_{k,l,\varrho}^*\mu_{k,l}(\operatorname{Tr}(\mathbf{H}_{k,l,\varrho}\mathbf{W}_{k,l,\varrho}^*))^{-\frac{1}{2}}\mathbf{H}_{k,l,\varrho} - \sum_{j=1}^{K}\rho_{j,l}^*\tau_{j,l}^2\mathbf{H}_{j,l,\varsigma} - \sum_{j=1}^{K}\vartheta_{j,l}^*\lambda_{j,l}^2\mathbf{H}_{j,l,\varrho} \\ - \sum_{j=1}^{K}\chi_{j,l}^*\tau_{j,l'}^2\mathbf{H}_{j,l',\varrho} - \sum_{j=1}^{K}\upsilon_{j,l}^*\lambda_{j,l'}^2\mathbf{H}_{j,l',\varrho}.\end{aligned} \tag{4.47}$$

Next, we can analyze the rank of $\mathbf{X}_{k,l,\varrho}^*$ in terms of (4.47). Let $e_{max}(\mathbf{R}_{M\times M})$ be the maximum eigenvalue of $\mathbf{R}_{M\times M}$. According to [84], due to the fact that the channels are randomly generated in this system, the probability for $\mathbf{R}_{M\times M}$ to have multiple maximum

eigenvalues is zero. Besides, the following relationship also exists that

$$u^* + y^2 \ln(2)\frac{1}{\eta} \geq e_{max}(\mathbf{R}_{M\times M}). \tag{4.48}$$

Otherwise, (4.44) cannot be guaranteed.

After the below demonstration, we find that the rank of $\mathbf{X}^*_{k,l,\varrho}$ should be greater or equal than $M-1$. According to the rank characteristics of the matrices multiplication, the rank of $\mathbf{W}^*_{k,l,\varrho}$ should be no more than 1. Since $\mathbf{W}^*_{k,l,\varrho}$ is a non-zero matrix, we show $\text{Rank}(\mathbf{W}^*_{k,l,\varrho}) = 1$. The above process also applies to the proof of $\text{Rank}(\mathbf{W}^*_{k,l,\varsigma}) = 1$.

Chapter 5

Joint Beamforming and Phase Shift Design in RIS Assisted CoMP-NOMA Systems

5.1 Introduction

5.1.1 Motivations of the Chapter

As is emphasised in **Chapter 1**, 6G requires the higher demands on the performance metrics. Firstly, a 6G network should provide ubiquitous and reliable connectivity for all nodes, which is one of the ultimate goal for the design of wireless communications. It is noted that, in wireless communication networks, the quality of service of cell-edge users is more difficult to satisfy compared to cell-central users, for the reason that cell-edge users usually suffer from weaker channel gains and severer inter-cell interference. To address this dilemma, CoMP technique mentioned in **Chapter 1** and **Chapter 2** is proposed to enhance the performance of the cell-edge users, by encouraging multiple base stations to cooperatively serve the cell-edge users [66]. However, conventional CoMP techniques were designed based on OMA, which means that when multiple BSs cooperatively serve a cell-edge user, each BS has to allocate one resource block to the cell-edge user which cannot be accessed by other users. As a consequence, when the number of cell-edge user increases, the spectrum resource which is inherently scarce becomes more stringent. Thus, conventional OMA based CoMP technique proposed is not favorable to support large connectivity for future wireless communication

networks. Fortunately, NOMA, as an innovative MA technique, has the capability to compensate for the shortage of the CoMP OMA networks. By introducing NOMA to CoMP systems, yielding the so called 'CoMP-NOMA' technique, the aforementioned dilemma in OMA based CoMP system can be significantly alleviated [87], [88]. In CoMP-NOMA, multiple BSs cooperatively serve the cell-edge users, while each BS serves additional cell-central users individually by applying NOMA. Thus, compared to conventional OMA based CoMP technique, CoMP-NOMA can serve more users and hence improve spectrum efficiency significantly, while the transmission reliability of the cell-edge users can be guaranteed.

More recently, the investigation of deploying reconfigurable intelligent surface to further improve the performance of wireless communications becomes an important research direction for 6G. In **Chapter 3**, it has been shown that RIS can further significantly improve the performance of many advanced wireless communication techniques. However, the application of RIS to CoMP-NOMA systems is still open, which motivates this chapter. Until now, RIS assisted NOMA has been extensively studied by existing works. In comparison to the RIS assisted NOMA, the number of papers associated with the RIS assisted CoMP-NOMA is much less in quantity. Therefore, in this stage, RIS assisted CoMP-NOMA technique processes a great value, which is worth exploiting to improve the performance of the networks. Meanwhile, it is noted that most existing works related to RIS assisted CoMP-NOMA mainly focused on the power allocation strategies and barely discussed the effect of the beamforming on the system performance. However, similar to the studies of power allocation, beamforming design of the systems is still a crucial topic in not only single NOMA but CoMP-NOMA systems [77]. However, two potential hazards of high-rank relaxed solutions and high-complexity still exist if apply SDR by Gaussian randomization in CoMP-NOMA networks, limiting the efficiency of the algorithms. Fortunately, because of the existence of the norm constraints of the phase shift elements, which are a type of manifold in Riemannian space, manifold optimization can be an efficient method with a lower complexity than SDR by Gaussian randomization [89]. However, most works related to manifold optimization did not introduce NOMA in their networks. Therefore, there is no inequality constraint in their optimization problems since no quality of service requirement is needed. This lack of research severely limits manifold optimization to be applied in the NOMA case. Motivated by the aforementioned issues, we

aim to propose a scheme to optimize the beamforming vectors of the users and phase shift elements, respectively in an RIS assisted CoMP-NOMA network. This network consists of two BSs, two NOMA users, one CoMP user and two RISs. The CoMP user is assumed to have a low rate requirement and used to share and bandwidth source with two NOMA users. The aim of this chapter is to maximize sum rate of two NOMA users since they are the users with high rate requirements.

5.1.2 Contributions of the Chapter

The contributions of this chapter can be summarized below:

- In this chapter, we effectively combine the NOMA with RIS and CoMP techniques and achieve the suboptimal solutions of the sum rate maximization problem by optimizing the beamforming vectors and phase shifts of two RISs. Since the beamforming and phase shifts are coupled with each other in the original problem. It is difficult to directly solve it by the existing optimization tools or develop the closed-formed solutions of it. In this chapter, the original problem is decomposed to the beamforming and phase shift sub-problems and these two problems are alternatively solved.
- In the beamforming optimization, we propose SDR to transform the non-convex problem to a convex one. However, the concern of this transformed problem is that the rank of the beamforming matrices might not be one. Fortunately, the beamforming matrices of NOMA users can be proven to be rank-one in mathematics. For the beamforming matrix of the CoMP user, the novel penalty-based SCA method is applied to tackle the rank-one issue of it.
- In the phase shift optimization, since SDR can have a high complexity, the novel penalty-based manifold optimization is leveraged to optimize the phase shifts of the RIS elements. Different from the previous works related to manifold optimization such as [89] and [90], the novelty of this method is that it can solve the problem which consists both inequality and norm equality constraints.
- With adequate experiments and analysis, the proposed scheme can achieve the better performance in sum rate than the other benchmark schemes such as SDR by the

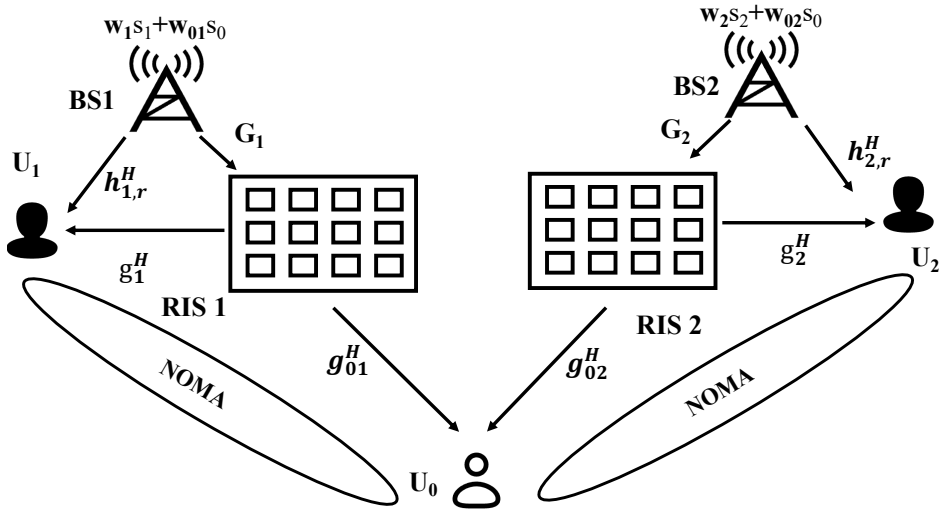

Fig. 5.1. RIS assisted CoMP-NOMA transmission system

Gaussian randomization and RIS assisted CoMP-OMA system. It demonstrates the great value of the combination among NOMA, RIS and CoMP techniques.

- From the perspective of system model, the proposed scheme can also apply to RIS assisted CoMP-NOMA system with the case of multiple NOMA users. From the perspective of mathematical tools, the proposed beamforming optimization method can also be used to solve the EE maximization, the similar mathematical proof can be proceeded to show the rank of beamforming matrices is one. Besides, the proposed phase shift optimization method can be considered as a universal scheme to tackle the problems related to the phase shift optimization in the RIS assisted NOMA systems.

5.2 System Model

Consider an RIS assisted downlink CoMP-NOMA network with two base stations, denoted by BS_1 and BS_2 and two RISs, denoted as RIS_1 and RIS_2. Each BS is equipped with M antennas. In this system, BS_k is quite near to the RIS_k and far away from another BS and RIS. A CoMP user, denoted by U_0, is located at the edge of two cells and is cooperatively served by two BSs and two RISs jointly. Besides, two users, denoted by U_1 and U_2, are located near to the BS_1 and BS_2, respectively. Each user is equipped with one single antenna. Different from the CoMP user, each NOMA user U_k, $k = 1, 2$ is only served by the BS_k and RIS_k. In this network, U_1 and U_2 are the device with high rate requirement while U_0 is the device with low rate requirement. To make U_1 and U_2 utilize the total bandwidth

resource sufficiently, NOMA is employed to achieve the downlinks between the BSs and $U_k, k = 0, 1, 2$. The detailed implementation of NOMA will be discussed later. The BS_1 transmits the signals of U_0 and U_1 with the beamforming of $\mathbf{w}_{01}$ and $\mathbf{w}_1$ and BS_2 transmits the signals of U_0 and U_2 with the beamforming of $\mathbf{w}_{02}$ and $\mathbf{w}_2$, respectively. Therefore, the transmitting signals of two BSs, denoted by $\mathbf{x}_1$ and $\mathbf{x}_2$ can be expressed as followed:

$$\mathbf{x}_1 = \mathbf{w}_1 s_1 + \mathbf{w}_{01} s_0, \tag{5.1}$$

$$\mathbf{x}_2 = \mathbf{w}_2 s_2 + \mathbf{w}_{02} s_0. \tag{5.2}$$

As is shown in the Fig. 5.1, each BS transmits the signals with the aid of two RISs. Each RIS is equipped with N reflecting elements. $\mathbf{G}_k \in \mathbb{C}^{N \times M}, k = 1, 2$, represents the channel gain between the BS_k and the RIS_k and $\mathbf{g}_{0k}^H \in \mathbb{C}^{1 \times N}, k = 1, 2$ is the channel gain between the RIS_k and U_0. $\mathbf{g}_k^H \in \mathbb{C}^{1 \times N}$ is the channel gain between the RIS_k and U_k, $k = 1, 2$. $\mathbf{h}_{k,r}^H \in \mathbb{C}^{1 \times M}, k = 1, 2$ refers to the direct link between the BS and U_k. The phase shift angle vector of the RIS_k is $\boldsymbol{\theta}_k = (e^{j\theta_{k,1}}, \cdots, e^{j\theta_{k,N}}), k = 1, 2$. $\theta_{k,n}$, $n = 1, \cdots N$ represents the phase shift angle of the n^{th} element of RIS_k. In this chapter, since two cooperative BSs are far away from each other, the interference from one cell to the NOMA user in another cell can be neglected. Therefore, the total channel gain from BS_k to U_0 can be expressed as

$$\mathbf{h}_{0k}^H = \mathbf{g}_{0k}^H \text{diag}(\boldsymbol{\theta}_k) \mathbf{G}_k, \tag{5.3}$$

and the total channel gain of $U_k, k = 1, 2$ can be expressed as

$$\mathbf{h}_k^H = \mathbf{g}_k^H \text{diag}(\boldsymbol{\theta}_k) \mathbf{G}_k + \mathbf{h}_{k,r}^H. \tag{5.4}$$

Consequently, the received signal of each user can be expressed as followed:

$$y_k = \mathbf{h}_k^H (\mathbf{w}_k s_k + \mathbf{w}_{0k} s_0), k = 1, 2, \tag{5.5}$$

$$y_0 = \mathbf{h}_{01}^H (\mathbf{w}_1 s_1 + \mathbf{w}_{01} s_0) + \mathbf{h}_{02}^H (\mathbf{w}_2 s_2 + \mathbf{w}_{02} s_0), \tag{5.6}$$

where $\mathbf{w}_k, k = 1, 2$ represents the beamforming vector of NOMA user U_k and $\mathbf{w}_{0k}$ represents

the beamforming vector of CoMP user U_0 from the BS$_k$, $k = 1, 2$. s_k, $k = 0, 1, 2$, represents the signal of U_k. To the principle of NOMA, the NOMA users carry out SIC to decode signals, the NOMA user has a higher decoding ability than the CoMP user so U_1 or U_2 can first decode the U_0's signal. If successfully, successive interference cancellation is employed to remove the signal of U_0, and then U_k is able to decode the signal of itself. Differently, the CoMP user will decode its signal by treating the signals from the NOMA users as interference due to its bad decoding capability. Based on the above decoding scheme, the SINR of each user can be expressed as:

$$\gamma_{0\to 0} = \frac{\|\mathbf{h}_{01}^H \mathbf{w}_{01} + \mathbf{h}_{02}^H \mathbf{w}_{02}\|^2}{\|\mathbf{h}_{01}^H \mathbf{w}_1\|^2 + \|\mathbf{h}_{02}^H \mathbf{w}_2\|^2 + \sigma^2}, \tag{5.7}$$

$$\gamma_{k\to 0} = \frac{\|\mathbf{h}_k^H \mathbf{w}_{0k}\|^2}{\|\mathbf{h}_k^H \mathbf{w}_k\|^2 + \sigma^2}, k = 1, 2, \tag{5.8}$$

$$\gamma_k = \frac{\|\mathbf{h}_k^H \mathbf{w}_k\|^2}{\sigma^2}, k = 1, 2, \tag{5.9}$$

where $\gamma_{0\to 0}$ is the SINR when decoding U_0's signal at itself. $\gamma_{k\to 0}$ is the SINR when decoding U_0's signal at the NOMA user U_k. γ_k is the SINR when decoding the NOMA user U_k's signal. As is mentioned above, the existence of the CoMP user is to share the bandwidth with the NOMA users, who have the high data requirement in order to improve the spectrum efficiency in the network. However, there is no need for the CoMP user to achieve such a high rate. Therefore, it is an inefficient strategy if trying to maximize the sum rate of all the users. Based on this viewpoint, the aim of this chapter is to maximize the sum rate of U_1 and U_2 with guaranteeing that each user's signal can be successfully decoded, which can be more efficient to meet the requirements of diverse types of users in the network. Therefore, the original optimization problem can be formulated as

$$\max_{\mathbf{w},\boldsymbol{\theta}} \quad \log_2(1+\gamma_1) + \log_2(1+\gamma_2), \tag{5.10a}$$

$$\text{s. t.} \quad \gamma_{0\to 0} \geq \gamma_{0,min}, \tag{5.10b}$$

$$\gamma_{1\to 0} \geq \gamma_{0,min}, \tag{5.10c}$$

$$\gamma_{2\to 0} \geq \gamma_{0,min}, \tag{5.10d}$$

$$\gamma_k \geq \gamma_{k,min}, k = 1, 2, \tag{5.10e}$$

$$\|\mathbf{w}_{01}\|^2 + \|\mathbf{w}_1\|^2 \leq P_{max}, \tag{5.10f}$$

$$\|\mathbf{w}_{02}\|^2 + \|\mathbf{w}_2\|^2 \leq P_{max}, \tag{5.10g}$$

$$\theta_{k,n} \in [0, 2\pi], k = 1, 2, n = 1, \cdots, N, \tag{5.10h}$$

where P_{max} is the maximum power budget of each BS. (5.10b)-(5.10d) mean that the achievable rate of each user must be greater than a target rate to satisfy the quality of service. (5.10f) and (5.10g) limit the transmitting power of the BSs. (5.10h) specifies that the magnitude of each RIS element should be equal to one.

5.3 Proposed Method

The aim of this section is to design the optimization methods and solve the formulated problem to achieve the sum rate maximization of NOMA users. It is noteworthy that (5.10) is not a convex optimization problem. The reasons are mainly in two folds. On the one hand, the beamforming variables and phase shifts are coupled with each other. On the other hand, the QoS constraints (5.10b)-(5.10d) and norm ball constraints of phase shift elements (5.10h) are also not convex. Thus, solving (5.10) is a challenging task. In this chapter, several optimization methods are applied to solve (5.10). Firstly, the coupling issue is resolved by alternating optimization. Specifically, the original problem is decoupled into two suboptimal problems, i.e., beamforming optimization and phase shift optimization suboptimal problems. In the beamforming suboptimal problem, the beamforming vectors are solved by treating phase shift vectors as constants and vice versa in the phase shift suboptimal problem. Then, these two problems can be solved alternatively. For the beamforming optimization problem, semidefinite programming is applied to transform the non-convex constraints to convex. For the phase shift optimization problem, the penalty-based manifold optimization is utilized to find the locally optimal solutions of the phase shifts.

5.3.1 Beamforming Optimization

In beamforming optimization problem, the phase shift vectors are treated as fixed constants. Therefore, the problem can be formulated as followed:

$$\max_{\mathbf{w}} \quad \log_2(1+\frac{\|\mathbf{h}_1\mathbf{w}_1\|^2}{\sigma^2})+\log_2(1+\frac{\|\mathbf{h}_2\mathbf{w}_2\|^2}{\sigma^2}), \tag{5.11a}$$

$$\text{s. t.} \quad \frac{\|\mathbf{h}_{01}^H\mathbf{w}_{01}+\mathbf{h}_{02}^H\mathbf{w}_{02}\|^2}{\|\mathbf{h}_{01}^H\mathbf{w}_1\|^2+\|\mathbf{h}_{02}^H\mathbf{w}_2\|^2+\sigma^2}\geq\gamma_{0,min}, \tag{5.11b}$$

$$\frac{\|\mathbf{h}_k^H\mathbf{w}_{0k}\|^2}{\|\mathbf{h}_k^H\mathbf{w}_k\|^2+\sigma^2}\geq\gamma_{0,min}, k=1,2, \tag{5.11c}$$

$$\frac{\|\mathbf{h}_k^H\mathbf{w}_k\|^2}{\sigma^2}\geq\gamma_{k,min}, k=1,2, \tag{5.11d}$$

$$\|\mathbf{w}_{01}\|^2+\|\mathbf{w}_1\|^2\leq P_{max}, \tag{5.11e}$$

$$\|\mathbf{w}_{02}\|^2+\|\mathbf{w}_2\|^2\leq P_{max}. \tag{5.11f}$$

It is noted that (5.11) is non-convex due to the existence of the quadratic expressions in (5.11a) and (5.11b)-(5.11d). To tackle it, semidefinite relaxation can be applied to transform the quadratic form to affine. However, in (5.11b), $\mathbf{w}_{01}$ and $\mathbf{w}_{02}$ are coupled with each other and this constraint cannot be transformed directly to the SDR form. To overcome this difficulty, a method, namely zero padding is utilized to rewrite this constraint. Specifically, Define $\mathbf{h}_0^H=[\mathbf{h}_{01}^H,\mathbf{h}_{02}^H]$, $\mathbf{w}_0=[\mathbf{w}_{01}^H,\mathbf{w}_{02}^H]^H$, $\overline{\mathbf{h}}_1^H=[\mathbf{h}_1^H,\mathbf{0}]$ and $\overline{\mathbf{h}}_2^H=[\mathbf{0},\mathbf{h}_2^H]$, where $\mathbf{0}$ is the $1\times M$ zero vector. Then, (5.11) can be transformed into the following form:

$$\max_{\mathbf{W}} \quad \log_2(1+\frac{\text{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\sigma^2})+\log_2(1+\frac{\text{Tr}(\mathbf{H}_2\mathbf{W}_2)}{\sigma^2}), \tag{5.12a}$$

$$\text{s. t.} \quad \frac{\text{Tr}(\mathbf{H}_0\mathbf{W}_0)}{\text{Tr}(\mathbf{H}_{01}\mathbf{W}_1)+\text{Tr}(\mathbf{H}_{02}\mathbf{W}_2)+\sigma^2}\geq\gamma_{0,min}, \tag{5.12b}$$

$$\frac{\text{Tr}(\overline{\mathbf{H}}_k\mathbf{W}_0)}{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)+\sigma^2}\geq\gamma_{0,min}, k=1,2, \tag{5.12c}$$

$$\frac{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)}{\sigma^2}\geq\gamma_{k,min}, k=1,2, \tag{5.12d}$$

$$\mathbf{I}_1\,\text{Tr}(\mathbf{W}_0)+\text{Tr}(\mathbf{W}_1)\leq P_{max}, \tag{5.12e}$$

$$\mathbf{I}_2\,\text{Tr}(\mathbf{W}_0)+\text{Tr}(\mathbf{W}_2)\leq P_{max}, \tag{5.12f}$$

$$\mathbf{W}_k\succeq\mathbf{0}, k=0,1,2, \tag{5.12g}$$

$$\text{Rank}(\mathbf{W}_k)=1, k=0,1,2, \tag{5.12h}$$

where $\mathbf{H}_k = \mathbf{h}_k\mathbf{h}_k^H$ and $\mathbf{W}_k = \mathbf{w}_k\mathbf{w}_k^H$. $\mathbf{I}_1$ and $\mathbf{I}_2$ are $2M \times 2M$ matrices which are given by:

$$\mathbf{I}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{5.13}$$

$$\mathbf{I}_2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{5.14}$$

$\mathbf{I}$ and $\mathbf{0}$ are $M \times M$ unit matrix and zero matrix, respectively. Until now, all the objective function and constraints of (5.12) except (5.12h) have been convex. Interestingly, the rank constraints for $\mathbf{W}_1$ and $\mathbf{W}_2$ can be removed without degrading the optimality of the original problem as highlighted in the following proposition:

Proposition 1 *The obtained optimal solutions* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *of* (5.12) *without* (5.12h) *are strictly rank-one. Proof: Please see Appendix 5.6.*

Based on the above proposition, no more steps are needed to tackle the obtained $\mathbf{W}_1$ and $\mathbf{W}_2$ and (5.12) can be further transformed as followed:

$$\max_{\mathbf{W}} \quad \log_2(1 + \frac{\text{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\sigma^2}) + \log_2(1 + \frac{\text{Tr}(\mathbf{H}_2\mathbf{W}_2)}{\sigma^2}), \tag{5.15a}$$

$$\text{s. t.} \quad (5.12\text{b}) - (5.12\text{g}), \tag{5.15b}$$

$$\text{Rank}(\mathbf{W}_0) = 1. \tag{5.15c}$$

For $\mathbf{W}_0$, whose rank might not be one after solving (5.15), Gaussian randomization is usually applied to get an approximated solution. However, to guarantee the quality of this solution, a large number of vectors is required to be sampled, which produces an extra complexity of the algorithm. Even worse, if the solution of $\mathbf{W}_0$ after solving (5.15) dropping (5.15c) is high-rank, it will be extremely difficult to generate the sampled vectors, which can satisfy the constraints of the problem. In this chapter, another novel method proposed in [84] is used to effectively avoid these two concerns. Specifically, in mathematics, the following equation holds for any hermitian semidefinite matrix $\mathbf{\Lambda}$, which is

$$\text{Tr}(\mathbf{\Lambda}) \geq \|\mathbf{\Lambda}\|_F, \tag{5.16}$$

where $\|\mathbf{\Lambda}\|_F$ is the Frobenius norm of $\mathbf{\Lambda}$. The equality holds in (5.16) if and only if $\mathbf{\Lambda}$ is rank-one. Inspired by the above relationship, the penalty-based method is leveraged to effectively tackle the rank constraint of $\mathbf{W}_0$. The idea of the penalty method is to transform the problems with constraints to the unconstrained ones by moving them to the objective functions. Due to the fact that the existing optimization toolboxes can accept the convex constraints, it is only needed to move the non-convex constraint to the objective function in this problem and the objective function is recast as

$$\min_{\mathbf{W}} \quad -\log_2(1+\frac{\text{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\sigma^2})-\log_2(1+\frac{\text{Tr}(\mathbf{H}_2\mathbf{W}_2)}{\sigma^2})+\rho_1(\text{Tr}(\mathbf{W}_0)-\|\mathbf{W}_0\|_F), \tag{5.17}$$

where ρ_1 is the penalty factor for the beamforming optimization problem.

