The University of Manchester

OPTIMISATION IN NOMA WIRELESS COMMUNICATION NETWORKS

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

Doctor of Philosophy

in the Faculty of Science and Engineering

2023

Ximing Xie

Department of Electrical and Electronic Engineering

Contents

Word count: 44263

List of figures

List of publications

X. Xie, F. Fang and Z. Ding, ”Joint Optimization of Beamforming, Phase-Shifting and Power Allocation in a Multi-Cluster IRS-NOMA Network,” in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7705-7717, Aug. 2021.

X. Xie, S. Jiao and Z. Ding, ”Maximizing the Sum-Rate of an RIS-Assisted Downlink NOMA Nnetwork: Active or Passive?,” submitted to *IEEE Transactions on Vehicular Technology*.

X. Xie and Z. Ding, “Backscatter-Assisted Non-orthogonal Multiple Access Network for Next Generation Vommunication,” *IET Signal Processing*, vol. 17, no. 4, e12211, 2023.

X. Xie, S. Jiao, K. Wang and Z. Ding, ”BAC-NOMA for Secondary Transmission,” accepted in *IEEE Communications Letters*.

S. Jiao, **X. Xie**, and Z. Ding, “Deep Reinforcement Learning-based Optimization for RIS based UAV-NOMA Downlink Networks (invited paper),” *Front. Sig. Proc*. 2:915567, 2022. (I mainly contributed to the algorithm design.)

S. Jiao, **X. Xie**, K. Wang and Z. Ding, “A BAC-NOMA Design for 6G umMTC with Hybrid SIC: Convex Optimization or Learning-based?,” major revision from *IEEE Transactions on Vehicular Technology* (I mainly contributed to the coding.).

Terms and abbreviations

AI	Artificial Intelligent
AmBC	Ambient Backscatter Communication
ANN	Artificial Neural Network
AR	Augmented Reality
AWGN	Additive White Gaussian Noise
B5G	Beyond 5-th Generation
BAC	Backscatter
BackCom	Backscatter Communication
BD	Backscatter Device
BS	Base Station
CDMA	Code Division Multiple Access
CEL	Cross Entropy Loss
CSI	Channel State Information
D2D	Device-to-Device
DDPG	Deep Deterministic Policy Gradient
DL	Deep Learning
DQN	Deep Q Learning
DRL	Deep Reinforcement Learning
ER	Extended Reality
FD	Full Duplex
FDMA	Frequency Division Multiple Access
FFT	Fast Fourier Transform
GAN	Generative Adversarial Network
IoT	Internet of Things
IRS	Intelligent Reflective Surface

KKT	Karush–Kuhn–Tucker
LoS	Line of Sight
LP	Linear Programming
MA	Multiple Access
MAE	Mean Absolute Error
MBB	Mobile Broadband
MIMO	Multiple-Input and Multiple-Output
ML	Machine Learning
MSE	Mean Square Error
NOMA	Non-orthogonal Multiple Access
OFDMA	Orthogonal Frequency-Division Multiple Access
OMA	Orthogonal Multiple Access
OTFS	Orthogonal Time Frequency Space
PAC	Principal Component Analysis
PG	Policy Gradient
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
SOCP	Second-order Cone Programming
STAR	Simultaneous Wireless Information and Power Transfer
STAR	Simultaneously Transmitting and Reflecting
TDMA	Time Division Multiple Access

THz Terahertz

umMTC Ultra-Massive Machine Type Communications

VR Virtual Reality

Abstract

With the exponential expansion of the Internet of Things (IoT) and the increasing demand for multimedia applications, the upcoming sixth-generation (6G) wireless communication network is poised to revolutionise connectivity. Non-orthogonal multiple access (NOMA) technique has been extensively studied in recent years because of the higher spectrum efficiency compared with orthogonal multiple access (OMA). NOMA enables multiple users and devices to share the same resource block, i.e., time slot, bandwidth and code, simultaneously, where the spectrum efficiency is improved. Furthermore, two innovative techniques, known as reconfigurable intelligent surface (RIS) and backscattering (BAC), have aroused people's interest. RIS has the capability to dynamically reconfigure the channel, enhancing signal quality, while BAC enables passive devices to transmit signals without consuming energy. Both of these techniques hold significant potential in IoT networks. This thesis focuses on exploring various optimisation problems arising from different NOMA scenarios to enhance the system's performance. First, a RIS-assisted NOMA downlink network, where multiple users receive signals from the base station (BS) with the help of multiple RISs, is investigated. Second, sum rate masmisation problem is formulated of multiple users in a RIS-assisted downlink NOMA network, where reinforcement learning is utilised as a tool to solve the this optimisation problem. Third, the combination of backscatter communication (BackCom) and NOMA is investigated. Finally, we verify the feasibility of introducing a BAC device into a legacy NOMA network without compromising its performance. The findings of this thesis not only underscore the critical significance of optimisation within the realm of wireless communication but also vividly illustrate the remarkable strides in spectrum efficiency realized through the deployment of NOMA technology.

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

Acknowledgements

First, I would like to express my deepest gratitude to my supervisor Prof. Zhiguo Ding for his unwavering guidance, invaluable insights, and continuous support throughout the entire journey of my doctoral research. His expertise, patience, and encouragement have been instrumental in shaping my ideas and refining my work. His strong enthusiasm for research, highly efficient ways, and hardworking attitude can also serve as guiding principles for my future career. It has been a joyful, unforgettable, and remarkable experience pursuing my Ph.D. under the guidance of Prof. Zhiguo Ding.

Second, I would like to thank all my colleagues, especially Dr. Fang Fang, Dr Kaidi Wang and Dr. Shiyu Jiao, at the University of Manchester. Their valuable comments and suggestions improve my research contributions and academic writing skills. I would also like to express my gratitude to them for their companionship and sharing during the Covid-19 pandemic, which greatly enriched and uplifted my life during that challenging time.

Last but most importantly, I would like to express my deepest gratitude to my parents, Mr. Diyu Xie and Ms. Jie Wu, for their unwavering support and unconditional encouragement during my Ph.D journey. Above all, I would like to express my deepest gratitude and love to my beloved wife, Ms. Jiaying Wang. Not only is she my best friend, but her steadfast support and invaluable inspiration have played a pivotal role in my research journey, enriching both my understanding of my work and my life as a whole.

Chapter 1

Introduction

1.1 Overview

With the 5-th generation (5G) of wireless communication has been successfully commercialised, people start to look forward to the next generation, the 6-th generation (6G) of wireless communication. Although the standard of 6G has not been released yet, academia and industry have placed the focus on it. First, a brief introduction of 6G is presented and the reason why 6G is necessary in the future is illustrated. Then, the introduction that how optimisation is important in wireless communication area is provided. The last part is about the development of multiple access (MA) techniques and the state-of-the-art technique in recent years.

1.1.1 The Evolution of 6G

Since the middle of 2019, commercial 5G networks have been widely deployed in many countries [1]. Follow the tradition, researchers starts to put more interest in research of the next generation wireless communication. The main technological driver is the explosively growing mobile traffic. A large amount of smart products, interactive services,and intelligent applications emerge, which introduce massive mobile traffic. A report from Ericsson reveals that 5G is hard to support such tremendous volume of mobile traffic in the next 10 years because of extensive growth of mobile traffic [2]. In particular, rapid increase of mobile broadband (MBB) leads to an exponential growth of smartphones. Meanwhile, rapid commercialisation of new types of electronic terminals such as virtual reality (VR) glasses and the autopilot sensors of vehicle also generate more traffic. On the other hand, the traffic demand of each electronic device, especially smartphones and smart-pads, continuously raises. The reason of that is streaming services like Youtube and Tik-Tok and mobile games have become the main form of entertainment in people's daily life. Low latency and high stability are critical factors of streaming services and mobile games. The Ericsson report also indicates that traffic from streaming services and mobile games has been dominant in all traffic and this phenomenon will become more pronounced in the next decade [2]. Another type of potential traffic arises from augmented reality (AR) and VR applications. An ITU report [3] reveals that the average data consumption is 5GB of every mobile user per month in 2020 and the estimated data consumption of every mobile user per month in 2030 will increase to 250GB. To better accommodate these tremendous traffic, the research and development of 6G are necessary. Hence, the rapid increase of mobile traffic has become a driving force of 6G.

To better understand characteristics and define requirements of 6G, some potential use cases are foreseen in [1]. Intelligent transport and logistics is a promising solution for international trades, where autonomous vehicles and drones are widely adopted to provide an efficient, safe and green environment of goods. Communications between autonomous vehicles and drones require the support from mobile networks thus the capacity of future mobile networks should be large enough to simultaneously support different types of scenarios. Extended reality (ER) is a new concept by combining AR and VR, which has stepped into practical applications. The main factor of ER is stable 360 field 3D video streaming, which requires a large bandwidth demand. The necessity of 6G is not only emphasized by two aforementioned user cases but also massive potential applications, which will request a lot of traffic and large bandwidth. Many advanced technologies have also laid the foundation for the advent of 6G.

1.1.2 Optimisation in a Wireless Communication Network

Optimisation as a fundamental mathematical theory is widely used in many engineering areas. Optimisation as a large subject includes many branches such as linear and non-linear programming, fractional programming, dynamic programming and convex optimisation. Going through the history of optimisation, the great ancient mathematician Euclid solved the earliest optimisation problem for geometry in his book The Elements [4]. Since then, optimisation theory has started thousands of years of development. Many important optimisation techniques were proposed by different mathematicians in the twentieth century. For example, William Karush, Harold William Kuhn, and Albert William Tucker analysed optimality conditions for nonlinear problems, respectively and these conditions are named by the Karush–Kuhn–Tucker (KKT) condition. The KKT condition has been widely used to decide whether an optimisation problem has an optimal solution and it is also an efficient method to calculate the closed-form optimal solution of a convex optimisation problem. In the last decade, with the rapid growth of computer computing power, artificial intelligence (AI) especially machine learning (ML) has been considered as a promising solution for complex multivariate nonlinear problems [5].

Wireless communication technology has achieved unprecedented development in the 21st century. In the past 30 years, wireless communication networks have evolved from the earliest 1G to the current 5G and the next generation 6G will be commercialised in the near future. However, the communication resource for each wireless communication network is limited. In particular, the frequency bandwidth available for wireless communication is not infinite and energy for transmitting signals is also limited. How to improve the performance of the network under limited resources has become an important research topic. There are a few popular types of optimisation problems, namely the transmit power minimisation problem, the sum rate maximisation problem and the energy efficiency problem, in wireless communication area. As the name indicates, the aim of a transmit power minimisation problem is to minimise the total transmit power, especially the transmit power of the base station or the transmitter, however, the QoS of each user and device is guaranteed. In a sum rate maximisation problem, the sum rate of all users and devices is the target that we want to maximally improve. The total transmit power in this scenario is usually limited. However, the balance of the total transmit power and the sum rate is considered in an energy efficiency problem, where the ratio of the total transmit power and the sum rate is maximised. In addition to the common optimisation problems mentioned above, there are many other types of optimisation problems in the field of wireless communication. Hence, a specific optimisation problem needs to be formulated according to the system model and the performance requirement.

1.1.3 The Development of Multiple Access Techniques

Multiple access techniques perform as an very important role in wireless communication networks to improve efficiency. According to the different ways of orthogonality, we divide multiple access techniques into two categories, namely orthogonal multiple access (OMA) techniques and non-orthogonal multiple access (NOMA) techniques [6]. There are some different types of OMA techniques, which are frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA) and orthogonal frequency division multiple access (OFDMA). Different multiple access techniques are adopted in different generations of wireless communication. The first generation (1G) mobile cellular system was designed to enable voice communications. FDMA was adopted in 1G network, where the spectrum was divided into segments, called channel and each user occupied the entire channel to transmit its signal. The second generation (2G) was designed to expand the capacity of voice communication and also integrate a data service. The multiple access technique in this generation was shifted from FDMA to TDMA in European community and CDMA in U.S. The third generation (3G) mobile cellular system was designed to expand the capacity of data service. TDMA was abandoned and the standard CDMA was selected as the multiple access technique in this generation. When the cellular system evolved to the forth generation (4G), which was designed for rapid increasing of data requirement, OFDMA was adopted to address the higher data rate demand [6].

Another category of multiple access technique is called non-orthogonal multiple access. NOMA allows different users and devices to share the same resource black such as time slots, spectrum segments and codes by introducing intra-cell interference. Intra-cell interference can be removed by successive interference cancellation (SIC). Compared with OMA, NOMA greatly improves resource efficiency, which has been a candidate of 5G and a potential solution of multiple access in beyond 5G (B5G) [6].

1.2 Organisation of the thesis

The organisation of this thesis is summarised as follows:

Chapter 2: In this chapter, some background information is introduced. First, some orthogonal multiple access techniques are described, e.g., frequency-division multiple access

(FDMA), time-division multiple access (TDMA), code-division multiple access (CDMA) and space-division multiple access (SDMA). An advanced FDMA named orthogonal frequency-division multiple access (OFDMA) is also introduced. Them, the power domain non-orthogonal multiple access (NOMA) is introduced. Furthermore, some wireless communication techniques such reconfigurable intellect surface (RIS) and backscattering communication (BackCom) are illustrated. Finally, optimisation tools, convex optimisation and machine learning, are presented.

Chapter 3: In this chapter, a downlink multi-cluster NOMA network is considered, where each cluster is supported by one RIS. This chapter aims to minimise the transmit power by jointly optimising the beamforming, the power allocation and the phase shift of each RIS. The formulated problem is non-convex and challenging to be solved due to the coupled variables, i.e., the beamforming vector, the power allocation coefficient and the phase shift matrix. To address this non-convex problem, an alternating optimisation based algorithm is proposed. Specifically, the primal problem is divided into two subproblems for beamforming optimisation and phase shifting feasibility, where the two subproblems are solved iteratively. Moreover, to guarantee the feasibility of the beamforming optimisation problem, an iterative algorithm is proposed to search the feasible initial points. To reduce the complexity, a simplified algorithm based on partial exhaustive search for this system model is also proposed. Simulation results demonstrate that the proposed alternating algorithm can yield a better performance gain than the partial exhaustive search algorithm, NOMA with random RIS phase shift scheme and OMA-RIS scheme. The content is related to the first publication ”Joint Optimization of Beamforming, Phase-Shifting and Power Allocation in a Multi-Cluster IRS-NOMA Network,” in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7705-7717, Aug. 2021.

Chapter 4: This chapter investigates a sum rate maximisation problem in a RIS-assisted NOMA downlink network. Specifically, the sum rate of all the users is maximised by jointly optimising the beams at the base station and the phase shift at the RIS. The deep reinforcement learning (DRL), which has achieved massive successes, is applied to solve this sum rate maximisation problem. In particular, an algorithm based on the deep deterministic policy gradient (DDPG) is proposed. Both the random channel case and the fixed channel case are studied in this chapter. The simulation result illustrates that the DDPG based algorithm has the competitive performance on both cases. The content is related to the second publication ”A Reinforcement Learning Approach for an IRS-Assisted NOMA Network,” ready to submit.

Chapter 5: NOMA technique introduces spectrum cooperation among different users and devices, which improves spectrum efficiency significantly. Energy-limited devices benefit from the backscatter (BAC) technique to transmit signals without extra energy consumption. The combination of NOMA and BAC provides a promising solution for Internet of Things (IoT) networks, where massive devices simultaneously transmit and receive signals. This chapter investigates a system model with two NOMA downlink users and an uplink device. The aim is to maximise the data rate of the uplink device by optimising the power allocation coefficient and the backscattering coefficient. Meanwhile the quality of service requirements of two NOMA users are guaranteed. The closed-form solution of two optimisation variables is derived, and an alternating algorithm is also proposed to solve the formulated optimisation problem efficiently. The proposed system verifies the feasibility of IoT devices being added into existing networks and provides a promising solution for wireless communication networks in the future. The content is related to the thrid publication “Backscatter-Assisted Non-orthogonal Multiple Access Network for Next Generation Vommunication,” *IET Signal Processing*, vol. 17, no. 4, e12211, 2023.

Chapter 6: Introducing BAC devices into a legacy NOMA network greatly improves spectrum efficiency, which provides a promising solution for the combination of IoT and wireless networks. Deep learning (DL) as an emerging optimisation tool gradually attracts people's interest in wireless communication area. In this chapter, a BAC-NOMA network is investigated, where a sum-rate maximisation problem is formulated and the closed form solution of backscattering coefficient is derived. The original problem is transformed and solved by a semi-definite relaxation (SDR) based algorithm and a learning based algorithm. The simulation results show that both algorithms have their own advantages and disadvantages and should be chosen wisely according to actual situations. The content is related to the fourth publication ”BAC-NOMA for Secondary Transmission,” accepted in *IEEE Communications Letters*.

Chapter 7: This chapter summaries the conclusion of this thesis and also provides some potential future research directions.

Chapter 2

Background Information

2.1 Orthogonal Multiple Access Techniques

In wireless communication, it is desirable to allow the base station send signals simultaneously to different users and devices. Multiple access techniques enables base stations and mobile users to a finite amount of radio spectrum. Orthogonal multiple access OMA techniques are a batch of multiple access techniques where signals to different users are mutually orthogonal. Orthogonality guarantees that there is no interference between different users. Several commonly used OMA techniques shown in Fig. 2.1 are summarized below.

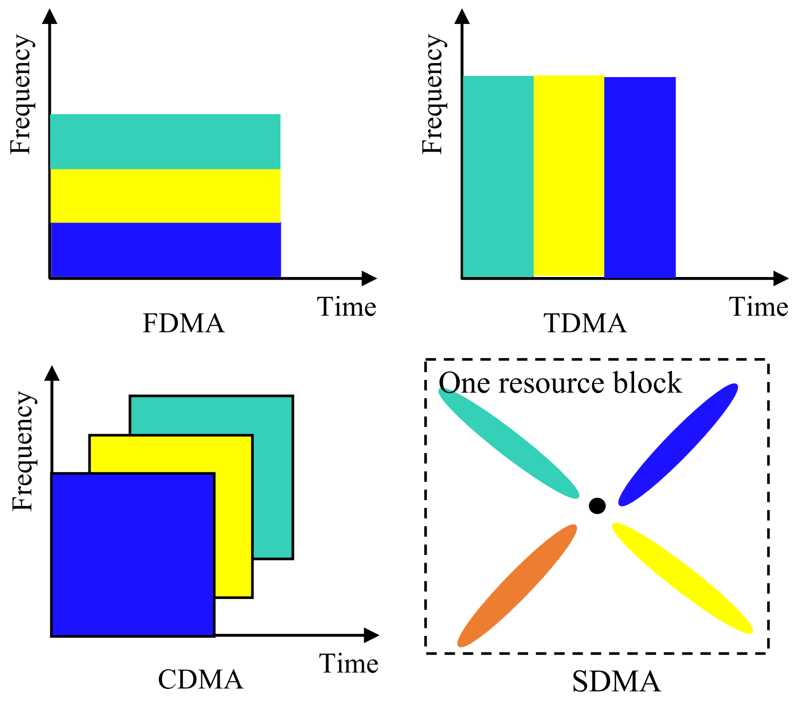

Figure 2.1. Four commonly used OMA techniques

- Frequency-division multiple access (FDMA)

FDMA allows multiple users to send their data simultaneously through the same commutation channel. The bandwidth is divided into several channels each of which is assigned to a user. Users can send data through a subchannel by modulating data on a carrier wave and the frequency of carrier wave is used to identify users. Although different users can send their signals at the same time, however, the bandwidth occupied by each user will decrease when the number of users increases. The crosstalk may happen between two adjacent users. Hence, the frequency gap between subchannels of two adjacent users is always inserted to avoid the crosstalk, which, obviously, cause a waste of spectrum resource. The mathematical expression is described as follows. We assume there are K users being serve by a base station. Let B denotes the whole bandwidth and T denotes the whole time slot. The data rate of each user can be expressed as follows:

$$R_{\mathrm{FDMA}} = \left(\frac{B}{K} - \Delta f\right) \log_2 \left(1 + \frac{S}{N_0}\right), \tag{2.1}$$

where Δf is the frequency gap and $\frac{S}{N_0}$ is the signal-noise ratio. Equation (2.1) indicates the data rate is affected by the bandwidth of the subchannel. When K increases, the bandwidth allocated to each subchannel decreases, which further makes the data rate decrease.

- Time-division multiple access (TDMA)

  TDMA allows multiple users to send their data through the same bandwidth by dividing the signal into different sub time slots. The users transmit in rapid succession, one after the other, each using its own sub time slot. Although each user can occupy the whole bandwidth to transmit its signal, however, the transmit time of each user is reduced. We still adopt the same notations as above, where K denotes the number of users, T denotes the whole time slot and B denotes the whole bandwidth. The instantaneous data rate of each user can be expressed as follows:

$$\tilde{R} = B \log_2 \left(1 + \frac{S}{N_0}\right). \tag{2.2}$$

  Since the whole time slot is divided into K subchannels and the duration of each subchannel is $\frac{T}{K}$, hence, a user can transmit $\frac{T\tilde{R}}{K}$ data in total. The average data rate of each user over the whole time slot T can be expressed as follows:

$$R_{\text{TDMA}} = \frac{B}{K}\log_2\left(1+\frac{S}{N_0}\right). \tag{2.3}$$

Equation (2.3) indicates the data rate of each user will decrease when the number of users increases.

- Code-division multiple access (CDMA)

  CDMA uses different codes to identity different users which allows multiple users to send their data through the same resource block, i.e., the same time slot and the same bandwidth. In a CDMA system, each user will be assigned a unique code and the signal sent by this user will be modulated on this unique code. All the modulated signals from different users will be sent to the receiver simultaneously. Let us consider a CDMA uplink scenario, where K users communicate with a base station. The signal sent by the $i-$th user is

$$x_i = c_i(t)s_i(t), i \in \{1, ..., K\} \tag{2.4}$$

where $c_i(t)$ is the unique code to identify this order. It is assumed that all signals are synchronised, hence, the signal received by the receiver is given by

$$y = \sum_{i=1}^{K} c_i(t)s_i(t). \tag{2.5}$$

The unique code usually has strong auto-correlation and weak cross-correlation, which can be described by

$$c_i c_j = \begin{cases} 1, j = i; \\ 0, j \neq i. \end{cases} \tag{2.6}$$

As a result, the receiver can remove interference caused by other users through orthogonality between codes. Since every user can fully utilise the entire bandwidth and the time slot, the data rate of each user can be expressed by follows:

$$R_{\text{CDMA}} = B\log_2\left(1+\frac{S}{N_0}\right). \tag{2.7}$$

It seems CDMA can bring a higher data rate compared with FDMA and TDMA. However, it still has some disadvantages. For example, the synchronisation is required, which will introduce extra complexity to equipment. The CDMA system performance degrades with an increase in the number of users because cross-correlation increases.

- Space-division multiple access (SDMA)

  Multi-antenna system can generate parallel spatial pipes, i.e., beamforming technique, to allow multiple users send their data through the same resource block. In the beamforming technique, the base station will generate a unique beam for each user and the beam decides the direction and the power. Let us assume that a network consists of K users and a M-antenna base station. Only when $M \geq K$, the base station can generate K linearly independent beams. The beam vector is given by

$$\mathbf{w}_i = [\beta_{i,1}e^{\theta_{i,1}}, \beta_{i,2}e^{\theta_{i,2}}, ..., \beta_{i,M}e^{\theta_{i,M}}], i \in \{1, 2, ..., K\}, \tag{2.8}$$

  where $\beta_{i,m}$ and $e^{\theta_{i,m}}, i \in \{1, 2..., K\}, m \in \{1, 2, ..., M\}$ denote the amplitude coefficient and the phase of the m-th antenna for the $i-$th user, respectively. Take zero-forcing (ZF) as an example, each two beam vectors are mutually orthogonal, which means

$$\mathbf{w}_i^H \mathbf{w}_j = \begin{cases} C_i, j = i; \\ 0, j \neq i, \end{cases} \tag{2.9}$$

  where C_i is a constant denoting the power allocated to the $i-$th beam. Each user can fully utilise the entire bandwidth and the time slot, the data rate of each user can be expressed by follows:

$$R_{\text{SDMA}} = B \log_2 \left(1 + \frac{S}{N_0} \right). \tag{2.10}$$

  . An important condition of the beamforming technique is that the number of antennas at the base station should be greater than then the number of users, otherwise, equation (2.9) cannot be satisfied.