Since (5.17) is not convex, the SCA is employed to transform the non-convex term $(\text{Tr}(\mathbf{W}_0)-\|\mathbf{W}_0\|_F)$ to affine. Let $D(\mathbf{W}_0)=\|\mathbf{W}_0\|_F$, the objective function can be recast as:

$$\begin{aligned}\min_{\mathbf{W}} \quad & -\log_2(1+\frac{\text{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\sigma^2})-\log_2(1+\frac{\text{Tr}(\mathbf{H}_2\mathbf{W}_2)}{\sigma^2}) \\ & +\rho_1^{(i)}(\text{Tr}(\mathbf{W}_0)-D(\mathbf{W}_0^{(i-1)})-\text{Tr}(\nabla D(\mathbf{W}_0^{(i-1)})(\mathbf{W}_0-\mathbf{W}_0^{(i-1)}))),\end{aligned} \tag{5.18}$$

where i is the index of iterations and $\nabla D(\mathbf{W}_0)$ is the first order derivative with respect to $\mathbf{W}_0$, which can be expressed as:

$$\nabla D(\mathbf{W}_0)=\frac{\mathbf{W}_0}{\sqrt{\text{Tr}(\mathbf{W}_0^H\mathbf{W}_0)}}. \tag{5.19}$$

As a result, the beamforming optimization problem becomes convex, and can be expressed as

$$\begin{aligned}\max_{\mathbf{W}} \quad & -\log_2(1+\frac{\text{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\sigma^2})-\log_2(1+\frac{\text{Tr}(\mathbf{H}_2\mathbf{W}_2)}{\sigma^2}) \\ & +\rho_1^{(i)}(\text{Tr}(\mathbf{W}_0)-D(\mathbf{W}_0^{(i-1)})-\text{Tr}(\nabla D(\mathbf{W}_0^{(i-1)})(\mathbf{W}_0-\mathbf{W}_0^{(i-1)}))), && \text{(5.20a)} \\ \text{s. t.} \quad & \text{Tr}(\mathbf{H}_0\mathbf{W}_0)\geq\gamma_{0,min}(\text{Tr}(\mathbf{H}_{01}\mathbf{W}_1)+\text{Tr}(\mathbf{H}_{02}\mathbf{W}_2)+\sigma^2), && \text{(5.20b)} \\ & \text{Tr}(\overline{\mathbf{H}}_k\mathbf{W}_0)\geq\gamma_{k,min}(\text{Tr}(\mathbf{H}_k\mathbf{W}_k)+\sigma^2), k=1,2, && \text{(5.20c)} \\ & \frac{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)}{\sigma^2}\geq\gamma_{k,min}, k=1,2, && \text{(5.20d)} \\ & \text{(5.12e), (5.12f), (5.12g).} && \text{(5.20e)}\end{aligned}$$

This problem can be solved iteratively. The optimal solution can be obtained when the algorithm converges. Then, beamforming vector of each user can be obtained by the eigenvalue decomposition.

5.3.2 Phase Shift Optimization

In the phase shift optimization suboptimal problem, the beamforming vectors are considered as the fixed constants. Until now, the phase shift optimization problem can be expressed as followed:

$$\begin{aligned}
\max_{\mathbf{v}} \quad & \log_2(1+\frac{\|\mathbf{v}_1^H\mathbf{a}_1+\mathbf{h}_{1,r}^H\mathbf{w}_1\|^2}{\sigma^2}) \\
& +\log_2(1+\frac{\|\mathbf{v}_2^H\mathbf{a}_2+\mathbf{h}_{2,r}^H\mathbf{w}_2\|^2}{\sigma^2}), && \text{(5.21a)}\\
\text{s. t.} \quad & \frac{\|\mathbf{v}_1^H\mathbf{a}_{01}+\mathbf{v}_2^H\mathbf{a}_{02}\|^2}{\|\mathbf{v}_1^H\mathbf{a}_{01}\|^2+\|\mathbf{v}_2^H\mathbf{a}_{02}\|^2+\sigma^2} \geq \gamma_{0,min}, && \text{(5.21b)}\\
& \frac{\|\mathbf{v}_1^H\mathbf{a}_{01}+\mathbf{h}_{1,r}^H\mathbf{w}_0\|^2}{\|\mathbf{v}_1^H\mathbf{a}_1+\mathbf{h}_{1,r}^H\mathbf{w}_1\|^2+\sigma^2} \geq \gamma_{0,min}, && \text{(5.21c)}\\
& \frac{\|\mathbf{v}_2^H\mathbf{a}_{02}+\mathbf{h}_{2,r}^H\mathbf{w}_0\|^2}{\|\mathbf{v}_2^H\mathbf{a}_2+\mathbf{h}_{2,r}^H\mathbf{w}_2\|^2+\sigma^2} \geq \gamma_{0,min}, && \text{(5.21d)}\\
& \frac{\|\mathbf{v}_1^H\mathbf{a}_1+\mathbf{h}_{1,r}^H\mathbf{w}_1\|^2}{\sigma^2} \geq \gamma_{1,min}, && \text{(5.21e)}\\
& \frac{\|\mathbf{v}_2^H\mathbf{a}_2+\mathbf{h}_{2,r}^H\mathbf{w}_2\|^2}{\sigma^2} \geq \gamma_{2,min}, && \text{(5.21f)}\\
& \|\mathbf{v}_k(n)\|=1, k=1,2, n=1,\cdots,N, && \text{(5.21g)}
\end{aligned}$$

where $\mathbf{v}_k=[e^{j\theta_{k,1}},\cdots,e^{j\theta_{k,N}}]^H$, $\mathbf{a}_{0k}=\text{diag}(\mathbf{g}_{0k}^H)\mathbf{G}_1\mathbf{w}_{0k}$, $\mathbf{a}_{k0}=\text{diag}(\mathbf{g}_k^H)\mathbf{G}_k\mathbf{w}_{0k}$ and $\mathbf{a}_k=\text{diag}(\mathbf{g}_k^H)\mathbf{G}_k\mathbf{w}_k$, $k=1,2$. Intuitively, the penalty-based SDR method in the beamforming optimization problem might be also applicable in the phase shift optimization problem. However, SDR can usually produce a high complexity of the algorithm. Fortunately, due to the existence of (5.21g), which is a manifold in Riemannian space, the novel penalty-based manifold optimization in this chapter is proposed to solve (5.21). Firstly, it is noted that there exist several inequality constraints. Therefore, manifold optimization should be combined with the penalty method. It is noted that in the beamforming suboptimal problem, the optimization toolboxes are used to find the optimal solutions. Unfortunately, the optimization toolboxes can only accept the problems which are convex. Therefore, the penalty term is required to be transformed to affine by SCA first and then added to the objective function.

Differently, in the phase shift suboptimal problem, the locally optimal solutions are obtained by manifold optimization without the use of optimization toolboxes. Therefore, all the constraints except (5.21g) are added to the objective function as the penalty terms directly without any transformation from non-convex to convex. Besides, since derivation is required in both manifold optimization and penalty method, for easily finding the derivative of the phase shift variables, $\mathbf{v}_1$ and $\mathbf{v}_2$ are combined to one variable, defined as $\mathbf{v} = \begin{bmatrix}\mathbf{v}_1, \mathbf{v}_2\end{bmatrix}^T$. Then, define $\bar{\mathbf{a}}_1 = \begin{bmatrix}\mathbf{a}_1, \mathbf{0}\end{bmatrix}^T$, $\bar{\mathbf{a}}_2 = \begin{bmatrix}\mathbf{0}, \mathbf{a}_2\end{bmatrix}^T$, $\bar{\mathbf{a}}_{01} = \begin{bmatrix}\mathbf{a}_{01}, \mathbf{0}\end{bmatrix}$ $\bar{\mathbf{a}}_{02} = \begin{bmatrix}\mathbf{0}, \mathbf{a}_{02}\end{bmatrix}^T$, where $\mathbf{0}$ here is an $N \times 1$ zero vector. Therefore, (5.21) can be transformed to the following form:

$$\min_{\mathbf{v}} \quad -\log_2(1+\frac{\|\mathbf{v}^H\bar{\mathbf{a}}_1+\mathbf{h}_{1,r}^H\mathbf{w}_1\|^2}{\sigma^2})-\log_2(1+\frac{\|\mathbf{v}^H\bar{\mathbf{a}}_2+\mathbf{h}_{2,r}^H\mathbf{w}_2\|^2}{\sigma^2}), \tag{5.22a}$$

$$\text{s.t.} \quad f_k(\mathbf{v}) \le 0, k=1,\cdots,5, \tag{5.22b}$$

$$\|\mathbf{v}(n)\| = 1, n=1,\cdots,2N, \tag{5.22c}$$

where

$$f_1(\mathbf{v}) = \gamma_{0,min}(\|\mathbf{v}^H\bar{\mathbf{a}}_{01}\|^2+\|\mathbf{v}^H\bar{\mathbf{a}}_{02}\|^2+\sigma^2)-\|\mathbf{v}^H\bar{\mathbf{a}}_{01}+\mathbf{v}^H\bar{\mathbf{a}}_{02}\|^2, \tag{5.23}$$

$$f_2(\mathbf{v}) = \gamma_{0,min}(\|\mathbf{v}^H\bar{\mathbf{a}}_1+\mathbf{h}_{1,r}^H\mathbf{w}_1\|^2+\sigma^2)-\|\mathbf{v}^H\bar{\mathbf{a}}_{01}+\mathbf{h}_{1,r}^H\mathbf{w}_{01}\|^2, \tag{5.24}$$

$$f_3(\mathbf{v}) = \gamma_{0,min}(\|\mathbf{v}^H\bar{\mathbf{a}}_2+\mathbf{h}_{2,r}^H\mathbf{w}_2\|^2+\sigma^2)-\|\mathbf{v}^H\bar{\mathbf{a}}_{02}+\mathbf{h}_{2,r}^H\mathbf{w}_{02}\|^2, \tag{5.25}$$

$$f_4(\mathbf{v}) = \gamma_{1,min}\sigma^2-\|\mathbf{v}^H\bar{\mathbf{a}}_1+\mathbf{h}_{1,r}^H\mathbf{w}_1\|^2, \tag{5.26}$$

$$f_5(\mathbf{v}) = \gamma_{2,min}\sigma^2-\|\mathbf{v}^H\bar{\mathbf{a}}_2+\mathbf{h}_{2,r}^H\mathbf{w}_2\|^2. \tag{5.27}$$

To tackle the problems with the inequality constraints by penalty method, the first step is to transform the original problems to the ones without constraints. Besides, in conventional penalty method, the equality constraints are also supposed to be added to the objective function as the penalty terms. However, the equality constraints in this problem are the unit norm ball, manifold optimization can be applied to solve it and (28b) is not needed to be added to the objective function.

$$\min_{\mathbf{v}} \quad f_0(\mathbf{v}), \tag{5.28a}$$

$$\text{s.t.} \quad \|\mathbf{v}(n)\| = 1, n=1,\cdots,2N, \tag{5.28b}$$

where

$$f_0(\mathbf{v}) = -\log_2(1+\frac{\|\mathbf{v}^H\mathbf{a}_1+\mathbf{h}_{1,r}^H\mathbf{w}_1\|^2}{\sigma^2})-\log_2(1+\frac{\|\mathbf{v}^H\mathbf{a}_2+\mathbf{h}_{2,r}^H\mathbf{w}_2\|^2}{\sigma^2})+\frac{\rho_2}{2}\sum_{i=1}^{5}(\bar{f}_i(\mathbf{v}))^2, \tag{5.29}$$

and the penalty term $\bar{f}_i$ is defined as followed:

$$\bar{f}_i(\mathbf{v}) = \max\{0, f_i(\mathbf{v})\}, i = 1, \cdots 5, \tag{5.30}$$

where ρ_2 is the penalty factor. In fact, the unit modulus constraints are considered as Riemannian manifold, which is very analogous to the Euclidean space. Therefore, the similar search methods for the optimization in the Euclidean space can be applied in manifold optimization. More detailed knowledge of manifold optimization can be found in [91]. In this chapter, we first define $\mathbf{A}_k = \bar{\mathbf{a}}_k\bar{\mathbf{a}}_k^H$ $\mathbf{A}_{0k} = \bar{\mathbf{a}}_{0k}\bar{\mathbf{a}}_{0k}^H, k = 1, 2$ and $\mathbf{A}_0 = (\bar{\mathbf{a}}_{01}+\bar{\mathbf{a}}_{02})(\bar{\mathbf{a}}_{01}+\bar{\mathbf{a}}_{02})^H$, and $b_k = \mathbf{h}_{k,r}^H\mathbf{w}_k, k = 1, 2$ and $b_{0k} = \mathbf{h}_{k,r}^H\mathbf{w}_{0k}, k = 1, 2$. Given a feasible point and $\mathbf{v}_t^{(l)}$ on the complex circle manifold $\mathcal{M} = \{\mathbf{v} \in \mathbb{C}^{2N\times 1}, \|\mathbf{v}(n)\| = 1, n = 1, \cdots, 2N\}$ at the t^{th} iteration of finding the minimum of (5.28), l represents the number of iterations for $f_0(\mathbf{v})$ to converge. The tangent space of this point is given by

$$\mathcal{T}_{\mathbf{v}_l}\mathcal{M} = \{\mathbf{z} \in \mathbb{C}^{2N\times 1}, \mathbf{z} \circ \bar{\mathbf{v}}_t^{(l)} = \mathbf{0}\}, \tag{5.31}$$

where $\bar{\mathbf{v}}$ is the conjugate of $\mathbf{v}$ and $\mathbf{z}$ is the vector which is on the tangent space through $\mathbf{v}_t^{(l)}$. In Euclidean space, one can find the globally or locally optimal solutions by finding the gradient $\nabla_{\mathbf{v}} f_0$ as the steepest decrease direction of $\mathbf{v}$, which can be expressed as

$$\nabla_{\mathbf{v}} f_0 = -\frac{2\sigma^2}{\ln 2}\frac{\mathbf{A}_1\mathbf{v}+\bar{\mathbf{a}}_1 b_1^H}{\sigma^2+\|\bar{\mathbf{a}}_1^H\mathbf{v}+b_1^H\|}\frac{2\sigma^2}{\ln 2}\frac{\mathbf{A}_2\mathbf{v}+\bar{\mathbf{a}}_2 b_2^H}{\sigma^2+\|\bar{\mathbf{a}}_2^H\mathbf{v}+b_2^H\|}+\rho_2\sum_{i=1}^{5}(f_i(\mathbf{v}))\nabla f_i(\mathbf{v}), \tag{5.32}$$

and when $f_k(\mathbf{v}) \le 0, k = 1, \cdots, 5$, the gradient $\nabla\bar{f}_k(\mathbf{v}) = 0$. Otherwise,

$$\nabla\bar{f}_1(\mathbf{v}) = 2(\gamma_{0,min}(\mathbf{A}_{01}+\mathbf{A}_{02})\mathbf{v}-\mathbf{A}_0\mathbf{v}), \tag{5.33}$$

$$\nabla\bar{f}_2(\mathbf{v}) = 2(\gamma_{0,min}(\mathbf{A}_1\mathbf{v}+\bar{\mathbf{a}}_1 b_1^H)-\mathbf{A}_{01}\mathbf{v}-\bar{\mathbf{a}}_{01}b_{01}^H)), \tag{5.34}$$

$$\nabla\bar{f}_3(\mathbf{v}) = 2(\gamma_{0,min}(\mathbf{A}_2\mathbf{v}+\bar{\mathbf{a}}_2 b_2^H)-\mathbf{A}_{02}\mathbf{v}-\bar{\mathbf{a}}_{02}b_{02}^H)), \tag{5.35}$$

$$\nabla \bar{f}_4(\mathbf{v}) = -2(\mathbf{A}_1\mathbf{v} + \bar{\mathbf{a}}_1 b_1^H), \tag{5.36}$$

$$\nabla \bar{f}_5(\mathbf{v}) = -2(\mathbf{A}_2\mathbf{v} + \bar{\mathbf{a}}_2 b_2^H). \tag{5.37}$$

Similarly, to find the steepest decrease direction of $\mathbf{v}_t^{(l)}$ in the manifold, we can obtain the Riemannian gradient, which is defined as:

$$\text{grad}_{\mathbf{v}_t^{(l)}} f_0 = \nabla_{\mathbf{v}_t^{(l)}} f_0 - \Re\{\nabla_{\mathbf{v}_t^{(l)}} f_0 \circ \bar{\mathbf{v}}_t^{(l)}\} \circ \mathbf{v}_t^{(l)}. \tag{5.38}$$

By the above transformation from $\nabla_{\mathbf{v}_t^{(l)}} f_0$ to $\text{grad}_{\mathbf{v}_t^{(l)}} f_0$, the Euclidean gradient can be projected to the tangent space of $\mathbf{v}_t^{(l)}$ and it is the steepest decrease direction of $\mathbf{v}_t^{(l)}$ in the manifold. Since (5.28) is a minimization problem, the search direction is usually the negative Riemannian gradient of the objective function at $\mathbf{v}^{(l)}$, which can be expressed as:

$$\eta(\mathbf{v}_t^{(l)}) = -\text{grad}_{\mathbf{v}_t^{(l)}} f_0. \tag{5.39}$$

Therefore, the next iteration point $\mathbf{v}_{t+1}^{(l)}$ is

$$\mathbf{v}_{t+1}^{(l)} = \mathbf{v}_t^{(l)} + \alpha\eta(\mathbf{v}_t^{(l)}), n = 1, \cdots, 2N, \tag{5.40}$$

where α is the step size and it is a positive value. However, it is noted that the updated $\mathbf{v}_{t+1}^{(l)}$ might not lie in the surface of the manifold. To tackle this issue, a process called retraction is required to map the $\mathbf{v}_{t+1}^{(l)}$ in the manifold, which is

$$\mathbf{v}_{t+1}^{(l)}(n) = \frac{\mathbf{v}_{t+1}^{(l)}(n)}{\|\mathbf{v}_{t+1}^{(l)}(n)\|}. \tag{5.41}$$

By this searching strategy, we can find the locally optimal solution when $\text{grad}_{\mathbf{v}_t^{(l)}} f_0$ tends to zero, denoted by $\mathbf{v}_{t^*}^{(l)}$ of (5.28). Let $\mathbf{v}_0^{(l+1)} = \mathbf{v}_{t^*}^{(l)}$. Then, the above steps are repeatedly executed when $f_0(\mathbf{v}_0^{(l+1)}) - f_0(\mathbf{v}_0^{(l)})$ is less than a threshold at the point $\mathbf{v}_0^{(l^*)}$. More detailed steps will be presented in the section of algorithm design.

5.3.3 Algorithm Design

Initial Point Selection Algorithm

The details of the proposed algorithm are shown in **Algorithm 3**. Firstly, the phase shifts of the RIS elements can be randomly generated, denoted by $\boldsymbol{\theta}^{(0)}$. Since the SCA is applied in the beamforming optimization, a feasible initial beamforming $\mathbf{w}^{(0)}$ is supposed to be searched. The idea of finding $\mathbf{w}^{(0)}$ is to transform the constraints in (5.10) to the forms in (5.42). Specifically, the constraints in (5.10) are not convex and cannot be accepted by the existing optimization toolboxes. However, we can make these constraints stricter but acceptable to the optimization toolboxes as shown in (5.42). In other words, if a $\mathbf{w}^{(0)}$ satisfies the constraints in (5.42), it can also satisfy the constraints of the original problem and a feasible initial point $\mathbf{w}^{(0)}$ can be given by the optimization toolboxes. The pseudocode of finding the initial point algorithm is presented in **Algorithm 3**.

Alternating Optimization Algorithm

The details of the proposed alternating algorithm are shown in **Algorithm 4**. This algorithm includes two sub-algorithms, namely beamforming and phase shift optimization algorithms. For beamforming optimization algorithm, the initial $\mathbf{W}^{(0)}$ has been obtained by **Algorithm 3**. Then, (5.21) can be iteratively solved by the optimization toolboxes. It is noted that the penalty factor ρ_1 should be updated as shown in **Algorithm 4** to guarantee the convergence of the algorithm. For the phase shift optimization problem, given the initial $\mathbf{v}_0^{(0)}$, let $e^{(0)} = f_0(\mathbf{v}_0^{(0)})$, the (5.28) can be solved by the manifold optimization. Different from the initial and beamforming algorithms, the phase shift algorithm does not utilize the existing optimization toolboxes. For a feasible point $\mathbf{v}_t^{(l-1)}$, where t is the iteration time for Riemannian gradient to converge, and l is the iteration time for $e^{(l)}$ to converge, $t \geq 0, l \geq 1$. the Riemannian gradient is updated by (5.38) and $\mathbf{v}_t^{(l-1)}$ is updated by (5.39)-(5.41). After the norm of Riemannian gradient converges at the point $\mathbf{v}_{t^*}^{(l-1)}$, make $\mathbf{v}_0^{(l)} = \mathbf{v}_{t^*}^{(l-1)}$ and $e^{(l)} = f_0(\mathbf{v}_0^{(l)})$ and check whether $e^{(l)} - e^{(l-1)}$ is less than a threshold or not. If not, increase the penalty factor by $\rho_2 = \beta_2\rho_2$ and repeat the above steps until $e^{(l)} - e^{(l-1)}$ is less than a threshold at the point $\mathbf{v}_0^{(l^*)}$. By alternatively solving the beamforming and phase shift optimization problems, we can obtain the suboptimal solution of the proposed optimization problem. The

Algorithm 3 The Initial Point Algorithm

1: **Initialize:** Generate random phase shift angles in the range of $[0, 2\pi]$ and for RIS elements.
2: **repeat:**
3: Find a $\mathbf{w}$, which satisfies

$$\Re\{\mathbf{h}_{01}^H\mathbf{w}_{01} + \mathbf{h}_{02}^H\mathbf{w}_{02}\} \geq \sqrt{\gamma_{0,min}}\|\mathbf{h}_{01}^H\mathbf{w}_1, \mathbf{h}_{02}^H\mathbf{w}_2, \sigma\|, \tag{5.42a}$$
$$\Re\{\mathbf{h}_1^H\mathbf{w}_{01}\} \geq \sqrt{\gamma_{0,min}}\|\mathbf{h}_1^H\mathbf{w}_1, \sigma\|, \tag{5.42b}$$
$$\Re\{\mathbf{h}_2^H\mathbf{w}_{02}\} \geq \sqrt{\gamma_{0,min}}\|\mathbf{h}_2^H\mathbf{w}_2, \sigma\|, \tag{5.42c}$$
$$\Re\{\mathbf{h}_1^H\mathbf{w}_1\} \geq \sqrt{\gamma_{1,min}}\sigma, \tag{5.42d}$$
$$\Re\{\mathbf{h}_2^H\mathbf{w}_2\} \geq \sqrt{\gamma_{2,min}}\sigma, \tag{5.42e}$$
$$(5.11e), (5.11f). \tag{5.42f}$$

4: **until:** (5.42) is solvable.
5: **output:** The initial feasible solution $\boldsymbol{\theta}^{(0)}$ and $\mathbf{w}^{(0)}$.

pseudocode of the alternating algorithm is presented in **Algorithm 4**.

5.3.4 Complexity Analysis

In the beamforming sub-algorithm, the SDR is applied to find the solutions of the beamforming matrices. According to the complexity analysis in chapter 4, a total complexity $\mathcal{O}(L^3)$ is required, where L is the size of the matrix variable. Since SCA is applied and a convergence threshold ϵ_1 is introduced to determine whether the algorithm stops or not, the total complexity for the beamforming algorithm is $\mathcal{O}(M^3 \log(\frac{1}{\epsilon_1}))$ in the worst case. In this chapter, M refers to the number of antennas. But according to [89], the manifold algorithm can cost a complexity of $\mathcal{O}((2N)^{1.5})$ in worst case. Since manifold optimization and penalty method is applied to make $\|\text{grad}(f_0)\|$ and $e^{(l)}$ converge with a threshold ϵ_2 and ϵ_3, respectively, the total complexity of the penalty-based manifold algorithm proposed is $\mathcal{O}((2N)^{1.5} \log(\frac{1}{\epsilon_2}) \log(\frac{1}{\epsilon_3}))$ in worst case. Numerically, the complexity of penalty-based manifold can be much lower than SDR as long as the N is not far less than M and there is no difference between ϵ_1 and ϵ_2. Unfortunately, despite the advantage of the manifold optimization in terms of complexity, it cannot be applied in the beamforming optimization problem since there is no specific equality constraint such as the norm ball constraint in the phase shift optimization and it cannot form a manifold.

Algorithm 4 Alternating Optimization Algorithm

Initialize: Given the initial $\mathbf{W}^{(0)} = \mathbf{w}^{(0)}\mathbf{w}^{(0),H}$.
repeat
Beamforming Optimization: Given the fixed beamforming vectors $\mathbf{W}^{(l)}$ and initialize $\rho_1 = \rho_1^0$.
repeat
Iteratively solve the Optimization Problem (5.21) and update the penalty factor $\rho_1 = \beta_1\rho_1$.
until R_{sum} convergence with a threshold ϵ_1.
Phase Shift Optimization: Given the optimal beamforming vector $\mathbf{w}^*$ by EVD of $\mathbf{W}^*$ and initialize $\rho_2 = \rho_2^0$.
repeat
repeat
Solve (5.28) by manifold optimization.
until: $\|\text{grad}_{\mathbf{v}} f_0\| \leq \epsilon_2$ convergence.
Update $\rho_2 = \beta_2\rho_2$.
until: $e^{(l)} - e^{(l-1)} \leq \epsilon_3$ convergence.
until: R_{sum} convergence.
output: The optimal $\mathbf{w}^*$ and $\mathbf{v}^*$.

5.4 Simulation Results

In this section, simulation results are presented to demonstrate the superior performance of the proposed scheme. The considered simulation settings are as followed: each user is located in circle region whose radius is 10 m. The centre of this circle of the NOMA user U_k is both 30 m away from the BS_k and RIS_k. The centre of the CoMP user is 40 m away from both of two RISs. Without the loss of generality, the total bandwidth of this system is considered as 10 MHz and the noise spectrum density is -150 dBm/Hz. The target SINR for each user is assumed to be the same. Besides, the Rayleigh fading is applied in the channels of the direct links between BS and NOMA users and Rician fading is applied in the other channels [92]. For example, channel $\mathbf{F}$, following the Rician fading can be expressed as

$$\mathbf{F} = \sqrt{\frac{\kappa}{1+\kappa}}\mathbf{F}_{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}}\mathbf{F}_{\text{nLoS}}, \tag{5.43}$$

where κ denotes the Rician factor and it is set to 2 in this chapter. $\mathbf{F}_{\text{LoS}}$ and $\mathbf{F}_{\text{nLoS}}$ denote the line of sight and non line of sight components, respectively. The path loss exponents for the Rayleigh fading channels and Rician fading channels are set to 2.8 and 2.2, respectively. Besides, for comparison, five benchmark schemes, namely SDR by Gaussian randomization, NOMA with random phase shifts, OMA with optimized phase shifts, OMA with random phase shifts and NOMA without RIS, respectively are proposed.

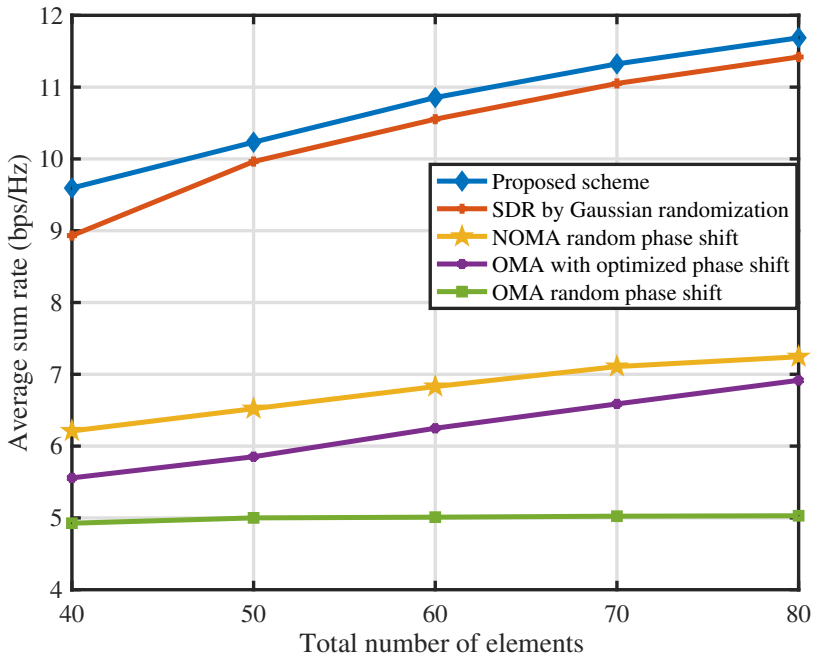

Fig. 5.2. Average sum rate versus number of elements. $M = 12$, P_{max}=30 dBm and γ_{min}=0 dB

5.4.1 Impact of the Number of RIS Elements

Fig. 5.2 shows the relationship between the sum rate and the total number of elements of two RISs. As shown in the figure, the proposed scheme outperforms the other four benchmark schemes. Besides, the increase of the RIS elements can lead to the better performance of the sum rate in this system for all these five schemes. It is because that a higher number of elements can produce higher beamforming gains. Besides, the schemes with the optimized phase shifts perform much better than those with random phase shift as the number of elements increases. For example, in the proposed scheme, the sum rate increases from 9.5948 bps/Hz to 11.6871 bps/Hz as the number of total elements increases from 40 to 80, while there is only 1.0333 bps/Hz increase in the scheme of NOMA random phase shifts scheme. Even worse, for the scheme of OMA random phase shift, the increase is only 0.1041 bps/Hz. These numerical results indicate the benefit of optimizing the phase shifts on improving the sum rate of the system. Furthermore, it is shown that the proposed scheme can perform better than the SDR by Gaussian randomization. For example, when the number of total elements is 40, the sum rate of the proposed scheme can be 0.6628 bps/Hz higher than that of Gaussian randomization.

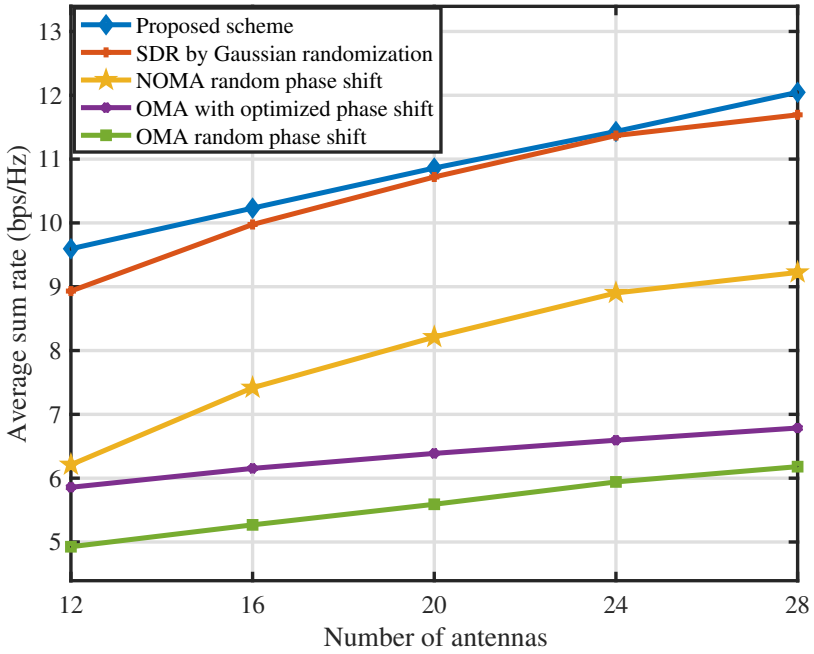

Fig. 5.3. Average sum rate versus number of antennas. $N = 20$, P_{max}=30 dBm and γ_{min}=0 dB

5.4.2 Impact of the Number of Antennas

Fig. 5.3 shows the relationship between the sum rate and the number of antennas. From the figure, it can be seen that the proposed scheme still has an overwhelming advantage than the other four schemes in terms of the sum rate. Meanwhile, from the figure, it can be found that the sum rate increases as the number of antennas increases. It is reasonable since a larger amount of antennas can achieve a higher active beamforming gain. Meanwhile, it can be observed that the number of antennas can have a more significant impact on sum rate in NOMA than that in OMA.

5.4.3 Impact of the Maximum Power Budget

Fig. 5.4 shows the relationship between the sum rate and the maximum power budget P_{max}. From the figure, it can be found that the sum rate achieved by all schemes increases with P_{max} from 30 dBm to 38 dBm. It is because different from the energy efficiency maximization problem, in sum rate maximization problem, all the power budget can be fully allocated to the users. With more allocated power, the achievable sum rate is more likely to increase. From the aspect of mathematics, the higher P_{max} expands the feasible domain of the variables. Therefore, a more optimal value can be obtained with an expanded

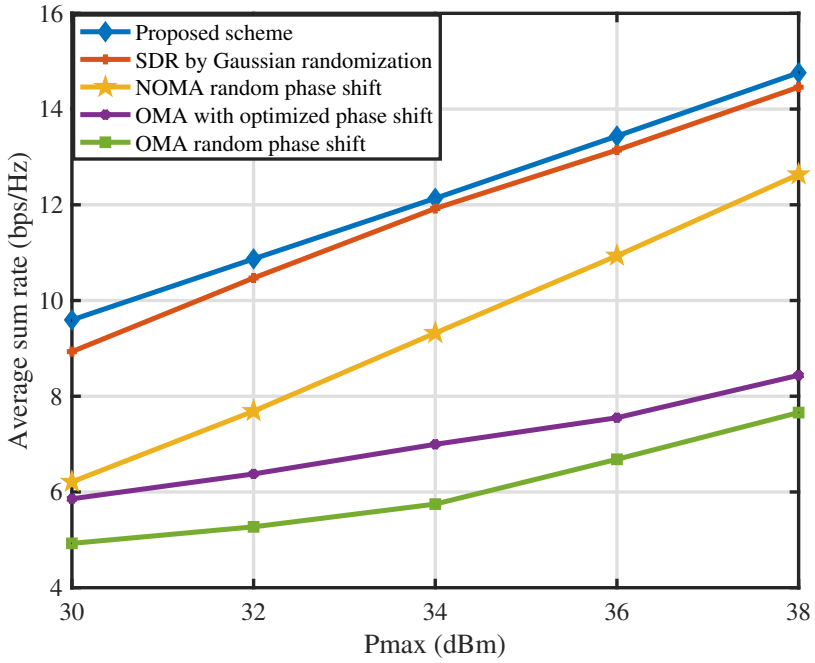

Fig. 5.4. Average sum rate versus the total power budget. $N = 20$, $M = 12$ and γ_{min}=30 dB

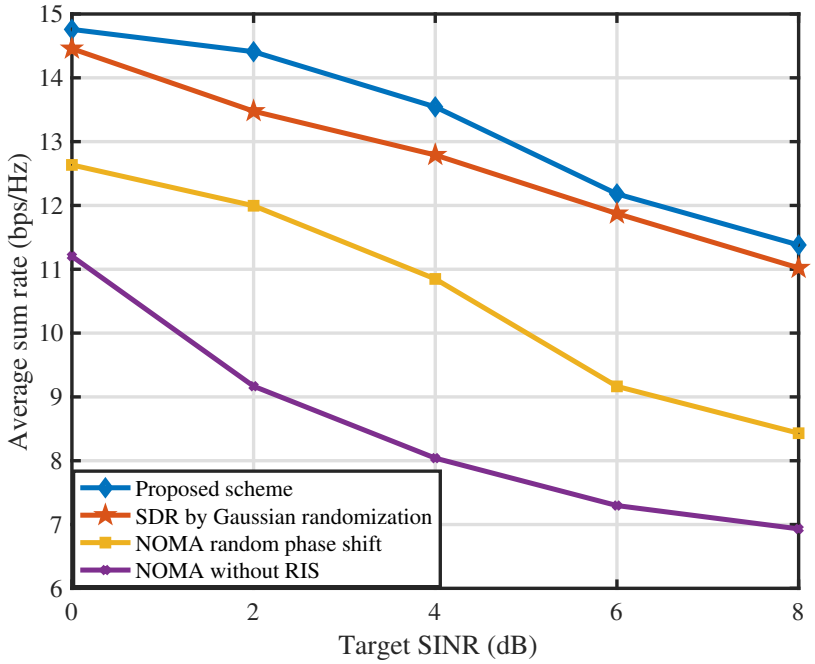

Fig. 5.5. Average sum rate versus the SINR

domain.

5.4.4 Impact of the Target SINR

Fig. 5.5 shows the relationship between the sum rate and the target SINR. From the figure, it can be seen that the sum rate will attenuate as the target SINR increases. It is because the higher target SINR makes the constraints stricter. More power has to be allocated to satisfy

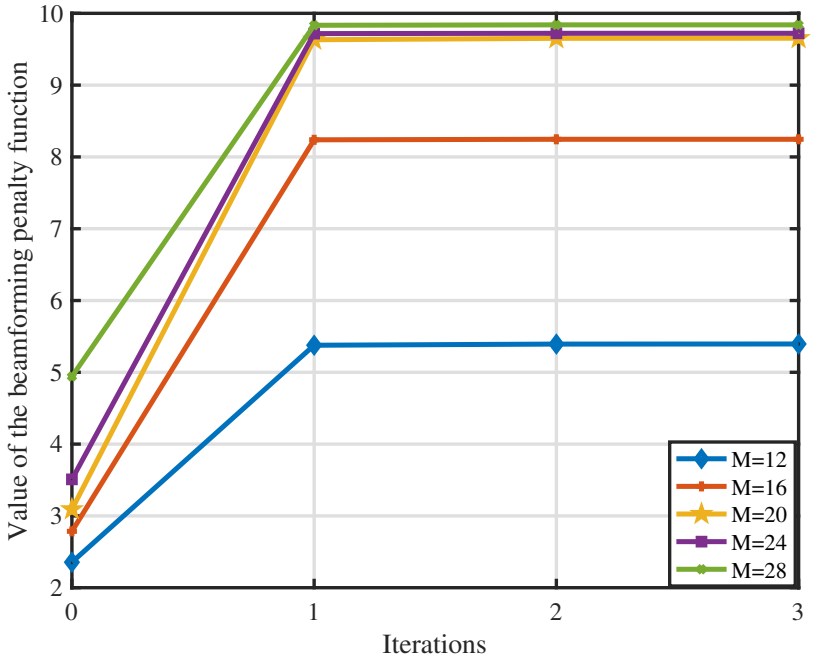

Fig. 5.6. Convergence of the beamforming sub-algorithm. $N = 20$ P_{max}=30 dBm and $\gamma_{min} = 0$ dB

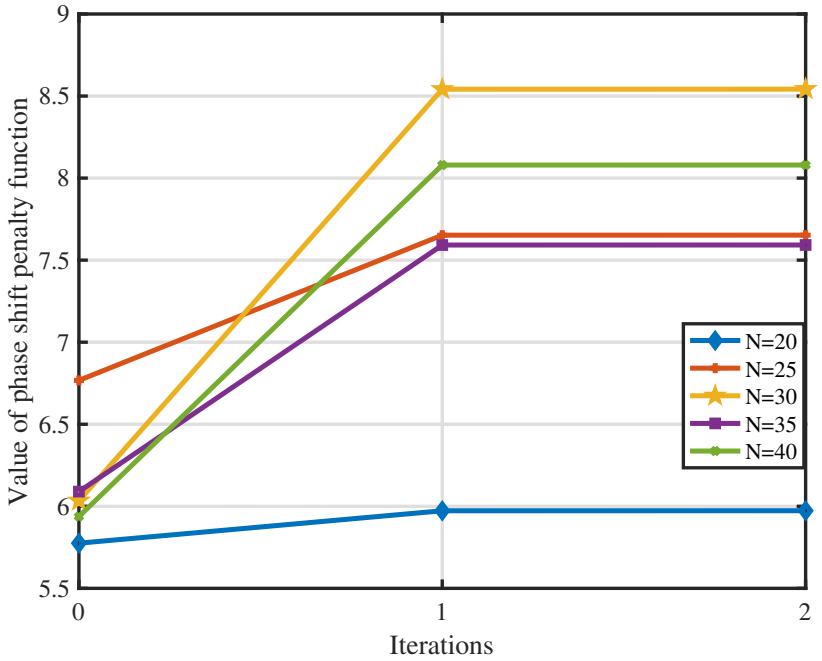

Fig. 5.7. Convergence of the phase shift sub-algorithm. $M = 12$ P_{max}=30 dBm and $\gamma_{min} = 0$ dB

the QoS constraints and less power is used to improve the sum rate of the NOMA users in the system. Besides, in this figure, the scheme of NOMA without RIS is introduced and it performs the worst among all the schemes. It shows that the RIS can be a useful and effective device to improve the sum rate of the system not only because elements can produce higher beamforming gains, but RIS can build a new links from the BS to the users, which provide more degree of freedom (DoF) with the users to adjust the beamforming vectors and phase shifts of the RISs.

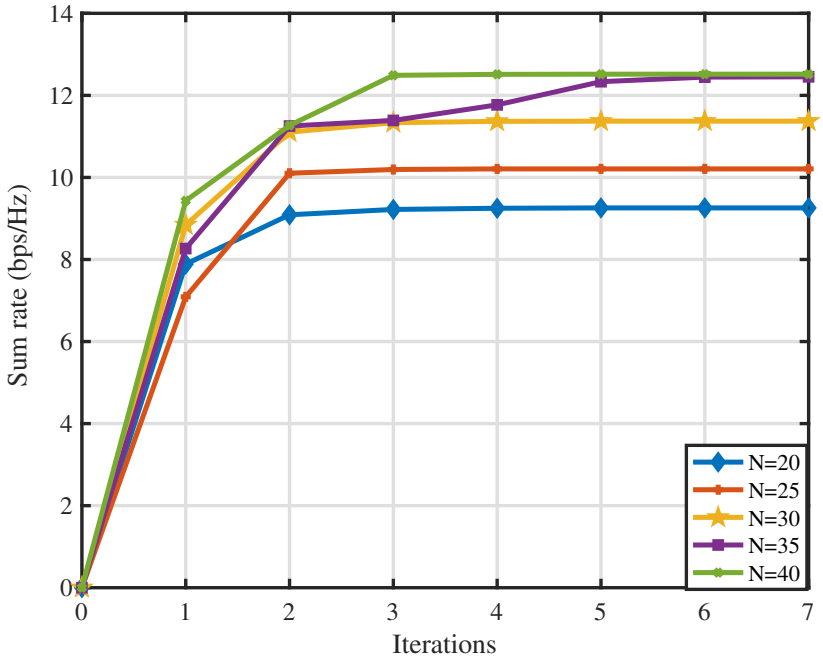

Fig. 5.8. Convergence of the alternating algorithm. $M = 12$, P_{max}=30 dBm and $\gamma_{min} = 0$ dB

5.4.5 Convergence of the Proposed Algorithm

In Fig. 5.6, Fig. 5.7 and Fig. 5.8, the convergence of two sub-algorithms and the alternating algorithm are demonstrated by the numerical analysis. In Fig. 5.6 and Fig. 5.7, the convergence of the penalty functions of two sub-algorithms are presented. For the convergence of the beamforming algorithm, N is set to 20. For the convergence of the phase shift algorithm, M is set to 12. It can be observed that all these two penalty functions can converge within a very low inner iterations. Interestingly, in the phase shift optimization algorithm, after convergence of these 5 cases, the value of the penalty function for $N = 30$ is the highest and the value for $N = 40$ only comes in second place. Intuitively, this phenomenon contradicts the results in Fig. 5.2 because for each curve in Fig. 5.2, the sum rate increases as the number of elements increases. However, it is reasonable because when discussing the convergence of the algorithm, the channels are generated once. There might exist the cases which the sum rate with a higher number of elements can outperform one with a lower number of elements since the quality of the generated channels cannot be guaranteed. However, when discussing the impact of elements on the performance of the system, the algorithm is executed many times to obtain the average sum rate for each case. In statistics, taking the average can weaken the influence caused by channels and relative fair results can be obtained. In Fig. 5.7, it can demonstrate the performance of the system can significantly be increased by the

alternating optimization.