- Orthogonal frequency-division multiple access (OFDMA)

  OFDMA is a special form of FDMA, which is designed to address the spectrum waste issue of FDMA caused by the frequency gap. Fig. 2.13 shows the frequency domain of OFDMA where all the signals are orthogonal at every sampling point. As a result,

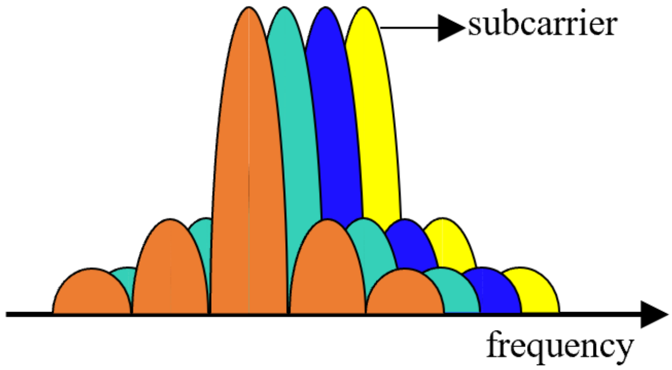

Figure 2.2. The frequency domain of OFDMA

crosstalk can be avoided without inserting frequency gaps and meanwhile the spectrum efficiency is improved. The mathematical description is presented below. We assume there are K users so K subcarriers in total are requred. Let $d_i, i \in \{1, ..., K\}$ denote the symbol of the $i-$th subcarrier and f_i denote the frequency of the $i-$th subcarrier. If time starts from t_s, the OFDM signal can be expressed as follows:

$$s(t) = \begin{cases} \sum_{i=0}^{K-1} d_i \text{rect}\left(t - t_s - \frac{T_{\text{OFDM}}}{2}\right) e^{j2\pi f_i (t-t_s)}, t_s \leq t \leq t_s + T_{\text{OFDM}}; \\ 0, t \leq t_s \text{ or } t \geq t_s + T_{\text{OFDM}}, \end{cases} \tag{2.11}$$

where T_{OFDM} is the duration of a OFDM symbol and rect$(.)$ is a rectangular function. f_i usually is designed as $\frac{i}{T_{\text{OFDM}}}$ to guarantee the orthogonality. If $s(t)$ is sampled with the sampling rate T_{OFDM}/K, the distract signal can be expressed as follows:

$$s_n(k) = \sum_{i=0}^{K-1} d_i e^{j\frac{2\pi jn}{K}}, 0 \geq n \geq K-1. \tag{2.12}$$

Note the expression of a discrete OFDM signal is the same at that of inverse discrete Fourier transform (IDFT). Hence, OFDM is realised by inverse fast Fourier transform (IFFT) in engineering. The data rate of each user can is given by

$$R_{\text{OFDMA}} = \left(\frac{B}{K}\right) \log_2 \left(1 + \frac{S}{N_0}\right). \tag{2.13}$$

Although OFDMA is a well-established and widely implemented technique, it does have certain disadvantages that need to be considered. Two prominent drawbacks are its sensitivity to frequency offsets and phase noise, as well as the increased complexity of

OFDMA electronics due to the inclusion of IFFT.

2.2 Non-orthogonal Multiple Access Technique

Non-orthogonal multiple access technique (NOMA) is a different type of multiple access technique compared with OMA. The name of NOMA directly indicates that transmit channels for all users are non-orthogonal. The paper [7] proposed NOMA for cellular future radio access towards the 2020s information society. In the context of NOMA, a groundbreaking concept is employed where all users' signals are combined and transmitted simultaneously over the same channel, resulting in a remarkable enhancement of spectrum efficiency. However, due to the non-orthogonal nature of the channels, each user's signal experiences interference from other users' signals. To overcome this challenge, a NOMA system utilizes a successive interference cancellation (SIC) receiver, which enables a user to eliminate the interference caused by other users before decoding its own signal. In power domain NOMA, a sophisticated technique used in modern communication systems, users are assigned different power levels based on their individual channel qualities. This approach allows for effective user identification within the system. By employing power allocation strategies, users with superior channel conditions are allocated lower power levels, while users experiencing weaker channel conditions are assigned higher power levels.

2.2.1 Downlink NOMA

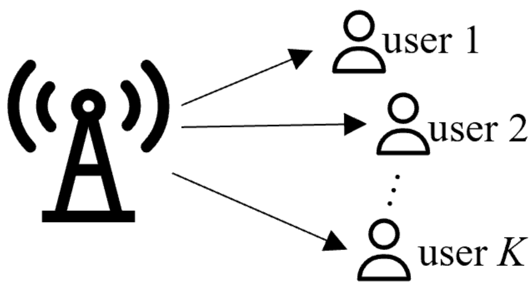

Figure 2.3. T
he downlink NOMA network

Fig. 2.3 illustrates the downlink NOMA network, consisting of multiple users. First, the decoding order needs to be decided. Without loss of generality, we assume the decoding order is $\Lambda_K \geq \Lambda_{K-1} \geq ... \geq \Lambda_2 \geq \Lambda_1$. Given the aforementioned decoding order, the $i-$ th user needs to decode $K - i$ users' signals before decoding its own signal. Let us define the set $S_j = \{1, 2, ..., j-1\}, j \in \{1, 2, ..., K\}$ as the collection of users who will cause interference when decoding the j-th user's signal. Hence, the $j-$th user's data rate observed at the $i-$th user can be expressed as follows:

$$R_{i \to j} = \log_2 \left(1 + \frac{\alpha_j P_0 |h_i s_j|^2}{\sum_{k \in S_j} \alpha_k P_0 |h_i s_k|^2 + \sigma_i^2} \right), j \geq i \tag{2.14}$$

where α_i, h_i and $\sigma_i^2, \in \{1, 2, ..., K\}$ denote the power allocation coefficient, channel coefficient and the power of AWGN of user i, respectively. The power allocation coefficient should satisfy $\sum_{i=1}^{K} \alpha_i = 1$. The user i's data rate is given by

$$R_{i \to i} = \log_2 \left(1 + \frac{\alpha_i P_0 |h_i s_i|^2}{\sum_{k \in S_i} \alpha_k P_0 |h_i s_k|^2 + \sigma_i^2} \right), i \in \{1, 2, ..., K\}. \tag{2.15}$$

From two equations above, we have the achievable data rate of user i, which is given by

$$R_i = \min\{R_{k \to i}, k \in \{1, 2, .., i\}\}. \tag{2.16}$$

The reason of the i- th user's acheivable data rate is defined by (2.16) is to guarantee the SIC can be proceed successfully. When QoS is considered, this definition guarantees that every user can decode user i's data rate successfully. It is important to note that the first user is required to decode the signals of $K-1$ other users before decoding its own signal. However, this approach is impractical due to the high time complexity involved. Hence, a downlink NOMA network typically cannot accommodate a large number of users.

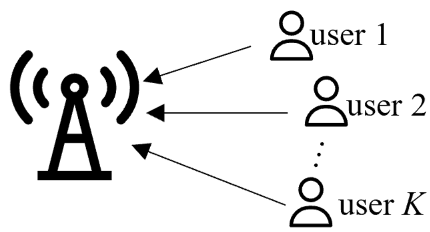

Figure 2.4. An uplink NOMA network

2.2.2 Uplink NOMA

Fig. 2.4 illustrates an uplink NOMA network where K users simultaneously transmit their signals to the base station. The signal sent by the $i-$ th user is given by $\sqrt{P_i} s_i$. We assume that signals from all the users are synchronised at the base station, hence, the received signal by the base station can be expressed as follows:

$$y = \sum_{i=1}^{K} \sqrt{P_i} s_i + w, \tag{2.17}$$

where w denotes AWGN at the base station.

The SIC process in an uplink NOMA network is significantly simpler compared to that of a downlink NOMA network. Firstly, in an uplink NOMA network, the SIC process is performed at the base station, which has access to the received signals from multiple users simultaneously. This allows for more efficient interference cancellation since the base station can exploit the knowledge of the user signals and their power levels. In contrast, in a downlink NOMA network, the SIC process is performed at the user devices, which have limited capabilities compared to the base station. Each user device must decode the signals from other users in a sequential manner, starting with the strongest interfering signal. This sequential decoding process increases the complexity and introduces potential errors. Furthermore, the uplink NOMA network benefits from having a centralised base station that can coordinate and manage the interference cancellation process. This centralisation enables more sophisticated algorithms and resource allocation strategies, simplifying the SIC process for individual user devices.

Equation (2.16) illustrates that a user's achievable data rate is is constrained by the decoding outcomes of other users, however, this issue does not exist in an uplink NOMA network. The achievable data rate of a user in an uplink NOMA network under the decoding order $\Lambda_K \geq \Lambda_{K-1} \geq ... \geq \Lambda_2 \geq \Lambda_1$ can be expressed as follows:

$$R_i = \log_2 \left(1 + \frac{P_i |h_i s_i|^2}{\sum_{k \in S_i} P_k |h_i s_k|^2 + \sigma^2} \right), i \in \{1, 2, ..., K\}. \tag{2.18}$$

2.3 Reconfigurable Intelligent Surface

Reconfigurable intelligent surfaces (RIS) also known as intelligent reflecting surfaces (IRS) have emerged as a revolutionary technology in wireless communication systems, which is shown in Fig. 2.5. There are multiple passive reflecting elements on the panel of a RIS,

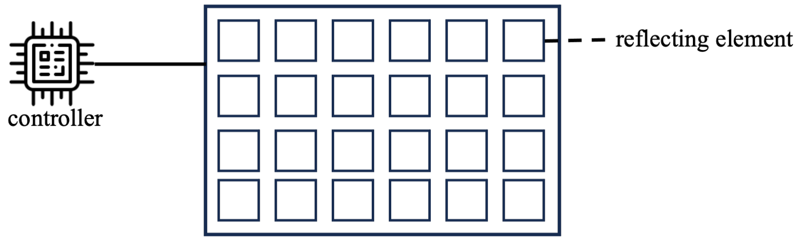

Figure 2.5. The structure of a RIS

which can manipulate and control electromagnetic waves. The entire RIS is controller by a smart controller to reconfigure the phase shift of each reflecting element [8], [9]. Unlike the traditional wireless communication relying on active antennas, the RIS introduces a new transmit model where the wireless propagation environment can be tuned in a desired manner. One of the most important features of the RIS is to reshape channels, enabling better signal quality and large signal propagation range. Moreover, RIS provides a cost-effective and energy-efficient solution. By adopting the passive nature of the elements, RIS requires minimal power consumption compared to traditional active transmitters. In addition, more properly designed algorithms can be applied to optimally adjust the phase shift of each reflecting element. The RIS offers more opportunities to a wireless network to improve its performance.

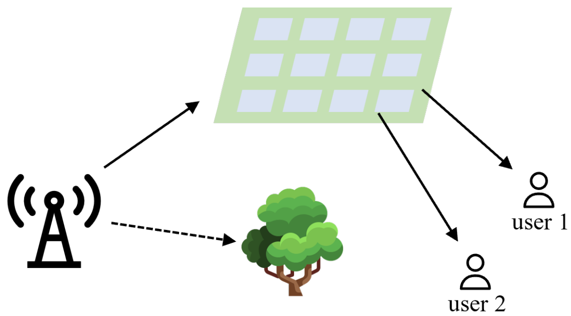

Figure 2.6. A RIS-assisted NOMA downlink network

Fig. 2.6 shows a RIS-assisted NOMA downlink network, where two NOMA users only can communicate with the base station with the help from RIS since the direct link between the base station and each user is unavailable due to heavy blockage. We assume the base station and users are all equipped with a single antenna and the RIS consists of N reflecting elements. The base station broadcasts the superimposed signal which is the same as (**??**). Let $\mathbf{h}_{\text{BR}} \in \mathbb{C}^{N\times 1}$ denote the channel vector between the base station and the RIS and $\mathbf{h}_1 \in \mathbb{C}^{N\times 1}$ and $\mathbf{h}_2 \in \mathbb{C}^{N\times 1}$ denote the channel vectors between the RIS and user 1 and user 2, respectively. The vector

$$\mathbf{p} = \left[\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, ..., \beta_N e^{j\theta_N}\right] \tag{2.19}$$

describes the phase shift of each reflecting elements, where β_i and $\theta_i, i \in \{1, 2, ..., N\}$ denote the amplitude coefficient and the phase angle of the $i-$th reflecting element, respectively. The signal received by users can be expressed as follows:

$$y_i = \mathbf{h}_i^H \Theta \mathbf{h}_{\text{BR}} \left(\sqrt{\alpha P_0} s_1 + \sqrt{(1-\alpha) P_0} s_2\right) + w_i, i \in \{1, 2\}, \tag{2.20}$$

where Θ denotes a diagonal matrix whose diagonal collects all the elements of $\mathbf{p}$ and w_i denotes AWGN at user i. Defining the decoding order prior to SIC is essential, therefore, the decoding order in this example is assumed to be $\Lambda_2 \geq \Lambda_1$. In particular, user 1 first decodes user 2's signal, eliminates it, and then proceeds to decode its own signal independently, free from any interference, while user 2 directly decodes its own signal by treating user 1's signal as interference. The date rate of user 2 observed at user 1 can be expressed as

$$R_{1\to 2} = \log_2 \left(1 + \frac{(1-\alpha) P_0 |\mathbf{h}_1^H \Theta \mathbf{h}_{\text{BR}} s_2|^2}{\alpha P_0 |\mathbf{h}_1^H \Theta \mathbf{h}_{\text{BR}} s_1|^2 + \sigma_1^2}\right), \tag{2.21}$$

where σ_1^2 is the power of AWGN at user 1. After user 2's signal is removed, the data rate of user 1 can be expressed as follows:

$$R_1 = \log_2 \left(1 + \frac{\alpha P_0 |\mathbf{h}_1^H \Theta \mathbf{h}_{\text{BR}} s_1|^2}{\sigma_1^2}\right). \tag{2.22}$$

For user 2, it decodes its own signal directly by treating user 1's signal as interference. Hence, the data rate of user 2 can be expressed as follows:

$$R_{2\to 2} = \log_2 \left(1 + \frac{(1-\alpha) P_0 |\mathbf{h}_2^H \Theta \mathbf{h}_{\text{BR}} s_2|^2}{\alpha P_0 |\mathbf{h}_2^H \Theta \mathbf{h}_{\text{BR}} s_1|^2 + \sigma_2^2}\right), \tag{2.23}$$

where σ_2^2 is the power of AWGN at user 2. The achievable data rate of user 2 is

$$R_2 = \min\{R_{1\to 2}, R_{2\to 2}\}. \tag{2.24}$$

According to the discussion above, it is noted that the phase shift matrix Θ and the power allocation coefficient α can be optimised to make the entire system achieve the best performance. Therefore, some typical optimisation problems are commonly investigated in a RIS-assisted network. For example, the sum rate optimisation problem can be summarised as follows:

$$\max_{\alpha,\Theta,P_0} \quad R_1 + R_2 \tag{2.25a}$$

$$\text{s.t.} \quad 0 \leq P_0 \leq P_{\text{max}} \tag{2.25b}$$

$$|\Theta_{i,i}|^2 = 1, i = 1, \cdots, N \tag{2.25c}$$

$$\theta_i \in [0, 2\pi], i = 1, \cdots, N \tag{2.25d}$$

The aim of this optimisation problem is to achieve the largest sum rate of two users. Constraint (2.25b) enforces the maximum power limit defined by the system, ensuring that the total transmitted power does not exceed this limit. Constraints (2.25c) and (2.25d) stem from inherent characteristics of the RIS, imposing restrictions on the reflecting elements' reflecting coefficient and phase shift. Another classic optimisation problem is the energy-efficiency maximisation problem, which can be summarised as follows:

$$\max_{\alpha,\Theta,P_0} \quad \frac{R_1 + R_2}{P_0} \tag{2.26a}$$

$$\text{s.t.} \quad R_i \geq R_t^i, i = 1, 2 \tag{2.26b}$$

$$0 \leq P_0 \leq P_{\text{max}} \tag{2.26c}$$

$$|\Theta_{i,i}|^2 = 1, i = 1, \cdots, N \tag{2.26d}$$

$$\theta_i \in [0, 2\pi], i = 1, \cdots, N \tag{2.26e}$$

where $R_t^i.i = 1, 2$ denotes the minimal data rate of user i. The objective function in this problem quantifies energy efficiency as the ratio of the sum rate to the transmit power. Constraint (2.26b) guarantees the successful SIC. There are many other optimisation problems such as multi-user interference management, security optimisation, cost-effectiveness optimisation and so on. Each of these optimisation problems provides promising research directions to study RIS-assisted NOMA in depth.

2.4 Backscattering Communication

Backscattering (BAC) communication is a wireless communication technique which allows BAC devices to transmit its own signal by utilising the principle of backscattering. According to my knowledge, the paper [10] first proposed the concept of backscattering communication. Unlike tradition transmission relying on active antennas, backscattering relies on reflecting and modulating existing signals to convey data without extra energy consumption. It well fits the requirement of the Internet of Things (IoT). A device having the function to backscatter signal is usually called as a backscatterring device (BD). A BD is equipped with a backscattering circuit. The circuit will be excited by the received signal first and modulate the BD's signal onto the received signal. The modulated signal is then reflected to the destination. The BD usually operates with minimal energy requirements. The passive feature allows a BD to be small, have high endurance or even operate without batteries. As a result, backscattering communication provides a promising future for wireless communication and IoT networks.

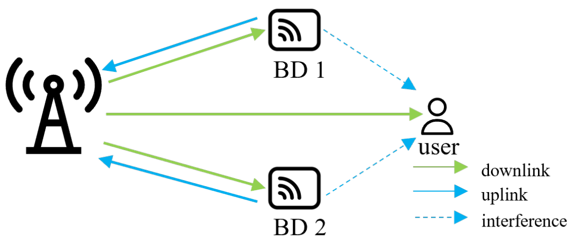

Figure 2.7. A RIS-assisted NOMA downlink network

The backscattering technique has a good combination with uplink NOMA. Fig. 2.7 shows a simple BAC-NOMA uplink network. which consists of one base station, two BDs and one downlink user. Two BDs only communicate with the BS and the downlink user only receives signal from the base station. We assume the base station and all devices in this network are all equipped with one single antenna. The base station operates in full-duplex mode, which means it can transmit and receive signals simultaneously. Let h_1 and h_2 denote the channel coefficients between the base station and two BDs, respectively. h_d denotes the channel coefficient between the base station and the downlink user and g_1 and g_2 denote the channel coefficients of two interference links. The base station sends a downlink signal $x = \sqrt{P_0}s_0$ to the downlink user, meanwhile, this signal will be received by two BDs. The signal received by each BD can be expressed as follows:

$$y_i^{\text{BD}} = \sqrt{P_0} h_i s_0 + w_i, i \in \{1, 2\}. \tag{2.27}$$

The BD modulates its own signal onto the received signal and reflects the modulated signal to the destination. In this way, the signal received by the base station is given by

$$y_u = \sqrt{P_0 \eta_1} |h_1|^2 s_0 c_1 + \sqrt{P_0 \eta_2} |h_2|^2 s_0 c_2 + w_u, \tag{2.28}$$

where $c_i, i \in \{1, 2\}$ denotes the signal from BD i, $\eta_i, i \in \{1, 2\}$ represents the backscattering coefficient of BD i and w_u refers to AWGN at the base station. The backscattering coefficient enables the adjustment of the energy utilised for backscattering its own signal. It is worth noting that we disregard the term $w_i c_i$ due to its negligible power contribution. The SIC process is operated by the base station, where BD 1's signal is decoded first, followed by the decoding of BD 2's signal. By assuming $\mathbb{E}(|c_i|^2) = 1, i \in \{1, 2\}$, the data rate of BD 1 is expressed as follows:

$$R_1 = \log_2 \left(1 + \frac{P_0 \eta_1 |h_1|^4 |s_0|^2}{P_0 \eta_2 |h_2|^4 |s_0|^2 + \sigma_u^2} \right), \tag{2.29}$$

where σ_u^2 is the power of AWGN. Once BD 1's signal is removed, the data rate of BD 2 is expressed as follows:

$$R_2 = \log_2 \left(1 + \frac{P_0 \eta_2 |h_2|^4 |s_0|^2}{\sigma_u^2} \right). \tag{2.30}$$

The sum rate of two BDs is given by

$$R_{\text{sum}} = \log_2 \left(1 + \frac{P_0 \eta_1 |h_1|^4 |s_0|^2 + P_0 \eta_2 |h_2|^4 |s_0|^2}{\sigma_u^2} \right). \tag{2.31}$$

Since two BDs will cause interference to the downlink user, the signal recieved by the downlink user can be expressed as follows:

$$y_d = \sqrt{P_0} h_d s_0 + \sqrt{P_0 \eta_1} h_1 g_1 s_0 c_1 + \sqrt{P_0 \eta_2} h_2 g_2 s_0 c_2 + w_d, \tag{2.32}$$

where σ_u^2 is the power of AWGN at the downlink user. The downlink user's data rate is given by

$$R_d = \log_2 \left(1 + \frac{P_0 |h_d|^2}{P_0 \eta_1 |h_1|^2 |g_1|^2 + P_0 \eta_2 |h_2|^2 |g_2|^2 + \sigma_d^2} \right), \tag{2.33}$$

where σ_d^2 is the power of AWGN at the downlink user.

A classic optimisation problem of backscatterring communication is the sum rate maximisation problem, which can be summarised as follows:

$$\max_{P_0, \eta_1, \eta_2} \quad R_{\text{sum}} \tag{2.34a}$$

$$\text{s.t.} \quad R_d \geq R_t \tag{2.34b}$$

$$0 \leq P_0 \leq P_{\text{max}} \tag{2.34c}$$

$$0 \leq \eta_i \leq 1, i = 1, 2. \tag{2.34d}$$

Constraint (2.34b) guarantees the quality of service (QoS) of the downlink user to make its data rate greater than the minimal target data rate R_t. Constraint (2.34c) is the transmit power limit and constraint (2.34d) arises from the passive feature of backscattering.

2.5 Convex Optimisation

Convex optimisation is a subfield of mathematical optimisation that studies the optimisation problem with a convex function over a convex set. Convex optimisation plays a crucial role to solve an optimisation problem due to the global optimal solution can be obtained [11]. The convex optimisation is also the foundation of machine learning (ML), which is very popular in every engineer field. Thus, the convex optimisation is compulsory for my research. In this section, we discuss some basic concepts in convex optimisation.

2.5.1 Basic Concepts of Convex Optimisation

Some basic concepts are introduced in this subsection to better understand convex optimisation.

- Affine set and convex set: An affine set is a set that preserves the property that any line passing through two points within the set remains entirely within the set. The mathematical express of an affine set is for any x_1, $x_2 \in C$ and $\theta \in \mathbf{R}$, C is an affine set if $\theta x_1 + (1-\theta)x_2 \in C$. A set is called a convex set if it satisfies the property that for any two points within the set, the line segment connecting those points lies entirely within the set. The mathematical express of a convex set is for any x_1, $x_2 \in C$ and $0 \le \theta \le 1$, C is a convex set if $\theta x_1 + (1-\theta)x_2 \in C$.

- Convex function: A function $f : \mathbf{R}^n \to \mathbf{R}$ is a convex function if the domain of f is a convex set and if for all x_1, $x_2 \in \mathbf{dom} f$ and $0 \le \theta \le 1$, we have

$$f(\theta x_1 + (1-\theta)x_2) \le \theta f(x_1) + (1-\theta)f(x_2), \tag{2.35}$$

which is shown in Fig. 2.8.

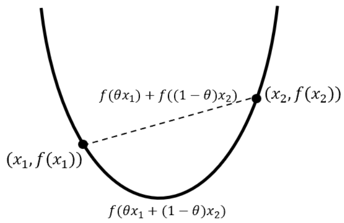

Figure 2.8. A convex function

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

$$f(x_2) \ge f(x_1) + \nabla f(x_1)^T (x_2 - x_1) \tag{2.36}$$

holds for all x_1, $x_2 \in \mathbf{dom} f$. Fig. 2.9 graphically illustrates the first-order condition of a convex function.

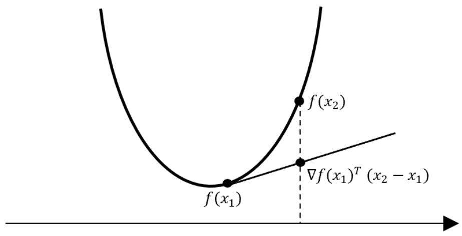

Figure 2.9. A graph to show the first-order condition

- Second-order condition: If f is twice differentiable, f is convex if and only if $\mathbf{dom} f$ is convex and the Hessian matrix of f is positive semidefinite. The mathematical express of the second-order condition is given by: for all $x \in \textbf{dom}f$, if and only if

$$\nabla^2 f(x) \succcurlyeq 0. \tag{2.37}$$

- α-sublevel set: the $\alpha-$ sublevel set of a function f: $\mathbf{R}^n \to \mathbf{R}$ is defined as

$$C_\alpha\{x \in \textbf{dom}f | f(x) \leq \alpha\}. \tag{2.38}$$

A sublevel set of a convex function is a convex set.

2.5.2 Definition of Convex Problem

Given the definitions of convex set and convex function above, we can finally define the convex problem. If the objective function of an optimisation problem is convex and the feasible set of this problem is also a convex set, then we name this problems as a convex problem. In general, we can write a standard form of a convex problem as:

$$\min \quad f(\mathbf{x}) \tag{2.39a}$$

$$\text{s.t.} \quad f_i(\mathbf{x}) \leq 0, i = 1, \cdots, m \tag{2.39b}$$

$$h_i(\mathbf{x}) = 0, i = 1, \cdots, p \tag{2.39c}$$

where $\mathbf{x} \in \mathbb{R}^n$ is an optimisation variable and functions $f(\mathbf{x}), f_i(\mathbf{x}), \forall i$ are all convex functions and $h_i(\mathbf{x}), \forall i$ are affine functions (i.e. $a_i\mathbf{x} + b_i$). Generally, a convex problem may have one or more solutions and also a convex function has many equivalent transformations. It is necessary to point out that the local optimal solution of a convex problem is also its global optional solution.