5.5 Conclusions of the Chapter

In this chapter, we have designed a novel scheme in an RIS-assisted CoMP-NOMA system and realized the sum rate maximization of the NOMA users. Firstly, to tackle the non-convexity of the original problem caused by the coupled variables, the proposed original problem has been decomposed to two suboptimal problems, beamforming and phase shift optimization problems. In the beamforming optimization problem, since there is no manifold in the constraints, SDR has been applied to transform the problem to one which had the rank-one constraints. Then, interestingly, the rank-one constraints of the beamforming matrices of NOMA users can be relaxed without degrading the optimality of the problem. For CoMP user's beamforming matrix, the penalty-based SCA method has been used to tackle its rank-one issue. For the phase shift optimization, the penalty-based manifold optimization has been applied to greatly decrease the complexity of the algorithm. The simulation results have demonstrated that the proposed scheme could outperform other five benchmark schemes such as SDR by Gaussian randomization and RIS assisted OMA systems in sum rate maximization.

5.6 Proof of Proposition 1

From the problem formulation of (5.12), it can be observed that it is a convex problem with each $\mathbf{W}_k, k = 1, 2$. Therefore, the strong duality holds and the solutions of $\mathbf{W}$ which satisfy the KKT conditions are optimal. To discuss the KKT condition of this problem, the

Lagrangian function is first formulated as followed:

$$\begin{aligned}\mathcal{L}(\mathbf{W},\boldsymbol{\lambda},\mathbf{X}) = &-\log_2(1+\frac{\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)}{\sigma^2})-\log_2(1+\frac{\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_2)}{\sigma^2})\\ &-\lambda_1(\mathrm{Tr}(\mathbf{H}_0\mathbf{W}_0)-\gamma_{0,min}(\mathrm{Tr}(\mathbf{H}_{01}\mathbf{W}_1)+\mathrm{Tr}(\mathbf{H}_{02}\mathbf{W}_2)+\sigma^2))\\ &-\lambda_2(\mathrm{Tr}(\overline{\mathbf{H}}_1\mathbf{W}_0)-\gamma_{0,min}(\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)+\sigma^2))\\ &-\lambda_3(\mathrm{Tr}(\overline{\mathbf{H}}_2\mathbf{W}_0)-\gamma_{0,min}(\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_2)+\sigma^2))\\ &-\lambda_4(\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)-\gamma_{1,min}\sigma^2)-\lambda_5(\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_2)-\gamma_{2,min}\sigma^2)\\ &+\lambda_6(\mathrm{Tr}(\mathbf{I}_1\mathbf{W}_0)+\mathrm{Tr}(\mathbf{W}_1)-P_{max})+\lambda_7(\mathrm{Tr}(\mathbf{I}_2\mathbf{W}_0)+\mathrm{Tr}(\mathbf{W}_2)-P_{max})\\ &-\mathrm{Tr}(\mathbf{X}_0\mathbf{W}_0)-\mathrm{Tr}(\mathbf{X}_1\mathbf{W}_1)-\mathrm{Tr}(\mathbf{X}_2\mathbf{W}_2),\end{aligned} \tag{5.44}$$

where $\boldsymbol{\lambda}$ and $\mathbf{X}$ represent the Lagrangian multipliers.

According to the KKT conditions, the following relationships should hold such that:

$$\lambda_i \geq 0, i = 1,\cdots,7, \tag{5.45}$$

$$\mathbf{X}_k \succeq \mathbf{0}, k = 0,1,2, \tag{5.46}$$

$$\mathbf{X}_k\mathbf{W}_k \succeq \mathbf{0}, k = 0,1,2, \tag{5.47}$$

$$\frac{\partial\mathcal{L}}{\partial\mathbf{W}_k} = \mathbf{0}, k = 0,1,2. \tag{5.48}$$

Then, (5.48) can be further developed as followed:

$$\mathbf{X}_0 = -\lambda_1\mathbf{H}_0-\lambda_2\overline{\mathbf{H}}_1-\lambda_3\overline{\mathbf{H}}_2+\lambda_6\bar{\mathbf{I}}_1+\lambda_7\bar{\mathbf{I}}_2, \tag{5.49}$$

$$\mathbf{X}_1 = -(\frac{1}{\ln 2(\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)+\sigma^2)}+\lambda_4-\lambda_2\gamma_{0,min})\mathbf{H}_1+\lambda_1\gamma_{0,min}\mathbf{H}_{01}+\lambda_6\mathbf{I}, \tag{5.50}$$

$$\mathbf{X}_2 = -(\frac{1}{\ln 2(\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_2)+\sigma^2)}+\lambda_5-\lambda_3\gamma_{0,min})\mathbf{H}_2+\lambda_1\gamma_{0,min}\mathbf{H}_{02}+\lambda_7\mathbf{I}. \tag{5.51}$$

Besides, $\mathbf{X}_0$ can be also be written as:

$$\mathbf{X}_0 = \begin{bmatrix}\lambda_6\mathbf{I}-\lambda_1\mathbf{H}_{01}-\lambda_2\mathbf{H}_1 & -\lambda_1\mathbf{h}_{01}\mathbf{h}_{02}^H\\ -\lambda_1\mathbf{h}_{02}\mathbf{h}_{01}^H & \lambda_7\mathbf{I}-\lambda_1\mathbf{H}_{02}-\lambda_3\mathbf{H}_2\end{bmatrix}. \tag{5.52}$$

From (5.52), $\lambda_6\mathbf{I}-\lambda_1\mathbf{H}_{01}-\lambda_2\mathbf{H}_1$ and $\lambda_7\mathbf{I}-\lambda_1\mathbf{H}_{02}\lambda_7\mathbf{H}_2$ are the principle submatrices of

$\mathbf{X}_0$, which is obtained by selecting the same subset of columns of rows. Since $\mathbf{X}_0$ is positive semidefinite, the two principle submatrices above should be both positive semidefinite. Besides, it is impossible for λ_1, λ_2 and λ_6 to be equal to zero simultaneously. Otherwise, $\mathbf{X}_1$ will be negative definite matrix, which contradicts (5.46). By the same logic, λ_1, λ_3 and λ_7 cannot be zero simultaneously as well. To guarantee these properties, λ_6 and λ_7 should be positive. Otherwise, the two submatrices are negative definite. Then, we return to analyze the properties of $\mathbf{X}_1$ and $\mathbf{X}_2$. Denote $\mathbf{Y}_1$ and $\mathbf{Y}_2$ as

$$\mathbf{Y}_1 = (\frac{1}{\ln 2(\text{Tr}(\mathbf{H}_1\mathbf{W}_1) + \sigma^2)} + \lambda_4 - \lambda_2\gamma_{0,min})\mathbf{H}_1 - \lambda_1\gamma_{0,min}\mathbf{H}_{01}, \tag{5.53}$$

$$\mathbf{Y}_2 = (\frac{1}{\ln 2(\text{Tr}(\mathbf{H}_2\mathbf{W}_2) + \sigma^2)} + \lambda_5 - \lambda_3\gamma_{0,min})\mathbf{H}_2 - \lambda_1\gamma_{0,min}\mathbf{H}_{02}. \tag{5.54}$$

Then,

$$\mathbf{X}_1 = \lambda_6\mathbf{I} - \mathbf{Y}_1, \tag{5.55}$$

$$\mathbf{X}_2 = \lambda_7\mathbf{I} - \mathbf{Y}_2. \tag{5.56}$$

From (5.55) and (5.56), since the channels are randomly generated, according to [84], the possibility that $\mathbf{Y}_1$ or $\mathbf{Y}_2$ can have two identical eigenvalues is zero. Thus, the following relationships should hold, where

$$\lambda_6 \geq e_{max}(\mathbf{Y}_1), \tag{5.57}$$

$$\lambda_7 \geq e_{max}(\mathbf{Y}_2), \tag{5.58}$$

where $e_{max}(\cdot)$ is the maximum eigenvalue of the matrix. Otherwise, $\mathbf{X}_1$ and $\mathbf{X}_2$ will not be positive semidefinite since they have at least one negative eigenvalue, which contradicts (5.46). Therefore, the rank of $\mathbf{X}_1$ and $\mathbf{X}_2$ should be greater than $M-1$. To guarantee (5.47), the following relationships should also hold, where

$$\text{Rank}(\mathbf{X}_1) + \text{Rank}(\mathbf{W}_1) \leq M, \tag{5.59}$$

$$\text{Rank}(\mathbf{X}_2) + \text{Rank}(\mathbf{W}_2) \leq M. \tag{5.60}$$

If $\mathbf{X}_1$ and $\mathbf{X}_2$ are full-rank, the rank of $\mathbf{W}_1$ and $\mathbf{W}_2$ is zero, meaning that (5.20d) can not be satisfied. Therefore, the rank of $\mathbf{X}_1$ and $\mathbf{X}_2$ is $M-1$ and the rank of $\mathbf{W}_1$ and $\mathbf{W}_2$ is one.

Chapter 6

Beamforming and Phase Shift Design in STAR-RIS Assisted ISAC NOMA Networks

6.1 Introduction

6.1.1 Motivations of the Chapter

As is demonstrated in **Chapter 1**, it becomes increasingly demanding that next generation wireless networks should also be capable to simultaneously achieve both sensing and communications [93]. Confronting this demand, ISAC has been proposed to achieve the sharing of spectrum resource and equipment between communications and sensing functionalities [94], [95], enabling high rate transmissions for wireless devices while realizing accurate sensing for targets in many application scenarios, i.e., autonomous driving [96]. Except ISAC, **Chapter 3**-**Chapter 5** have sufficiently demonstrated the superior performance of RIS/STAR-RIS assisted NOMA on improving EE and sum rate of the networks. In fact, NOMA, as a innovative and promising technique is still able to empower the communication systems of ISAC networks due to its high spectrum efficiency. More recently, as the development of ultra-large scale RIS/STAR-RIS, it is possible for RIS/STAR-RIS to participate in ISAC networks to empower both sensing and communications simultaneously. Meanwhile, the benefits of applying RIS/STAR-RIS can also be leveraged to ISAC networks, i.e., low power consumption, low hardware investment, wide coverage

and simple installation. It is noted that RIS assisted ISAC NOMA systems have been widely studied as mentioned in **Chapter 1**. However, on the one hand, the ISAC NOMA networks with the assistance of STAR-RIS have not received sufficient attentions. To a certain extent, the STAR-RIS assisted ISAC networks are also a very promising topic due to the limitations of RIS communications and sensing. By using STAR-RIS in the wireless networks, both communications and sensing are able to be realized in full space. On the other hand, most existing RIS assisted ISAC NOMA works focus on how to improve the quality of sensing by maximizing the minimum baeampattern gain. However, in certain specific ISAC cases, the priority of communication quality is much higher than that of sensing, i.e., a network with a great number of AR/VR devices and limited number of sensing targets. In such scenarios, the quality of communications should be the major objective to consider, and the importance of sensing is far less critical than communications. Based on the aforementioned views, this chapter is motivated to investigate a sum rate maximization problem in a STAR-RIS ISAC NOMA network by jointly optimizing the beamforming and phase shifts in the 3D dimension space, which addresses the current void in research of RIS/STAR-RIS assisted ISAC NOMA. From the perspective of mathematical methods, recall in **Chapter 2** and **Chapter 4** that it is common to apply SDR to transform a QCQP to LP problem with non-convex rank-one constraints in non-NOMA case i.e., OMA. Then, the Gaussian randomization is applied to get approximated solutions based the solutions of LP dropping rank-one if the number of sampling points are large enough [76]. Usually, the approximated optimal values can be at least $\frac{\pi}{4}$ of the values of LP by dropping rank-one [78]. However, as is illustrated in **Chapter 4**, Gaussian randomization has the following drawbacks. On the one hand, in NOMA, the transformed problems by using SDR are not LP, causing that the quality of solutions cannot be guaranteed. On the other hand, although the transformed problems can become convex by using additional mathematical techniques, if the problems have too many constraints or the relaxed solutions are high rank, it will be extremely difficult to generate the points which satisfy all the constraints in these optimization problems [83]. To solve this dilemma, in this chapter, the original problem is first decoupled to two suboptimal problems. A penalty-based SDR is used to tackle the beamforming optimization problem. Then, a penalty-based SDR method is also proposed to tackle the rank-one issue of phase shifts.

6.1.2 Contributions of the Chapter

The contributions of this chapter are summarized as followed:

- A STAR-RIS assisted downlink ISAC NOMA network is proposed, where multiple primary users are located at the R region of STAR-RIS and one secondary user is located at the T region of STAR-RIS. Sensing targets are located at both regions of STAR. Such this network effectively breaks the limitation of RIS in the field of communication and sensing coverage. The sum rate maximization problem for primary users is studied by optimizing the beamforming vectors and phase shifts of STAR-RIS elements.
- To effectively solve the original non-convex problem, this chapter first decouples it into two suboptimal problems, namely beamforming and phase shift optimization problems. The beamforming problem is solved by SDR and fractional programming [68], [82] Similarly, SDR is also applied in phase shift optimization. In these problems, rank-one constraints should be taken into consideration. Rather than using inefficient Gaussian randomization, a penalty-based SDR and fractional programming method are implemented and the rank-one suboptimal solutions can be obtained.
- By sufficient simulation results, the proposed scheme can perform better than the benchmark schemes such as RIS assisted ISAC NOMA, STAR-RIS assisted ISAC NOMA with random phase shifts and STAR-RIS assisted ISAC OMA. Besides, the performance of communications only STAR-RIS assisted NOMA is also shown as a benchmark and compared to that of the proposed scheme. Besides, the proposed method is also applicable in other RIS/STAR-RIS networks related to the beamforming and phase shift optimization.

6.1.3 Particular Notations of the Chapter

Except the common mathematical notations mentioned in **Chapter 1**, some other particular notations in this chapter is also presented in Table. 6.1:

Notations	Description
T & R	Transmission & reflection regions of STAR-RIS
K	Number of primary users.
Q_T & Q_R	Number of sensing targets in T and R regions
$\mathbf{h}_k^H$	Channel gain between STAR-RIS to k^{th} user
$\mathbf{G}$	Channel gain between BS to STAR-RIS
K	Number of users
$\mathbf{w}_k$	Beamforming vector of the k^{th} user
$\mathbf{W}_k$	Transformed matrix of $\mathbf{w}_k$ by SDR
$\boldsymbol{\theta_l}$	Phase shift vector of elements of STAR-RIS.
$\mathbf{V_l}$	Transformed matrix of $\boldsymbol{\theta_l}$ by SDR
$\mathbf{a}_{q_l,l}(\psi_{q_l,l})$	Steering vector of the target
$\mathcal{P}_{q_l,l}$	Beampattern gain of the target
ς & $\boldsymbol{\mu}$	Auxiliary variables of beamforming and phase shift optimization problems

Table 6.1. Parameter Settings

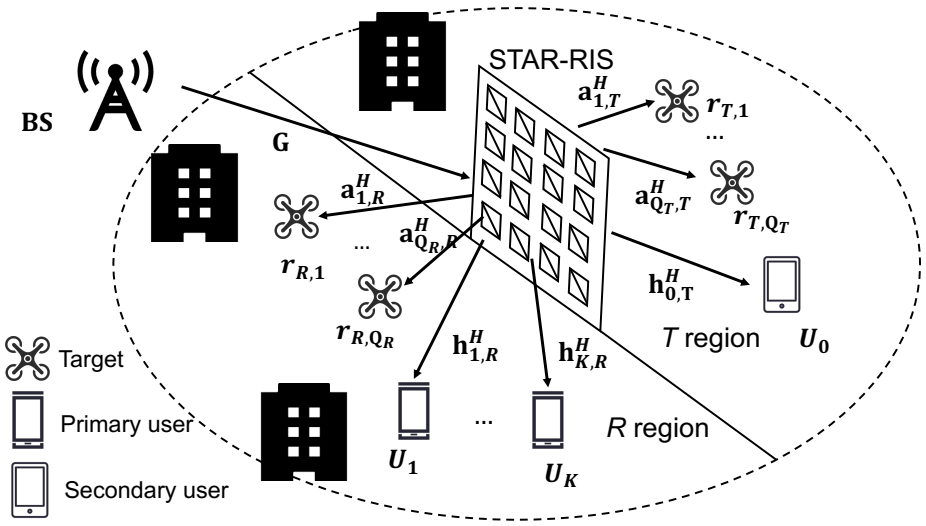

Fig. 6.1. STAR-RIS assisted ISAC-NOMA transmission system

6.2 System Model

In this chapter, a STAR-RIS assisted ISAC NOMA network is proposed, where one dual-functional BS is equipped with M antennas and K single-antenna primary users, denoted as $U_k, k = 1, \cdots, K$ are located in the reflection region (R region) and one single-antenna secondary user, denoted as U_0 is located in the transmission region (T region). A uniform linear array (ULA)-STAR-RIS consists of N transmitting and reflecting elements, respectively. Meanwhile, there are Q_R and Q_T targets in R and T regions, respectively. r_{l,q_l} represents the q_l^{th} target in l region, $q_l \leq Q_l$, $l \in \{T, R\}$. For example, r_{R,q_R} refers to the q_R^{th} target in R region.

6.2.1 Communication Model

As is shown in Fig. 6.1, BS transmits the signals of the users with the beamforming of $\mathbf{w}_0$ and $\mathbf{w}_1 \cdots \mathbf{w}_K$. Therefore, the superimposed signal transmitted by the BS $\mathbf{x}$ can be expressed as followed:

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{w}_k s_k + \mathbf{w}_0 s_0, \tag{6.1}$$

where $s_0, s_k, k = 1, \cdots, K$ refer to the signals of U_0 and U_k, respectively. Assume there is no direct link between BS and each communication user. Denote $\mathbf{G} \in \mathbb{C}^{N \times M}$, as the channel gain from the BS to STAR-RIS. $\mathbf{h}_{k,R}^H, k = 1, \cdots, K$ refers to the channel gain between STAR-RIS to the k^{th} primary user at R region. $\mathbf{h}_{0,T}^H$ refers to the channel gain between STAR-RIS to the secondary user in the T region. $\boldsymbol{\theta}_l = (\sqrt{\beta_{1,l}} e^{j\theta_{l,1}}, \cdots, \sqrt{\beta_{N,l}} e^{j\theta_{l,N}})$, $l \in \{T, R\}$ refers to the vector of the phase shift angles of the l elements. It is noted that $\beta_{n,l}$ is the amplitude coefficient of the n^{th} l element and $\beta_{n,T} + \beta_{n,T} = 1$ since passive STAR-RIS elements are applied. Therefore, the total channel gain of the $U_k, k = 1, \cdots, K$ can be expressed as

$$\mathbf{h}_k^H = \mathbf{h}_{k,R}^H \text{diag}(\boldsymbol{\theta}_R) \mathbf{G}, \tag{6.2}$$

and the total channel gain of U_0 can be expressed as

$$\mathbf{h}_0^H = \mathbf{h}_{0,T}^H \text{diag}(\boldsymbol{\theta}_T) \mathbf{G}. \tag{6.3}$$

The received signal of each user $y_i, i = 0, \cdots, K$ can be expressed as

$$y_i = \mathbf{h}_i^H (\sum_{i=0}^{K} \mathbf{w}_i s_i) + n_i, \tag{6.4}$$

where n_i is the additional white Gaussian noise (AWGN) of each user.

In this chapter, the primary users are assumed to have a higher decoding capability than the secondary user. Therefore, by employing NOMA in each user, the secondary user's signal is decoded before each primary user. This will cause that the signal of secondary user can be decoded at the primary users but the primary users' signals cannot be decoded at the secondary user. SIC is able to be employed to remove the interference of U_0's signal when decoding U_k's signals, $k = 1, \cdots, K$. However, conversely, U_k's signals should be

considered as the interference and cannot be cancelled by SIC when decoding U_0's signal. By such an SIC scheme, the SINR of the primary user U_k can be expressed as

$$\gamma_k = \frac{\|\mathbf{h}_k^H \mathbf{w}_k\|^2}{\sum_{i=1, i\neq k}^{K} \|\mathbf{h}_k^H \mathbf{w}_i\|^2 + \sigma^2}, k, i = 1, \cdots, K, \tag{6.5}$$

and $\gamma_{i\to 0}$, which means the SINR when decoding U_0's signal at U_i should be formulated as followed:

$$\gamma_{i\to 0} = \frac{\|\mathbf{h}_i^H \mathbf{w}_0\|^2}{\sum_{k=1}^{K} \|\mathbf{h}_i^H \mathbf{w}_k\|^2 + \sigma^2}, i = 0, \cdots, K, k = 1, \cdots K. \tag{6.6}$$

6.2.2 Sensing Model

In this chapter, there are Q_T and Q_R targets at the transmission and reflection regions of the STAR-RIS, respectively. For the convenience in simulation, Q_T and Q_R are set to be equal. Beampattern gain, also known as effective sensing power of each target is considered as the sensing performance metric, indicating the accuracy of sensing. It means that, the more the beampattern gain of a target is, the higher accuracy of the sensing can achieve. In this chapter, the transmitting power of the signals is used to detect the targets. To get the expression of beampattern gain of each target, the transmitting or reflected signal at the STAR-RIS from BS at the l region can be first expressed as

$$\bar{\mathbf{x}}_l = \mathbf{\Theta}_l \mathbf{G}(\mathbf{w}_0 s_0 + \sum_{k=1}^{K} \mathbf{w}_k s_k), k = 1, \cdots, K, \tag{6.7}$$

where $\mathbf{\Theta}_l = \text{diag}(\boldsymbol{\theta}_l)$. The corresponding covariance matrix, which affects the sensing quality is given by

$$\mathbf{R} = \mathbf{\Theta}_l \mathbf{G}(\mathbf{W}_0 + \sum_{k=1}^{H} \mathbf{W}_k)\mathbf{G}^H \mathbf{\Theta}_l^H, \tag{6.8}$$

and the power of the probing signal of q_l^{th} target in l region is

$$\mathcal{P}_{q_l,l} = \mathbf{a}_{q_l,l}^H(\psi_{q_l,l})\mathbf{\Theta}_l \mathbf{G}(\mathbf{W}_0 + \sum_{k=1}^{H} \mathbf{W}_k)\mathbf{G}^H \mathbf{\Theta}_l^H \mathbf{a}_{q_l,l}(\psi_{q_l,l}), \tag{6.9}$$

where $\mathbf{a}_{q_l,l}(\psi_{q_l,l}) = [1, e^{j\frac{2\pi d}{\lambda}\sin(\psi_{q_l,l})}, \cdots, e^{j\frac{2\pi d}{\lambda}(N-1)\sin(\psi_{q_l,l})}]^T$ is the steering vector to describe the position of targets relative to STAR-RIS [97]. $\psi_{q_l,l}$ denotes the angle of departure from STAR-RIS to each target. λ and d denote the wavelength and the space

between the elements, respectively.

6.2.3 Problem Formulation

The aim of this chapter is to maximize the sum rate of the primary users with guaranteeing the effective sensing power for each target is higher than the minimum requirement of sensing and each user's signal can be successfully decoded. Therefore, the original problem can be formulated as followed:

$$\max_{\mathbf{w},\boldsymbol{\theta}} \quad \sum_{k=1}^{K} \log_2(1+\gamma_k), \tag{6.10a}$$

$$\text{s. t.} \quad \gamma_{i\to 0} \geq \gamma_{0,min}, i = 0, \cdots, K, \tag{6.10b}$$

$$\gamma_k \geq \gamma_{k,min}, k = 1, \cdots, K, \tag{6.10c}$$

$$\sum_{i=0}^{K} \|\mathbf{w}_i\|^2 \leq P_{max}, \tag{6.10d}$$

$$\theta_{l,n} \in [0, 2\pi], n = 1, \cdots, N, l \in \{T, R\}, \tag{6.10e}$$

$$\beta_{n,T} + \beta_{n,R} = 1, n = 1, \cdots, N, \tag{6.10f}$$

$$\mathcal{P}_{q_l,l} \geq \mathcal{P}_{min}, \tag{6.10g}$$

where (6.10b) and (6.10c) are to satisfy the quality of service constraints of communications. (6.10d) is the power constraint, which limits the total power consumption of the transmitting signals. (6.10e) and (6.10f) are the constraints of the transmitting and reflecting elements. (6.10g) is the requirement of sensing accuracy. Specifically, in this chapter, $\mathcal{P}_{min}$ is the minimum beampattern gain for each target to achieve the sensing accuracy. In other words, if the beampattern gains of the targets are lower than this minimum requirement, the sensing function will fail in this network.

6.3 Proposed Method

In this section, a novel optimization method is designed to solve the original problem. Firstly, it is noted that (6.10) is not a convex optimization problem. The reasons are mainly in two folds. On the one hand, (6.10a)-(6.10c) contain the expressions with fractions, which

are not convex. On the other hand, the beamforming vectors and phase shift angles are coupled with each other. Even worse, the equality constraint (6.10e) is not affine, which makes this problem even more challenging to be solved. To effectively solve it, a novel optimization method is developed to transform the original problem to convex and the transformed problem can be solved by the existing optimization toolboxes, i.e., CVX. Firstly, to tackle the coupling issue of the original problem, the original problem is decoupled into two suboptimal problems. The suboptimal problems of beamforming and phase shifts are alternatively solved, respectively. In the beamforming suboptimal problem, SDR is first applied to transform the quadratic terms to affine. Then, several slack variables are introduced to move the SINR of each primary user out of the objective function and make the objective function concave. After the above transform, fractional programming are applied to transform the non-convex constraints to convex by introducing several auxiliary variables. Besides, fractional programming is then used to transform the constraints to convex. To tackle the rank constraints, the penalty-based SCA is used and transforms the problem to the one which is accepted by CVX. The similar method is also implemented in the phase shift optimization.

6.3.1 Beamforming Optimization

In the beamforming optimization problem, the phase shifts are considered as constant. Therefore, the original problem for beamforming suboptimal problem can be reformulated as followed:

$$\max_{\mathbf{w}} \quad \sum_{k=1}^{K} \log_2(1+\gamma_k), \tag{6.11a}$$

$$\text{s. t.} \quad \gamma_{i\to 0} \geq \gamma_{0,min}, i = 0, \cdots, K, \tag{6.11b}$$

$$\gamma_k \geq \gamma_{k,min}, k = 1, \cdots, K, \tag{6.11c}$$

$$\sum_{i=0}^{K} \|\mathbf{w}_i\|^2 \leq P_{max}, \tag{6.11d}$$

$$\mathcal{P}_{q_l,l} \geq \mathcal{P}_{min}. \tag{6.11e}$$

From the problem formulation, it is observed that all the objective function and constraints are not convex due to the existence of quadratic expressions in (6.11a)-(6.11e). In fact, SDR

is usually an effective tool to solve such a problem. Specifically, by defining $\mathbf{W} = \mathbf{w}\mathbf{w}^H$, $\mathbf{H} = \mathbf{h}\mathbf{h}^H$ and $\mathbf{A}_{q_l,l} = (\mathbf{a}_{q_l,l}^H \boldsymbol{\Theta}_l \mathbf{G})^H (\mathbf{a}_{q_l,l}^H \boldsymbol{\Theta}_l \mathbf{G})$, the original problem can be first rewritten in an SDR form as followed:

$$\max_{\mathbf{W}} \quad \sum_{k=1}^{K} \log_2(1+\gamma_k), \tag{6.12a}$$

$$\text{s. t.} \quad \gamma_{i\to 0} \geq \gamma_{0,min}, \ = 0, \cdots, K, \tag{6.12b}$$

$$\gamma_k \geq \gamma_{k,min}, k = 1, \cdots, K, \tag{6.12c}$$

$$\sum_{i=0}^{K} \text{Tr}(\mathbf{W}_i) \leq P_{max}, \tag{6.12d}$$

$$\mathcal{P}_{q_l,l} \geq \mathcal{P}_{min}, \tag{6.12e}$$

$$\mathbf{W}_i \succeq \mathbf{0}, i = 0 \cdots, K, \tag{6.12f}$$

$$\text{Rank}(\mathbf{W}_i) = 1, i = 0, \cdots K, \tag{6.12g}$$

where

$$\gamma_k = \frac{\text{Tr}(\mathbf{H}_k \mathbf{W}_k)}{\sum_{i=1, i\neq k}^{K} \text{Tr}(\mathbf{H}_k \mathbf{W}_i) + \sigma^2}, k, i = 1, 2, \tag{6.13}$$

and

$$\gamma_{i\to 0} = \frac{\text{Tr}(\mathbf{H}_i \mathbf{W}_0)}{\sum_{k=1}^{K} \text{Tr}(\mathbf{H}_i \mathbf{W}_k) + \sigma^2}, i = 0, \cdots, K, k = 1, \cdots K, \tag{6.14}$$

and

$$\mathcal{P}_{q_l,l} = \text{Tr}((\mathbf{W}_0 + \sum_{k=1}^{H} \mathbf{W}_k)\mathbf{A}_{q_l,l}). \tag{6.15}$$

From (6.12), it can be seen that (6.12b), (6.12d) and (6.16e)-(6.12f) can be directly transformed to convex. The next step is to transform the objective function to a concave one. Actually, it is noted that the expressions inside the logarithms in the objective function contain the fractional terms, which are not convex. To make the objective function concave, several slack variables are ingeniously introduced to replace the fractions in the logarithms and move them to the constraints. Specifically, by introducing several variables $\boldsymbol{\alpha}$, (6.12) can be equivalently transformed to

$$\max_{\mathbf{W}, \boldsymbol{\alpha}} \quad \sum_{k=1}^{K} \log_2(1+\alpha_k), \tag{6.16a}$$

$$\text{s. t.} \quad \frac{\text{Tr}(\mathbf{H}_i\mathbf{W}_0)}{\sum_{k=1}^{K}\text{Tr}(\mathbf{H}_i\mathbf{W}_k)+\sigma^2} \geq \gamma_{0,min}, i=0,\cdots,K, \tag{6.16b}$$

$$\frac{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)}{\sum_{i=1,i\neq k}^{K}\text{Tr}(\mathbf{H}_k\mathbf{W}_i)+\sigma^2} \geq \alpha_{k,min}, k=1,\cdots,K, \tag{6.16c}$$

$$\sum_{i=0}^{K}\text{Tr}(\mathbf{W}_i) \leq P_{max}, \tag{6.16d}$$

$$\text{Tr}((\mathbf{W}_0+\sum_{k=1}^{H}\mathbf{W}_k)\mathbf{A}_{q_l,l}) \geq \mathcal{P}_{min}, \tag{6.16e}$$

$$\alpha_k \geq \gamma_{k,min}, k=1,\cdots K, \tag{6.16f}$$

$$\mathbf{W}_i \succeq \mathbf{0}, i=0\cdots,K, \tag{6.16g}$$

$$\text{Rank}(\mathbf{W}_i)=1, i=0,\cdots K, \tag{6.16h}$$

Until now, (6.16a)-(6.16b) and (6.16d)-(6.16g) have been converted to convex or concave. The next step is to transform (6.16c) to convex. According to [68] and [82], fractional programming is applied. Specifically, by introducing real auxiliary $\boldsymbol{\varsigma}$, $\frac{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)}{\sum_{i=1,i\neq k}^{K}\text{Tr}(\mathbf{H}_k\mathbf{W}_i)+\sigma^2}$ is equivalent to the following form:

$$\max_{\boldsymbol{\varsigma}} \quad 2\varsigma_k\sqrt{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)}-\varsigma_k^2(\sum_{i=1,i\neq k}^{K}\text{Tr}(\mathbf{H}_k\mathbf{W}_i)+\sigma^2). \tag{6.17}$$

By finding the partial derivative of the above function with respect to $\boldsymbol{\varsigma}$, the optimal $\boldsymbol{\varsigma}$ can be expressed as

$$\varsigma_k=\frac{\sqrt{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)}}{\sum_{i=1,i\neq k}^{K}\text{Tr}(\mathbf{H}_k\mathbf{W}_i)+\sigma^2}, k=1,\cdots K. \tag{6.18}$$

If (6.18) is substituted to (6.17), the same expression with the one at the left side of the inequality in (6.16c) can be obtained.

It is noted that (6.17) is jointly convex with respect to $\mathbf{W}$ and $\boldsymbol{\alpha}$ since the expression is concave at the left side of $\geq$ and convex at the right hand side. Inspired by this idea, the non-convex (6.16c) can be transformed as followed:

$$2\varsigma_k\sqrt{\text{Tr}(\mathbf{H}_k\mathbf{W}_k)}-\varsigma_k^2(\sum_{i=1,i\neq k}^{K}\text{Tr}(\mathbf{H}_k\mathbf{W}_i)+\sigma^2) \geq \alpha_k, \tag{6.19}$$

Therefore, $\mathbf{W}$ can be updated by optimization tools considering $\boldsymbol{\varsigma}$ is fixed and $\boldsymbol{\varsigma}$ is then updated by (6.18). After these, the beamforming optimization problem is finally converted to

convex except (6.16h), which can be expressed as

$$\max_{\mathbf{W},\boldsymbol{\alpha}} \quad \sum_{k=1}^{K} \log_2(1+\alpha_k), \tag{6.20a}$$

$$\text{s. t.} \quad (6.16b), (6.16d)-(6.16g), (6.19), \tag{6.20b}$$

$$\text{Rank}(\mathbf{W}_i) = 1, i = 0, \cdots, K. \tag{6.20c}$$

However, it is noticed that (6.20c) is still not convex. More actions are required to be taken to transform the constraints to convex ones. The novel penalty method is employed to transform the rank one constraints to another convex means. Specifically, a positive semidefinite matrix $\mathbf{W}_k$ has the following relationship as followed:

$$\text{Tr}(\mathbf{W}_k) \geq \|\mathbf{W}_k\|_F, \tag{6.21}$$

where $\|\mathbf{W}_k\|_F$ is the Frobenius norm of $\mathbf{W}_k$.The equality holds if and only if $\mathbf{W}_k$ is rank-one. Based on the above relationship, we can add (6.21) to the objective function as followed:

$$\min_{\mathbf{W},\boldsymbol{\alpha}} \quad -\sum_{k=1}^{K} \log_2(1+\alpha_k) + \rho \sum_{i=0}^{K} f_i(\mathbf{W}_i), \tag{6.22a}$$

$$\text{s. t.} \quad (6.20b), \tag{6.22b}$$

where

$$f_i(\mathbf{W}_i) = \text{Tr}(\mathbf{W}_i) - \|\mathbf{W}_i\|_F, \tag{6.23}$$

and it is also known as the penalty term. However, it is noted that (6.22a) is not convex and it cannot be solved by the existing optimization tools. SCA can be employed to transform this to convex, which can be solved by CVX. Denote $D(\mathbf{W}_i) = \|\mathbf{W}_i\|_F$. By finding the first order Taylor expansion of $D(\mathbf{W}_i)$, (6.22a) can be transformed to

$$\min_{\mathbf{W},\boldsymbol{\alpha}} \quad -\sum_{k=1}^{K} \log_2(1+\alpha_k) + \rho \sum_{i=0}^{K} (\text{Tr}(\mathbf{W}_i) - D(\mathbf{W}_i^{(t)}) - \text{Tr}(\nabla D(\mathbf{W}_i^{(t)})(\mathbf{W}_i - \mathbf{W}_i^{(t)}))), \tag{6.24a}$$

$$\text{s. t.} \quad (6.20b), \tag{6.24b}$$

where

$$\nabla D(\mathbf{W}_i) = \frac{\mathbf{W}_i}{\sqrt{\text{Tr}(\mathbf{W}_i^H \mathbf{W}_i)}}, \tag{6.25}$$

is the first order derivative of $D(\mathbf{W}_i)$.

Therefore, (6.24) is finally transformed to a convex optimization problem and the suboptimal solutions of $\mathbf{W}$ can be found by iteratively solving it.

6.3.2 Phase Shift Optimization

After the beamforming has been optimized, it can be considered as the constant value and the phase shift optimization problem can be solved by the similar method applied in the beamforming optimization. Firstly, the phase shift optimization problem can be rewritten in another manner, let $\mathbf{v}_l = (\beta_{1,l} e^{j\theta_{l,1}}, \cdots, \beta_{n,l} e^{j\theta_{l,N}})^H$ and find the beamforming vectors $\mathbf{w}$ by performing EVD on the optimized $\mathbf{W}$. The phase shift optimization problem can be formulated as:

$$\max_{\boldsymbol{\theta}} \quad \sum_{k=1}^{K} \log_2(1+\gamma_k), \tag{6.26a}$$

$$\text{s. t.} \quad \gamma_{i\to 0} \geq \gamma_{0,min}, i = 0, \cdots, K, \tag{6.26b}$$

$$\gamma_k \geq \gamma_{k,min}, k = 1, \cdots, K, \tag{6.26c}$$

$$\|\mathbf{v}_{l,n}\| = \sqrt{\beta_{n,l}}, n = 1, \cdots, N, l \in \{T, R\}, \tag{6.26d}$$

$$\beta_{n,T} + \beta_{n,R} = 1, n = 1, \cdots, N, \tag{6.26e}$$

$$\mathcal{P}_{q_l,l} \geq \mathcal{P}_{min}, \tag{6.26f}$$

where

$$\gamma_{0\to 0} = \frac{\|\mathbf{v}_T^H \text{diag}(\mathbf{h}_{0,T}^H)\mathbf{G}\mathbf{w}_0\|^2}{\sum_{k=1}^{K} \|\mathbf{v}_T^H \text{diag}(\mathbf{h}_{0,T}^H)\mathbf{G}\mathbf{w}_k\|^2 + \sigma^2}, \tag{6.27}$$

$$\gamma_{i\to 0} = \frac{\|\mathbf{v}_R^H \text{diag}(\mathbf{h}_{i,R}^H)\mathbf{G}\mathbf{w}_0\|^2}{\sum_{k=1}^{K} \|\mathbf{v}_R^H \text{diag}(\mathbf{h}_{i,R}^H)\mathbf{G}\mathbf{w}_k\|^2 + \sigma^2}, i = 1, \cdots K, \tag{6.28}$$

and

$$\gamma_k = \frac{\|\mathbf{v}_R^H \text{diag}(\mathbf{h}_{k,R}^H)\mathbf{G}\mathbf{w}_k\|^2}{\sum_{i=1, i\neq k}^{K} \|\mathbf{v}_R^H \text{diag}(\mathbf{h}_{k,R}^H)\mathbf{G}\mathbf{w}_i\|^2 + \sigma^2}, \tag{6.29}$$

and

$$\mathcal{P}_{q_l,l} = \mathbf{v}_l^H \text{diag}(\mathbf{a}_{q_l,l}^H)\mathbf{G}(\mathbf{W}_0 + \sum_{k=1}^{K} \mathbf{W}_k)\mathbf{G}^H \text{diag}(\mathbf{a}_{q_l,l})\mathbf{v}_l. \tag{6.30}$$

Let $\mathbf{C}_{q_l,l} = \text{diag}(\mathbf{a}_{q_l,l}^H)\mathbf{G}(\mathbf{W}_0 + \sum_{k=1}^K \mathbf{W}_k)\mathbf{G}^H\text{diag}(\mathbf{a}_{q_l,l})$ and $\mathbf{b}_{0,i} = \text{diag}(\mathbf{h}_{0,T}^H)\mathbf{G}\mathbf{w}_i, i = 0, \cdots, K$, $\mathbf{b}_{k,i} = \text{diag}(\mathbf{h}_{k,R}^H)\mathbf{G}\mathbf{w}_i$, $k = 1, \cdots, K$, $i = 0, \cdots, K$ and $\mathbf{B} = \mathbf{b}\mathbf{b}^H$. Meanwhile, similar to $\boldsymbol{\alpha}$, a series of slack variables $\boldsymbol{\delta}$ is introduced to move the fractional SINR terms out of the logarithms and make the objective function concave. By using SDR, the problem can be equivalently transformed to

$$\max_{\mathbf{V},\boldsymbol{\delta}} \quad \sum_{k=1}^K \log_2(1+\delta_k), \tag{6.31a}$$

$$\text{s. t.} \quad \frac{\text{Tr}(\mathbf{V}_T\mathbf{B}_{0,0})}{\sum_{k=1}^K \text{Tr}(\mathbf{V}_T\mathbf{B}_{0,k}) + \sigma^2} \geq \gamma_{0,min}, k = 1, \cdots K, \tag{6.31b}$$

$$\frac{\text{Tr}(\mathbf{V}_R\mathbf{B}_{i,0})}{\sum_{k=1}^K \text{Tr}(\mathbf{V}_R\mathbf{B}_{i,k}) + \sigma^2} \geq \gamma_{0,min}, i, k = 1, \cdots, K, \tag{6.31c}$$

$$\frac{\text{Tr}(\mathbf{V}_R\mathbf{B}_{k,k})}{\sum_{i=1,i\neq k}^K \text{Tr}(\mathbf{V}_R\mathbf{B}_{k,i}) + \sigma^2} \geq \delta_k, k = 1, \cdots, K, \tag{6.31d}$$

$$\text{Tr}(\mathbf{V}_l\mathbf{C}_{q_l,l}) \geq \mathcal{P}_{min}, \tag{6.31e}$$

$$\delta_k \geq \gamma_{k,min}, k = 1, \cdots K, \tag{6.31f}$$

$$\mathbf{V}_{l,n} = \beta_{n,l}, \tag{6.31g}$$

$$\beta_{n,T} + \beta_{n,R} = 1, \tag{6.31h}$$

$$\mathbf{V}_l \succeq \mathbf{0}, \tag{6.31i}$$

$$\text{Rank}(\mathbf{V}_l) = 1, \tag{6.31j}$$

where $\mathbf{V}_{l,n}$ is the n^{th} diagonal element of $\mathbf{V}_l$. It can be observed that (6.31a)-(6.31c) and (6.31e)-(6.31i) are convex or concave. However, the QoS constraints for primary users (6.31d) and rank one constraints (6.31j) are still not convex. Similar to the purpose of introducing ς in the beamforming optimization, by introducing several auxiliary variables $\boldsymbol{\mu}$, (6.31d) can be converted to fractional programming forms as followed:

$$2\mu_k\sqrt{\text{Tr}(\mathbf{V}_R\mathbf{B}_{k,k})} - \mu_k^2(\sum_{i=1,i\neq k}^K \text{Tr}(\mathbf{V}_R\mathbf{B}_{k,i}) + \sigma^2) \geq \delta_k, \tag{6.32}$$

and μ_k is updated by

$$\mu_k = \frac{\sqrt{\text{Tr}(\mathbf{V}_R\mathbf{B}_{k,k})}}{\sum_{i=1,i\neq k}^K \text{Tr}(\mathbf{V}_R\mathbf{B}_{k,i}) + \sigma^2}, \tag{6.33}$$

and $\mathbf{V}$ and $\boldsymbol{\mu}$ can be alternatively updated in each iteration. After such a transform, all the constraints are convex except (6.31j). The penalty method can be still employed to tackle this

issue. Similar to the beamforming part, a positive semidefinite matrix $\mathbf{V}_l$ has the following relationship as followed:

$$\text{Tr}(\mathbf{V}_l) \geq \|\mathbf{V}_l\|_F, \tag{6.34}$$

where $\|\mathbf{V}_l\|_F$ is the Frobenius norm of $\mathbf{V}_l$.The equality holds if and only if $\mathbf{V}_l$ is rank-one. Based on the above relationship, we can add (6.34) to the objective function as followed:

$$\min_{\mathbf{V},\boldsymbol{\delta}} \quad -\sum_{k=1}^{K} \log_2(1+\delta_k) + \rho \sum_{l\in\{T,R\}} f_l(\mathbf{V}_l), \tag{6.35a}$$

$$\text{s. t.} \quad (6.31\text{b}), (6.31\text{c}), (6.31\text{e}) - (6.31\text{h}), (6.33), \tag{6.35b}$$

where

$$f_l(\mathbf{V}_l) = \text{Tr}(\mathbf{V}_l) - \|\mathbf{V}_l\|_F. \tag{6.36}$$

However, it is noted that (6.35a) is not convex and it cannot be solved by the existing optimization tools. SCA can be still employed to transform this to convex, which can be solved by CVX,

$$\min_{\mathbf{V},\boldsymbol{\delta}} \quad -\sum_{k=1}^{K} \log_2(1+\delta_k) + \rho \sum_{l\in\{T,R\}} (\text{Tr}(\mathbf{V}_l) - D(\mathbf{V}_l^{(t)}) - \text{Tr}(\nabla D(\mathbf{V}_l^{(t)})(\mathbf{V}_l - \mathbf{V}_l^{(t)})), \tag{6.37a}$$

$$\text{s. t.} \quad (6.31\text{b}), (6.31\text{c}), (6.31\text{e}) - (6.31\text{h}), (6.33), \tag{6.37b}$$

where

$$\nabla D(\mathbf{V}_l) = \frac{\mathbf{V}_l}{\sqrt{\text{Tr}(\mathbf{V}_l^H \mathbf{V}_l)}}, \tag{6.38}$$

is the derivative of $D(\mathbf{V}_l)$.