2.5.3 Classic Optimisation Problems

In a linear program (LP), the objective function and all constraints are affine, which is obviously a convex problem. A general linear program has the form

$$\min \quad c^T x + d \tag{2.40a}$$

$$
\begin{aligned}
\text{s.t.} \quad & Gx \preccurlyeq h && (2.40b) \\
& Ax = b, && (2.40c)
\end{aligned}
$$

where $G \in \mathbf{R}^{m \times n}$ and $A \in \mathbf{R}^{p \times n}$. There are special cases of a LP problem, namely the standard form LP and the inequality form LP. A standard form LP is expressed as

$$
\begin{aligned}
\min \quad & c^T x && (2.41a) \\
\text{s.t.} \quad & x \succcurlyeq 0 && (2.41b) \\
& Ax = b, && (2.41c)
\end{aligned}
$$

where only one component-wise non-negativity constraint $x \succcurlyeq 0$ exists. An inequality form LP is expressed as

$$
\begin{aligned}
\min \quad & c^T x && (2.42a) \\
\text{s.t.} \quad & Ax \preccurlyeq b, && (2.42b)
\end{aligned}
$$

where only inequality constraints exist.

The linear-fractional programming is sightly different from LP, where the objective function is not affine. A linear-fractional program usually has a form

$$
\begin{aligned}
\min \quad & f_0(x) = \frac{c^T x + d}{e^T x + f} && (2.43a) \\
\text{s.t.} \quad & Gx \preccurlyeq h && (2.43b) \\
& Ax = b. && (2.43c)
\end{aligned}
$$

whose objective function is a quasilinear function so we also call this problem quasiconvex optimisation problems. However, this kind of problem can be easily transformed to a convex form as

$$
\begin{aligned}
\min \quad & c^T y + dz && (2.44a) \\
\text{s.t.} \quad & Gy - hz \preccurlyeq 0 && (2.44b) \\
& Ay - bz = 0 && (2.44c)
\end{aligned}
$$

$$e^T y + fz = 1 \tag{2.44d}$$

$$z \geq 0. \tag{2.44e}$$

If x is feasible in (2.43) then the pair

$$y = \frac{x}{e^T x + f}, \qquad z = \frac{1}{e^T x + f}$$

is feasible in (2.44), with the same objective value.

If the objective function is quadratic and convex, and other constraints are affine, this kind of problem is quadratic program (QP), which can be summarised as:

$$\begin{aligned} \min \quad & (1/2)x^T P x + q^T x + r && (2.45a) \\ \text{s.t.} \quad & Gx \preccurlyeq h && (2.45b) \\ & Ax = b, && (2.45c) \end{aligned}$$

where $P \in \mathbf{S}_+^n$, $G \in \mathbf{R}^{m \times n}$ and $A \in \mathbf{R}^{p \times n}$. According to the definition of a convex problem, QP is a convex problem. If the objective function and the inequality constraint functions are all quadratic, QP becomes quadratically constrained quadratic programming (QCQP), which is given by

$$\begin{aligned} \min \quad & (1/2)x^T P_0 x + q_0^T x + r_0 && (2.46a) \\ \text{s.t.} \quad & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \cdot, m && (2.46b) \\ & Ax = b, && (2.46c) \end{aligned}$$

It is noted that the QCQP problem is not a convex problem due to its feasible set may not be a convex set. However, there are many existing method to deal with this QCQP problems.

There is another common form named second-order cone programming (SOCP), which is closely related to quadratic programming.

$$\begin{aligned} \min \quad & f^T x, && (2.47a) \\ \text{s.t.} \quad & ||A_i x + b_i||_2 \leq c_i^T x + d_i, i = 1, \cdots, m && (2.47b) \end{aligned}$$

$$Fx = g, \tag{2.47c}$$

where $A \in \mathbf{R}^{n_i \times n}$ and $F \in \mathbf{R}^{p \times n}$. Constraint (2.47b) is a second-order cone constraint since it is the same as requiring the affine function $(Ax + b, c^T x + d)$ to lie in the second-order cone.

The semidefinite programming (SDP) problem has the form

$$\min \quad f^T x, \tag{2.48a}$$

$$\text{s.t.} \quad x_1 F_1 + \cdots + x_n F_n + G \preccurlyeq 0 \tag{2.48b}$$

$$Ax = b, \tag{2.48c}$$

where $F, F_1, \cdots F_n \in \mathbf{S}^k$ and $A \in \mathbf{R}^{p \times n}$. The inequality constraint of a SDP problem is a linear matrix inequality.

2.5.4 Duality and KKT Condition

The duality is a very important concept in convex optimisation. The optimisation problem we formulated in a project is non-convex at most time but its dual problem may be convex or have a simpler form to be solved. The Karush-Kuhn-Tucher (KKT) condition is a powerful tool to derive out the closed-form optimal solution of a convex problem. We consider an optimisation problem in the standard form (2.39), we define the Lagrangian function L as

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x), \tag{2.49}$$

where λ_i and ν_i are the Lagrangian multiplier associated to the i-th inequality constraint $f_i(x) \le 0$ and i-th equality constraint $h_i(x) = 0$ respectively. We further define the Lagrangian dual function based on Lagrangian function (2.49) as

$$g(\lambda, \nu) = \underset{x \in \mathcal{D}}{inf}\, L(x, \lambda, \nu) = \underset{x \in \mathcal{D}}{inf} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right) \tag{2.50}$$

where $\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom} f_i \cap \bigcap_{i=0}^{p} \mathbf{dom} h_i$ is non-empty. It is noted that even problem (2.39) is not convex, the dual function is concave since it is the pointwise infimum of a family of affine functions of λ, ν.

From the definition of the dual function, we can see that the Lagrange dual function provides a lower bound of optimal value p^* (assume the optimal value of problem (2.39) is p^*). We always want to find the best lower bound so we can formulate a optimisation problem as

$$\min \quad g(\lambda, \nu) \tag{2.51a}$$

$$\text{s.t.} \quad \lambda \succcurlyeq 0. \tag{2.51b}$$

We name this problem as the Lagrange dual problem associated with the problem (2.39). We define (λ^*, ν^*) as the optimal Lagrange multiplier and d^* as the optimal solution of this dual problem. If the equality

$$d^* = p^* \tag{2.52}$$

holds, then we can alternatively solve the dual problem to obtain the optimal solution of the related primal problem if the dual problem has a simpler form than the primal one. We also call this problem has a strong duality. Now, we introduce the Slater's condition. If a optimisation satisfies the Slater's condition, then the strong duality holds. The Slater's condition can be expressed as: There exists an $x \in$ **relint**$\mathcal{D}$

$$f_i(x) \leq 0, \quad i = 1, ..., m, \qquad Ax = b. \tag{2.53}$$

Consider a convex optimisation problem holding a strong duality. we define x^* and (λ^*, ν^*) are the optimal valuable of the primal problem and the associated dual problem. Then, x^*, λ^*, ν^* satisfy the KKT conditions.

$$\begin{aligned} f_i(x^*) \leq 0, &\quad i = 1, \cdots, m \\ h_i(x^*) = 0, &\quad i = 1, \cdots, p \\ \lambda_i^* \geq 0, &\quad i = 1, \cdots, m \\ \lambda_i^* f_i(x^*) = 0, &\quad i = 1, \cdots, m \\ \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0, & \end{aligned} \tag{2.54}$$

In some cases, we can derive the closed-form of the optimal solution from the KKT conditions.

2.6 Algorithms

2.6.1 Semidefinite Relaxation

Semidefinite relaxation (SDR) was first introduced to solve a nonconvex QCQP problem in [12]. There is an important equation in SDR, which are given by

$$\mathbf{x}^T\mathbf{C}\mathbf{x} = \text{Tr}(\mathbf{x}^T\mathbf{C}\mathbf{x}) = \text{Tr}(\mathbf{C}\mathbf{x}\mathbf{x}^T), \tag{2.55}$$

where $\mathbf{x}$ is a n dimension vector and $\mathbf{C}$ is a $n \times n$ symmetric matrix. A QCQP problem can be expressed as follows:

$$\min_{\mathbf{x}} \quad \mathbf{x}^T\mathbf{C}\mathbf{x}, \tag{2.56a}$$

$$\text{s.t.} \quad \mathbf{x}^T\mathbf{A}_i\mathbf{x} \ \trianglerighteq_i \ b_i, i = 1, \cdots, m, \tag{2.56b}$$

where "$\trianglerighteq_i$" represents either "$\geq$", "$\leq$" or "$=$" for each i and $\mathbf{C}, \mathbf{A}_1, \cdots, \mathbf{A}_m$ are $n \times n$ symmetric matrices and $b_i, \cdots, b_m$ are constant numbers. By introducing a new variable $\mathbf{X} = \mathbf{x}^T\mathbf{x}$, problem 2.56 has the following equivalent form

$$\min_{\mathbf{X}} \quad \text{Tr}(\mathbf{C}\mathbf{X}), \tag{2.57a}$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{A}_i\mathbf{X}) \ \trianglerighteq_i \ b_i, i = 1, \cdots, m, \tag{2.57b}$$

$$\mathbf{X} \succcurlyeq 0, \tag{2.57c}$$

$$\text{Rank}(\mathbf{X}) = 1. \tag{2.57d}$$

Note that constraint (2.57c) and (2.57d) guarantee $\mathbf{X} = \mathbf{x}^T\mathbf{x}$ is a rank one symmetric positive semidefinite (PSD) matrix. Except constraint (2.57d), all other constraints including the objective function are convex. Problem 2.57 can be recast into a SDP problem by ignoring the rank one constraint (2.57d). Since the rank one constraint is ignored, how to recover the original optimisation variable $\mathbf{x}$ from $\mathbf{X}$ is important. Denote the optimal solution of the SDP problem with $\mathbf{X}^*$, there are two cases when retrieving $\mathbf{x}^*$:

- Rank($\mathbf{X}^*$) $= 1$: Eigenvalue decomposition is applied to $\mathbf{X}^*$ to retrieve $\mathbf{x}^*$.
- Rank($\mathbf{X}^*$) $\neq 1$: Gaussian randomisation procedure described in [12] is utilised to find an approximation of $\mathbf{x}^*$.

2.6.2 Successive Convex Approximation

Successive convex approximation (SCA) in this thesis is utilised to approximate a non-convex function with its first-order Taylor series at a specific point and iteratively optimise the approximated function.[13]. Given a non-convex function $f(x)$, its first-order Tayler series at point x_0 in the $i-$th iteration is expressed as follows:

$$\tilde{f}_{x_0}(x) = f(x_0^{(i)}) + f'(x_0^{(i)})(x - x_0^{(i)}). \tag{2.58}$$

$\tilde{f}_{x_0}(x)$ is a linear function and convex. The updating criteria is given by $x_0^{(i+1)} = x^{*(i)}$, where $x^{*(i)}$ is the optimised result in the previous iteration. If the original problem is a convex problem, SCA can also provide an optimal solution, otherwise, SCA can only provide a suboptimal solution.

2.7 Machine Learning

Machine learning as a powerful and popular tool to solve optimisation problems has been applied in every engineering field. Learning is an update process, in this process the ML algorithm will gradually approach the optimal solution. In many cases, ML only can find the local optimal solution due to the high complexity of the objective function. However, this local optimal solution will satisfy the engineer requirement in most time and will be better than non-ML algorithm [14]. Deep learning (DL) as a subfield of ML is considered to solve a problem with artificial neural networks (ANNs) [15]. The structure of a neural network is inspired by the structure of the human brain, specifically the interconnected networks of neurons. There are three main types of learning in the field of DL, which are supervised learning, unsupervised learning and reinforcement learning. Each type of them has its applicable fields.

2.7.1 Neural Networks

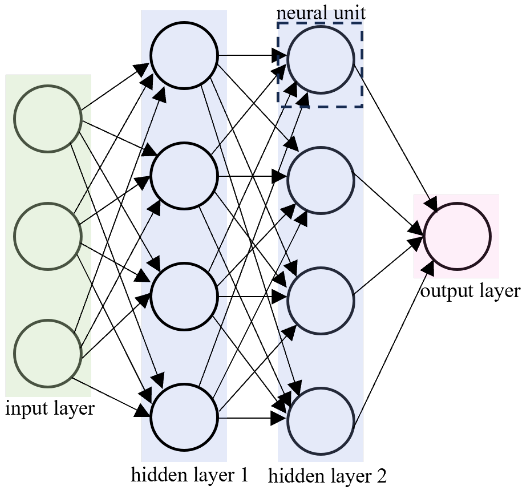

Figure 2.10. The structure of a fully connected neural network

Fig. 2.10 illustrates the structure of a fully connected neural network, which consists of four layers. The name of this network directly indicates that every two neurons located at two adjacent layers are mutually connected. There are three types of layers, namely the input layer, the hidden layer and the output. It is important to note that the number of neurons in each layer of this network varies. The dimension of the input data and the output data is determined by the number of neurons in the input layer and the output layer, respectively. By adjusting the number of neurons in these layers, the dimension of data can be easily controlled. Fig. 2.11 shows a classic structure of a neuron but may be changed in some special networks. The input vector is denoted by $\mathbf{x} = [x_1, x_2, \cdots x_N]$. the weight vector is denoted by $\mathbf{w} = [w_1, w_2, \cdots w_N]$ and the bias is denoted by b. $\sigma(\cdot)$ denotes the activation function. The whole process of a neuron can be expressed as

$$y = \sigma(\mathbf{w}^T\mathbf{x} + b). \tag{2.59}$$

The reason why we introduce activation functions into a neural network is that activation functions can introduce non-linearity. The neural network without activation functions can only approximate linear functions, where the neural network without activation functions can understand the non-linear relationship present in data. Two commonly adopted activation functions are

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.60}$$

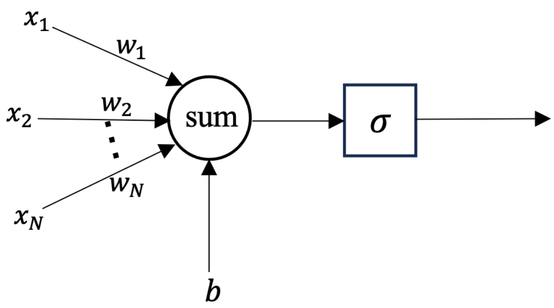

Figure 2.11. The structure of a neuron

and

$$f(x) = \max(0, x). \tag{2.61}$$

(2.60) is named by *sigmoid* function and (2.61) is named by *relu* function. It is worth noting that the sigmoid function normalises the output within the range 0 to 1, which is commonly adopted to produce a probabilistic result. Another feature of the *sigmoid* function is differentiable everywhere, whose gradient can be calculated in the back propagation. However, the *sigmoid* function is not commonly used between hidden layers due to vanishing gradient issue because the gradients usually become very small with the backpropagate proceeds. To address this issue, *relu* is preferred to be the activation function between hidden layers. The *relu* function does not only introduce non-linearity to the network but also introduce sparsity in the network because the negative values are set to be zero. Compared with the *sigmoid* function, the *relu* function is easier to be implemented in the network.

2.7.2 Loss Function

The loss function is the gateway to train a neural network by adjusting the weights of neurons to minimise it. The loss function in the supervised learning is normally designed to represent the distance between the correct label and the predict result. However, the loss function in unsupervised learning is more flexible. It can be Euclidean distance between each data or the objective function of a specific optimisation problem. Some commonly used loss functions are summarised below [15].

- Mean squared error loss (MSE): MSE loss is the most common loss function in deep learning, which is calculated by squaring the difference between the correct label and the predict result. The equation of MSE loss is expressed as follows:

$$F_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2, \tag{2.62}$$

where we assume there are N samples and y_i and $\hat{y}_i$ denote the predicted result and the label of the i-th sample, respectively. MSE loss is usually utilised in a regression problem to train a neural network.

- Mean absolute error loss (MAE): MAE loss is another common loss function, which is quite similar to MSE loss. The equation of MAE loss can be expressed as follows:

$$F_{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|. \tag{2.63}$$

  The advantage of MAE compared with MSE is that MAE is more robust at abnormal points because MSE squares the error, making the error of the abnormal point too large. However, MAE converges slower than MES because the gradient of MSE changes with the size of the error, while the gradient of MAE remains at 1.

- Huber loss: Huber loss is a combination of MSE and MAE, taking advantages of both. The principle underlying the selection of MSE or MAE is based on the magnitude of the error. MSE is typically employed when the error is relatively small or close to zero, whereas MAE is more suitable when the error is larger in magnitude. The equation of Huber loss can be expressed as follows:

$$F_{huber} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}_{|y_i-\hat{y}_i|\leq\delta}\left(\frac{(y_i - \hat{y}_i)^2}{2}\right) + \mathbb{I}_{|y_i-\hat{y}_i|\geq\delta}\left(\delta|y_i - \hat{y}_i| - \frac{1}{2}\delta^2\right), \tag{2.64}$$

  where δ is a hyper-parameter to define the position where MSE and MAE are connected and $\mathbb{I}$ is an indicator function.

- Cross entropy loss (CEL): The cross entropy loss is most commonly utilised in the classification problem. For a binary classification problem, the cross entropy loss can be expressed as

$$F_{CEL}^{2} = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i). \tag{2.65}$$

2.7.3 Supervised Learning

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs [16]. It will infer a function from a training data set where each data sample has a associated label to form a pair [17]. The most important work in supervised learning is to generate the training data set, where training data will consist of inputs paired with the correct outputs.

There are two subcategories in supervised learning which are classification and regression. The aim of classification is to assign the input data to a specific class. A very simple example of classification is to determine the gender of a person. In this case, the input data will be an image of a person, while the network's output is a set of probabilities indicating the likelihood of that person being classified as either male or female. The gender having the larger probability will be the predicted result of the neural network. This predicted result will be compared with the correct label to train the network. The goal of regression is to predict a continuous number such as temperature and profit. For example, we want to predict a farmer's revenue based on weather condition. There is a positive correlation between the agricultural revenue and the weather condition. Regression algorithms try to find these relationships between two variables.

2.7.4 Unsupervised Learning

Unsupervised learning is looking for a previously undetected patterns from a non-label dataset [18]. Unlike supervised learning, unlabeled data is utilised to train the model.

There are several popular techniques used in unsupervised learning, including clustering, dimensionality reduction, and generative modeling. Clustering algorithms aim to cluster data having the similarity into the same group. Dimensionality reduction techniques, such as principal component analysis (PCA) [19], aim to capture the most relevant features within a low dimension. Generative models, such as autoencoders [20] or generative adversarial networks (GANs) [21], are designed to learn the underlying probability distribution of the data. A generative model is able to generate new sample based on features extracted from training data.

There is no specific loss function for unsupervised learning. If unsupervised learning is utilised to solve an optimisation problem, the objective function is usually the loss function. The neural network will directly apply gradient decent on the objective function to train the neural network. A very classic unsupervised learning algorithm is K-means, which is utilised to allocate data into different clusters [22]. Let assume the training set is $x_1, x_2, \cdots, x_N$ and $x_i, i \in \{1, 2, \cdots, N\}$ is a data sample. There are K clusters for each data sample to be allocated. The loss function in this algorithm is expressed as follows:

$$F_{KMEANS} = \sum_{i=1}^{N} \sum_{k_1}^{K} w_{ik} ||x_i - \mu_k||^2, \tag{2.66}$$

where μ_k is centroid of the $k-$th cluster and $w_{ik} = 1$ means the data sample x_i belongs to cluster k, otherwise, $w_{ik} = 0$. Note that equation (2.66) is non-convex and it is difficult to find a global optimal solution. Therefore, gradient-based optimisation strategy is commonly utilised to find a local optimal solution. If unsupervised learning is adopted to solve an optimisation problem in wireless communication field, the loss function is normally the objective function of the optimisation problem. For example, if we want to maximise the sum rate of two mobile users, the loss function can be expressed as $-(R_1 + R_2)$, where R_1 and R_2 denote the data rate of user 1 and user 2, respectively. Note that minimising $-(R_1 + R_2)$ is equivalent to maximising $R_1 + R_2$. Moreover, if the optimisation problem is a constrained problem, the loss function should be modified accordingly based on all constraints.

2.7.5 Reinforcement Learning

Basic knowledge of reinforcement learning

Reinforcement learning (RL) is a kind of learning types to obtain rewards through an agent which constantly takes actions in an environment and maximize the total reward [23]. The reinforcement learning is an area of machine learning that handles with sequential decision-making [24]. There are two critical parts for a RL system, where are the agent and the environment. The key idea of RL is to train an agent can generate good actions based on the environment. A few factors fully characterize the RL processing.

- *Policy* π: $\pi(s^{(t)}, a^{(t)})$ reflects the probability of an action chose by an agent conditioned to $s^{(t)}$.

- *State* s: The state is a set of observations from the environment. The state will continuously change with time. We utilize $s^{(t)}$ to represent the state in the t-th time slot.
- *Action* a: The action is the choices made by the agent adopting a policy π based on the current state. $a^{(t)}$ denotes the action correlating to the state $s^{(t)}$. The state will switch to $s^{(t+1)}$ after the agent makes an action.
- *Reward* r: The reward is an evaluation of the action. Once the agent generates an action, the environment will give feedback to evaluate this action. $r^{(t)}$ is utilized as the reward for the action $a^{(t)}$.

There are two kinds of RL algorithms in the RL family, one is value-based RL and another is policy gradient RL.

The value-based RL algorithm aims to build a value function and by minimizing or maximizing this function to define a policy. Q learning is a value based algorithm. Under a policy π,, the Q function for a certain action a_t and state s_t can be defined as

$$Q_\pi(s_t, a_t) = \mathbb{E}_\pi \left[\sum_{k=0}^{+\infty} \xi^k r_{t+k} \mid s_t, a_t \right], \tag{2.67}$$

where $\xi \in (0, 1]$ is the reward discount factor. (2.67) indicates that Q value this the expectation of the cumulative reward and the future reward has a smaller impression of the current Q value. Deep Q learning (DQN) is also a value based learning algorithm. The Q function is also defined as the expectation of the cumulative reward shown in (2.67). The difference between QL and DQN is that the deep neural network is utilized to estimate the Q function. According to the Bellman's equation, the Q function can be also expressed as follows:

$$Q(s_t, a_t|\theta) = r_t + \xi \max_{a'} Q'\left(s_{t+1}, a'|\theta'\right) \tag{2.68}$$

where Q' is another neural network named target network to predict the Q value of the next state. To train the neural network, we define the loss function as the expectation of the mean square error (MSE) between (2.67) and (2.68), which can be expressed as follows:

$$L(\theta) = \mathbb{E}\left[\left(r_t + \xi \max_{a'} Q'\left(s_{t+1}, a'|\theta'\right) - Q(s_t, a_t|\theta) \right)^2 \right]. \tag{2.69}$$

It is worth to point out the parameters of the target network θ' are fixed during the training. Both QL and DQL aim to address discrete action problems.

The policy gradient (PG) RL algorithm aims to optimize a performance objective by finding a good policy [24]. PG algorithms are also designed for the continuous action space. There is one classic PG algorithm named stochastic policy gradient. The parameter updating rule can be expressed as follows:

$$\theta_{t+1} = \theta_t + l\mathbb{E}[\nabla_\theta(\log\ \pi_\theta(s_t, a_t))Q^{\pi_\theta}(s_t, a_t)], \tag{2.70}$$

where l is the learning rate and Q^{π_θ} is the Q function under the current policy π_θ to evaluate the current policy.

Deep deterministic policy gradient

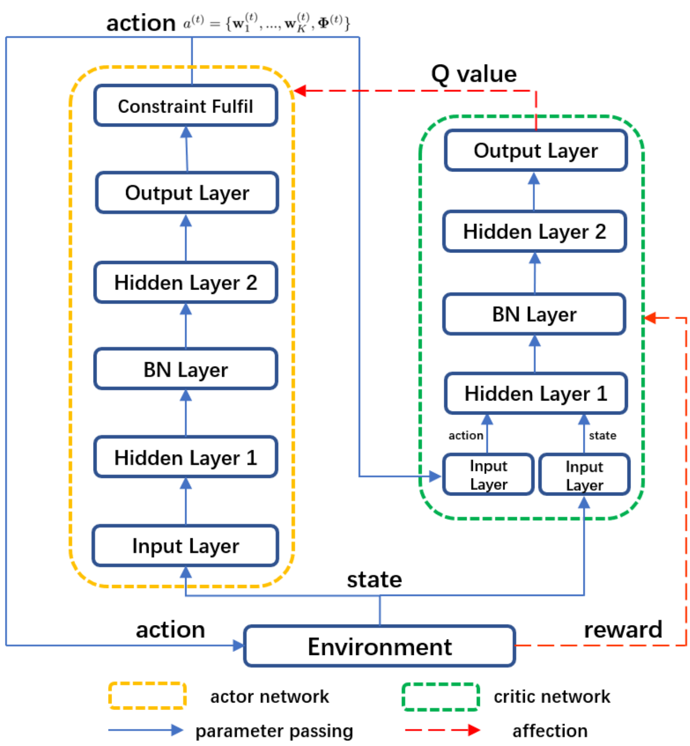

Figure 2.12. The block diagram of a DDPG model

According to the above description, QL and DQN are not suitable to solve the problem having a continuous action space. Although PG algorithms can be applied to continuous problems, the convergence performance is unsatisfactory in a wireless communication environment [25]. The idea of deep deterministic policy gradient (DDPG) is to combine DQN and PG together which can handle with the problem with a continues action space and also has a better convergence performance.