6.3.3 Algorithm Design

Initial Point Algorithm Design

Before alternatively solving the beamforming and phase shift optimization problems, the initial point for beamforming $\mathbf{W}^{(0)}$ and phase shifts $\boldsymbol{\theta}^{(0)}$ are supposed to be selected since iterative algorithm is applied in the beamforming optimization. Actually, in this chapter,

the initial point can still be found by using optimization tools, i.e., CVX. Firstly, random phase shift angles for $\theta_{n,T}$ and $\theta_{n,R}$ can be generated from 0 to 2π. Similarly, amplitude coefficients of transmitting and reflecting elements can also be randomly generated from 0 to 1 and satisfy $\beta_{n,T} + \beta_{n,R} = 1$. After the phase shifts have been fixed, CVX is utilized to find the initial beamforming. It is noted that the original beamforming optimization problem (6.10) is non-convex since the objective function is non-convex. Therefore, a new convex optimization problem with different objective function but same constraints with (6.10) can be constructed as followed:

$$\max_{\mathbf{w}} \quad f_0, \tag{6.39a}$$

$$\text{s.t.} \quad (6.11\text{b}) - (6.11\text{e}), \tag{6.39b}$$

where the objective function can be defined by $f_0 = \sum_{i=0}^{K} \log_2(1 + \frac{\text{Tr}(\mathbf{H}_i \mathbf{W}_i)}{\sigma^2}) - \sum_{i=0}^{K} \text{Tr}(\mathbf{W}_i)$. On the one hand, (6.39) is a standard convex optimization problem and it can be directly solved by CVX. On the other hand, according to the proof logic in Appendix A of [84], all the solutions of (6.39) can be guaranteed to be rank-one and satisfy the constraints in original optimization problem.

Beamforming Algorithm Design

After the initial point of beamforming is obtained, the initial auxiliary variables $\boldsymbol{\varsigma}^{(0)}$ can be calculated by (6.18). Then, (6.24) is iteratively solved and $\mathbf{W}$ and $\boldsymbol{\varsigma}$ are iteratively updated until the algorithm converges. Meanwhile, the penalty factor should be increased by $\rho_1 = \eta\rho_1$ in each iteration, where η is a real value and $\eta > 1$. The pseudo-code of the beamforming algorithm is shown in **Algorithm 5**.

Phase Shift Algorithm Design

After the suboptimal $\mathbf{W}$ is obtained. The beamforming vector $\mathbf{w}$ can be found by performing EVD on $\mathbf{W}$. The value of $\boldsymbol{\varsigma}$ is directly assigned to the initial auxiliary variable $\boldsymbol{\mu}^{(0)}$ and $\mathbf{V}_l^{(0)} = \mathbf{v}_l^{(0)} \mathbf{v}_l^{(0),H}$ when the beamforming algorithm converges. Then, (6.37) is iteratively solved by CVX. It is noted that during the iterations, the penalty factor should be updated by $\rho_2 = \eta\rho_2$, where η is a real value and $\eta > 1$. The pseudo-code of the phase shift algorithm is

Algorithm 5 Beamforming Optimization Algorithm

1: **Initialize:** Given the initial $\mathbf{W}^{(0)}$ and $\boldsymbol{\theta}_l^{(0)}$, ρ_1 is initialized as ρ_0, $\boldsymbol{\varsigma}^{(0)}$ is calculated by (6.18).
2: **repeat**
3: Iteratively solve (6.24) and update $\mathbf{W}^{(t)} = \mathbf{W}^{(t-1)}$ and $\boldsymbol{\varsigma}$ by (6.18). $\rho_1^{(t+1)} = \eta\rho_1^{(t)}$.
4: **until** convergence.
5: **Output:** The optimal $\mathbf{W}$ and $\boldsymbol{\varsigma}$.

Algorithm 6 Phase shift Optimization Algorithm

1: **Initialize:** Given the optimized $\mathbf{w}$ by EVD of $\mathbf{W}$ and initial and $\boldsymbol{\mu}^{(0)} = \boldsymbol{\varsigma}$, ρ_2 is initialized as ρ_0.
2: **repeat**
3: Update $\mathbf{V}_l$ by $\mathbf{V}_l^{(t)} = \mathbf{V}_l^{(t-1)}$.
4: Iteratively solve (6.31) and update $\boldsymbol{\mu}$ by (6.33). ρ is updated by $\rho_2^{(t+1)} = \eta\rho_2^{(t)}$.
5: **until** convergence.
6: **Output:** The optimal $\mathbf{V}$ and $\boldsymbol{\mu}$.

shown in **Algorithm 6**.

Alternating Algorithm Design

In this chapter, alternating algorithm contains two sub-algorithms, **Algorithm 5** and **Algorithm** 6. When two algorithms are each completed sequentially one time, one outer iteration is completed. Specifically, by alternatively solving **Algorithm** 5 and **Algorithm 6**, i.e., the optimal $\mathbf{V}_l$ can be decomposed to $\boldsymbol{\theta}_l$ by EVD and $\boldsymbol{\theta}_l$ is the initial phase shifts of the beamforming optimization in next outer iteration, the suboptimal solutions for the beamforming and phase shifts can be obtained until the convergence of the outer iterations.

Parameter	Description	Value
κ	Rician factor	2
ω	Path loss exponent	2
$d_{\text{BS-STAR-RIS}}$	Distance between BS and STAR-RIS	40 m
$d_{\text{STAR-RIS-pu}}$	Distance between STAR-RIS and each primary user	20-30 m
$d_{\text{STAR-RIS-su}}$	Distance between STAR-RIS and the secondary user	30-40 m
$d_{\text{STAR-RIS-target}}$	Distance between STAR-RIS and each target	20-30 m
K	Number of users	4
$\frac{d}{\lambda}$	Ratio of two adjacent antennas and wavelength	0.3
B	Total bandwidth	100 MHz
σ^2	Noise power spectrum	-140 dBm/Hz
$\mathcal{P}_{min}$	Minimum requirement of beampattern gain	1×10^{-6}
$\gamma_{k,min}$	Minimum QoS requirement of each primary user	0 dB

Table 6.2. Parameter Settings for Simulation

6.3.4 Complexity Analysis

According to [85], the computational complexity is $\mathcal{O}(LRP)$ when calculating the multiplication of $L \times R$ and $R \times P$ matrices. An optimization problem is solved by transforming the problem with constraints to the one without constraints, which is to put the constraints to the objective function. Thus, $2K + 1$ constraints of QoS constraints and $Q_T + Q_R$ sensing requirement constraints should be taken into consideration. Since the beamforming variable is an $M \times M$ matrix, the complexity of the beamforming optimization problem is $\mathcal{O}(\log(\frac{1}{\epsilon_1})(2KM^3 + M^3 + Q_T M^3 + Q_R M^3))$ and it is equivalent to $\mathcal{O}(\log(\frac{1}{\epsilon_1})(KM^3 + Q_T M^3 + Q_R M^3))$ in the worst case, where ϵ_1 is convergence threshold of the beamforming optimization problem. Similarly, the complexity of the phase shift optimization is $\mathcal{O}(\log(\frac{1}{\epsilon_2})(KN^3 + Q_T N^3 + Q_R N^3))$ in the worst case, where ϵ_2 is convergence threshold of phase shift optimization.

6.4 Simulation Results

In this section, the simulation results are presented to demonstrate the performance of the proposed scheme and convergence of the proposed algorithms. In this chapter, Rician fading is applied to all the channels in this network. Specifically, the Rician fading channel $\mathbf{F}$ can be modelled as followed:

$$\mathbf{F} = \sqrt{\frac{\kappa}{1+\kappa}}\mathbf{F}_{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}}\mathbf{F}_{\text{nLoS}}, \tag{6.40}$$

where κ represents the Rician factor and $\mathbf{F}_{\text{LoS}}$ and $\mathbf{F}_{\text{nLoS}}$ denote the line of sight and non line of sight components, respectively. The fixed parameters are shown in the Table 6.2. It is noted that STAR-RIS communications only with optimized phase shifts, RIS assisted NOMA with optimized phase shifts, STAR-RIS assisted NOMA with random phase shifts and STAR-RIS OMA with optimized phase shifts are shown as the benchmark schemes. The detailed analysis of the simulation results is demonstrated as followed:

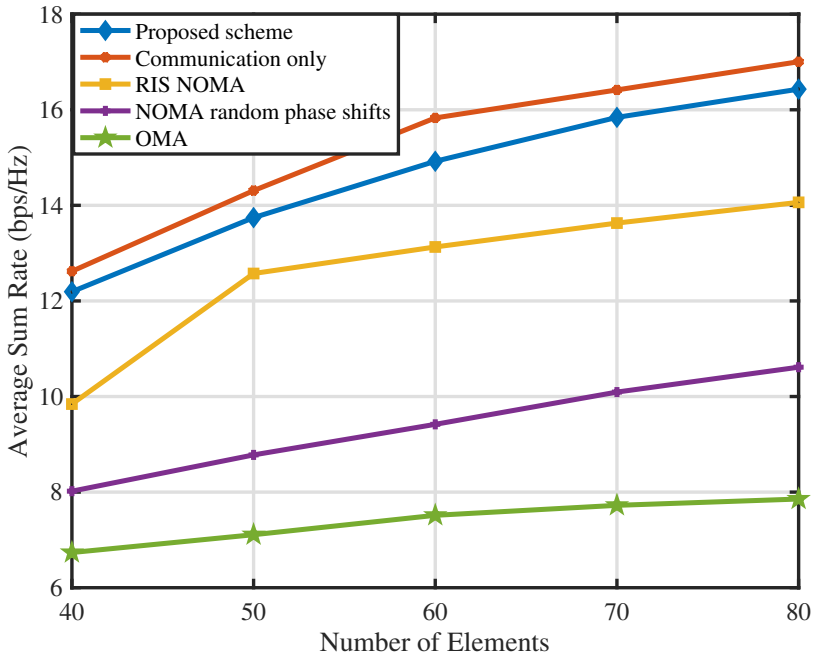

Fig. 6.2. Average sum rate versus the number of elements with $P_{max} = 34$ dBm, $\gamma_{0,min} = -10$ dB and $M = 12$

6.4.1 Impact of the Number of Elements

In Fig. 6.2, the relationship between the average sum rate and number of elements is presented. It is noted that the RIS assisted NOMA scheme is that there are one transmitting-only RIS and one reflecting-only RIS adjacent to the original position of the STAR-RIS. The distance between two RISs are neglected. In this way, the full space communication coverage can also be realized in the network. To guarantee the same number of elements, each RIS is equipped with N elements. Besides, to guarantee the fairness of the amplitude coefficient allocation between STAR-RIS and two RISs. The amplitude coefficients corresponding two elements of two RISs should be fixed and satisfy $\beta_{n,T} + \beta_{n,R} = 1$. From the figure, it can be seen that for all 5 schemes, the average sum rate can increase as the number of elements grows. For example, when the number of elements grows from 40 to 80, the average sum rate increases from 12.1929 bps/Hz to 16.4312 bps/Hz. It is because a higher number of elements can have a higher transmitting and reflecting gains. In particular, it can be seen that the proposed scheme can outperform that of RIS assisted NOMA. It is because from the side of RIS, only phase shifts participate in the optimization process and the amplitude coefficient for each element is fixed in parameter setting and cannot be tuned. However, from the side of STAR-RIS, both the phase shifts and amplitude coefficients participate

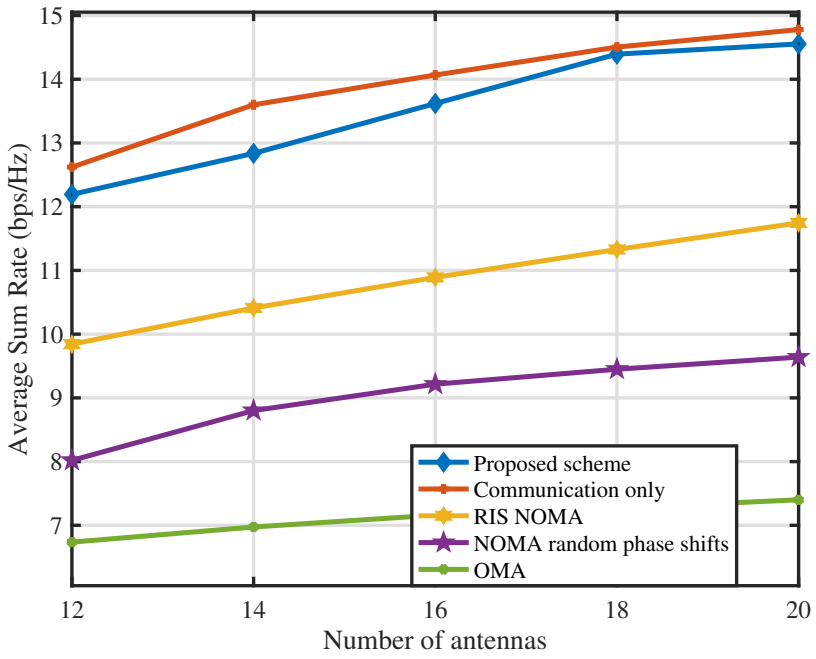

Fig. 6.3. Average sum rate versus the number of antennas with $P_{max} = 34$ dBm, $\gamma_{0,min} = -10$ dB and $N = 20$

in optimizing the sum rate of the system. Therefore, STAR-RIS provides higher degree of freedom (DoF) than two RISs to mitigate interference of the users and beampattern gains of the sensing targets. In other words, two RISs assisted NOMA is a special case of one STAR-RIS assisted NOMA one. Therefore, the performance under STAR-RIS should be always greater than or equal to the performance under RIS. It can also been seen that the proposed scheme can always perform better than STAR-RIS assisted NOMA with random phase shifts. No matter the number of elements, both STAR-RIS assisted NOMA and RIS assisted NOMA with optimized phase shifts schemes are clearly inferior to STAR-RIS with random phase shifts, proving the necessity of phase shift optimization in improving the performance of RIS/STAR-RIS networks.

6.4.2 Impact of the Number of Antennas

In Fig. 6.3, the relationship between the average sum rate and number of antennas is shown. From the figure, it can be seen that the average sum rate can also increase as the number of antennas grows. It is because higher number of antennas can produce higher beamforming gains, it directly leads to the increase of the sum rate, However, it is difficult that the sum rate can continuously increase since the power budget is limited. It is more likely that the sum rate can increases rapidly when the number of the antennas is small. However, as the

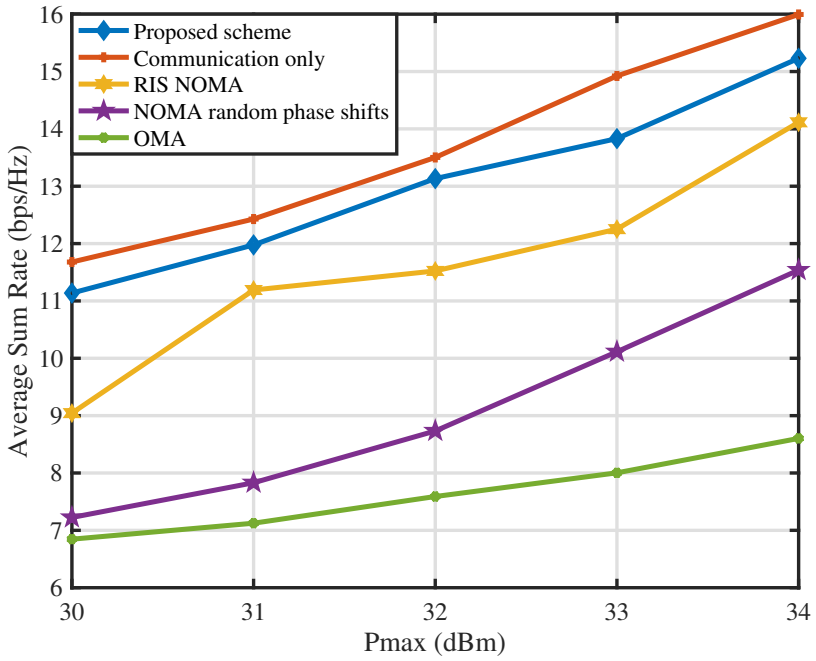

Fig. 6.4. Average sum rate versus the maximum power budget with $M = 16$, $N = 30$ and $\gamma_{0,min} - 10$ dB

number of antennas further grows, the effect of beamforming gains will become miner and miner. When the number of antennas is large enough, there will be almost no increase if further deploying more antennas. Besides, it can be observed that for all NOMA cases in the figure, the increases in sum rate are all more significant than the increase in OMA. For example, there are 2.3613 bps/Hz, 2.1584 bps/Hz, 1.9040 bps/Hz, and 1.62 bps/Hz increases in STAR-RIS assisted communications only NOMA, proposed scheme, RIS assisted ISAC NOMA with optimized phase shifts and STAR-RIS assisted ISAC NOMA with random phase shifts while only 0.6646 bps/Hz increase occurs in ISAC OMA. It demonstrate the superior performance of NOMA than OMA since NOMA can share the bandwidth resource and the spectrum efficiency can be significantly improved.

6.4.3 Impact of the Maximum Power Budget

In Fig. 6.4, the relationship between the average sum rate and maximum power budget is shown. From the figure, it can be seen that the average sum rate can increase as the P_{max} increases, it is because that more power can lead to higher transmitting power of signals. On the one hand, it is more easier to satisfy the requirement beampattern gains of the sensing targets. On the other hand, more transmitting power can also provide more DoF in the domain of the beamforming. a wider domain is expected to produce a higher

sum rate of the network. From the figure, it can be also observed that the sum rate of communications only NOMA case is always higher than the proposed scheme. It is because in ISAC NOMA networks, the beamforming and phase shifts should participate in both improving the sum rate and guaranteeing the sensing requirements. It will greatly limit the effect of beamforming and phase shifts on improving sum rate. From the perspective of mathematics, the number of constraints of the optimization problem of the communication only STAR-RIS assisted NOMA case is lower than that of the STAR-RIS NOMA ISAC case. Therefore, it is more likely that an optimal value with less constraints outperform one with additional constraints. One more interesting insight can be provided. It is noted that all the cases perform a exponent increase after P_{max} grows from 32 dBm. But this trend is different from Fig. 4.8 in which EE was studied in a STAR-RIS NOMA network. It is because in this chapter, the budget power is fully used to obtain the best performance of sum rate. Therefore, the optimal sum rate will be infinite when the transmitting power grows to infinite. However, in EE study, to achieve the optimal tradeoff between sum rate and power consumption, not all the power budget is needed to be consumed. When power budget grows from a low value, EE is likely to increase since higher DoF mentioned above is obtained. However, as the further increase of the power budget, the effect of DoF becomes miner and miner to improve the EE. Therefore, in EE study, the objective function is more likely to go saturation instead of limitless increasing as P_{max} grows.

6.4.4 Impact of the Number of Targets

In Fig. 6.5, the relationship between the average sum rate and number of targets is shown. From the figure, it can be seen that the sum rate decreases as the number of targets grows. It is because more number of targets greatly restricts the effect of beamforming and phase shifts on sum rate maximization. In other words, more targets mean more constraints in the formulated problem. Therefore, domain of the beamforming and phase shifts may be contracted and the optimal sum rate of higher targets is likely to be less than that of lower number of targets. Meanwhile, it is noted that the number of targets has a greater impact on NOMA networks. However, it is not as significant for OMA networks. For example, in Fig. 6.5, the average sum rate decreases from 18.7714 bps/Hz to 15.5972 bps/Hz for proposed scheme while only 0.0839 bps/Hz is lost in STAR-RIS assisted OMA scheme.

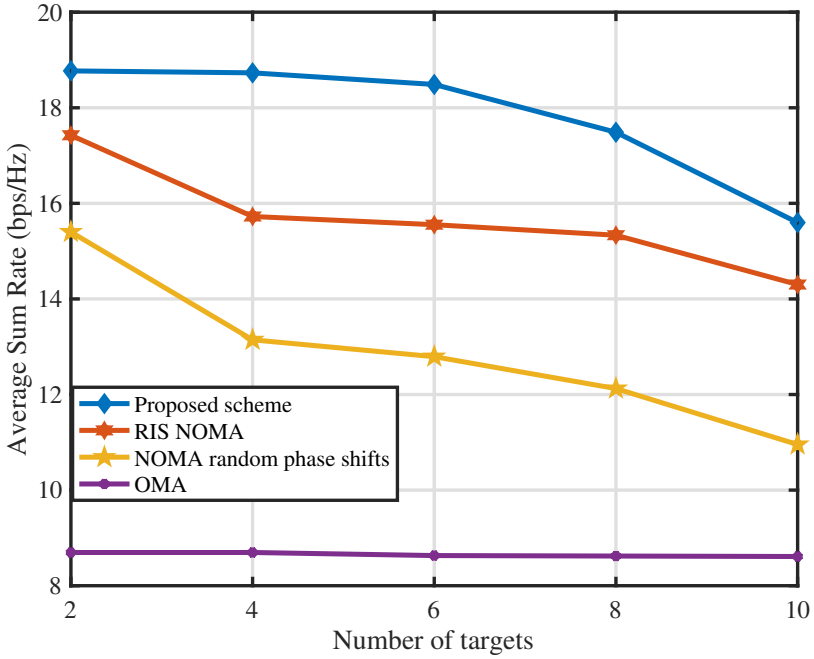

Fig. 6.5. Average sum rate versus the number of targets $M = 20$, $N = 40$, $P_{max} = 34$ dBm and $\gamma_{0,min} = -10$ dB

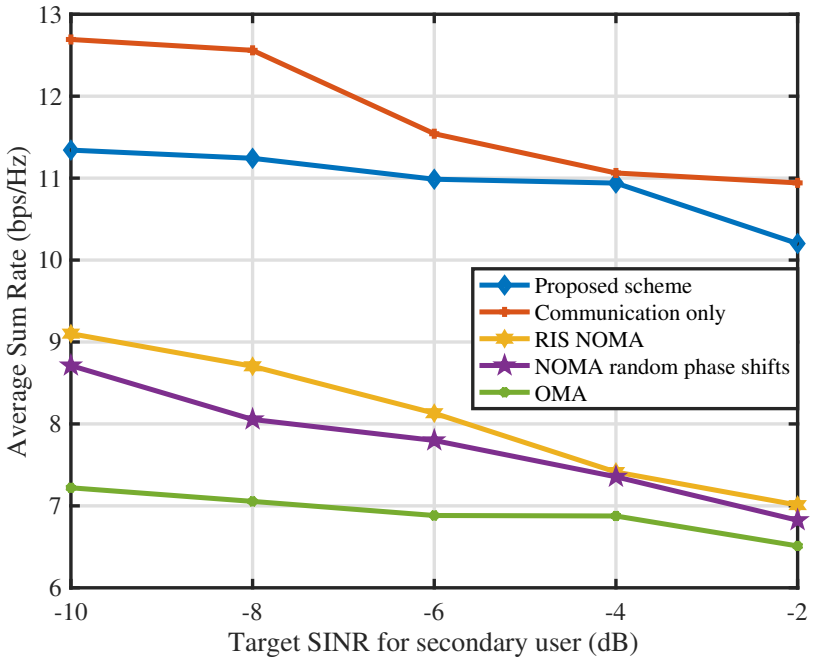

Fig. 6.6. Average sum rate versus the target SINR of the secondary user with $M = 12$, $N = 40$, $P_{max} = 30$ dBm

6.4.5 Impact of the Target SINR of Secondary User

In Fig. 6.6, the relationship between the average sum rate and target SINR of secondary user is shown. From the figure, it can be seen that the average sum rate decreases as the target SINR of secondary user grows. It is because the higher SINR make the domain of the optimization problem more strict, the optimal sum rate will be lower than that with the higher

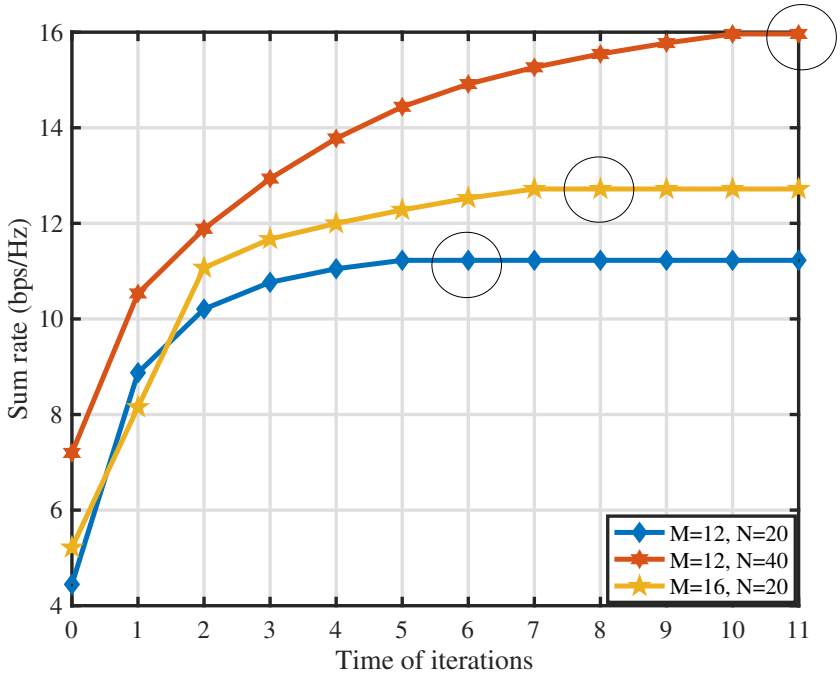

Fig. 6.7. Convergence of proposed algorithm with $P_{max} = 34$ dBm, $\gamma_{0,min} = -10$ dB.