A DDPG framework shown in Fig. 2.12 is constructed by two neural networks. One named critic network estimates the Q function to evaluate the policy. Another named actor network generates policy and makes actions based on this policy. In training process, DDPG model updates the critic network through Bellman equation to minimising the MSE between the estimated Q value and the observed rewards plus the discounted value of the next state. The actor network is then updated using the PG.

The replay buffer is a crucial aspect in DDPG. The replay buffer stores the training samples, consisting of state, action, reward, next state tuples, which are randomly selected during training. This replay buffer helps break the temporal correlations in the data, making the learning process more stable.

DDPG has been successfully applied to a wide range of continuous control problems. DDPG is utilised in the chapter 4 in this thesis to maximise the sum rate of multiple users in a RIS-NOMA network.

Chapter 3

Joint Optimisation of Beamforming, Phase-Shifting and Power Allocation in a Multi-cluster RIS-NOMA Network

3.1 Introduction

The 5G communication system has been commercialised world-widely, and the beyond 5G (B5G) system starts attracting more and more researchers' attention due to its low energy consumption, high spectrum efficiency and massive multi-device interconnections [26]–[28]. In order to satisfy the increasing demand caused by the fast-growing number of users, various techniques, including millimetre wave [29], massive multi-inputs and multi-outputs (MIMO) system [30], and small cell [31], have been investigated and extensively used in practice. As a potential technique of B5G, non-orthogonal multiple access (NOMA) has received widespread attention due to its high spectral efficiency [6], [32]. Different from conventional orthogonal multiply access (OMA), such as frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA), and orthogonal frequency-division multiple access (OFDMA), NOMA allows multiple users to share the same time slot, frequency block and channel code, which dramatically increases the spectral efficiency. In particular, the users in a NOMA network usually adopt successive inference cancellation (SIC) to remove the inference from other NOMA users, which can efficiently improve the signal to interference and noise ratio (SINR) and reception reliability [33]. Recently, intelligent reflective surface (RIS) has also been proposed as a potential solution to further improve the performance of wireless networks, including enlarging the communication coverage, and improving transmission robustness. Specifically, the RIS can reflect the electromagnetic wave to extend the cover rage of the base station (BS). It also has the ability to tune the channel by adjusting the phase shift of each element, which will greatly improve the quality of users' received signal[34]. The typical architecture of RIS consists of a reflecting panel and a smart controller. The reflecting panel is composed of many reflecting elements and a control circuit. The control circuit is responsible for tuning the phase shift of each reflecting element. Moreover, the smart controller determines the reflection adaptation and also performs as a gateway to communicate with the BS. The smart controller can receive the control signal from the BS and then adjust the phase shift of each reflecting element [35].

3.1.1 Related Works

In literature, extensive research has been carried out for the NOMA technique, which has been combined with various state-of-the-art techniques such as MIMO and orthogonal time-frequency space modulation (OTFS)[36]–[40]. Recently, RIS has emerged as a kind of powerful equipment for wireless communication networks [8], [41], [42]. Among these works, RIS was proved as a perfect solution for a wireless communication network, where the channel will be intelligently reconfigured by the RIS[39], [43].

Motivated by the benefits from NOMA and RIS, the combination of NOMA and RIS has been recently proposed as a promising solution to improve the communication systems. There have been some ongoing works studying the combination of NOMA and RIS. Some recent research works such as [44], [45] considered a simple scenario where a single RIS serves two users in a downlink NOMA network. In [44], the authors minimised the transmit power at the BS by optimizing beamforming and RIS phase shift and also considered an improved quasi-degradation condition to guarantee that NOMA can achieve the capacity region with high possibility. In [45], the authors analysed two kinds of phase shift designs, namely random phase shifting and coherent phase shifting.

Moreover, there are many works considering an RIS-assisted NOMA network where a signal RIS serves multiple users [46]–[50]. The problems which have been researched can be divided into two categories, one is about the transmit power minimisation [46], [47] and the other is about the the sum-rate maximization[48]–[50]. For the transmit power minimisation problem, the authors in [48] minimised the total transmit power by optimizing beamforming vectors of each user and the phase shift design of the RIS in an RIS-empowered downlink NOMA network. [47] considered a single RIS assisted downlink NOMA network and adopted reinforcement learning to design the beamforming vectors which minimised the transmit power at the BS. Regarding to the sum rate maximization problem, the authors of [48] optimised the beamforming design to maximise the sum rate in a downlink MISO RIS aided NOMA system. [49] discussed a multi-channel downlink communications RIS-NOMA framework, where the sum rate of multiple NOMA users served by one RIS was maximised by optimizing resource allocation to each user and jointly considering channel assignment and decoding order. [50] considered an RIS-assisted uplink NOMA system where multiple

NOMA users can only transmit data through an RIS to the BS.

There are also some works considering a multi-cluster system mode, i.e., users are divided into different clusters [51], [52]. In [51], the authors discussed a downlink RIS-assisted NOMA network where two types of users named the central user and the cell edge user were assigned to different clusters. Each cluster had one central user, one cell edge user and one RIS serving all users. The authors minimised the transmit power at the BS by jointly optimizing the beamforming vectors of each user and the phase shift design of the RIS. In [52], the authors considered a multi-cluster and multi-BS RIS-aided NOMA network, where each cluster is served by its associated BS and one RIS serves all clusters. The sum rate was minimised by jointly optimizing power allocation and phase shift.

3.1.2 Motivation and Challenges

All the above works only consider one RIS. However the channel state of each user is related to its particular surrounding environment. Therefore, one single RIS might not be enough to reconfigure all users' channels simultaneously. Thus, multiple RISs are deployed to assist the users whose channel conditions are bad. One RIS can adjust its phase shift dedicatedly for its associated user to generate a better channel condition. This chapter considers a multi-cluster NOMA network, where each cluster has one RIS and the BS generates an unique beam for each cluster to serve all users located in this cluster.

With the considered scenario, there are a few challenges which need to be overcome. We consider a multi-user and multi-RIS scenario which increases the number of optimisation variables and hence make the optimisation more complicated than the case with a single RIS in the network. The joint optimisation problem contains three coupled variables, which is a non-convex problem and highly intractable. Therefore, the primal problem is divided into subproblems. Those subproblems are approximately transformed to the convex form but the feasibility of these subproblems cannot be guaranteed during the transformation. Moreover, due to the high quality of variables, the computing time of algorithms will be extensive.

3.1.3 Contributions

Different from the above mentioned works [51], [52], a new system model assisted with multiple RISs is adopted in this chapter. Then, a power minimisation problem is formulated, which is non-convex and highly intractable. A novel alternating algorithm is proposed to solve this non-convex problem efficiently. Finally, a low-complexity algorithm, which achieves a reasonable performance, is also provided. The contributions are summarised as follows:

- A multi-cluster RIS-NOMA system is considered, where each cluster contains two users served by one RIS. The transmit power minimisation problem with respect to the beamforming vector, the phase shifting matrix of RISs and the power allocation coefficient of each cluster is formulated. Each RIS can accomplish channel reconfiguration according to the channel condition between the BS and the cell edge user it serves, which intuitively yields a better performance than the scenario with the single RIS serving the whole system.
- The formulated problem is non-convex because three variables are highly coupled together. To solve the proposed optimisation problem, a novel alternating algorithm is proposed, where the primal problem is divided into the beamforming optimisation problem and the phase shift feasibility problem. However, the beamforming optimisation problem still has two variables coupled together, which causes the intractability. To address this challenge, the arithmetic and geometric means inequality is utilised to approximately transform the non-convex set to its convex upper bound. Then, the equivalence between Schur complement larger than zero and the positive semidefinite matrix and successive convex approximation (SCA) are applied to transfer another non-convex constraint to a convex form. Finally, the proposed alternating algorithm is utilised to iteratively solve those two subproblems.
- Some fixed points are introduced during the approximation. It is essential to obtain the initial choice of the fixed points to guarantee the feasibility of the beamforming optimisation problem. Therefore, a feasible initial points search algorithm is proposed, where an auxiliary variable forces all constraints to be feasible. Minimising this auxiliary variable until it equals to zero will find the feasible fixed points. The values of the fixed points when this auxiliary variable equals to zero can be the initial choice of the

fixed points for the proposed alternating algorithm.

- The complexity of the proposed alternating algorithm is high. To reduce the complexity, a simplified system model, where each cluster shares the same power allocation coefficient, is considered. With this assumption, the previous problem will be degraded into a simpler one with two coupled variables. A partial exhaustive search algorithm is proposed to solve this new problem. Compared with the alternating algorithm, the complexity is reduced but the performance is still reasonable.

3.1.4 Organisation

The rest of chapter is organised as follows. In Section II, a multi-cluster RIS-assisted NOMA downlink network is introduced and a power minimisation problem is formulated. In Section III, the solution to solve the formulated problem is introduced. In Section IV, the simplified optimisation problem and the partial exhaustive search based algorithm are introduced. In Section V, the convergence analysis of the algorithms and the simulation results are provided. Finally, a conclusion is summarised in Section VI.

3.2 System Model and Problem Formulation

3.2.1 System Model

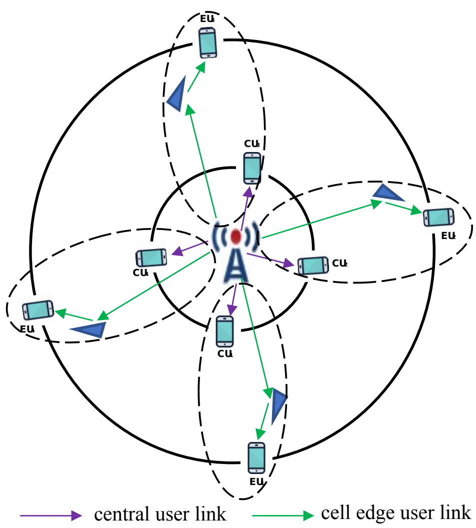

Figure 3.1. An RIS NOMA sytem model.

As shown in Fig.3.1, the multi-user downlink network contains two types of users, namely the central user and the cell edge user. They are served by the BS simultaneously. Generally, the central users are much closer to the base station than the cell edge users. It is assumed that there are K clusters and each cluster contains a central user, a cell edge user and an RIS. We use CU_k, EU_k and RIS_k to represent the central user, the cell edge user and the RIS in the k-th cluster, respectively. Each RIS is equipped with N passive reflecting elements and assists the cell edge user receiving signal from the BS. The BS is equipped with M $(K \leq M \leq 2K)$ antennas and generates K beams to serve K clusters. It is assumed that the direct links between the BS and all the cell edge users are not available due to blockage, and the RISs are implemented to reflect the signals sent by the BS to the cell edge users. Each cluster is far from others so the interference caused by the RISs serving the other clusters can be reasonably ignored. In practice, the surface area of the RIS hardware is very limited, which can only reflect partial electromagnetic waves sent by the BS. The energy of the reflected signal will be greatly attenuated if there is severe path loss or fading attenuation [35]. In each cluster, an RIS can be deployed carefully to ensure that it has strong connection to the cell edge user which does not have line-of-sight with the base station . As such, it is very likely that this RIS also has weak connections to those central users due to potential blockages [53]. The study for the case with direct links to those central users will be beyond the scope of this work, but it will be an important direction for future research. The locations of each RIS and each user will also affect the total transmit power. For simplicity, the system model will be presented by assuming that the distances between the RISs and the BS are the same. In Section V, simulation results will be presented to demonstrate the impact on the performance of the proposed algorithm with different BS-RIS distances. We note that the locations of the RISs provide another dimension of system optimisation, which is beyond the scope of this work and will be treated as an important direction for future research.

To improve the spectrum efficiency, each cluster will use the same frequency-time resource block but with different beams, which is similar to the principle of spatial division multiple access (SDMA). NOMA is adopted within each mean to further improve the spectrum efficiency. The BS assigns different power levels to the signals being sent to the users in each cluster. The base station broadcasts the superposition signal $\sum_{i=1}^{K} \mathbf{w}_i(\sqrt{\alpha_i}s_{i,c}+\sqrt{1-\alpha_i}s_{i,e})$, where $\mathbf{w}_i \in \mathbb{C}^M$ denotes the beamforming vector in the i-th cluster and $i \in 1, 2, ..., K$. $s_{i,c}$ and $s_{i,e}$ denote the signals to be sent to the central user and the cell edge user in the i-th cluster, respectively, and α_i is the power allocation coefficient of CU_i, thus $1 - \alpha_i$ is the power coefficient of EU_i.

Since the RIS only reflects the signal to users, it is assumed there is no line-of-sight (LoS) component. Therefore, the channel between the RIS to the cell edge user follows Rayleigh fading in each cluster. When it comes to the base station, which serves as the signal source, it is essential to consider the LoS component. As a result, we employ a Rician fading channel model for both the link between the base station and the RIS and the link between the BS and the central user within each cluster. This modeling approach can be represented as follows:

$$\mathbf{f} = \sqrt{\frac{\kappa}{1+\kappa}}\mathbf{f}^{\mathrm{LoS}} + \sqrt{\frac{1}{1+\kappa}}\mathbf{f}^{\mathrm{nLoS}}, \tag{3.1}$$

where κ is the Rician factor, $\mathbf{f}^{\mathrm{LoS}}$ is the LoS component and $\mathbf{f}^{\mathrm{nLoS}}$ is the non-Los (nLoS) component following the Rayleigh distribution.

Channel estimation is crucial for an RIS-assisted network to realise the beamforming design and phase shift design. If perfect channel state information (CSI) is available, RIS is able to properly adjust the phase shift under the aid of a smart controller and the BS can generate the proper beams. In an RIS-assisted network, there are two types of channels, namely the BS-RIS channel and the RIS-user channel. These two channels are always coupled together at the receiver end. Individually estimating these two channels is the main challenge for RIS channel estimation. Typically, an RIS is implemented two operational modes, which are the estimation mode and the reflecting mode [46]. The RIS can be switched between these two modes. On the estimation mode, the RIS will adjust each reflecting element to a particular phase and then RIS channel estimation algorithms, e.g. passive channel estimation based on machine learning [54], are applied to acquire CSI. On the reflecting mode, the RIS performs like a mirror to reflect the signal sent by the BS. This work mainly focuses on the beamforming, power allocation and phase shift design, and it is assumed that the perfect CSI is available at every node. Therefore, the signal received at CU_k is given by

$$y_{k,c} = \underbrace{\mathbf{h}_{k,c}^H \mathbf{w}_k \sqrt{\alpha_k} s_{k,c}}_{\text{signal}} + \underbrace{\mathbf{h}_{k,c}^H \mathbf{w}_k \sqrt{1-\alpha_k} s_{k,e}}_{\text{intra-cluster interference}} + \underbrace{\mathbf{h}_{k,c}^H \sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{w}_i (\sqrt{\alpha_i} s_{i,c} + \sqrt{1-\alpha_i} s_{i,e})}_{\text{inter-cluster interference}} + w_{k,c}, \tag{3.2}$$

where $\mathbf{h}_{k,c} \in \mathbb{C}^{M\times 1}$ denotes the channel vector between the base station and CU_k, and $w_{k,c} \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN). Meanwhile, the signal received at EU_k is given by

$$\begin{aligned} y_{k,e} = & \underbrace{(\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k) \mathbf{w}_k \sqrt{1-\alpha_k} s_{k,e}}_{\text{signal}} + \underbrace{(\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k) \mathbf{w}_k \sqrt{\alpha_k} s_{k,c}}_{\text{intra-cluster interference}} \\ & \underbrace{+ (\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k) \sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{w}_i (\sqrt{\alpha_i} s_{i,c} + \sqrt{1-\alpha_i} s_{i,e})}_{\text{inter-cluster interference}} + w_{k,e}, \end{aligned} \tag{3.3}$$

where $\mathbf{G}_k \in \mathbb{C}^{N\times M}$ denotes the channel matrix between the BS and RIS$_k$, $w_{k,e} \sim \mathcal{CN}(0, \sigma^2)$ denotes AWGN, $\mathbf{h}_{k,e} \in \mathbb{C}^{N\times 1}$ denotes the channel vector between RIS$_k$ and EU_k and $\Theta_k = \text{diag}(\beta e^{j\theta_1^k}, ..., \beta e^{j\theta_n^k})$ is the phase shift matrix of RIS$_k$, where $\theta_n^k \in [0, 2\pi), n \in \{1, ..., N\}$ and $\beta \in [0, 1]$ denote the phase shift of each reflecting element n and amplitude coefficient on the signal, respectively. It is assumed that $\beta = 1$ given the fact that each reflecting element can only change the phase but not the amplitude of the reflected signal [47]. It is worth to point out that the RIS with $\beta = 1$ may not be the optimal choice in some scenarios which are related to physical layer security. Therefore, the design of amplitudes is still crucial. More detailed discussions about the choices of the reflecting amplitude and the phase shift can be found in [55]. Due to path loss, we consider that the signal can be only efficiently reflected by the RIS once. Moreover, the long distance that geographically separates each cluster justifies the assumption that the RIS in one cluster will not infect other clusters. SIC strategies will directly affect the power allocation level. Since the cell edge user is far from the BS whose signal will suffer severe large scale path loss, the strategy that SIC is only performed at the central user to eliminate the interference from its intra-cluster edge user and the cell edge user decodes its data directly is adopted. In this case, the cell edge user will be allocated more power. It is necessary to point out that other SIC strategies will result in different power allocation, which will be studied in the future research. Hence, the SINR of EU_k is given by

$$\text{SINR}_{k,e} = \frac{|\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k \mathbf{w}_k|^2 (1-\alpha_k)}{|\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k \mathbf{w}_k|^2 \alpha_k + \sum\limits_{\substack{i=1 \\ i \neq k}}^{K} |\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k \mathbf{w}_i|^2 + \sigma^2}, \tag{3.4}$$

where $|\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k \mathbf{w}_k|^2 \alpha_k$ is intra-cluster interference and $\sum_{\substack{i=1 \\ i \neq k}}^{K} |\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k \mathbf{w}_i|^2$ is inter-cluster interference. For the central users, they need to apply SIC to decode $s_{k,e}$ first and then remove it. Thus, the SINR of signal $s_{k,e}$ observed at CU_k can be expressed as follows:

$$\text{SINR}_{k,c \to e} = \frac{|\mathbf{h}_{k,c}^H \mathbf{w}_k|^2 (1-\alpha_k)}{|\mathbf{h}_{k,c}^H \mathbf{w}_k|^2 \alpha_k + \sum\limits_{\substack{i=1 \\ i \neq k}}^{K} |\mathbf{h}_{k,c}^H \mathbf{w}_i|^2 + \sigma^2}. \tag{3.5}$$

The SINR of CU_k to decode its own signal is given by

$$\text{SINR}_{k,c} = \frac{|\mathbf{h}_{k,c}^H \mathbf{w}_k|^2 \alpha_k}{\sum\limits_{\substack{i=1 \\ i \neq k}}^{K} |\mathbf{h}_{k,c}^H \mathbf{w}_i|^2 + \sigma^2}. \tag{3.6}$$

The design of beamforming vector is critical. Some existing works adopted block-diagonalization-based beamforming, e.g. using vectors from a FFT matrix, to cancel the inter-cluster interference[51]. However, in this chapter, beamforming vectors are deigned by applying convex optimisation directly, thus, the inter-cluster interference exists as noise which cannot be ignored.

3.2.2 Problem Formulation

In this section, a transmit power minimisation problem is formulated by jointly optimizing the beamforming vector $(\mathbf{w}_k, k \in \{1, ...K\})$, power allocation coefficients $(\alpha_k, k \in \{1, ...K\})$ and phase shifting matrix $(\Theta_k, k \in \{1, ...K\})$, while considering the quality of service (QoS) requirement and the constraints of reflecting elements. The considered transmit power minimisation problem can be formulated as follows:

$$\text{P0}: \min_{\boldsymbol{\alpha},\mathbf{w},\Theta} \quad \sum_{k=1}^{K} ||\mathbf{w}_k||^2 \tag{3.7a}$$

$$\text{s.t. } \log_2(1+\text{SINR}_{k,c}) \geq R_{k,c}, \quad \forall k \tag{3.7b}$$

$$\log_2(1+\min(\text{SINR}_{k,e}, \text{SINR}_{k,c\to e})) \geq R_{k,e}, \forall k \tag{3.7c}$$

$$0 \leq \theta_{k,n} \leq 2\pi, \quad \forall\, k,n \tag{3.7d}$$

$$|\Theta_{k,n,n}| \leq 1, \quad \forall\, k,n \tag{3.7e}$$

where $||\mathbf{w}_k||^2$ is the transmit power allocated to the cluster k, $R_{k,c}$ and $R_{k,e}$ denote the required minimum data rate of CU_k and EU_k, respectively. The constraints (3.7b) and (3.7c) indicate the QoS requirements of the central users and the cell edge users, (3.7d) defines the phase shift range of the reflecting elements and (3.7e) ensures that the RIS is a passive component.

Note that for many important applications for RIS, such as the next-generaiton Internet of Things (IoT), users, such as IoT sensors, can be severely energy constrained, which motivates the formulated power minimisation problem. In particular, this formulated optimisation problem can reduce the energy consumption at the IoT sensors, while guaranteeing their communication requirements [46], [51]. Alternatively, the energy efficiency optimisation problem can also be formulated for the addressed RIS-NOMA network, which is beyond the scope of this work but is an important direction for future research.

However, problem P0 is highly intractable due to the non-convex constraints (3.7b) and (3.7c). The non-convexity is caused by three highly coupled variables (i.e. $\mathbf{w}$, α and Θ). To efficiently solve this problem, SCA, SDR and the inequality approximation are adopted to develop an alternating algorithm to iteratively solve it.

3.3 Optimisation Solution

As discussed in the previous section, it is difficult to find the optimal solution of P0 due to its non-convexity. In this section, an alternating optimisation algorithm is proposed to solve P0 efficiently. The main idea of this algorithm is to divide the primal problem into two subproblems and solve them alternatively. In particular, P0 is divided to a beamforming optimisation subproblem and a feasible phase shifting matrix search subproblem. As shown later, each of the two subproblems is non-convex, and we will approximately transform them into the convex form, which can be solved efficiently by convex solvers, e.g., CVX in Matlab.

3.3.1 Beamforming Optimisation

For a given phase shifting matrix Θ, the concatenated channel respond $\mathbf{h}_{k,e}^H \Theta_k \mathbf{G}_k \in \mathbb{C}^{1\times M}$ is fixed. Thus, the beamforming optimisation problem can be formulated as

$$\text{P1}: \min_{\alpha,\mathbf{w}} \quad \sum_{k=1}^{K} ||\mathbf{w}_k||^2 \tag{3.8a}$$

$$\text{s.t.} \quad \log_2(1+\text{SINR}_{k,c}) \geq R_{k,c}, \quad \forall k \tag{3.8b}$$

$$\log_2(1+\text{SINR}_{k,e}) \geq R_{k,e}, \quad \forall k \tag{3.8c}$$

$$\log_2(1+\text{SINR}_{k,c\to e}) \geq R_{k,e}, \quad \forall k \tag{3.8d}$$

$$0 \leq \alpha_k \leq 1, \quad \forall k. \tag{3.8e}$$

P1 is non-convex because the beamforming vector and the power allocation coefficient are still coupled together in all constraints except (3.8e), which is challenging to be solved. It is noted that the rank-constrained semidefinite programming (SDP) problem can be approximated to a convex form. Therefore, after converting P1 into a SDP form, SDR can be applied to solve this problem.