SINR.

6.4.6 Convergence of the Proposed Algorithms

In Fig. 6.7, the convergence of the proposed alternating algorithm is demonstrated. Three types of parameter settings are compared in this subsection. From the figure, it can be seen that the algorithm can converge for all these three cases. Besides, a new insight about the time of iterations can be observed in this chapter. It is noted that with the same number of elements, the algorithm with a higher number of antennas usually needs more time of iterations than that with lower number of antennas. The similar phenomenon also occurs in two cases with the same number of antennas but different number of elements. It is because the algorithm with higher number of antennas or elements produce higher complexity of the algorithm, making it harder to converge.

6.5 Conclusions of the Chapter

In this chapter, we have studied a sum rate maximization problem in a STAR-RIS assisted ISAC NOMA network. To efficiently solve this problem, the original problem with respect to beamforming and phase shifts has been first decoupled to two sub-problems. In these

two problems, both SDR and fractional programming have been leveraged to transform non-convex sub-problems to SDR ones. However, the rank constraints were inevitably introduced to the transformed SDR problems. To effectively tackle this issue, a novel penalty-based method was employed. By increasing the penalty factors in each iteration, the algorithm can converge to a suboptimal value. After this, by alternatively solving the beamforming and phase shift problems, the suboptimal solutions of the original problem were obtained. In simulation, STAR-RIS communications only scheme, RIS assisted ISAC NOMA, STAR-RIS assisted NOMA with random phase shifts and STAR-RIS assisted ISAC OMA have been proposed as the benchmarks. Through numerical analysis, it was found that the performance of the proposed scheme outperforms the other ones except communications only case. Though the proposed scheme cannot perform better than the communication only NOMA one, considering the important role of sensing in further 6G networks, it is worthwhile sacrificing a certain sum rate improvement in order to achieve the purpose of both sensing and communications.

Chapter 7

Conclusions and Future Directions

7.1 Conclusions

The thesis has comprehensively demonstrated the benefits of a NOMA based RIS assisted multi-user network, STAR-RIS assisted multi-cluster network, RIS assisted CoMP network and STAR-RIS assisted ISAC network. In this thesis, the performance of networks is evaluated by energy efficiency and sum rate. Energy efficiency maximization can provide the best tradeoff between sum rate and power consumption. Therefore, this strategy is more likely to be adopted by telecommunication companies. Sum rate maximization is aimed to improve the communication quality of the users regardless of the power consumption of the systems. Therefore, this strategy is usually more friendly to consumers.

All the chapters of the thesis have fully illustrated the extraordinary performance of NOMA compared to OMA since all the users can share the same bandwidth resources, which improves the spectrum efficiency. Besides, from chapter 3 to chapter 6, it can be also observed that the dynamic adjustment of RIS/STAR-RIS phase shifts plays the important role on improving the performance of the networks. The numerical results in the thesis have demonstrated the optimal EE/sum rate with optimized phase shifts can always outperform those with random phase shifts. More interestingly, the STAR-RIS networks always outperform the two independent RISs networks both in EE and sum rate among the chapters, showing that the STAR-RIS can have a higher DoF than two RISs. This DoF reflects on the coordination between T elements and R elements. In STAR-RIS, both the phase shift angles and amplitude coefficients can be smartly tuned. However, in two RISs networks, amplitude coefficients of the two RISs are independent and they cannot be coordinated as the whole system. As the development of manufacturing process, both RIS and STAR-RIS are expected

to be widely deployed and become the indispensable techniques in future wireless networks. From the perspective of mathematics, the thesis has demonstrated that convex optimization can effectively be implemented to solve communication problems and provide solutions on the designs of beamforming and phase shifts. Relaxation, fractional programming, SDR and manifold method have been applied on the beamforming and phase shift optimization problems. Specifically, the non-convex original problems have been decoupled into two suboptimal problems and they have been alternatively solved. Then, the problems have transformed to convex ones by relaxation, fractional programming and SDR. Especially, mathematical proof, SROCR, and SCA-based penalty method have been adopted to tackle the rank-one issues introduced by SDR. In particular, a manifold method can be applied on the designs of phase shifts in a lower computational complexity manner. The proposed mathematical methods in this thesis are also universal on other beamforming and phase shift designs in RIS/STAR-RIS NOMA networks. With the further progress on optimization theory and algorithms, it is expected that more novel methods and high-efficiency algorithms can be developed to provide new solutions on the designs of communication networks.

7.2 Future Directions

7.2.1 RIS Assisted Terahertz Communications

Terahertz (THz) band will be the main transmission medium of 6G networks to satisfy the requirements of 6G such as super faster data speed, ultra lower latency and more massive connections of mobile devices [98]. As the frequency of the bandwidth is further expanded in THz, Terahertz NOMA can have a higher spectrum efficiency. The deployment of RIS is also an effective strategy in THz NOMA networks. According to [18], In cluster-based THz NOMA scenario, which the direct links between BS and users are highly correlated, there is no enough DoFs for BS to provide user clustering with beamforming. To solve it, the deployment of centralized RIS may produce additional DoFs and provide beamforming for the user clustering. In fact, convex optimization theory and the proposed mathematical tools in this thesis can help to tackle the optimization problems in THz NOMA networks. For example, the authors in [99] proposed capacity region and deployment strategy in

RIS assisted THz NOMA networks, in which SDR was applied to find the upper bound of the proposed problem. In fact, SDR by SROCR in **Chapter 4** and penalty based SDR in **Chapter 6** can also be employed to find the locally optimal solutions of the above problem. Meanwhile, it should also be noticed that some challenges about the implementation of THz into the wireless communications are still inevitable. For example, more power consumption and investments of hardware for Terahertz communications are required, which is not economical to telecommunication companies. Although there are several technical shortages about THz, it is still promising that the introduction of RIS and NOMA can effectively alleviate the aforementioned issues [100].

7.2.2 RIS Assisted Near-Field Communications

Near-field communications have been acknowledged as a potential direction of 6G. As is known that characteristics of electromagnetic (EM) radiation field are divided by a Rayleigh distance $\frac{2A^2}{\lambda}$, where A is the distance of antenna array and λ is the wavelength of signals [101], [102]. The region within this boundary is named as near-field region and the region outside this boundary is named far-field region. Currently, limited by the dimensions of the antenna arrays, far-field channel models are usually adopted in communication networks in most existing works since the Rayleigh distance is extremely small, only a few meters. However, to meet the requirements of rapid development of next generation multiple access (NGMA), higher frequency carriers and larger number of antennas are more likely to be deployed. Therefore, the Rayleigh distance can become much larger than before and the near-field channel models can be adopted in 6G networks. In near-field networks, beam focusing is feasible to be implemented to increase the quality of communications and prevent unnecessary power loss of the transmitting signals . In [103], four types of architectures including fully-digital architectures, hybrid phase shifter-based precoders, and the emerging dynamic metasurface antenna (DMA) architectures are proposed for beam focusing designs. RIS/STAR-RIS is also capable to effectively assist the near-field networks. RIS/STAR-RIS can produce high intensity regions for high efficiency energy transfer, which can achieve beam focusing in a lower power consumption manner [104]. Besides, the proposed mathematical methods can still apply to the beam focusing design in near field networks. For instance, the authors applied the classic methods to optimize beams in fully-digital

architectures proposed in [103]. On the design of phase-shifters based hybrid antenna architecture, beamforming of users and analog precoding matrix were alternatively solved. Interestingly, manifold optimization proposed in **Chapter 5** was employed in optimizing analog precoding matrix, indicating the universality of the proposed mathematical methods in this thesis.

In the future, it is convincing that more works related to RIS assisted near-field NOMA networks will be studied.

7.2.3 RIS Assisted Federated Learning

As is illustrated in **Chapter 1**, distributed learning will be the core characteristic of 6G. federated learning, as a type of distributed learning plays an important role on the model sharing and training. In 6G, FL has the following advantages. Firstly, the data privacy can be protected by FL. In a distributed system, each mobile device trains its model locally and the data produced by it is not directly collected by central servers. Only the parameters on the neural networks of mobile devices are uploaded to servers. After the centralized training at the servers, the aggregated model will send from serves to each device for the purpose of next local training. Besides, FL can greatly save the computing resources of the central servers and fully take advantages of the local computing resources. It is because the computing mainly occurs at the local. Central servers only need to collect the parameters of the neural networks of each user instead of their great amount of row data.

The quality of communications directly determines the effect of FL. Therefore, for a better performance of FL, RIS can be deployed to assist FL as an innovative technique [105]. Firstly, RIS can greatly increase the rate of the transmission between the mobile devices and servers. It is crucial to speed up models exchanging in FL. Besides, since RIS allows the mobile devices in the service dead zone to participate in the FL, it is effective to construct a universal framework and further improve the accuracy of FL [106]. Except the above views, as the deeper research is being carried out, it is promising that more advantages of RIS will be exploited on improving the performance of FL, not only in training time but accuracy.

7.2.4 Active RIS Assisted NOMA Networks

Although the passive elements of RIS can be smartly tuned to improve the quality of channels and enable the communications in service dead zone. However, energy deficiency of the received signals caused by long-distance path loss still exists despite of the tuning of RIS elements. Active RIS provides a new resolution on this issue. Specifically, active RIS consists of multiple active elements. Active RIS can both amplify the amplitude and modify the phase shifts of the elements, thereby performing more DoF than the passive one. Although active RIS consumes more power than passive one, active RIS is still superior to relay both in power consumption, hardware cost and installation complexity. The difference between active RIS and passive RIS is that additional noise by active RIS is inevitably introduced [107]. Hence, in the problem formulation, the constraints related to noise should be taken into consideration. Although two different types of RISs adopt different ways of the problem formulation, the proposed methods can also apply to active RIS assisted NOMA networks. For example, the authors in [107] proposed a joint transmit beamforming and reflect precoding design by using alternating optimization. Especially, fractional programming proposed in **Chapter 4** and **Chapter 6** was also applied similarly in the beamforming design.

Nowadays, there has been some research on active RIS [107], [108]. In the future, studies related to active RIS assisted NOMA networks are promising to become a popular direction as the research value of passive elements has been fully exploited.

7.2.5 RIS Assisted Age of Information Reduction

As exponential increase in mobile data traffic and emerging applications which are time sensitive, i.e., autonomous driving [109], the latency of signals is not the only time measure to evaluate the communication performance of networks. It is because latency mainly focuses on the transmission time from transmitter to receiver. However, timeliness of information is also important to keep the stable performance of some specific scenarios. For example, as for an AR device, even though the latency is small, if the frequency for data freshness is low, it is still difficult for the device to provide people with high quality experience since the received signals might not reflect the most recent data. Age of information (AoI) is an measure,

defined as elapsed time from the generation time to received time of most recent update of status to describe the timeliness and freshness [110], [111]. Therefore, lower AoI means the faster freshness time and the AoI minimization is meaningful in wireless networks. In fact, RIS and NOMA can help to reduce AoI. It is because demodulation rate is one of the impacts to determine the freshness of data [109]. RIS assisted NOMA enables users to reach high SINR to meet the requirement of successful transmission of the status-update information [109]. Therefore, low AoI can be guaranteed in RIS assisted NOMA networks. The AoI reduction problems can also be tackled by the mathematical methods, which are similar to those in the thesis. For instance, the authors in [112] proposed an AoI minimization problem in RIS assisted NOMA SWIPT systems. When optimizing beamformers, SCA was implemented on the beamformers by finding the first order Taylor expansion of them. Moreover, alternating optimization was also utilized for multiple variables formulated in the proposed optimization problem.

In the future, AoI is expected to be an indispensable metrics to evaluate the performance of the systems and RIS assisted NOMA will play the important role on the reduction of AoI.

References

[1] J. Butler, "5G spectrum challenges," in *5G Radio Technology Seminar. Exploring Technical Challenges in the Emerging 5G Ecosystem.London*, 2015, pp. 1–9. DOI: 10.1049/ic.2015.0028 (cited on p. 20).

[2] *6G, The Next Horizon: From Connected People and Things to Connected Intelligence*. [Online]. Available: https://www-file.huawei.com/-/media/corp2020/pdf/tech-insights/1/6g-white-paper-en.pdf?la=en (cited on pp. 20, 21).

[3] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6G Wireless Communication Systems: Applications, Requirements, Technologies, Challenges, and Research Directions," *IEEE open j. Commun. Soc.*, vol. 1, pp. 957–975, 2020. DOI: 10.1109/OJCOMS.2020.3010270 (cited on p. 20).

[4] X. Xie, "OPTIMISATION IN NOMA WIRELESS COMMUNICATION NETWORKS," Ph.D. dissertation, University of Manchester, 2023 (cited on p. 21).

[5] F. Tang, X. Chen, M. Zhao, and N. Kato, "The Roadmap of Communication and Networking in 6G for the Metaverse," *IEEE Wirel Commun*, vol. 30, no. 4, pp. 72–81, 2023. DOI: 10.1109/MWC.019.2100721 (cited on p. 21).

[6] T. Hewa, G. Gür, A. Kalla, M. Ylianttila, A. Bracken, and M. Liyanage, "The Role of Blockchain in 6G: Challenges, Opportunities and Research Directions," in *2020 2nd 6G Wireless Summit (6G SUMMIT).Levi*, 2020, pp. 1–5. DOI: 10.1109/6GSUMMIT49458.2020.9083784 (cited on p. 21).

[7] L. U. Khan, W. Saad, D. Niyato, Z. Han, and C. S. Hong, "Digital-Twin-Enabled 6G: Vision, Architectural Trends, and Future Directions," *IEEE Commun Mag*, vol. 60, no. 1, pp. 74–80, Jan.2022. DOI: 10.1109/MCOM.001.21143 (cited on p. 21).

[8] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The Road Towards 6G: A Comprehensive Survey," *IEEE open j. Commun. Soc.*, vol. 2, pp. 334–366, 2021. DOI: 10.1109/OJCOMS.2021.3057679 (cited on p. 21).

[9] K. Wang, Z. Ding, D. K. C. So, and Z. Ding, "Age-of-Information Minimization in Federated Learning based Networks with Non-IID Dataset," *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2024. DOI: 10.1109/TWC.2024.3357208 (cited on p. 21).

[10] H. Shi, R. Ma, D. Li, and H. Guan, "Hierarchical Adaptive Collaborative Learning: A Distributed Learning Framework for Customized Cloud Services in 6G Mobile Systems," *IEEE Netw.*, vol. 37, no. 2, pp. 44–53, 2023. DOI: 10.1109/MNET.001.2200382 (cited on p. 21).

[11] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, 2020. DOI: 10.1109/MNET.001.1900287 (cited on p. 21).

[12] H. G. Myung, "Introduction to single carrier FDMA," in *2007 15th European Signal Processing Conference.Poznan*, 2007, pp. 2144–2148 (cited on p. 21).

[13] H. G. Myung, J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," *IEEE Veh. Technol. Mag.*, vol. 1, no. 3, pp. 30–38, 2006. DOI: 10.1109/MVT.2006.307304 (cited on p. 21).

[14] N. Abramson, "Multiple access in wireless digital networks," *Proc IEEE Inst Electr Electron Eng*, vol. 82, no. 9, pp. 1360–1370, 1994. DOI: 10.1109/5.317082 (cited on p. 22).

[15] C. Ciochina and H. Sari, "A review of OFDMA and single-carrier FDMA," in *2010 European Wireless Conference (EW).Lucca*, 2010, pp. 706–710. DOI: 10.1109/EW.2010.5483464 (cited on p. 22).

[16] M. Morelli, C.-C. J. Kuo, and M.-O. Pun, "Synchronization Techniques for Orthogonal Frequency Division Multiple Access (OFDMA): A Tutorial Review,"

Proc IEEE Inst Electr Electron Eng, vol. 95, no. 7, pp. 1394–1427, 2007. DOI: 10.1109/JPROC.2007.897979 (cited on p. 22).

[17] B. Makki, K. Chitti, A. Behravan, and M.-S. Alouini, "A Survey of NOMA: Current Status and Open Research Challenges," *IEEE open j. Commun. Soc.*, vol. 1, pp. 179–189, Jan.2020. DOI: 10.1109/OJCOMS.2020.2969899 (cited on p. 22).

[18] Z. Ding, L. Lv, F. Fang, *et al.*, "A State-of-the-Art Survey on Reconfigurable Intelligent Surface-Assisted Non-Orthogonal Multiple Access Networks," *Proc IEEE Inst Electr Electron Eng.*, vol. 110, no. 9, pp. 1358–1379, 2022. DOI: 10.1109/JPROC.2022.3174140 (cited on pp. 22, 147).

[19] Z. Ding, "Resolution of Near-Field Beamforming and Its Impact on NOMA," *IEEE Wireless Commun. Lett.*, vol. 13, no. 2, pp. 456–460, 2024. DOI: 10.1109/LWC.2023.3331954 (cited on p. 22).

[20] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A Survey of Non-Orthogonal Multiple Access for 5G," *IEEE COMMUN SURV TUT*, vol. 20, no. 3, pp. 2294–2323, 2018. DOI: 10.1109/COMST.2018.2835558 (cited on p. 22).

[21] Z. Ding, R. Schober, and H. V. Poor, "NOMA-Assisted Grant-Free Transmission: How to Design Pre-Configured SNR Levels?" *IEEE Wireless Commun. Lett.*, vol. 13, no. 2, pp. 412–416, 2024. DOI: 10.1109/LWC.2023.3330785 (cited on p. 22).

[22] D. Lin, K. Cumanan, and Z. Ding, "Beamforming Design for BackCom assisted NOMA Systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 9, pp. 1494–1498, 2023. DOI: 10.1109/LWC.2023.3279668 (cited on p. 22).

[23] Wu, Qingqing and Zhang, Shuowen and Zheng, Beixiong and You, Changsheng and Zhang, Rui, "Intelligent Reflecting Surface-Aided Wireless Communications: A Tutorial," *IEEE Trans Commun.*, vol. 69, no. 5, pp. 3313–3351, 2021. DOI: 10.1109/TCOMM.2021.3051897 (cited on pp. 22, 80).

[24] V. Drozdova and R. Akhpashev, "The Usage of Load Intensity Balance Approach For 5G MAC Protocol Chatacteristics Investigation," in *2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT).Yekaterinburg*, 2020, pp. 292–294. DOI: 10.1109/USBEREIT48449.2020.9117725 (cited on p. 23).

[25] *Reconfigurable Intelligent SurfaceTechnology White Paper*. [Online]. Available: https://www.researchgate.net/publication/369039356_RISTA_-_Reconfigurable_Intelligent_Surface_Technology_White_Paper_2023 (cited on p. 23).

[26] *China Telecom and ZTE Harness Cutting-edge Dynamic RIS Technology to Empower the Smart Asian Games*, Nov. 2023. [Online]. Available: https://www.mobileworldlive.com/china-telecom-and-zte-harness-cutting-edge-dynamic-ris-technology-to-empower-the-smart-asian-games/#:~:text=Providing%20a%20superior%20network%20experience,key%20technologies%20of%205G%2DA. (cited on p. 23).

[27] M. Ahmed, A. Wahid, S. S. Laique, *et al.*, "A Survey on STAR-RIS: Use Cases, Recent Advances, and Future Research Challenges," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14 689–14 711, 2023. DOI: 10.1109/JIOT.2023.3279357 (cited on p. 23).

[28] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously Transmitting and Reflecting (STAR) RIS Aided Wireless Communications," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 5, pp. 3083–3098, 2022. DOI: 10.1109/TWC.2021.3118225 (cited on p. 23).

[29] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: from theory to practice*. Cambridge University Press, 2011 (cited on p. 24).

[30] J. Lee, Y. Kim, H. Lee, *et al.*, "Coordinated multipoint transmission and reception in LTE-advanced systems," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 44–50, 2012. DOI: 10.1109/MCOM.2012.6353681 (cited on p. 24).

[31] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced [Coordinated and Distributed MIMO]," *IEEE Wirel. Commun.*, vol. 17, no. 3, pp. 26–34, 2010 (cited on p. 24).

[32] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint Radar and Communication Design: Applications, State-of-the-Art, and the Road Ahead," *IEEE Trans Commun.*, vol. 68, no. 6, pp. 3834–3862, 2020. DOI: 10.1109/TCOMM.2020.2973976 (cited on p. 24).