First, the constraint (3.8c) needs to be approximately transformed into a convex form. According to (3.4), the constraint (3.8c) can be rewritten as follows:

$$\frac{|e_k^H \mathbf{D}_{k,e}\mathbf{G}_k\mathbf{w}_k|^2(1-\alpha_k)}{|e_k^H \mathbf{D}_{k,e}\mathbf{G}_k\mathbf{w}_k|^2\alpha + \sum_{\substack{i=1\\i\neq k}}^{K} |e_k^H \mathbf{D}_{k,e}\mathbf{G}_k\mathbf{w}_i|^2 + \sigma^2} \geq r_{k,e}, \tag{3.9}$$

where $r_{k,e} = 2^{R_{k,e}} - 1$, e_k is an $N \times 1$ vector containing all the diagonal elements of Θ_k^H, and $\mathbf{D}_{k,e}$ is a diagonal matrix, whose main diagonal elements are from the channel vector $\mathbf{h}_{k,e}^H$. After some algebraic transformations, (3.9) can be equivalently expressed as follows:

$$|e_k^H \mathbf{D}_{k,e}\mathbf{G}_k\mathbf{w}_k|^2(1+r_{k,e})\alpha_k \leq |e_k^H \mathbf{D}_{k,e}\mathbf{G}_k\mathbf{w}_k|^2 - \sum_{\substack{i=1\\i\neq k}}^{K} |e_k^H \mathbf{D}_{k,e}\mathbf{G}_k\mathbf{w}_i|^2 r_{k,e} - \sigma^2 r_{k,e}. \tag{3.10}$$

Since the CSI is perfectly known by the BS, the channel $e_k^H\mathbf{D}_{k,e}\mathbf{G}_k$ is fixed with a given phase shifting matrix. For simply notation, we replace $e_k^H\mathbf{D}_{k,e}\mathbf{G}_k$ with $\mathbf{z}_{k,e}^H$ and rewrite (3.10) as follows:

$$\alpha_k|\mathbf{z}_{k,e}^H\mathbf{w}_k|^2 \le \frac{|\mathbf{z}_{k,e}^H\mathbf{w}_k|^2}{1+r_{k,e}} - \left(\sum_{\substack{i=1\\i\neq k}}^{K}|\mathbf{z}_{k,e}^H\mathbf{w}_i|^2+\sigma^2\right)\frac{r_{k,e}}{1+r_{k.e}}, \tag{3.11}$$

where $\mathbf{z}_{k,e}^H = e_k^H\mathbf{D}_{k,e}\mathbf{G}_k$. The quadratic form $|\mathbf{z}_{k,e}^H\mathbf{w}_k|^2$ in (3.11) can be rewritten as $\mathbf{w}_k^H\mathbf{Z}_{k,e}\mathbf{w}_k$, where $\mathbf{Z}_{k,e} = \mathbf{z}_{k,e}\mathbf{z}_{k,e}^H$. A slack matrix $\mathbf{W}_k = \mathbf{w}_k\mathbf{w}_k^H$ is introduced, which is a rank-one positive semidefinite (PSD) matrix. It is known that $\mathbf{w}_k^H\mathbf{Z}_{k,e}\mathbf{w}_k = \mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)$ from SDR. Moreover, the rank of $\mathbf{W}_k$ has to be 1 and $\mathbf{W}_k$ is a PSD because of $\mathbf{W}_k = \mathbf{w}_k\mathbf{w}_k^H$. Then the constraint (3.11) can be equivalently rewritten as follows:

$$\alpha_k\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k) \le \frac{\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{1+r_{k,e}} - \left(\sum_{\substack{i=1\\i\neq k}}^{K}\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_i)+\sigma^2\right)\frac{r_{k,e}}{1+r_{k.e}} \tag{3.12}$$

$$\mathbf{W}_k \succcurlyeq 0 \tag{3.13}$$

$$\mathrm{Rank}(\mathbf{W}_\mathrm{k}) = 1. \tag{3.14}$$

From (3.12), it is noted that the right hand side of (3.12) is a liner combination of two convex terms with respect to $\mathbf{W}_k$, which is convex. The only obstacle is the left hand side, which is a bilinear term constructed by α_k and $\mathrm{Tr}(\mathbf{Z}_{\mathrm{k,e}}\mathbf{W}_\mathrm{k})$. To make this constraint a convex set, we need to approximately transform the non-convexity function $\alpha_k\mathrm{Tr}(\mathbf{Z}_{\mathrm{k,e}}\mathbf{W}_\mathrm{k})$ to a convex form. Inspired by the inequality of arithmetic and geometric means $2ab \le a^2+b^2$, where a and b are both non-negative numbers, we have the inequality that

$$2\alpha_k\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k) \le (\alpha_k c_k)^2 + \left(\frac{(\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{c_k}\right)^2, \tag{3.15}$$

where c_k is a fixed point. From (3.15), it is noted that $(\alpha_k c_k)^2 + \left(\frac{(\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{c_k}\right)^2$ is an upper bound of $2\alpha_k\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)$ and is a liner combination of two convex terms. Therefore, the non-convex feasible set of the left hand side term can be upper bounded by a convex set $\frac{1}{2}(\alpha_k^2 + \mathrm{Tr}(\mathbf{Z}_{\mathrm{k,e}}\mathbf{W}_\mathrm{k})^2)$. To tighten this upper bound in each iteration of the proposed iterative algorithm, updating rule of c_k in each iteration needs to be defined.

Lemma 1. *The fixed point c_k at the m-th iteration can be updated by:*

$$c_k^{(m)} = \sqrt{\frac{\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k^{(m-1)})}{\alpha_k^{(m-1)}}} \tag{3.16}$$

Proof. A difference function of the original function $2\alpha_k \mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)$ and its approximated upper bound is defined as follows:

$$\mathcal{F}(c_k) = 2\alpha_k \mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k) - (\alpha_k c_k)^2 - \left(\frac{(\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{c_k}\right)^2. \tag{3.17}$$

When the function (3.17) equals to 0, the equality of (3.15) holds, which tightens the upper bound. From (3.15), it is noted that the maximum value of function $\mathcal{F}(c_k)$ is 0. Since

$$\frac{\partial^2 \mathcal{F}(c_k)}{\partial c_k^2} = -2\alpha_k - \frac{6\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{c_k^4} \leq 0, \tag{3.18}$$

when $\alpha_k \geq 0$ and $\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k) \geq 0$, the function $\mathcal{F}(c_k)$ is a concave function with respect to c_k. According to the Karush–Kuhn–Tucker (KKT) conditions, the maximum value of a concave function is obtained by letting the first order derivative equal to 0. Thus, the optimal value of c_k, defined as c_k^*, can be obtained by $\frac{\partial \mathcal{F}(c_k)}{\partial c_k} = 0$, then c_k^* can be given by

$$c_k^* = \sqrt{\frac{\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{\alpha_k}}. \tag{3.19}$$

□

Hence, the constraint (3.12) can be approximated as follows:

$$(\alpha_k c_k)^2 + \left(\frac{(\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{c_k}\right)^2 \leq 2\frac{\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{1+r_{k,e}} - 2\left(\sum_{\substack{i=1\\i\neq k}}^{K} \mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_i) + \sigma^2\right)\frac{r_{k,e}}{1+r_{k.e}}. \tag{3.20}$$

It is noted that the left hand side of (3.20) is convex and the right hand side of (3.20) is an affine function, which means that the constraint (3.20) is a convex set. Eventually, (3.8c) can be approximated by (3.13), (3.14) and (3.20).

For handling with the next non-convex constraint (3.8d), after some algebraic manipulations, (3.8d) can be rewritten as follows:

$$\alpha_k|\mathbf{h}_{k,c}^H\mathbf{w}_k|^2 \leq \frac{|\mathbf{h}_{k,c}^H\mathbf{w}_k|^2}{1+r_{k,e}} - (\sum_{\substack{i=1\\i\neq k}}^{K}|\mathbf{h}_{k,c}^H\mathbf{w}_i|^2+\sigma^2)\frac{r_{k,e}}{1+r_{k.e}}. \tag{3.21}$$

It is worth to point out that (3.21) has the same form as (3.11). Similarly, the method allied to (3.11) can be efficiently applied to (3.21) to yield a convex form. Therefore, (3.21) can be approximately transformed to

$$(\alpha_k d_k)^2 + \left(\frac{(\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)}{d_k}\right)^2 \leq 2\frac{\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)}{1+r_{k,e}} - 2\left(\sum_{\substack{i=1\\i\neq k}}^{K}\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i)+\sigma^2\right)\frac{r_{k,e}}{1+r_{k.e}}, \tag{3.22}$$

$$(3.13),\quad (3.14),$$

where $\mathbf{H}_{k,c} = \mathbf{h}_{k,c}\mathbf{h}_{k,c}^H$, and d_k is a fixed point. At the m-th iteration, d_k can be updated as follows:

$$d_k^{(m)} = \sqrt{\frac{\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k^{(m-1)})}{\alpha_k^{(m-1)}}}. \tag{3.23}$$

Eventually, (3.8d) can be approximated by (3.13), (3.14) and (3.22).

Now, we focus on the last non-convex constraint (3.8b). First, it can be rewritten as follows:

$$\alpha_k\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k) \geq \sum_{\substack{i=1\\i\neq k}}^{K}\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i)r_{k,c}+\sigma^2 r_{k,c} \tag{3.24}$$

where $r_{k,c} = 2^{R_{k,c}}-1$. Though (3.24) also has a bilinear term $\alpha_k\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)$, the method which has been successfully applied to the constraint (3.8c) and (3.8d) cannot be straightforwardly applied. Replacing $\alpha_k\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)$ with the upper bound through the inequality of arithmetic and geometric means does not work because it is located at the left hand side of $\geq$ sign, which causes this inequality to be concave not convex. Hence, another method is proposed to deal with this constraint. First, we introduce a slack variable t_k and (3.8b) can be transformed to

$$\alpha_k\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k) \geq t_k^2 \tag{3.25}$$

$$t_k^2 \geq \sum_{\substack{i=1 \\ i \neq k}}^{K} \mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) r_{k,c} + \sigma^2 r_{k,c}. \tag{3.26}$$

It can be straightforwardly shown that neither of (3.25) and (3.26) is convex. According to the Schur complement theory [11], it is known that the sufficient and necessary condition for a matrix to be PSD is that its Schur complement is greater than zero. Moreover, a PSD matrix is a convex constraint. After a simple transformation, (3.25) can be rewritten as follows:

$$\alpha_k - \frac{t_k^2}{\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)} \geq 0, \tag{3.27}$$

which is equivalent to

$$\begin{bmatrix} \alpha_k & t_k \\ t_k & \mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) \end{bmatrix} \succcurlyeq 0. \tag{3.28}$$

Constraints (3.27) and (3.28) are mutually sufficient, and constraint (3.28) is convex. Now, we deal with the constraint (3.26). It is noted that t_k^2 is on the left hand side of the greater sign, which makes the whole constraint a non-convex set. SCA is utilised to deal with this, where the first order Taylor series approximation is adopted to approximate the quadratic form (3.26) as

$$t_{k,0}^2 + 2t_{k,0}(t_k - t_{k,0}) \geq \sum_{\substack{i=1 \\ i \neq k}}^{K} \mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) r_{k,c} + \sigma^2 r_{k,c} \tag{3.29}$$

where $t_{k,0}$ is a fixed point introduced by SCA. The updating rule of $t_{k,0}$ at the m-th iteration is given by

$$t_{k,0}^{(m)} = t_k^{(m-1)}. \tag{3.30}$$

The final obstacle to deal with this problem arises from the rank-one constraint (3.14). By applying SDR, the rank-one constraint is omitted to make the whole problem tractable. Thus, P1 is eventually transformed to

$$\mathrm{P2} : \min_{\alpha, \mathbf{W}, t} \quad \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k) \tag{3.31a}$$

$$\text{s.t.}\quad (\alpha_k c_k)^2 + \left(\frac{(\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{c_k}\right)^2 \leq$$

$$2\frac{\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{1+r_{k,e}} - 2\left(\sum_{\substack{i=1\\i\neq k}}^{K}\mathrm{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_i)+\sigma^2\right)\frac{r_{k,e}}{1+r_{k.e}}, \forall k \tag{3.31b}$$

$$(\alpha_k d_k)^2 + \left(\frac{(\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)}{d_k}\right)^2 \leq$$

$$2\frac{\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)}{1+r_{k,e}} - 2\left(\sum_{\substack{i=1\\i\neq k}}^{K}\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i)+\sigma^2\right)\frac{r_{k,e}}{1+r_{k.e}}, \forall k \tag{3.31c}$$

$$\begin{bmatrix}\alpha_k & t_k\\ t_k & \mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i)\end{bmatrix} \succcurlyeq 0, \forall k \tag{3.31d}$$

$$t_{k,0}^2 + 2t_{k,0}(t_k - t_{k,0}) \geq \sum_{\substack{i=1\\i\neq k}}^{K}\mathrm{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i)r_{k,c} + \sigma^2 r_{k,c}, \forall k \tag{3.31e}$$

$$0 \leq \alpha_k \leq 1, \forall k. \tag{3.31f}$$

Algorithm 1 Initial Point Search Algorithm

1: **Initialise:** $c_k^{(0)}$, $d_k^{(0)}$, $t_{k,0}^{(0)}$ $\quad\forall k$, $\epsilon = 0.00001$, $i = 0$, $q^{(0)} = 100$.
2: **while** $q^{(i)} > \epsilon$ **do**
3: $\quad i = i + 1$.
4: $\quad$ Update $\mathbf{W}_k^{(i)}$, $\alpha_k^{(i)}$ and $q^{(i)}$ with fixed $c_k^{(i-1)}$, $d_k^{(i-1)}$, $t_{k,0}^{(i-1)}$ by solving P3.
5: $\quad$ Update $c_k^{(i)}$, $d_k^{(i)}$, and $t_{k,0}^{(i)}$ based on (3.16), (3.23) and (3.30) respectively.
6: **end while**
7: **Output** $c_k^{(i)}$, $d_k^{(i)}$, and $t_{k,0}^{(i)}$.

Since the restriction of rank one is removed, P2 is a convex problem and can be efficiently solved by convex optimisation toolboxes, for instance, CVX. It is noted that P1 and P2 have different optimisation valuables. It is crucial to extract the optimal solution of P1 from the optimal solution of P2. We define the optimal solution of P2 as $\mathbf{W}_k^*, \forall k$, and each $\mathbf{W}_k^*$ is a positive semidefinite matrix. However, the optimal solution of P1 will not be obtained from the optimal solution of P2 unless the rank of $\mathbf{W}_k^*, \forall k$ is 1. If the rank of $\mathbf{W}_k^*$ is not 1, Gaussian randomization [12] is applied to alternatively obtain a suboptimal solution of P1. Specifically, several random vectors $\xi_k \sim \mathcal{N}(0, \mathbf{W}_k^*)$ will be generated and stored in a vector set. The one from this set which can satisfy all the constraints in P1 and also yield the best objective of P1 will be the suboptimal solution of P1.

Before solving P2, three fixed points, c_k, d_k and $t_{k,0}, \forall k$ need to be initialised. It is noted that initializing them randomly will make the formulated problem infeasible. Hence, a feasible initial points search algorithm is proposed to find the feasible fixed points to make P2 solvable. From P2, it is noted that the fixed points c_k. d_k and $t_{k,0}$ must satisfy the constraints (3.31b), (3.31c) and (3.31d). An auxiliary variable q, which intentionally relaxes the constraints to enlarge the feasible set, is introduced to address this problem. The initial point search problem can be formulated as follows:

$$\text{P3}: \min_{\alpha, \mathbf{W}, t, q} \quad q \tag{3.32a}$$

$$\text{s.t.} \quad (\alpha_k c_k)^2 + \left(\frac{(\text{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{c_k}\right)^2 - q \leq$$

$$2\frac{\text{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{1+r_{k,e}} - 2(\sum_{\substack{i=1 \\ i\neq k}}^{K} \text{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_i) + \sigma^2)\frac{r_{k,e}}{1+r_{k.e}}, \forall k \tag{3.32b}$$

$$(\alpha_k d_k)^2 + \left(\frac{(\text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)}{d_k}\right)^2 - q \leq$$

$$2\frac{\text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k)}{1+r_{k,e}} - 2(\sum_{\substack{i=1 \\ i\neq k}}^{K} \text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) + \sigma^2)\frac{r_{k,e}}{1+r_{k.e}}, \forall k \tag{3.32c}$$

$$\begin{bmatrix} \alpha_k & t_k \\ t_k & \text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) \end{bmatrix} \succcurlyeq 0, \forall k \tag{3.32d}$$

$$t_{k,0}^2 + 2t_{k,0}(t_k - t_{k,0}) + q \geq \sum_{\substack{i=1 \\ i\neq k}}^{K} \text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) r_{k,c} + \sigma^2 r_{k,c}, \forall k \tag{3.32e}$$

$$0 \leq \alpha_k \leq 1, \forall k \tag{3.32f}$$

$$q \geq 0. \tag{3.32g}$$

Specifically, when q equals to 0, all constraints in P3 are exactly the same as the constraints in P2 and the obtained values of c_k, d_k and $t_{k,0}$ can be the initial points of P2, which will guarantee the feasibility. It is noted that the objective function is an affine function and all constraints are convex so it can be solved easily by CVX. To solve P3 efficiently, an iterative algorithm shown as Algorithm 1 is proposed. It is worth to point out that, unlike P2, the initial points $c_k^{(0)}$, $d_k^{(0)}$ and $t_{k,0}^{(0)}$ in P3 can be generated randomly because the feasibility of P3 can be always guaranteed by q.

After deciding the fixed points, the last challenge for solving the beamforming optimi-

Algorithm 2 The Beamforming Optimisation Algorithm

1: **Initialise:** fixed feasible points $\{c_k^{*(0)}, d_k^{*(0)}, t_{k,0}^{*(0)},\}$ $\forall k$, $\epsilon = 0.001$, $m = 0$.
2: **while** $\sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k^{(m-1)}) - \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k^{(m)}) \geq \epsilon$ **do**
3: Update beamforming matrix $\{\mathbf{W}_k^{(m)}, \alpha_k^{(m)}\}, \forall k$ by solving P2 with the fixed feasible point$\{c_k^{*(m)}, d_k^{*(m)}, t_{k,0}^{*(m)}\}, \forall k$.
4: Update $\{c_k^{*(m)}, d_k^{*(m)}, t_{k,0}^{*(m)}\}, \forall k$ based on (3.16), (3.23) and (3.30) respectively.
5: $m = m + 1$.
6: **end while**
7: Update $\alpha_k^* = \alpha_k^{(m)}, \forall k$
8: Update beamforming vector $\mathbf{w}_k^*, \forall k$ by decomposing $\mathbf{W}_k^{(m)}, \forall k$ based on Gaussian Randomization method.
9: **Output** $\{\mathbf{w}_k^*, \alpha_k^*\}, \forall k$

sation problem has been removed. To solve this problem efficiently, an iterative algorithm is designed to solve P2 iteratively. The details of the algorithm are shown in Algorithm 2. Specifically, the fixed initial points $\{c_k^{*(0)}, d_k^{*(0)}, t_{k,0}^{*(0)}\}$ $\forall k$ are obtained from Algorithm 1.

3.3.2 Phase Shifting Optimisation

In this section, we focus on the phase shifting optimisation. The phase shifting optimisation can be transformed to a feasibility problem since the objective function in the primal problem does not contain the phase shifting parameter $\Theta_k, \forall k$. Only the constraints (3.7c), (3.7d) and (3.7e) in the primal problem contain the phase shifting parameter and (3.7c) can be equivalently divided into (3.8c) and (3.8d), where only (3.8c) contains the phase shifting parameter. Therefore, given the beamforming vectors, the phase shift feasibility problem can be written as follows:

$$\text{P4 : find} \quad \Theta \tag{3.33a}$$

$$\text{s.t.} \quad \log_2(1 + \mathrm{SINR}_{k,e}) \geq R_{k,e}, \forall k \tag{3.33b}$$

$$0 \leq \theta_{k,n} \leq 2\pi, \forall k, n \tag{3.33c}$$

$$|\Theta_{k,n,n}| = 1, \forall k, n. \tag{3.33d}$$

It is straighforward to find out that the non-convexity arises from the constraint (3.33b). The first step is to transform this non-convex constraint to be a convex constraint. Thus, (3.33b) can be rewritten as follows:

$$|\mathbf{h}_{k,e}^{H}\Gamma_{\mathbf{p}_k}\mathbf{e}_k|^2(1+r_{k,e})\alpha_k \leq |\mathbf{h}_{k,e}^{H}\Gamma_{\mathbf{p}_k}\mathbf{e}_k|^2 - \sum_{\substack{i=1\\ i\neq k}}^{K}|\mathbf{h}_{k,e}^{H}\Gamma_{\mathbf{p}_i}\mathbf{e}_k|^2 - \sigma^2 r_{k,e}, \tag{3.34}$$

where $\Gamma_{\mathbf{p}_i}$ is a diagonal matrix whose main diagonal elements are from $\mathbf{p}_i = \mathbf{G}_k\mathbf{w}_i$ and $\mathbf{e}_k$ is the phase shifting vector. However, with $\mathbf{W}_k, \alpha_k, \forall k$ already obtained from the beamforming optimisation problem, the constraint (3.34) is a quartic form with respect to $\mathbf{e}_k$. For simplicity, we substitute $\mathbf{h}_{k,e}^{H}\Gamma_{\mathbf{p}_i}$ with $\mathbf{r}_{k,e}^{iH}$. From [12], it is known that a quartic form can be equivalently transformed to a linear form with a rank-one constraint. Thus, (3.34) can be expressed as follows:

$$\mathrm{Tr}(\mathbf{R}_{k,e}^{k}\mathbf{V}_k)(1+r_{k,e})\alpha_k \leq \mathrm{Tr}(\mathbf{R}_{k,e}^{k}\mathbf{V}_k) - \sum_{\substack{i=1\\ i\neq k}}^{K}\mathrm{Tr}(\mathbf{R}_{k,e}^{i}\mathbf{V}_i) - \sigma^2 r_{k,e} \tag{3.35}$$

$$\mathbf{V}_k \succcurlyeq 0 \tag{3.36}$$

$$\mathrm{Rank}(\mathbf{V}_k) = 1, \tag{3.37}$$

where $\mathbf{R}_{k,e}^{i} = \mathbf{r}_{k,e}^{i}\mathbf{r}_{k,e}^{iH}$ and $\mathbf{V}_i = \mathbf{e}_i\mathbf{e}_i^{H}$. Given $\mathbf{w}_k, \alpha_k, \forall k$, (3.35) is an affine constraint. The rank-one constraint will make the whole problem intractable, thus SDR is adopted again to remove this rank-one constraint. Then, P4 can be transformed as follows:

$$\mathrm{P5}: \text{find} \quad \mathbf{V}_k, \quad \forall k \tag{3.38a}$$

$$\text{s.t.} \qquad (3.35), \quad \forall k \tag{3.38b}$$

$$\mathbf{V}_k \succcurlyeq 0, \quad \forall k \tag{3.38c}$$

$$|\mathbf{V}_{k,n,n}| = 1, \quad \forall k, n. \tag{3.38d}$$

P5 is a convex problem, which can be solved by CVX efficiently. Since the rank-one constraint is removed, the optimal solution of P5 may not be the optimal solution of P4. Therefore, Gaussian randomization will be applied to achieve a suboptimal solution for P4.

Algorithm 3 The Proposed Alternating Algorithm

1: **Initialise:** $\epsilon = 0.001, j = 0$.

2: **while** $\sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k^{*(j-1)}) - \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k^{*(j)}) \geq \epsilon$ **do**

3: Searching initial fixed feasible point $\{c_k^{*(j)}, d_k^{*(j)}, t_{k,0}^{*(j)}\}, \forall k$ based on Algorithm 1.

4: Update $\{\mathbf{W}_k^{*(j)}, \mathbf{w}_k^{*(j)}, \alpha_k^{*(j)}\}, \forall k$ based on Algorithm 2.

5: Update $\mathbf{V}_k^{*(j)}, \forall k$ by solving P5 with $\{\mathbf{w}_k^{*(j)}, \alpha_k^{*(j)}\}, \forall k$

6: Update phase shift vector $\mathbf{e}_k^{*(j)}, \forall k$ by decomposing $\mathbf{V}_k^{*(j)}, \forall k$ based on Gaussian Randomization method.

7: $j = j + 1$

8: **end while**

9: **Output** $\{\mathbf{w}_k^{*(j)}, \alpha_k^{*(j)}, \mathbf{e}_k^{*(j)}\}, \forall k$.

3.3.3 Algorithm Design

The detail of the proposed alternating algorithm are illustrated in Algorithm 3, where P2 and P5 are alternately solved until the convergence metric is satisfied. At the i-th iteration of Algorithm 3, first, the initial points are obtained by Algorithm 1. Then, the algorithm begins to solve the beamforming optimisation problem by solving P2 through Algorithm 2. Then, the algorithm starts to solve phase shifting feasibility problem by solving P5 (step 5 and step 6) to obtain a feasible phase shift vector $\mathbf{e}_k^{*(i)}, \forall k$. The feasible phase shifting vector of this current iteration will be used as a given phase shift for the beamforming optimisation in the next iteration. It is worth to point out that after each iteration, the channel state will change with the new obtained phase shifting vector $\mathbf{e}_k, \forall k$, so it is critical to search new feasible fixed points (step 3) before solving P2, which necessarily guarantees that P2 is always feasible.

It is worth to point out that there are three optimisation variables coupled together in constraints (3.7b) and (3.7c), which are non-convex as well. Therefore, P0 is a NP-hard problem , i.e., it is difficult to solve it in polynomial time. It is difficult to find the global optimal solution of P0 by applying convex optimisation. In Algorithm 3, the alternating algorithm and a few approximations are adopted to transform P0 to a solvable convex problem. Therefore, Algorithm 3 only provides a suboptimal solution for P0.