[33] J. A. Zhang, F. Liu, C. Masouros, *et al.*, "An Overview of Signal Processing Techniques for Joint Communication and Radar Sensing," *IEEE J. Sel. Top. Signal Process.*, vol. 15, no. 6, pp. 1295–1315, 2021. DOI: 10.1109/JSTSP.2021.3113120 (cited on p. 24).

[34] S. Kiranyaz, T. Ince, and M. Gabbouj, "Optimization Techniques: An Overview," in *Multidimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 13–44, ISBN: 978-3-642-37846-1. DOI: 10.1007/978-3-642-37846-1_2. [Online]. Available: https://doi.org/10.1007/978-3-642-37846-1_2 (cited on p. 25).

[35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004 (cited on p. 25).

[36] M. Wright, "The interior-point revolution in optimization: history, recent developments, and lasting consequences," *Bulletin of the American mathematical society*, vol. 42, no. 1, pp. 39–56, 2005 (cited on p. 25).

[37] D. P. Palomar and Y. C. Eldar, *Convex optimization in signal processing and communications*. Cambridge university press, 2010 (cited on p. 25).

[38] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1426–1438, 2006 (cited on p. 25).

[39] Y. Zhang, A. d'Aspremont, and L. E. Ghaoui, "Sparse PCA: Convex relaxations, algorithms and applications," *Handbook on Semidefinite, Conic and Polynomial Optimization*, pp. 915–940, 2012 (cited on p. 26).

[40] J. C. Ye, *Geometry of Deep Learning*. Springer, 2022 (cited on p. 26).

[41] F. Fang, Y. Xu, Q.-V. Pham, and Z. Ding, "Energy-Efficient Design of IRS-NOMA Networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14 088–14 092, 2020. DOI: 10.1109/TVT.2020.3024005 (cited on pp. 26, 56, 57, 62, 78, 88).

[42] S. Jiao, F. Fang, X. Zhou, and H. Zhang, "Joint Beamforming and Phase Shift Design in Downlink UAV Networks with IRS-Assisted NOMA," *Journal of Communications and Information Networks*, vol. 5, no. 2, pp. 138–149, 2020. DOI: 10.23919/JCIN.2020.9130430 (cited on pp. 26, 64).

[43] X. Xie, F. Fang, and Z. Ding, "Joint Optimization of Beamforming, Phase-Shifting and Power Allocation in a Multi-Cluster IRS-NOMA Network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7705–7717, 2021. DOI: 10.1109/TVT.2021.3090255 (cited on pp. 26, 57, 66, 70).

[44] Y. Li, M. Jiang, Q. Zhang, and J. Qin, "Joint Beamforming Design in Multi-Cluster MISO NOMA Reconfigurable Intelligent Surface-Aided Downlink Communication Networks," *IEEE Trans Commun.*, vol. 69, no. 1, pp. 664–674, 2021. DOI: 10.1109/TCOMM.2020.3032695 (cited on pp. 26, 70, 87, 92).

[45] S. Jiao, X. Xie, and Z. Ding, *Deep Reinforcement Learning Based Optimization for IRS Based UAV-NOMA Downlink Networks, arXiv: 2106.09616*, 2021. arXiv: 2106.09616 [eess.SP] (cited on p. 26).

[46] X. Xie, S. Jiao, and Z. Ding, "A Reinforcement Learning Approach for an IRS-assisted NOMA Network," *arXiv preprint arXiv:2106.09611*, 2021 (cited on p. 26).

[47] Y. Guo, F. Fang, D. Cai, and Z. Ding, "Energy-Efficient Design for a NOMA Assisted STAR-RIS Network With Deep Reinforcement Learning," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5424–5428, 2023. DOI: 10.1109/TVT.2022.3224926 (cited on pp. 26, 27).

[48] C. Wu, Y. Liu, X. Mu, X. Gu, and O. A. Dobre, "Coverage Characterization of STAR-RIS Networks: NOMA and OMA," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3036–3040, 2021. DOI: 10.1109/LCOMM.2021.3091807 (cited on p. 26).

[49] J. Zuo, Y. Liu, Z. Ding, L. Song, and H. V. Poor, "Joint Design for Simultaneously Transmitting and Reflecting (STAR) RIS Assisted NOMA Systems," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 1, pp. 611–626, Jan. 2023. DOI: 10.1109/TWC.2022.3197079 (cited on p. 27).

[50] C. Wu, X. Mu, Y. Liu, X. Gu, and X. Wang, "Resource Allocation in STAR-RIS-aided Networks: OMA and NOMA," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7653–7667, 2022. DOI: 10.1109/TWC.2022.3160151 (cited on p. 27).

[51] Y. Sun, Z. Ding, X. Dai, M. Zhou, and Z. Ding, "On the Application of Quasi-Degradation to Network NOMA in Downlink CoMP Systems," *IEEE Trans Commun.*, pp. 1–1, 2023. DOI: 10.1109/TWC.2023.3284834 (cited on p. 27).

[52] M. S. Ali, E. Hossain, A. Al-Dweik, and D. I. Kim, "Downlink Power Allocation for CoMP-NOMA in Multi-Cell Networks," *IEEE Trans Commun.*, vol. 66, no. 9, pp. 3982–3998, 2018. DOI: 10.1109/TCOMM.2018.2831206 (cited on p. 27).

[53] S. Mirbolouk, M. Valizadeh, M. C. Amirani, and S. Ali, "Relay Selection and Power Allocation for Energy Efficiency Maximization in Hybrid Satellite-UAV Networks With CoMP-NOMA Transmission," *IEEE Trans. Veh. Technol.*, vol. 71, no. 5, pp. 5087–5100, 2022. DOI: 10.1109/TVT.2022.3152048 (cited on p. 27).

[54] M. Elhattab, M. A. Arfaoui, and C. Assi, "Power Allocation in CoMP-Empowered C-NOMA Networks," *IEEE Netw. Lett.*, vol. 3, no. 1, pp. 10–14, 2021. DOI: 10.1109/LNET.2020.3045070 (cited on p. 27).

[55] H. Wang, C. Liu, Z. Shi, Y. Fu, and R. Song, "Power Minimization for Uplink RIS-Assisted CoMP-NOMA Networks With GSIC," *IEEE Trans Commun.*, vol. 70, no. 7, pp. 4559–4573, 2022. DOI: 10.1109/TCOMM.2022.3177773 (cited on p. 27).

[56] M. Elhattab, M. A. Arfaoui, C. Assi, and A. Ghrayeb, "RIS-Assisted Joint Transmission in a Two-Cell Downlink NOMA Cellular System," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1270–1286, 2022. DOI: 10.1109/JSAC.2022.3143211 (cited on p. 27).

[57] Q. Liu, Y. Zhu, M. Li, R. Liu, Y. Liu, and Z. Lu, "DRL-Based Secrecy Rate Optimization for RIS-assisted Secure ISAC Systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16 871–16 875, 2023. DOI: 10.1109/TVT.2023.3297602 (cited on p. 28).

[58] J. Zuo, Y. Liu, C. Zhu, Y. Zou, D. Zhang, and N. Al-Dhahir, "Exploiting NOMA and RIS in Integrated Sensing and Communication," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 12 941–12 955, Jan.2023. DOI: 10.1109/TVT.2023.3272036 (cited on p. 28).

[59] W. Lyu, Y. Xiu, X. Li, *et al.*, "Hybrid NOMA assisted Integrated Sensing and Communication via RIS," *IEEE Trans. Veh. Technol.*, pp. 1–6, 2023. DOI: 10.1109/TVT.2023.3343379 (cited on pp. 28, 43).

[60] K. Wang, F. Fang, D. B. d. Costa, and Z. Ding, "Sub-Channel Scheduling, Task Assignment, and Power Allocation for OMA-Based and NOMA-Based MEC Systems," *IEEE Trans Commun.*, vol. 69, no. 4, pp. 2692–2708, 2021. DOI: 10.1109/TCOMM.2020.3047440 (cited on p. 35).

[61] K. Wang, H. Li, Z. Ding, and P. Xiao, "Reinforcement Learning Based Latency Minimization in Secure NOMA-MEC Systems With Hybrid SIC," *IEEE Trans.*

Wirel. Commun., vol. 22, no. 1, pp. 408–422, Jan. 2023. DOI: 10.1109/TWC.2022.3194685 (cited on p. 35).

[62] R. Zhong, Y. Liu, X. Mu, Y. Chen, and L. Song, "AI Empowered RIS-assisted NOMA Networks: Deep Learning or Reinforcement Learning?" *IEEE J Sel Areas Commun.*, vol. 40, no. 1, pp. 182–196, 2022. DOI: 10.1109/JSAC.2021.3126068 (cited on p. 37).

[63] Z. Ding and H. Vincent Poor, "A Simple Design of IRS-NOMA Transmission, year=Feb.2020," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1119–1123, DOI: 10.1109/LCOMM.2020.2974196 (cited on p. 37).

[64] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users," *IEEE Signal Process Lett.*, vol. 21, no. 12, pp. 1501–1505, 2014. DOI: 10.1109/LSP.2014.2343971 (cited on p. 38).

[65] J. Xu, Y. Liu, X. Mu, and O. A. Dobre, "STAR-RISs: Simultaneous Transmitting and Reflecting Reconfigurable Intelligent Surfaces," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3134–3138, 2021. DOI: 10.1109/LCOMM.2021.3082214 (cited on p. 39).

[66] Y. Sun, Z. Ding, X. Dai, and O. A. Dobre, "On the Performance of Network NOMA in Uplink CoMP Systems: A Stochastic Geometry Approach," *IEEE Trans Commun.*, vol. 67, no. 7, pp. 5084–5098, 2019. DOI: 10.1109/TCOMM.2019.2906307 (cited on pp. 41, 97).

[67] W. Yi, Y. Liu, and Z. Ding, "Developing NOMA to Next-Generation Multiple Access," in *Fundamentals of 6G Communications and Networking*, Springer, 2023, pp. 291–316 (cited on p. 42).

[68] K. Shen and W. Yu, "Fractional Programming for Communication Systems—Part I: Power Control and Beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, 2018. DOI: 10.1109/TSP.2018.2812733 (cited on pp. 55, 71, 79, 84, 124, 131).

[69] F. Fang, J. Cheng, Z. Ding, and H. V. Poor, "Energy Efficient Resource Optimization for a Downlink NOMA Heterogeneous Small-Cell Network," in *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2018, pp. 51–55. DOI: 10.1109/SAM.2018.8448840 (cited on p. 56).

[70] F. Fang, H. Zhang, J. Cheng, and V. C. Leung, "Energy efficiency of resource scheduling for non-orthogonal multiple access (NOMA) wireless network," in *2016 IEEE International Conference on Communications (ICC).Kuala Lumpur*, 2016, pp. 1–5. DOI: 10.1109/ICC.2016.7511239 (cited on p. 56).

[71] Z. Song, Q. Ni, and X. Sun, "Spectrum and Energy Efficient Resource Allocation With QoS Requirements for Hybrid MC-NOMA 5G Systems," *IEEE Access*, vol. 6, pp. 37 055–37 069, 2018. DOI: 10.1109/ACCESS.2018.2851609 (cited on p. 56).

[72] I. Abu Mahady, E. Bedeer, S. Ikki, and H. Yanikomeroglu, "Sum-Rate Maximization of NOMA Systems Under Imperfect Successive Interference Cancellation," *IEEE Commun. Lett.*, vol. 23, no. 3, pp. 474–477, Jan.2019. DOI: 10.1109/LCOMM.2019.2893195 (cited on p. 56).

[73] H. Al-Obiedollah, K. Cumanan, J. Thiyagalingam, A. G. Burr, Z. Ding, and O. A. Dobre, "Energy Efficiency Fairness Beamforming Designs for MISO NOMA Systems," in *2019 IEEE Wireless Communications and Networking Conference (WCNC).Marrakesh*, 2019, pp. 1–6. DOI: 10.1109/WCNC.2019.8886009 (cited on p. 57).

[74] M. Zeng, X. Li, G. Li, W. Hao, and O. A. Dobre, "Sum Rate Maximization for IRS-Assisted Uplink NOMA," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 234–238, 2021. DOI: 10.1109/LCOMM.2020.3025978 (cited on p. 57).

[75] B. Fang, Z. Qian, W. Zhong, and W. Shao, "Iterative precoding for MIMO wiretap channels using successive convex approximation," in *2015 IEEE 4th Asia-Pacific Conference on Antennas and Propagation (APCAP).Bali*, 2015, pp. 65–66. DOI: 10.1109/APCAP.2015.7374273 (cited on p. 62).

[76] Z.-q. Luo, W.-k. Ma, A. M.-c. So, Y. Ye, and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems," *IEEE Signal Process Mag*, vol. 27, no. 3, pp. 20–34, 2010. DOI: 10.1109/MSP.2010.936019 (cited on pp. 71, 123).

[77] Z. Ding, "NOMA Beamforming in SDMA Networks: Riding on Existing Beams or Forming New Ones?" *IEEE Commun. Lett.*, vol. 26, no. 4, pp. 868–871, Jan.2022. DOI: 10.1109/LCOMM.2022.3146583 (cited on pp. 71, 98).

[78] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network: Joint Active and Passive Beamforming Design," in *2018 IEEE Global Communications Conference (GLOBECOM).Abu Dhabi*, 2018, pp. 1–6. DOI: 10.1109/GLOCOM.2018.8647620 (cited on pp. 71, 80, 90, 123).

[79] J. Zhu, J. Wang, Y. Huang, K. Navaie, Z. Ding, and L. Yang, "On Optimal Beamforming Design for Downlink MISO NOMA Systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3008–3020, Jan.2020. DOI: 10.1109/TVT.2020.2966629 (cited on p. 71).

[80] Y. Huang and D. P. Palomar, "Rank-Constrained Separable Semidefinite Programming With Applications to Optimal Beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, 2010. DOI: 10.1109/TSP.2009.2031732 (cited on p. 71).

[81] Y. Xu, C. Shen, Z. Ding, *et al.*, "Joint Beamforming and Power-Splitting Control in Downlink Cooperative SWIPT NOMA Systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874–4886, 2017. DOI: 10.1109/TSP.2017.2715008 (cited on p. 71).

[82] Shen, Kaiming and W. Yu, "Fractional Programming for Communication Systems—Part II: Uplink Scheduling via Matching," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2631–2644, 2018. DOI: 10.1109/TSP.2018.2812748 (cited on pp. 71, 124, 131).

[83] P. Cao, J. Thompson, and H. V. Poor, "A sequential constraint relaxation algorithm for rank-one constrained problems," in *2017 25th European Signal Processing*

Conference (EUSIPCO).Kos, 2017, pp. 1060–1064. DOI: 10.23919/EUSIPCO.2017.8081370 (cited on pp. 72, 80, 123).

[84] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource Allocation for IRS-Assisted Full-Duplex Cognitive Radio Systems," *IEEE Trans Commun.*, vol. 68, no. 12, pp. 7376–7394, 2020. DOI: 10.1109/TCOMM.2020.3020838 (cited on pp. 80, 95, 105, 121, 136).

[85] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2022 (cited on pp. 87, 138).

[86] J. Zuo, Y. Liu, Z. Ding, L. Song, and H. V. Poor, *Joint Design for Simultaneously Transmitting And Reflecting (STAR) RIS Assisted NOMA Systems*, 2021. DOI: 10.48550/ARXIV.2106.03001. [Online]. Available: https://arxiv.org/abs/2106.03001 (cited on pp. 87, 91).

[87] Y. Sun, Z. Ding, X. Dai, and G. K. Karagiannidis, "A Feasibility Study on Network NOMA," *IEEE Trans Commun.*, vol. 66, no. 9, pp. 4303–4317, 2018. DOI: 10.1109/TCOMM.2018.2825420 (cited on p. 98).

[88] Y. Sun and X. Dai, "Stochastic Geometry Based Modeling and Analysis on Network NOMA in Vehicular Networks," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops).Dublin*, 2020, pp. 1–6. DOI: 10.1109/ICCWorkshops49005.2020.9145390 (cited on p. 98).

[89] X. Yu, D. Xu, and R. Schober, "MISO Wireless Communication Systems via Intelligent Reflecting Surfaces : (Invited Paper)," in *IEEE/CIC International Conference on Communications in China (ICCC), Changchun, China, Aug*, 2019, pp. 735–740. DOI: 10.1109/ICCChina.2019.8855810 (cited on pp. 98, 99, 112).

[90] W. Jiang, B. Chen, S. Garg, J. Nie, J. Zhao, and Z. Xiong, "Joint Transmit Precoding and Reflect Beamforming for IRS-Assisted MIMO-OFDM Secure Communications," in *IEEE Global Communications Conference (GLOBECOM), Madrid, Spain,*, 2021, pp. 1–6. DOI: 10.1109/GLOBECOM46510.2021.9685734 (cited on p. 99).

[91] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. USA: Princeton University Press, 2007, ISBN: 0691132984 (cited on p. 109).

[92] F. Fang, B. Wu, S. Fu, Z. Ding, and X. Wang, "Energy-Efficient Design of STAR-RIS Aided MIMO-NOMA Networks," *IEEE Trans Commun.*, vol. 71, no. 1, pp. 498–511, Jan. 2023. DOI: 10.1109/TCOMM.2022.3223706 (cited on p. 113).

[93] X. Mu, Z. Wang, and Y. Liu, "NOMA for Integrating Sensing and Communications towards 6G: A Multiple Access Perspective," *Wireless Commun. Lett.*, pp. 1–8, 2023. DOI: 10.1109/MWC.015.2200559 (cited on p. 122).

[94] Z. Wang, Y. Liu, X. Mu, Z. Ding, and O. A. Dobre, "NOMA Empowered Integrated Sensing and Communication," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 677–681, Jan.2022. DOI: 10.1109/LCOMM.2022.3140271 (cited on p. 122).

[95] Z. Yang, D. Li, N. Zhao, Z. Wu, Y. Li, and D. Niyato, "Secure Precoding Optimization for NOMA-Aided Integrated Sensing and Communication," *IEEE Trans Commun.*, vol. 70, no. 12, pp. 8370–8382, 2022. DOI: 10.1109/TCOMM.2022.3216636 (cited on p. 122).

[96] C. Zhang, W. Yi, Y. Liu, and L. Hanzo, "Semi-Integrated-Sensing-and-Communication (Semi-ISaC): From OMA to NOMA," *IEEE Trans Commun.*, vol. 71, no. 4, pp. 1878–1893, 2023. DOI: 10.1109/TCOMM.2023.3241940 (cited on p. 122).

[97] P. Stoica, J. Li, and Y. Xie, "On Probing Signal Design For MIMO Radar," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4151–4161, 2007. DOI: 10.1109/TSP.2007.894398 (cited on p. 127).

[98] Z. Ding and H. V. Poor, "Joint Beam Management and Power Allocation in THz-NOMA Networks," *IEEE Trans. Commun.*, vol. 71, no. 4, pp. 2059–2073, 2023. DOI: 10.1109/TCOMM.2023.3240432 (cited on p. 147).

[99] S. Zhang and R. Zhang, "Intelligent reflecting surface aided multi-user communication: Capacity region and deployment strategy," *IEEE Trans Commun.*, vol. 69, no. 9, pp. 5790–5806, 2021. DOI: 10.1109/TCOMM.2021.3079128 (cited on p. 147).

[100] Z. Ding and H. V. Poor, "Design of THz-NOMA in the Presence of Beam Misalignment," *IEEE Commun. Lett.*, vol. 26, no. 7, pp. 1678–1682, 2022. DOI: 10.1109/LCOMM.2022.3171439 (cited on p. 148).

[101] Z. Ding, R. Schober, and H. V. Poor, "NOMA-Based Coexistence of Near-Field and Far-Field Massive MIMO Communications," *IEEE Wirel. Commun. Lett.*, vol. 12, no. 8, pp. 1429–1433, 2023. DOI: 10.1109/LWC.2023.3277469 (cited on p. 148).

[102] Y. Liu, C. Ouyang, Z. Wang, J. Xu, X. Mu, and A. L. Swindlehurst, *Near-Field Communications: A Comprehensive Survey*, 2024. arXiv: 2401.05900 [eess.SP] (cited on p. 148).

[103] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam Focusing for Near-Field Multiuser MIMO Communications," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 9, pp. 7476–7490, 2022. DOI: 10.1109/TWC.2022.3158894 (cited on pp. 148, 149).

[104] D. Shen, L. Dai, X. Su, and S. Suo, "Multi-Beam Design for Near-Field Extremely Large-Scale RIS-Aided Wireless Communications," *IEEE trans. green commun. netw.*, vol. 7, no. 3, pp. 1542–1553, 2023. DOI: 10.1109/TGCN.2023.3259579 (cited on p. 148).

[105] H. Liu, X. Yuan, and Y.-J. A. Zhang, "Joint Communication-Learning Design for RIS-assisted Federated Learning," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops).Montreal*, 2021, pp. 1–6. DOI: 10.1109/ICCWorkshops50388.2021.9473672 (cited on p. 149).

[106] L. Hu, Z. Wang, H. Zhu, and Y. Zhou, "RIS-Assisted Over-the-Air Federated Learning in Millimeter Wave MIMO Networks," *J. Commun. Inf. Netw.*, vol. 7, no. 2, pp. 145–156, 2022. DOI: 10.23919/JCIN.2022.9815198 (cited on p. 149).

[107] Z. Zhang, L. Dai, X. Chen, *et al.*, "Active RIS vs. Passive RIS: Which Will Prevail in 6G?" *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1707–1725, 2023. DOI: 10.1109/TCOMM.2022.3231893 (cited on p. 150).

[108] K. Zhi, C. Pan, H. Ren, K. K. Chai, and M. Elkashlan, "Active RIS Versus Passive RIS: Which is Superior With the Same Power Budget?" *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1150–1154, 2022. DOI: 10.1109/LCOMM.2022.3159525 (cited on p. 150).

[109] Z. Guo, Y. Niu, S. Mao, *et al.*, "Sum Rate Maximization Under AoI Constraints for RIS-Assisted Mmwave Communications," *IEEE Trans. Veh. Technol.*, pp. 1–17, 2023. DOI: 10.1109/TVT.2023.3331707 (cited on pp. 150, 151).

[110] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of Information: An Introduction and Survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, 2021. DOI: 10.1109/JSAC.2021.3065072 (cited on p. 151).

[111] Z. Ding, R. Schober, and H. V. Poor, "Unveiling the Importance of NOMA for Reducing AoI," in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 2233–2238. DOI: 10.1109/ICC45041.2023.10279047 (cited on p. 151).

[112] G. Zhang, Y. Lu, Y. Lin, Z. Zhong, Z. Ding, and D. Niyato, "AoI Minimization in RIS-Aided SWIPT Systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 2, pp. 2895–2900, 2024. DOI: 10.1109/TVT.2023.3316257 (cited on p. 151).