3.3.4 Complexity analysis

The worst complexity of solving a SDR problem through CVX provided by [12] is

$$\mathcal{O}(\max\{m, n\}^4 n^{1/2} \log(1/\epsilon_c)), \tag{3.39}$$

where n is the problem size, and m is the number of constraints and ϵ_c is the accuracy of the algorithm that CVX adopts. It is assumed that the problem size is greater than the number of constraints, then the complexity of CVX to solve a SDR problem can be expressed as

$$\mathcal{O}(n^{4.5} \log(1/\epsilon_c)). \tag{3.40}$$

Algorithm 1 is essentially to solve a SDR problem multiple times until the accuracy is satisfied. Thus, the complexity of Algorithm 1 is

$$\mathcal{O}\left(n_1^{4.5} \log\left(\frac{1}{\epsilon_c}\right) \log\left(\frac{1}{\epsilon_1}\right)\right), \tag{3.41}$$

where n_1 is the problem size of P3 and ϵ_1 is the accuracy of Algorithm 1. Algorithm 2 is similar to Algorithm 1, which is also to solve a SDR problem multiple times and hence P2 has the same size as P3. Thus, the complexity of Algorithm 2 can be expressed as follows:

$$\mathcal{O}\left(n_1^{4.5} \log\left(\frac{1}{\epsilon_c}\right) \log\left(\frac{1}{\epsilon_2}\right)\right), \tag{3.42}$$

where ϵ_2 is the accuracy of Algorithm 2. Now, we have the complexities of step 3 and step 4 in Algorithm 3. The last one is the complexity of step 5. It is easy to find out that a single SDR problem is solved in the step 5, so the complexity is

$$\mathcal{O}(n_2^{4.5} \log(1/\epsilon_c), \tag{3.43}$$

where n_2 is the problem size of P5. Finally, the complexity of the proposed algorithm is given by

$$\mathcal{O}(\mathcal{O}_1 \log(1/\epsilon_3)), \tag{3.44}$$

where

$$\mathcal{O}_1 = n_1^{4.5}\left(\log\left(\frac{1}{\epsilon_c}\right) \log\left(\frac{1}{\epsilon_1}\right) + \log\left(\frac{1}{\epsilon_c}\right) \log\left(\frac{1}{\epsilon_2}\right)\right) + n_2^{4.5} \log(1/\epsilon_c).$$

It is noted that the complexity of the proposed algorithm is quite high. In this case, the computing time to run the algorithm is long. In a time sensitive system, the algorithm needs to be simplified to reduce the computing time. In the next section, a partial exhaustive search algorithm with lower complexity is proposed.

3.4 Partial Exhaustive Search Algorithm

In this section, a simple algorithm based on partial exhaustive search is proposed, which can significantly reduce computation complexity. The main idea of this partial exhaustive search algorithm is to assume that all the clusters share the same power allocation coefficient, of which the optimal value can be obtained by an exhaustive search within the range [0, 1]. The primal problem can also be divided into the beamforming optimisation problem and the phase shifting feasibility problem.

Since each cluster shares the same power coefficient, the power coefficient is first fixed in each searching progress so only the beamforming vector and the phase shifting vector need to be optimised in these two subproblems. It is noted that these two subproblems can be reduced to the QCQP problem, which is a classic form in convex optimisation theory. SDR is widely used as one of the most common methods to efficiently solve the QCQP problem. Two subprobelms are formulated as P6 and P7. P6 and P7 can be obtained through the basic SDR theory and some simple algebraic transformations, where the derivation is omitted in this chapter due to space limitations.

$$\text{P6}: \quad \min_{\mathbf{w}} \quad \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k) \tag{3.45a}$$

$$\text{s.t. } \alpha\text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k) \geq \sum_{\substack{i=1\\i\neq k}}^{K} \text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) r_{k,c} + \sigma^2 r_{k,c}, \forall k \tag{3.45b}$$

$$\alpha\text{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k) \leq \frac{\text{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{1+r_{k,e}} - (\sum_{\substack{i=1\\i\neq k}}^{K} \text{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_i) + \sigma^2)\frac{r_{k,e}}{1+r_{k.e}}, \forall k \tag{3.45c}$$

$$\alpha\text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_k) \leq \frac{\text{Tr}(\mathbf{Z}_{k,e}\mathbf{W}_k)}{1+r_{k,e}} - (\sum_{\substack{i=1\\i\neq k}}^{K} \text{Tr}(\mathbf{H}_{k,c}\mathbf{W}_i) + \sigma^2)\frac{r_{k,e}}{1+r_{k.e}}, \forall k \tag{3.45d}$$

$$\mathbf{W}_k \succcurlyeq 0, \forall k \tag{3.45e}$$

where $\mathbf{Z}_{k,e}, \mathbf{W}_k, \forall k$ in P6 are the same as those in P2.

$$\text{P7} : \text{find} \quad \mathbf{V}_k, \quad \forall k \tag{3.46a}$$

$$
\begin{aligned}
\text{s.t.}\quad & \mathrm{Tr}(\mathbf{R}_{k,e}^{k}\mathbf{V}_k)(1+r_{k,e})\alpha \leq \\
& \mathrm{Tr}(\mathbf{R}_{k,e}^{k}\mathbf{V}_k) - \sum_{\substack{i=1\\ i\neq k}}^{K} \mathrm{Tr}(\mathbf{R}_{k,e}^{i}\mathbf{V}_i) - \sigma^2 r_{k,e}, \forall k && \text{(3.46b)}\\
& \mathbf{V}_k \succcurlyeq 0, \quad \forall k && \text{(3.46c)}\\
& \mathbf{V}_{k,n,n} = 1, \quad \forall k, n && \text{(3.46d)}
\end{aligned}
$$

where $\mathbf{R}_{k,e}^{i}, \forall i, k$ and $\mathbf{V}_k, \forall k$ are the same as those in P5. The detail of the partial exhaustive search algorithm is illustrated in Algorithm 4.

In each search progress, the algorithm will solve two SDR problems with different sizes n_1 and n_2, which are the problem sizes of P6 and P7. Therefore, the complexity of Algorithm 4 can be expressed as follows:

$$
\mathcal{O}\left(I\left(n_1^{4.5}\log(1/\epsilon_c) + n_2^{4.5}\log(1/\epsilon_c)\right)\right). \tag{3.47}
$$

I is the number of searches, which depends on the search step α. Obviously, the partial exhaustive search algorithm has a lower complexity than the complexity of the proposed alternating algorithm. It is worth to point out that the performance of this partial exhaustive search algorithm is related to the step size $\Delta\alpha$. A smaller step size will yield a better performance. However, according to (3.47), when the step size decreases, the complexity of the algorithm will increase. It is important to find a balance between performance and complexity. In numerical results, the performance of the partial exhaustive search algorithm with different step sizes is provided.

3.5 Simulation Results

In this section, we evaluate all simulation results of the proposed algorithms. In simulations, channel gains are generated by

$$
\mathbf{h}_{k,e} = \frac{\mathbf{h}_{k,e}^{*}}{\sqrt{d_0^{\alpha_0}}} \quad \mathbf{G}_k = \frac{\mathbf{G}_k^{*}}{\sqrt{d_1^{\alpha_1}}} \quad \mathbf{h}_{k,c} = \frac{\mathbf{h}_{k,c}^{*}}{\sqrt{d_2^{\alpha_2}}} \tag{3.48}
$$

Algorithm 4 The Partial Exhaustive Search Algorithm

Initialization $P_{opt} = 10000, \alpha_{opt} = 0, \mathbf{w}_k^*, \mathbf{e}_k^*, \forall k$
for $\alpha = 0.1 : 0.1 : 0.9$ **do**
  Initialization $\epsilon = 0.001, i = 0, \mathbf{e}_k^{(0)}$
  while $\sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k^{(i-1)}) - \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k^{(i)}) > \epsilon$ **do**
    Update $\mathbf{W}_k^{(i)}, \forall k$ by solving P6.
    Update $\mathbf{w}_k^{(i)}, \forall k$ by decomposing $\mathbf{W}_k^{(i)}, \forall k$ based on Gaussian Randomization method.
    Update $\mathbf{V}_k^{(i)}, \forall k$ by solving P7 based on given $\mathbf{w}_k^{(i)}, \forall k$.
    Update $\mathbf{e}_k^{(i)}, \forall k$ by decomposing $\mathbf{V}_k^{(i)}, \forall k$ based on Gaussian Randomization method.
    $i = i + 1$.
  end while
  if $P_{opt} > \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k^{(i)})$ **then**
    $P_{opt} = \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k^{(i)})$.
    $\alpha_{opt} = \alpha, \mathbf{w}_k^* = \mathbf{w}_k^{(i)}, \mathbf{e}_k^* = \mathbf{e}_k^{(i)}, \forall k$.
  end if
end for
Output $\alpha_{opt}, \mathbf{w}_k^*, \mathbf{e}_k^*, \forall k$.

where $k = 1, 2, ...K$, $\mathbf{h}_{k,e}^*$ and $\mathbf{h}_{k,c}^*$ are complex Reyleigh channel coefficients and $\mathbf{G}_k^*$ is complex Rician channel coefficient based on (3.1). $d_0 = 10$ m, $d_1 = 50$ m, $d_2 = 10$ m, respectively denote the distances between the RIS and the cell edge user, the distance between the BS and the RIS, and the distance between the BS and the cell center user. $\alpha_0, \alpha_1, \alpha_2$ are the path loss exponents of the corresponding links. It is assumed that all the cell central users are at the same distance from the BS, all the cell edge users are at the same distance from the related RIS and all the RISs are at the same distance from the BS. We set $\alpha_0 = \alpha_2 = 1.8$ and $\alpha_1 = 2$. The noise power is $\sigma^2 = BN_0$, where the bandwidth $B = 100$ MHz and the noise power spectral density is $N_0 = -80$ dBm. The number of clusters is $K = 4$, which means there are 8 users in the system. For the OMA benchmark scheme, OFDM is adopted where each user will occupy a specific spectrum and will not cause interference to other users. It is assumed that the channel bandwidth is normalised, thus each user in the OFDM scheme can occupy $1/2K$ bandwidth.

Fig. 3.2 shows the transmit power at the BS versus the number of each RIS's reflecting elements. We provide the performance of the proposed schemes compared with the random phase scheme in NOMA and OFDM. In Fig. 3.2, the number of antennas at the BS is $M = 6$, and the date rate requirement of all the users is 1 bps/Hz. Obviously, the transmit power at the BS of all schemes decreases with the increasing of the number of RIS's reflecting elements. From Fig. 3.2, we can see that both proposed algorithms requires a less transmit power than the benchmarks. Comparing the two proposed algorithms, the performance gap

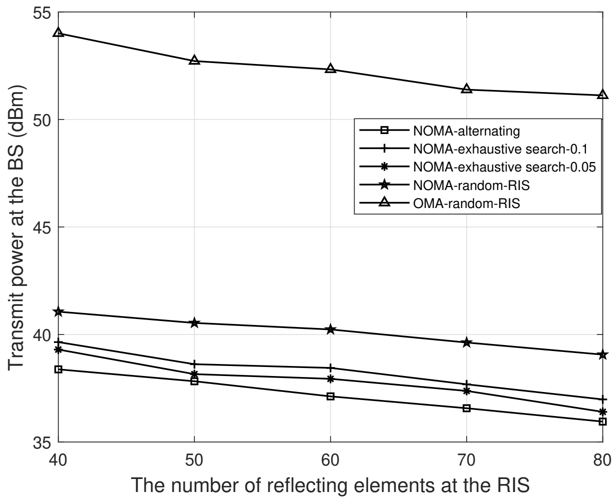

Figure 3.2. The transmit power versus the number of reflecting elements at the RIS.

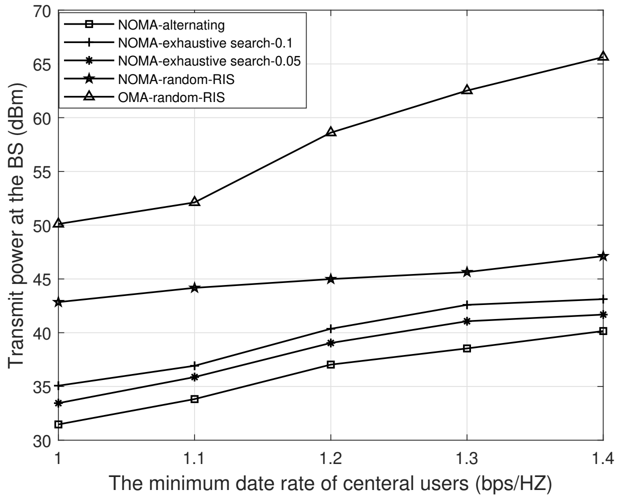

Figure 3.3. The transmit power versus the minimum date rate of the central users.

is very small and this gap will get smaller if the step size of the partial exhaustive search algorithm decreases. The result in Fig. 3.2 demonstrates that the alternating algorithm can yield the best performance among all the schemes but the partial exhaustive search can also yield competitive performance.

Fig. 3.3 shows the transmit power at the BS versus the minimum data rate of the central users. In this figure, it is assumed that each central user has the same date rate requirement, and all the cell edge users' date rate requirement is 1bps/Hz. In this figure, we set the number of antennas at the BS as $M = 6$ and the number of reflecting elements at each RIS as $N = 32$, respectively. According to the Shannon's capacity formula, it is well known that a higher date rate requires a higher transmit power at the BS. All schemes in Fig. 3.3 have the same trend, where the transmit power at the BS increases with the increasing of the central users' minimum date rate requirement. From Fig. 3.3, it is noted that the proposed alternating

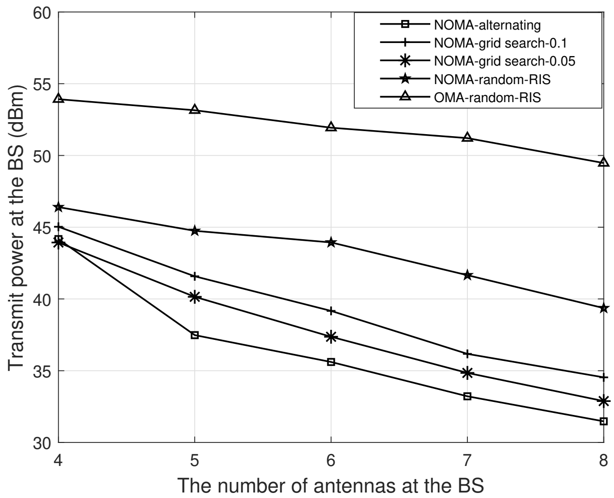

Figure 3.4. The transmit power versus the number of antennas at the BS.

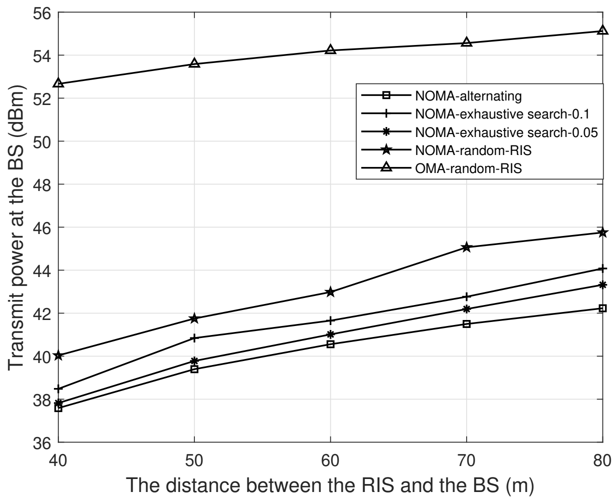

Figure 3.5. The transmit power versus the distance between the RIS and the BS.

algorithm needs less power consumption under the same date rate requirement. Although, the partial exhaustive search algorithm cannot achieve the same performance as the proposed alternating algorithm, it has low complexity and still yields a better performance than NOMA with random RIS scheme and OFDM scheme.

Fig. 3.4 shows the transmit power versus the number of antennas at the BS. In this simulation, the relationship between the algorithm performance and the number of antennas at the BS is illustrated. Fig. 3.4 shows the performances of two proposed algorithms with NOMA and OFDM with random phase RIS. In Fig. 3.4, we set the number of reflecting elements at each RIS as $N = 32$ and the date rate requirement of all the users as 1 bps/Hz. From Fig. 3.4, it is noted that the alternating algorithm achieves better performance gain than other algorithms with the number of antennas at the BS increasing.

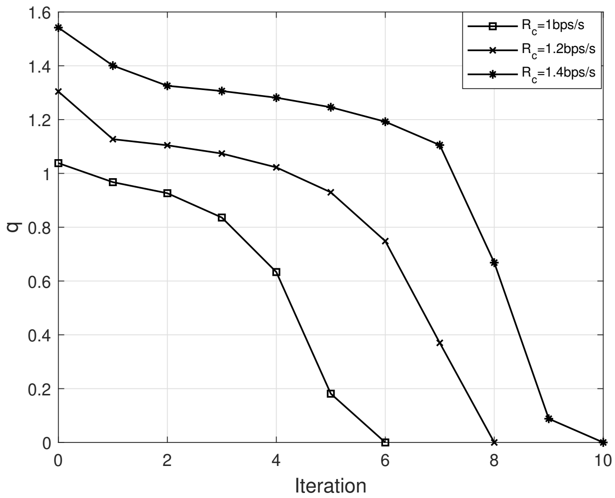

Figure 3.6. The value of q versus the iterative number.

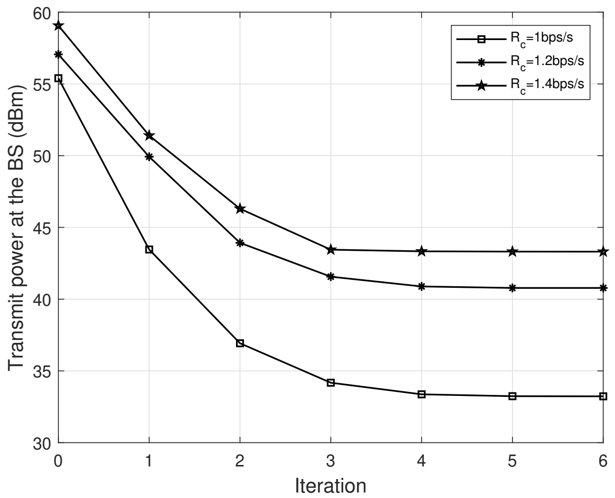

Figure 3.7. The transmit power at the BS versus the iterative number.

Fig. 3.5 shows the transmit power versus the distance between the RIS and the BS in each cluster. In Fig. 3.5, we set the number of antennas at the BS as $M = 6$, the number of each RIS's reflecting elements as $N = 32$. Each user's date rate requirement is 1bps/Hz. It has assumed that the RIS will not cause interference to the central user in each cluster. Since the central user in each cluster is close to the BS, thus the distance between the RIS and the BS cannot be short. Therefore, the starting point of the simulation is set as 40m. As expected, the transmit power of all schemes increases when the distance between the RIS in each cluster and the BS gets large. Similar to Fig. 3.4, the proposed alternating algorithm consumes less energy compared with all other schemes.

Fig. 3.6 shows the the value of q in the initial point search algorithm versus the iterative number. As previous discussion, the q represents the distance between the current problem and a feasible problem and q can enforce the current problem to be a feasible one. R_c denotes the data rate requirement of all the central users. Fig. 3.6 also shows that the value of q in the $R_c = 1.4$ bps/Hz scheme is larger than that in the $R_c = 1$ bps/Hz and $R_c = 1.2$ bps/Hz schemes at each iteration. Moreover, the scheme with $R_c = 1.4$ bps/Hz needs more iterations to converge, which indicates that a higher date rate requirement makes all constraints more difficult to be fulfilled.

Fig. 3.7 shows the transmit power at the BS versus the iterative number in Algorithm 2. R_c denotes the data rate requirement of all the central users. We evaluate the transmit power in different scenarios with the different data rate requirements of the central user. The data rate of all the cell edge users is 1 bps/Hz, the number of antennas at the BS is $M = 6$ and the number of each RIS is $N = 32$. From Fig. 3.7, we notice that the transmit power at the BS decreases with the number of iterations increasing, which also means this algorithm can converge with the algorithm proceeding.

3.6 Conclusion

The joint optimisation of beamforiming, power allocation and RIS phase shift in a NOMA-RIS assisted multi-cluster network is investigated in this chapter. By introducing inequality approximation, SCA and SDR, an alternating algorithm is proposed to minimise the transmit power by iteratively solving beamforming optimisation and phase shifting feasibility until the algorithm converges. Furthermore, an initial point search algorithm is proposed to guarantee the feasibility of the beamforming optimisation subproblem. Moreover, a low-complexity solution is also provided for this scenario based on the partial exhaustive search. The simulation results demonstrated the alternating algorithm outperforms the partial exhaustive search algorithm but has a higher complexity.

This chapter considered traditional convex optimisation to solve the optimisation problem. The approximation makes the solution loss optimality. To solve optimisation problems more efficiently and achieve better optimality, deep learning is considered in the next chapter to solve a sum rate problem in a RIS-NOMA network.

Chapter 4

A Reinforcement Learning Approach for an RIS-assisted NOMA Network

4.1 Introduction

The ultra-massive machine type communication (umMTC) is a key scenario of the next generation mobile communication [56]. An umMTC network always consists of massive communication devices, e.g. mobiles and sensors. Each device will communicate with other devices and cause massive traffic. Since the spectrum resource is extremely restricted, it is a challenge to support such heavy traffic. The non-orthogonal multiple access (NOMA) is being considered as a potential candidate of 6G communication system [32]. The main feature of NOMA is that it allows multiple devices share the same spectrum resource to communicate simultaneously, which greatly improves the spectrum efficiency. In particular, to satisfy the individual quality of service (QoS) requirement, each user or device in a NOMA network always adopts successive interference cancellation (SIC) to improve the signal to interference and noise ratio (SINR) and reception reliability [33].

Recently, the intelligent reflective surface (RIS) is proposed as a potential auxiliary device to improve the channel quality and help users in the area with heavy blockage to receive signals[42]. A RIS consists of many passive reflecting elements and a smart controller which can adjust the phase shift of each reflecting element. The RIS can adapt the channel between the transmitter and the receiver to increase the channel gain. Moreover, the RIS performs as a mirror to redirect the signal and enlarge the signal prorogation range.

Since the artificial intelligence has achieved great success, especially deep learning (DL), in wireless communication field [57], many works have studied the application of DL algorithms in a wireless communication network [58]. Deep reinforcement learning (DRL) as a type of DL has been attracting more and more attentions since there is no training data requirement [59]. Unlike the supervised learning and unsupervised learning, for which the training data is crucial, DRL adopts an agent to continuously interact with the environment and fetch feedbacks and utilizes the feedback to train the agent.

In this chapter, a sum rate maximisation problem is investigated and an algorithm based on DDPG is proposed. Due to the non-convexity of the formulated problem, it is very challenging to find the optimal solution via conventional convex optimisation. The DDPG based algorithm is proposed as an alternative method to efficiently solve this problem. This work also investigates two types of channels, which are the fixed channel and the time-varying channel. The simulation results demonstrate that the DDPG algorithm can adapt both types of channels and has the superior performance on both cases.

4.2 System Model and Problem Formulation

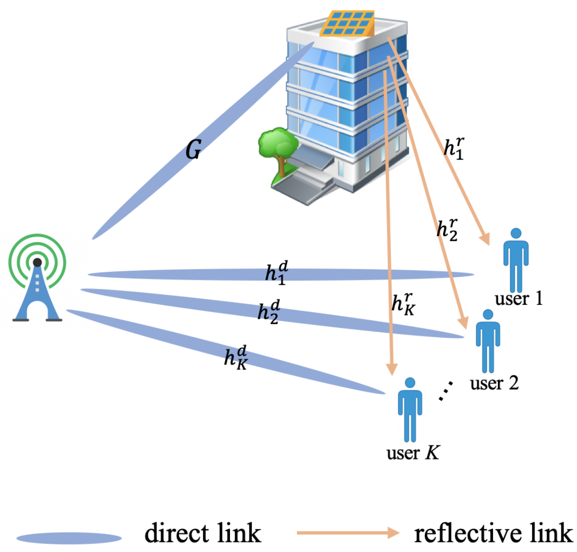

Figure 4.1. RIS NOMA sytem model.

We consider a NOMA-MISO downlink system comprising a BS and a RIS as shown in Fig. . The BS is equipped with M antennas and RIS has N reflective elements and a controller. The BS communicates with K ($M \leq K$) single antenna users. We define the user set as $\mathcal{K} = \{1, 2, ..., K\}$. We assume that the RIS is deployed on the surface of a building, thus the distance between the BS and the RIS is fixed. As shown in Fig. 6.1, each user will receive the reflective signal from the RIS and the direct signal from the BS. The BS will generate an unique beam for each user and superimpose each user's signal. The superimposed signal sent by the BS can be expressed as follows:

$$\mathbf{y} = \sum_{i=1}^{K} \mathbf{w_i} x_i, \tag{4.1}$$

where $\mathbf{y}$ denotes the superimposed signal sent to all users, $\mathbf{w_i} \in \mathbb{C}^M$ denotes the beamforming vector for the i-th user and x_i denotes the signal symbol of i-th user. We assume the power of signal symbol is unity which means $\mathbb{E}(x_i^2) = 1, \forall i$. There are two links between the BS and all the users. One is the reflective link where the signal sent by the BS will arrive at the RIS first and then the RIS reflects it to the users. The other is the direct link where the users can receive the signal directly from the BS. Thus, the received signal at each user is constructed with reflective signal and direct signal.We assume that the channel matrix between the BS and the RIS, $\mathbf{G} \in \mathbb{C}^{N\times M}$, the channel vector between the BS and the k-th user, $\mathbf{h}_k^d \in \mathbb{C}^{M\times 1}$ for all k, and the channel vector between the RIS and the k-th user, $\mathbf{h}_k^r \in \mathbb{C}^{N\times 1}$ for all k are perfectly known by the BS via ray tracing and channel estimation technology [42]. The received signal of k-th user can be expressed as follows:

$$y_k = \underbrace{\mathbf{h}_k^{dH}\mathbf{y}}_{\text{direct link}} + \underbrace{\mathbf{h}_k^{rH}\Phi\mathbf{G}\mathbf{y}}_{\text{reflective link}} + w_k, k \in \mathcal{K}, \tag{4.2}$$

where $\Phi \triangleq \text{diag}[\phi_1, \phi_2, ..., \phi_N]$ denotes the phase shift matrix of the RIS and w_k is the zero mean additive white Gaussian noise (AWGN) with variance σ_n^2.

Noted that Φ is a diagonal matrix and each element on the main diagonal describes the state of a reflective element on the RIS. The element on the main diagonal is expressed as $\phi_n = \beta_n e^{j\theta_n}$, where $\beta_n \in [0, 1]$ is the reflective coefficient describing the signal energy loss at the RIS and $\theta_n \in [0, 2\pi]$ is the phase shift introduced by the reflective element. We assume the RIS performs ideal reflection that we ignore the energy loss caused by reflection. Thus, the reflective signal has the same power with the original signal which means $\beta_n = 1, \forall n$ or $||\phi_n|| = 1, \forall n$.

Compared with the relaying communication system, RIS only performs as a passive mirror to reflect signals incident on it without decoding and encoding process. RIS will not consume power on signal processing. (6.2) indicates that RIS will not introduce any AWGN compared with relay. (6.2) also points out that the user experience is decided by the composite channel $\mathbf{h}_k^{dH} + \mathbf{h}_k^{rH}\Phi\mathbf{G}$. We can artificially adjust Φ to adapt to changes in the environment which means the channel is tunable by changing the phase shift continuously. Compared with the conventional wireless communication system, RIS-assisted system is more robust against the channel fading.

Noted that we use power domain NOMA to improve the spectrum efficiency, thus the BS will superimpose all the users' signals and broadcast the superimposed signal to all users. All the users in the NOMA system can share the same frequency channel, the same time slot and the same channel precode compared with FDMA, TDMA and CDMA and all the users will also receive the same signal. One of the challenges of NOMA is how each user detects and decodes their own signal from the superimposed signal. The technique named successive interference cancellation (SIC) can assist NOMA users to decode their own signal with less interference. For the simple notation, we use $\hat{\mathbf{h}}_k$ to represent the composite channel of k-th user, thus we have $\hat{\mathbf{h}}_k = \mathbf{h}_k^{dH} + \mathbf{h}_k^{rH}\Phi\mathbf{G}$. We first decide the weak user and the strong user. In a two users case, if the composite channel power of the user 1 is larger than the user 2's composite channel power, $||\hat{\mathbf{h}}_1||^2 \geq ||\hat{\mathbf{h}}_2||^2$, user 1 is the strong user compared with the weak user, user 2. In NOMA system, the BS will assign different powers to different users based on their composite channel power levels. Normally, the weak user will be assigned more power to compensate the worse channel condition. On the contrary, the strong user will be given less power. The power domain NOMA uses the power allocated to each user to distinguish users.

To successfully apply SIC, we need to decide the decoding order first. In NOMA system, the decode order is usually decided by the channel quality. However, in a RIS-assisted network, the channels are tuned by adjusting the phase shift of the RIS. Therefore, we use the composite channel power, which takes the impact of RIS on the channels into consideration, to decide the decoding order. We define Ω as the set of all possible decoding orders and ϵ is a certain decoding order in Ω. Without losing generality, we assume the composite channels are sorted as $||\hat{\mathbf{h}}_1||^2 \leq ||\hat{\mathbf{h}}_2||^2 \leq ... \leq ||\hat{\mathbf{h}}_{K-1}||^2 \leq ||\hat{\mathbf{h}}_K||^2$, then we have the decoding order $\bar{\epsilon} = (1, 2, ..., K)$. Based on $\bar{\epsilon}$, the weakest user, user 1, will decode its own signal directly and treat other users' signals as interference. The second weakest user, user 2 will decode user 1's signal locally and then omit user 1's signal, an interference to user 2, from the received signal. After that, user 2 starts to decode its own data without the interference caused by user 1 and meanwhile treats the rest of users' signals as interference. The SIC will proceed until the strongest user, user K, finishes the decoding of its own data. The strongest user, user K, will decode the first $K-1$ users' signals first and then remove all of them from the received signal. Thus, user K can decode its own data without any interference. To clearly describe the relationship between NOMA users, we adopt $i \preccurlyeq_\epsilon j$ to describe that the user j's composite channel power is larger than user i's and they are under the decoding order ϵ. Noted that the decoding order is decided by the composite channel power and the stronger user will decode weaker users' signal first, thus user j needs to decode user i's signal locally and user i do not need to decode user j's signal in ϵ.

In this chapter, we consider to maximize the sum rate of all the users. In order to maintain the total energy consumed by the entire system, we have the constraint as follows:

$$\sum_{i=1}^{K} ||\mathbf{w}_i||^2 + P_r \leq P_t, \tag{4.3}$$

where $\sum_{i=1}^{K} ||\mathbf{w}_i||^2$ is the total transmit power at the BS, P_r is the power supplying the RIS to control the phase shift and P_t is the maximum power threshold. (4.3) indicates that the energy consumed by the entire system cannot be unlimited. To satisfy the quality of service (QoS) requirement of each user, we have the constraint

$$R_i \geq R_t, i \in \mathcal{K}, \tag{4.4}$$

where R_t is the minimum data requirement of each user. Without losing generality, we assume that all the user in our system have the same minimum data requirement. We also have the constraints introduced by RIS's hardware limitation, which are

$$||\Phi_{n,n}|| = 1, \tag{4.5}$$

$$0 \leq \theta_n \leq 2\pi. \tag{4.6}$$

$\Phi_{n,n}$ denotes the n-th element on the main diagonal of the phase shift matrix Φ and (4.5) indicates that the RIS is an ideal reflective mirror and will not introduce any noise. Noted that we use NOMA, we need to introduce some constraints to guarantee each user can successfully finish SIC. The SIC progress points out that the date rate one NOMA user can achieve is affected by others users. To further explain this, we define the interference set of user k as κ_k, which contains the index of all the users cauing interference to user k. For example, considering the decoding order $\overline{\epsilon}$, user k will decode the first $k-1$ users' singals and remove them from the received signal. Thus, only users after user k will cause interference to user k when user k decodes its own data. Therefore, the interference set of user k based on $\overline{\epsilon}$ is $\{k+1, ..., K\}$. It is worth to point out that the interference set of each user is various according to different decoding orders. We assume $i \preccurlyeq_{\epsilon} j$, then we use γ_{ij} to represent the user i's SINR observed at user j, which can be expressed as follows:

$$\gamma_{ij} = \frac{|\hat{\mathbf{h}}_j^H \mathbf{w}_i|^2}{\sum\limits_{k \in \kappa_i} |\hat{\mathbf{h}}_j^H \mathbf{w}_k|^2 + \sigma_n^2}, \tag{4.7}$$

and the user i's data rate observed at user j is given by:

$$R_{ij} = \log_2(1 + \gamma_{ij}). \tag{4.8}$$

The user i's SINR when it decodes its own data can be expressed as follows:

$$\gamma_{i} = \frac{|\hat{\mathbf{h}}_i^H \mathbf{w}_i|^2}{\sum\limits_{k \in \kappa_i} |\hat{\mathbf{h}}_i^H \mathbf{w}_k|^2 + \sigma_n^2}, \tag{4.9}$$

and the user i's data rate is given by:

$$R_{i} = \log_2(1 + \gamma_{i}). \tag{4.10}$$

From (4.7), we find that one user's signal may be decoded multiple times at different users. Thus, one user's data rate is affected by other users. Then, we introduce the constraint

$$\min(R_{ij}, R_i) \geq R_t, \forall j \in \kappa_i, i \in \mathcal{K} \tag{4.11}$$

to guarantee SIC successful. (4.11) indicates that the signal of user $i, i \in \mathcal{K}$ can be decoded successfully at all other NOMA users. If the constraint (4.11) is violated that means one or some users fail to decode user i's signal, $i \in \mathcal{K}$. Then, SIC cannot proceed smoothly. We also note that once the constraint (4.11) is satisfies then the QoS constraint (4.4) is definitely satisfied.

In this chapter, we consider a sum rate maximisation problem where we maximize the sum rate of all the users. The optimisation problem can be formulated as follows:

$$\text{P8} : \max_{\{\mathbf{w}, \Phi\}} \quad \sum_{i=1}^{K} R_i \tag{4.12a}$$

$$\text{s.t.} \quad \min(R_{ij}, R_i) \geq R_t, \forall j \in \kappa_i, i \in \mathcal{K} \tag{4.12b}$$

$$\sum_{i=1}^{K} ||\mathbf{w}_i||^2 \leq P_t \tag{4.12c}$$

$$||\Phi_{n,n}|| = 1, \forall n \tag{4.12d}$$

$$0 \leq \theta_n \leq 2\pi, \forall n. \tag{4.12e}$$

4.3 DDPG-based Joint Optimisation of Phase Shift and Beamforming

4.3.1 Basic knowledge of DDPG

The reinforcement learning (RL) is an area of machine learning that handles with sequential decision-making [24]. There are two critical parts for a RL system, which are the agent and the environment. The key idea of RL is to train an agent to generate good actions based on the environment. A few factors fully characterize the RL processing. Policy π reflects the probability of an action chosen by an agent. State s is an observation from the environment. Action a is the decision made by the agent. Reward r is the feedback of an action. There are two kinds of RL algorithms in the RL family, one is value-based RL and another is policy gradient RL.

The value-based RL algorithm aims to build a value function and then decide a policy by minimizing or maximizing this function. The deep Q learning (DQN) is a typical value-based RL. In DQN, the loss function is defined as follows:

$$L(\theta) = \mathbb{E}\left[\left(r_t + \xi \max_{a'} Q'\left(s_{t+1}, a'|\theta'\right) - Q(s_t, a_t|\theta)\right)^2\right], \tag{4.13}$$

where Q' is the target network and Q is the training network. It is worth to point out the parameters of the target network θ' are fixed during the training. DQL is designed to address discrete action problems.

The policy gradient (PG) RL algorithm aims to optimize a performance objective by finding a good policy [24]. PG algorithms are also designed for the continuous action space. There is one classic PG algorithm named stochastic policy gradient. The parameter updating rule can be expressed as follows:

$$\theta_{t+1} = \theta_t + l\mathbb{E}[\nabla_a Q(s_t, a_t)|\ \nabla_\theta \mu_\theta(s_t)], \tag{4.14}$$

where l is the learning rate, Q^{π_θ} is the Q function to evaluate the current policy and μ_θ is the current policy.

The idea of DDPG combines DQN and PG together which can handle with the problem with a continues action space and also has a better convergence performance. A DDPG framework is constructed by two training networks and two target networks. The training actor network performs as a policy to generate actions. The training critic network estimates the Q function to evaluate the action. The target actor network generates the estimated action a' in (4.13). The target critic network generates the estimated Q value Q' in (4.13). (4.13) and (4.14) are utilized to train the training critic network and the training actor network, respectively. Note that the target network shares the same structure with its associated training network. Therefore, the target networks are updated through soft update, which is given by

$$\theta^{(target)} = \tau\theta^{(train)} + (1-\tau)\theta^{(target)}, \tag{4.15}$$

where τ is the updating rate .

4.3.2 Proposed DDPG framework

The corresponding elements are defined as follows:

- action $a^{(t)} = \left[\mathbf{w}_1^{(t)}, ..., \mathbf{w}_K^{(t)}, \Phi^{(t)}\right]$
- state $s^{(t)} = \left[\gamma_1^{(t-1)}, ..., \gamma_K^{(t-1)}, a^{(t-1)}, ||\mathbf{w}_1^{(t-1)}||^2, ..., ||\mathbf{w}_K^{(t-1)}||^2\right]$
- reward $r^{(t)} = \sum\limits_{i=1}^{K} R_i^{(t)}$

Note that the state space contains the square norm of each beamforming vector. It is reasonable because the neural network should take energy consumption into consideration due to the power control constraint (4.12c). It is worth to point out that the neural network can only take real number, therefore, the real part and the image part of a complex number should be separately input into neural networks.

In this chapter, both the actor network and the critic network are fully connected neural networks, comprised of input layer, hidden layer, batch normalisation layer and output layer. The size of the actor network's input layer is determined by the dimension of the state tuple. The critic network has two input layers for the action and the state specifically. The outputs of these two input layers will be horizontally stacked together as the input of the next hidden layer. The size of the hidden layer is related to the number of antennas at the BS, the number of reflecting elements at the RIS and the number of users. In this chapter, we adopt 300 neurons in every hidden layer. The batch normalisation layer is utilized between two hidden layers to contribute to faster convergence and shorter training time. The active functions utilized in the proposed DDPG are $tanh$ and $relu$, which make back propagation and gradient decent easier. Adam optimizer is utilized for both two networks with the learning rates 0.001 for the actor network and 0.002 for the critic network, respectively.

The experience replay is adopted in this DDPG framework to reduce the correlation of different training samples. A replay buffer $\mathcal{M}$ with the capacity $\mathcal{C}$ is implemented at the beginning of training. The training sample of each step, which is constructed by $\{a^{(t)}, r^{(t)}, s^{(t)}, s^{(t+1)}\}$, is stored into the replay buffer. The training progress will only start until the replay buffer is full. If the replay buffer is full, the newest training sample will replace the earliest one. In each training step, a mini-batch training samples are randomly selected from the replay buffer as the training data, which guarantees that the DDPG model has a good view of every training step.

A constraint fulfil layer is implemented as the output layer of the actor network, which guarantees that the action output from the actor network must satisfy all the constraints in (4.12). The next subsection introduces how this layer handles with all constraints.

4.3.3 Constraint handling

In the conventional optimisation algorithms, for instance the SDR algorithm and the alternating algorithm, we usually convert a non-convex problem to a convex problem and then solve it via convex optimisation solver, like CVX in Matlab. CVX can directly handle with convex constraints. However, this is not acceptable in deep neural networks. We manipulate the output layer of the actor network to make the action satisfy all the constraints in (4.12).

It is known that power and the data rate are positively related from Shannon's Theory. Higher data rate requires more energy consumption. Note that the maximum transmit power at the BS is P_t from (4.12c). Therefore, only letting the total transmit power at the BS as the maximum power allowed by the system can obtain the maximum sum rate. Hence, the constraint (4.12c) becomes an equality, which is given by:

$$\sum_{i=1}^{K} ||\mathbf{w}_i||^2 = P_t. \tag{4.16}$$

$\mathbf{w}_i^{(l)}$ denotes the beamforming vector for i-th user obtained by the actor network in the l-th training step. Since there is no constraint handling inside the neural network, thus $\mathbf{w}_i^{(l)}, \forall i$ may not satisfy the constraint (4.16). Therefore, normalisation is utilized to guarantee that constraint (4.16) is not violated. In the l-th step, we first calculate the total transmit power directly obtained by the actor network, which is

$$P^{(l)} = \sum_{i=1}^{K} ||\mathbf{w}_i^{(l)}||^2. \tag{4.17}$$

Then, we reassign power to user i's beam based on the following rule:

$$\mathbf{w}_i^{*(l)} = \mathbf{w}_i^{(l)} \sqrt{\frac{P_t}{P^{(l)}}}. \tag{4.18}$$

The key idea of (4.18) is reassigning the power based on the ratio $\frac{|\mathbf{w}_i^{(l)}|^2}{P^{(l)}}$ and setting the total transmit power as P_t. Meanwhile, the new beamforming vector $\mathbf{w}_i^{*(l)}$ has the same direction with $\mathbf{w}_i^{(l)}$. After the operation of (4.18), all the new beamforming vector satisfy

$$\sum_{i=1}^{K} ||\mathbf{w}_i^{*(l)}||^2 = P_t. \tag{4.19}$$

As for the constraint (4.12d), normalisation is used again to handle it. $^{(l)} = \{\phi_1^{(l)}, \phi_2^{(l)}, ..., \phi_N^{(l)}\}$ containing all the elements on the main diagonal of the phase shift matrix denotes the optimized result directly obtained by the actor network in the l-th step. To satisfy the constraint (4.12d), each element can be normalized by the following

$$\phi_i^{*(l)} = \frac{\phi_i^{(l)}}{||\phi_i^{(l)}||}, \quad i = \{1, 2, ..., N\}. \tag{4.20}$$

It is known that the function $f(\theta) = e^{j\theta}$ is a periodic function with a period of 2π. Therefore, the phase can be always mapped within $[0, 2\pi]$.

In the conventional optimisation, it is difficult to handle with the constraint (4.12b) since it is non-convex and contains two optimisation valuables coupled together. Conventionally, SDR is utilized to approximate it and then the alternating algorithm is applied to find a sub-optimal solution. In this DPL-based algorithm, we have $\mathbf{w}_i^{*l}, \forall i$ and $\Theta^{*(l)}$ from (4.18) and (4.20), respectively. It is known that one user has an original data rate and several observed date rates in a NOMA system. Once a specific decoding order $\epsilon^{'}$, $\mathbf{w}_i^{*l}, \forall i$ and $\Theta^{*(l)}$ are decided, the original data rate and the observed data rates can be calculated via (4.10) and (4.8) for each user, respectively. Then, we check if each user's original data rate and observed data rates satisfy the constraint (4.12b). If the constraint (4.12b) is violated, it means the actor networks outputs an invalid action. In order to reduce the invalid output as much as possible, a punishment mechanism is adopted on the reward. If the constraint (4.12b) is violated, the reward is set as

$$r^{(l)} = \sum_{i=1}^{K} R_i^{(l)} \sum_{i \in \mathcal{K}} \sum_{j \in \kappa_i} \min(\min(R_{ij}^{(l)}, R_i^{(l)}) - R_t, 0) \tag{4.21}$$

instead of the sum rate. Note that the reward is negative when constraint (4.12b) cannot be satisfied in the l-th training step. Meanwhile, if the action seriously violates the (4.12b), the reword will be assigned a smaller negative value, which is a more severe punishment to the neural network. Hence, the neural network will adjust the output in the following training steps to avoid such invalid action as much as possible.

4.3.4 Algorithm

The detail of the algorithm is shown in **Algorithm** 5. At the beginning of the algorithm, four networks and the replay buffer are initialized. At the beginning of each training episode, CSI, beamforming vectors and phase shift matrix are initialized. We simply adopt an identity matrix to initialize all the beamforming vectors. Note that CSI in each episode is various, which indicates that this DDPG model is compatible with the varying channel scenario. In each training step, a random process $\mathcal{N}_t$ is adopted to enlarge the action exploration. $\mathcal{N}_t$ in **Algorithm** 5 is a complex Gaussian noise vector with the same size of the action a. It is worth

Algorithm 5 DDPG-based Joint Beamforming and Phase Shift Optimisation

1: **Initialisation:** Generate two actor networks and two critic networks. Make the training network and the target network have the identical parameters, $\theta_a^{(train)} = \theta_a^{(target)}$ and $\theta_c^{(train)} = \theta_c^{(target)}$.
   Initialize the experience replay buffer $\mathcal{M}$ with the storage capacity C and the mini-batch size N_b.
2: **Output:** Optimal beamforming vector $\mathbf{w}_k, k \in \mathcal{K}$ and phase shift matrix Φ.
3: **for** episode $i = 1, 2, ..., I$ **do**
4: Randomly generate CSI $\mathbf{G}^{(i)}$, $\mathbf{h}_k^{d(i)}, \mathbf{h}_k^{r(i)}, k \in \mathcal{K}$ and the phase shift matrix $\Phi^{(i)}$. The beamforming vectors are initialized by a $M \times K$ identity complex matrix.
5: Calculate the decoding order ϵ_0 according the the composite channel power.
6: Obtain the initial state s_1
7: **for** step $t = 1, 2, ..., T$ **do**
8: Initialize a random process $\mathcal{N}_t$.
9: Choose an action from the actor training network $a_t = \mu^{(train)}(s_t|\theta_a^{(train)}) + \mathcal{N}_t$.
10: Normalisation the beamforming vectors and the phase shift matix by (4.18) and (4.20).
11: Calculate the new decoding order ϵ_t.
12: If constraint (4.12b) is not violated, the reward r_t is set as the sum of all users' original data rate. Otherwise, r_t is given by (4.21). Obtain the new state s_{t+1}.
13: Store $\{s_t, a_t, r_t, s_{t+1}\}$ to the buffer $\mathcal{M}$.
14: Sample a minibatch with the batch size N_b from $\mathcal{M}$ to train networks.
15: Given reword discount factor ξ, set the target Q value based on (4.13).
16: Update $Q^{(train)}(s, a|\theta_c^{(target)})$ by minimizing the loss function.
17: Update $\mu^{(train)}(s|\theta_a^{(train)})$ by the policy gradient.
18: Update two target networks by using soft update.
19: $s_t = s_{t+1}$.
20: **end for**
21: **end for**
22: **Output** $\{\mathbf{w}_k^{*(j)}, \alpha_k^{*(j)}, \mathbf{e}_k^{*(j)}\}, \forall k$.

to point out that the composite channel of each user is changing due to various phase shifts in one training episode. As mentioned before, the decoding order is decided by the composite channel gain, therefore, the new decoding order needs to be calculated based on the current phase shift, which is shown in step 11.

4.4 Simulation Results

This section demonstrates the performance of the proposed algorithm. It is assumed that the channels between the BS and all the users are Rayleigh channel, which indicates that the line-of-sight (LoS) signal is blocked and the users mainly receive signals through the RIS. This channel is model as follows:

$$\mathbf{h}_k^d = \widetilde{\mathbf{h}}_k^d / \sqrt{d_{d,k}^{\alpha}}, k \in \mathcal{K}, \tag{4.22}$$

where $\widetilde{\mathbf{h}}_{d,k} \in \mathbb{C}^{M\times 1}$ contains M independent and identical elements following complex $\mathcal{CN}(0,1)$ distribution. $d_{d,k}$ denotes the distance between the BS and the k-th user, which is generated randomly within the range (45,50). α denotes the path loss exponent, which is set as 2. The channel between the BS and the RIS and the channels between the RIS and all the users are assumed to be Rician channel, which can be modelled as follows:

$$\mathbf{G} = \left(\sqrt{\frac{\upsilon}{1+\upsilon}} \mathbf{G}^{\mathrm{LoS}} + \sqrt{\frac{1}{1+\upsilon}} \mathbf{G}^{\mathrm{nLoS}} \right) / \sqrt{d^{\alpha}}, \tag{4.23}$$

$$\mathbf{h}_k^r = \left(\sqrt{\frac{\upsilon}{1+\upsilon}} \mathbf{h}_k^{r\mathrm{LoS}} + \sqrt{\frac{1}{1+\upsilon}} \mathbf{h}_k^{r\mathrm{nLoS}} \right) / \sqrt{d_{r,k}^{\alpha}}, k \in \mathcal{K}, \tag{4.24}$$

where υ denotes the Rician factor, $\mathbf{G}^{\mathrm{LoS}}$ and $\mathbf{h}_k^{r\mathrm{LoS}}$ are the LoS component, $\mathbf{G}^{\mathrm{nLoS}}$ and $\mathbf{h}_k^{r\mathrm{nLoS}}$ are the non-LoS component and d and $d_{r,k}$ denote the distance between the BS and the RIS and the distance between the RIS and the k-th user. In this chapter, the Rician factor is set as 1 and the LoS components $\mathbf{G}^{\mathrm{LoS}}$ and $\mathbf{h}_k^{r\mathrm{LoS}}$ are assumed to be 1. The non-LoS components $\mathbf{G}^{\mathrm{nLoS}}$ and $\mathbf{h}_k^{r\mathrm{nLoS}}$ follow the Rayleigh fading. d is set as 50 and $d_{r,k}$ is randomly generated within the range (5,10).

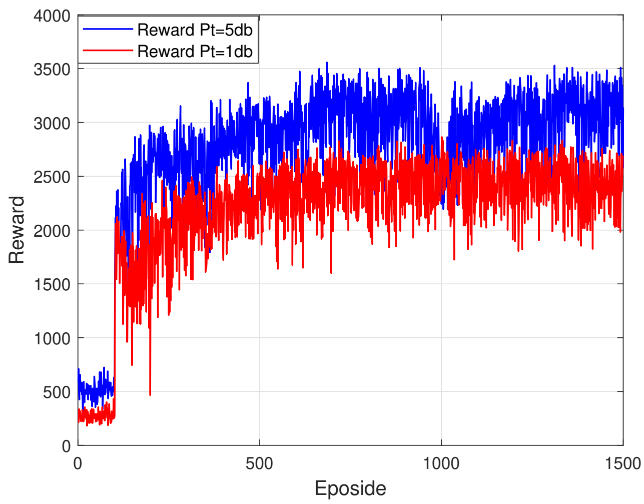

Figure 4.2. The accumulative reward in the varying channel scenario

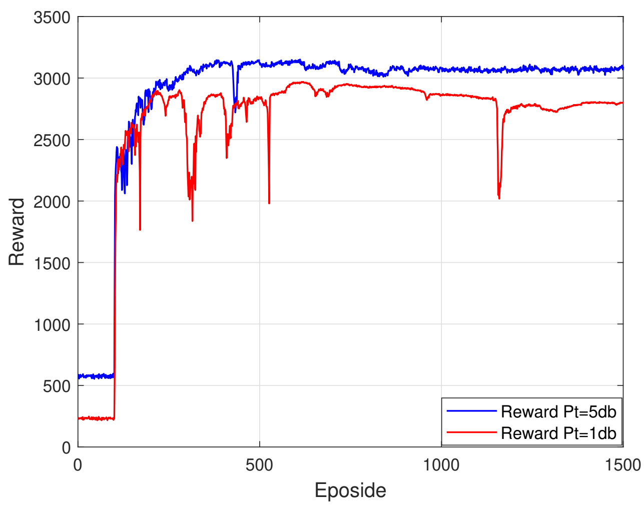

Figure 4.3. The accumulative reward in the fixed channel scenario

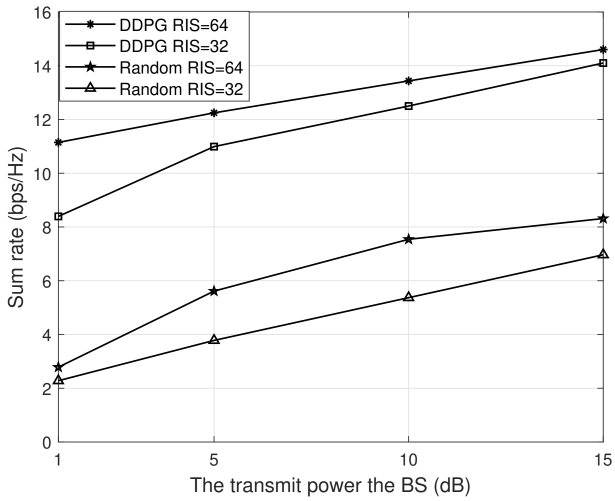

Figure 4.4. The sum rate as a function of the transmit power at the BS

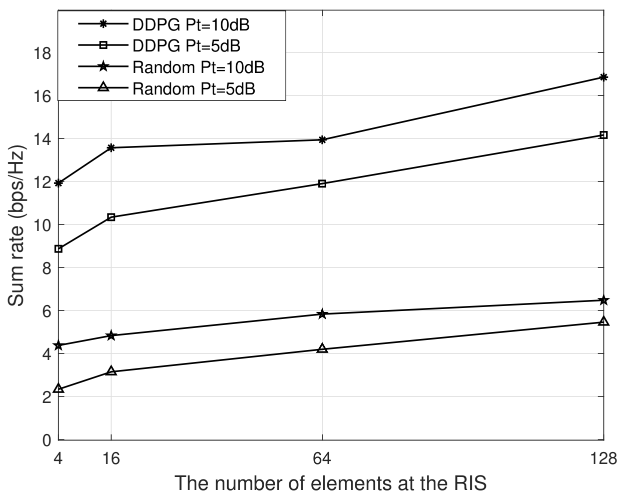

Figure 4.5. The sum rate as a function of the number of elements at the RIS

Fig. 4.2 and Fig. 4.3 show how the accumulative reward changes during training. In this simulation, the number of users is 4, the number of RIS elements is 32 and the number of anteenas at the BS is 4. The noise power is $\sigma^2 = -10$ dBm. The total transmit power is set as 10dB and 5dB. The RIS with random phase shift serves as the benchmark of the DDPG algorithm. Fig. 4.2 illustrates the scenario that the channel is varying. In particular, the channel is randomly generated before each training episode. It shows the accumulative reward converges with the training proceeds, which indicates that the proposed DDPG algorithm has a good adaptation to time-varying channels. Fig. 4.3 illustrates the scenario that the channel is fixed, which means the channel is remain the same during the whole training process. Compared with the varying channel case, the reward is more stable. Note that the training samples are randomly selected from the replay buffer at the beginning of each episode. If the selected samples are not good experience, i.e., action violates the constraint, the reward will drop suddenly at this episode. It is noted that solving a optimisation problem by DDPG is a online progress especially when the channel is time-varing because the DDPG model needs to constantly interact with the environment to update the action in time. The updated channel information needs to be input into the DDPG model once it is obtained.

Fig. 4.4 and Fig. 4.5 show the performance of the proposed algorithm. In these two figures, the X-axis represents the total transmit power at the BS and the number of the RIS elements, respectively. The Y-axis represents the sum rate of all the users for both two figures. It is obvious that the perforce is getting better with the increasing of total transmit power and the number of RIS elements. It also shows that the beamforming and the phase shift optimized through DDPG outperform the random beamforming and phase shift.

4.5 Conclusion

In this chapter, a sum rate maximisation problem in an RIS assisted NOMA downlink network was investigated. A DDPG based algorithm was proposed to jointly optimize beamforming and phase shift. The proposed DDPG algorithm can not only achieve competitive performance but also adapt to the varying channel scenario, however, the conventional convex optimisation is mainly suitable for the fixed channel scenario. More specifically, machine learning provided a new solution for wireless communication problems and also can be applied for more complicated scenarios, which could be a powerful tool for developing the next generation communication network.

The performance improvement from the RIS has been proven. The RIS reflects the incident signal and meanwhile reconfigure the channel. However, another technique named backscattering can module a new signal onto the incident signal and then reflect it to the receiver, which enables passive devices to send their own signals without any energy consumption. The next two chapters mainly investigated how to improve the performance of a backscattering-NOMA system.

Chapter 5

Backscatter-Assisted NOMA Network for the Next Generation Communication

5.1 Introduction

Recently, technological innovations in wireless communication have driven emerge rapidly, Internet of Things (IoT) has been considered as the next generation network [60]–[62]. One IoT network promises to support massive IoT devices[63], which provides opportunities to develop novel applications. However, there are still many challenges to operating an IoT network. An IoT network usually consists of massive devices, and massive spectrum resources will be consumed when many devices transmit signals simultaneously [64]. Therefore, one challenge is to support the massive IoT devices within a limited spectrum resource block. Another challenge arises from the energy constraint because most IoT devices are passive and cannot be equipped with a battery [65]. Thus, energy cooperation among different devices has become an important topic in IoT networks.

In recent years, non-orthogonal multiple access (NOMA) has emerged as a promising multiple access technique for the next generation of wireless communication because it allows multiple devices to share the same resource block simultaneously and increases spectrum efficiency [66]. In a particular NOMA downlink network, the base station (BS) superimposes signals for different users and broadcasts the superimposed signal. Typically, successive interference cancellation (SIC) is adopted to remove interference caused by other users' signals at each user's end [33], [67], [68].

Various energy cooperation techniques have been proposed to address the energy constraint. For example, simultaneous wireless information and power transfer (SWIPT) enables an energy-limited device to be powered by harvesting energy from the signal sent by the BS or other non-energy-constrained devices [69], [70]. Backscatter (BAC) is a more mature and practised technique for achieving energy cooperation among different devices [71]. Backscatter devices (BDs) transmit their own signals through backscatter circuits, which can be excited by signals from another device [72]–[74].

In this study, a legacy downlink NOMA system containing two NOMA users is considered, and an uplink device equipped with a backscatter circuit is added to this existing system. A full-duplex (FD) model BS is deployed to support the uplink and downlink transmission simultaneously[75]–[77]. This study aims to maximise the uplink device's data rate at the BS while guaranteeing the quality of service (QoS) of two NOMA users. The closed-form solution of the proposed system model is provided.

5.1.1 Related works

Recently, backscatter studies have emerged due to the benefit of enabling energy-limited devices to transmit signals without consuming extra energy. Ambient backscatter communication (AmBC), as a family member of backscatter technology, is attracting increasing research interest [78]–[80]. The key idea of AmBC is to adjust the amplitude and phase of the received signal to make this signal carry information before backscattering. In recent works, the combination of AmBC and NOMA achieved a promising performance by fully utilizing the spectrum and greatly improving energy efficiency. The authors in [81], [82] introduced a novel application based on backscatter, which generates multi-path signals to make the device-to-device (D2D) communication more reliable. Particularly, [81] considered a backscatter communication network containing multiple backscatter transceivers, powered by a power beacon station. This work aims to maximise the throughput performance by cooperative transmission. The results showed that backscatter radios and cooperative transmission significantly improve the throughput performance. In [82], authors considered an IoT network containing several devices and a power beacon station. These IoT devices can switch two modes: active radio frequency mode and passive backscatter mode. They receive signals from the BS in the first phase and then backscatter their own signals to the receiver in the second phase. The quality of the received signal is improved due to multi-path. The results indicate that the backscatter scheme can significantly improve the throughput of the system.

NOMA has been introduced to the backscatter network to ensure that multiple backscatter devices can mutually communicate with one access point in the same resource block. A novel system model was proposed in recent work, which contains both the uplink and the downlink [71]. Particularly, this system model consists of multiple BDs and a downlink user. The FD base station can transmit and receive signals simultaneously. The BS transmits the signal to the downlink user, whereas all BDs backscatter their own signals based on the received signal to the BS. Signals from all BDs will be superimposed together, and then the uplink NOMA technique will be applied to decode them individually. However, the system only contains one downlink user, and NOMA is applied to the uplink. The algorithm proposed in this work cannot be used to solve a downlink NOMA scenario. The authors in [83] considered an IoT network containing one transmitter, multiple BDs and one backscatter receiver. Each BD receives the same signal from the transmitter and backscatters its own signal to the receiver. The receiver receives signals from multiple BDs, which can be treated as an uplink NOMA scenario. Thus, the algorithm proposed in this work cannot be applied to a downlink NOMA scenario. Although a simple two-user downlink NOMA system was proposed in [84], BAC realised communications between these two users. However, the author only analysed the outage performance among different scenarios but did not provide an optimal solution. In [85], imperfect SIC was investigated in a backscatter-enabled NOMA network, and two readers received signals from a source and a backscatter tag. Although, the closed-form solutions of the power allocation coefficient and the backscatter coefficient were provided, the QoS constraint of the backscatter tag was not considered in the optimisation problem.

5.1.2 Motivations and Contributions

The motivation of this study is to enlarge the capacity of an existing network. Each of the previous works proposed a new system specifically designed for backscatter devices. However, this study investigates how a BD can be added to a legacy NOMA system while maintaining the original system's performance. In contrast to [71] and [83], the downlink NOMA is adopted, and the BD's excitation is a superimposed signal, posing additional challenges to SIC. Although [84] and [85] studied a downlink NOMA system, they did not consider any uplink device. In this study, the downlink user and the uplink device co-exist in one system.

First, a novel IoT system model is proposed, which is transformed from a legacy system by adding an uplink BD. Then, an uplink data rate maximisation problem is formulated, which is non-convex and difficult to be solved. The originally formulated problem is first transformed into two convex sub-problems by successive convex approximation (SCA). An alternating algorithm is proposed to solve the two sub-problems iteratively, and a sub-optimal solution is obtained. To further study, we derive closed-form solutions of the power allocation and the backscattering coefficients by analysing the monotonicity of the objective function and all the constraints. Finally, an exhaustive search algorithm is applied to verify the closed-form solution. The contributions are summarised as follows:

- A novel concept that an uplink BD is added into a legacy NOMA network is proposed.

The performance of legacy users is not affected, however, the total throughput of the new system is improved because a new uplink device is introduced. This concept is important to IoT networks, as it allows more IoT devices to be deployed into existing networks without extra bandwidth consumption.

- An uplink data rate maximization problem is formulated. Since two optimisation variables, the power allocation and the backscattering coefficients, are highly coupled together, the original optimisation problem, which is non-convex and NP hard, is difficult to be solved. The non-convex problem is split into two sub-problems and convex transformation is applied to both two sub-problems. An alternating algorithm is proposed to iteratively solve two sub-problems and a suboptimal solution is achieved.
- The study also provides a closed-form optimal solution of the power allocation and the backscattering coefficients. It is very difficult to obtain the closed-form of a non-convex optimisation problem, where two optimisation variables are coupled together, by directly using Karush–Kuhn–Tucker (KKT) conditions. A novel strategy to find the closed-form solution is proposed. The optimality is validated by an exhaustive search algorithm.

5.1.3 Organisation

The rest of the chapter is organised as follows. In Section II, a BAC-NOMA system is introduced and an uplink data rate maximisation problem is formulated. In Section III, an alternating algorithm is proposed to solve the problem iteratively and a sub-optimal solution is acquired. In Section IV, the closed-form solution is derived. In Section V, simulation results are provided. Finally, a conclusion is summarised in Section VI.

5.2 System Model and Problem Formulation

We consider a legacy NOMA network, where two NOMA users simultaneously receive signals from the BS. Meanwhile, a backscatter device is added to the existing NOMA system to communicate with the BS. It is assumed that the BS operates in FD mode, which allows the BS to receive and transmit signals simultaneously. The entire system is a single-input-single-output (SISO) system, which means every node has a single antenna. The BS broadcasts the

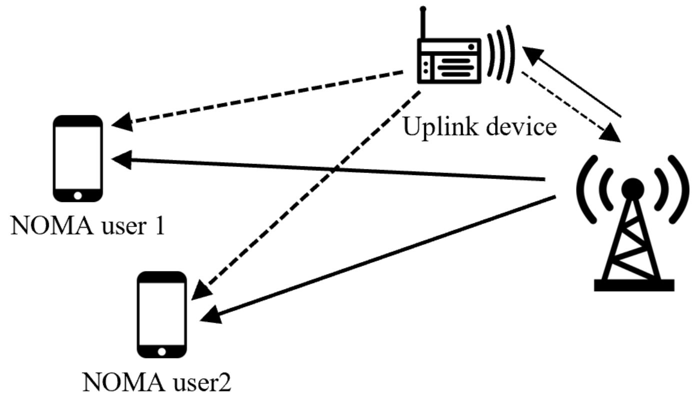

Figure 5.1. The system model.

superimposed signal, which is given by:

$$s_d = \sqrt{\alpha P_0} s_1 + \sqrt{(1-\alpha)P_0} s_2, \tag{5.1}$$

where α denotes the power allocation coefficient, P_0 denotes the maximum transmit power at the BS and s_1 and s_2 denote the signal sent to user 1 and user 2, respectively. We assume that the uplink device is energy limited, thus, the backscatter technique is adopted to support the communication with the BS. The uplink signal received at the BS can be expressed as follows:

$$y_u = |h_u|^2 \sqrt{\eta} s_d s_u + s_{\mathrm{SI}} + n_b, \tag{5.2}$$

where h_u denotes the channel coefficient between the BS and the uplink device, η denotes the backscattering coefficient, which is intelligently decided by the backscatter circuit, and s_u denotes the uplink signal sent to the BS, which is assumed to be normalised, i.e., $\mathbb{E}(|s_u|^2) = 1$, where $\mathbb{E}(\cdot)$ denotes an expectation operation. s_{SI} denotes the self-interference and is assumed to follow the complex Gaussian distribution $\mathcal{CN}(0, \beta P_0 |h_{\mathrm{SI}}|^2)$ [86], where h_{SI} denotes the self-interference channel and is assumed to be complex Gaussian distributed, i.e., $s_{\mathrm{SI}} \sim \mathcal{CN}(0,1)$ and $0 \leq \beta \leq 1$ denotes FD residual self-interference coefficient[71]. n_b is the noise at the BS, which is assumed to follow the complex Gaussian distribution $\mathcal{CN}(0, \sigma_b^2)$.

Given that multiple devices will work cooperatively within one IoT network, transmitting resources such as time, frequency, and space is always a challenge for such a dense network. NOMA is highly appreciated by resource-limited networks. We must admit that adopting OMA in this model may decrease the complexity and also obtain a competitive outcome. However, this study examines how to enlarge the system capacity, particularly by adding another uplink device into an existing NOMA network, when the current resources cannot satisfy the OMA scheme. It is also a valuable research direction to investigate the OMA's performance on this system model when sufficient transmit resources exist. This will be discussed in our future studies.

The uplink data rate achieved by this system can be expressed as follows:

$$R_u = \log_2\left(1 + \frac{|h_u|^4\eta|s_d|^2}{\beta P_0|h_{\mathrm{SI}}|^2 + \sigma_b^2}\right). \tag{5.3}$$

Two NOMA users receive signals from the BS and the BD; however, the signal from the BD is treated as interference. The received signal of the NOMA user i, $i \in \{1, 2\}$, is denoted by y_i, which can be expressed as follows:

$$y_i = h_i s_d + \sqrt{\eta} g_i h_u s_d s_u + n_i \tag{5.4}$$

The channel gain between the BS and NOMA user i is denoted as h_i, g_i denotes the channel gain between the uplink device and NOMA user i and n_i is assumed to be Gaussian distributed, i.e., $\mathcal{CN}(0, \sigma_i^2)$, additive white Gaussian noise (AWGN). In this study, we assume that all the perfect channel information (CSI) between every two nodes is fixed and well known by the BS. For a NOMA system, applying successful SIC based on a certain decoding order is critical. We must admit that CSI and SIC, as two necessary factors in a NOMA network, cannot be perfectly estimated and proceeded. Imperfect CSI and SIC will introduce extra interference when signals are being decoded, thereby further degrading system performance. Some existing works [85], [87] have already investigated the performance of a backscatter-NOMA system under imperfect CSI and SIC. The investigation of the case with imperfect CSI and SIC is beyond the scope of this work, but it is a promising direction for future research. Without loss of generality, it is assumed that user 1 is closer to the BS than user 2, and user 1 has a better channel quality than user 2, i.e., $|h_1|^2 \geq |h_2|^2$. Under this assumption, the decoding order is defined as follows: user 1 decodes user 2's signal first, then its own signal, whereas user 2 decodes its own signal directly while treating user 1's signal as interference. Referring to (5.4), the received signal at user 1 can be expressed as follows:

$$y_1 = h_1\sqrt{\alpha P_0} s_1 + h_1\sqrt{(1-\alpha)P_0} s_2 + \sqrt{\eta} g_1 h_u s_d s_u + n_1. \tag{5.5}$$

Given the decoding order above, user 2's achievable data rate observed at user 1 can be expressed as follows:

$$R_{12} = \log_2\left(1 + \frac{(1-\alpha)P_0|h_1|^2}{\alpha P_0|h_1|^2 + \eta|g_1 h_u|^2|s_d|^2 + \sigma_1^2}\right). \tag{5.6}$$

User 2's signal can be removed from user 1's received signal after user 1 successfully decodes user 2's signal. User 1 can further decode its own signal without interference caused by user 2. User 1's data rate can be expressed as follows:

$$R_1 = \log_2\left(1 + \frac{\alpha P_0|h_1|^2}{\eta|g_1 h_u|^2|s_d|^2 + \sigma_1^2}\right). \tag{5.7}$$

Referring to (5.4) again, the received signal at user 2 can be expressed as follows:

$$y_2 = h_2\sqrt{\alpha P_0}s_1 + h_2\sqrt{(1-\alpha)P_0}s_2 + \sqrt{\eta}g_2 h_u s_d s_u + n_2. \tag{5.8}$$

User 2 directly decodes its own signal by treating user 1's signal as interference so that user 2's data rate is given by

$$R_2 = \log_2\left(1 + \frac{(1-\alpha)P_0|h_2|^2}{\alpha P_0|h_2|^2 + \eta|g_2 h_u|^2|s_d|^2 + \sigma_2^2}\right). \tag{5.9}$$

In this system model, the uplink device enjoys the premium service, which requires the maximum data rate, whereas two NOMA users take basic service where only a basic data rate is guaranteed. Hence, the objective function is the uplink device's data rate, which needs to be maximised. The optimisation problem is formulated as follows:

$$\text{P9}: \max_{\{\alpha,\eta\}} \quad R_u \tag{5.10a}$$

$$\text{s.t.} \quad \min(R_{12}, R_2) \geq R_{t2} \tag{5.10b}$$

$$R_1 \geq R_{t1} \tag{5.10c}$$

$$0 \leq \alpha \leq 1 \tag{5.10d}$$

$$0 \leq \eta \leq 1 \tag{5.10e}$$

Constraint (5.10b) ensures the success of SIC, where R_{t2} is the target data rate of user 2.

Constraint (5.10c) guarantees the QoS requirement of user 1, where R_{t1} denotes user 1's data rate. Constraints (5.10d) and (5.10e) are the range limitation of the power allocation and the backscattering coefficients, respectively.

The formulated problem is difficult to be solved due to the non-convexity of objective function (5.10a) and constraints (5.10b) and (5.10c). In the next section, a transformation is provided to make the original problem convex.

5.3 Convex Transformation and Algorithm

As the discussion above, the formulated problem (5.10) is non-convex. The main challenge arises from the non-convexity in the objective function (5.10a) and constraints (5.10b) and (5.10c). Convex transformation is necessary to solve this problem efficiently. In this section, we first transform the original problem into a convex problem and then provide an algorithm to solve it.

The uplink device uses s_d as the excitement to backscatter its signal so that s_u is modulated onto s_d. From (5.6), (5.7) and (5.9), it is noted that $|s_d|^2$ is treated as a type of fading affecting s_u. Note that s_d is the superimposed signal of user 1 and user 2 which is given by (5.1), the expression of $|s_d|^2$ is

$$|s_d|^2 = P_0 + P_0\sqrt{\alpha - \alpha^2}(s_2 s_1^* + s_1 s_2^*). \tag{5.11}$$

$|s_d|^2$ consists of two terms and the second tern is conditioned on the choice of s_1 and s_2. Note that $s_2 s_1^*$ is conjugate with $s_1 s_2^*$, the sum of these two terms is a real number denoted by P_s Symbols of s_1 and s_2 affect the energy of s_d which further affect the throughput capacity of the system. The ingenious design of symbols will positively contributes to the performance of the system. However, the symbol design is beyond the scope of this study, which is still a very critical research direction for future studies. This study only focuses on optimising power allocation and backscattering coefficients. Therefore, it is assumed that the optimal choice of s_1 and s_2 has already been decided. $|s_d|^2$ can be further written as

$$|s_d|^2 = P_0(1 + P_s\sqrt{\alpha - \alpha^2}). \tag{5.12}$$

The original problem consists of two coupled variables. The alternating algorithm is an efficient way to solve multi-variable optimisation problems, whose key idea is to optimise only one variable and fix others in each iteration. Following this idea, we first fix α to obtain a sub-problem with an assumption that $\sigma_1 = \sigma_2 = \sigma$.

$$\text{P10}: \max_{\{\eta\}} \quad \eta \tag{5.13a}$$

$$\text{s.t.} \quad (1-\alpha)P_0|h_1|^2 \geq \epsilon_2(\alpha P_0|h_1|^2 + \eta|g_1 h_u|^2|s_d|^2 + \sigma^2) \tag{5.13b}$$

$$(1-\alpha)P_0|h_2|^2 \geq \epsilon_2(\alpha P_0|h_2|^2 + \eta|g_2 h_u|^2|s_d|^2 + \sigma^2) \tag{5.13c}$$

$$\alpha P_0|h_1|^2 \geq \epsilon_1(\eta|g_1 h_u|^2|s_d|^2 + \sigma^2) \tag{5.13d}$$

$$0 \leq \eta \leq 1, \tag{5.13e}$$

where ϵ_1 denotes $2^{R_{t1}}-1$ and ϵ_2 denotes $2^{R_{t2}}-1$. Since α is fixed, $|s_d|^2$ is also fixed. Since the log function is a monotonously increasing function, the objective function can be equivalently replaced by η. All constraints and the objective function in P10 are affine; therefore, P10 is a convex problem. It is easy to be solved via a convex optimisation tool box, such as CVX in Matlab. η^* is assumed to be the optimal solution of P10.

The second sub-problem is obtained by fixing η, which can be formulated as follows:

$$\text{P11}: \max_{\{\alpha\}} \quad \sqrt{\alpha-\alpha^2} \tag{5.14a}$$

$$\text{s.t.} \quad (1-\alpha) \geq \epsilon_2(\alpha + \eta A_1 P_0(1 + P_s\sqrt{\alpha-\alpha^2}) + \widetilde{\sigma}_1^2) \tag{5.14b}$$

$$(1-\alpha) \geq \epsilon_2(\alpha + \eta A_2 P_0(1 + P_s\sqrt{\alpha-\alpha^2}) + \widetilde{\sigma}_2^2)) \tag{5.14c}$$

$$\alpha \geq \epsilon_1(\eta A_1 P_0(1 + P_s\sqrt{\alpha-\alpha^2} + \widetilde{\sigma}_1^2) \tag{5.14d}$$

$$0 \leq \alpha \leq 1, \tag{5.14e}$$

where $A_k = \frac{P_0|h_k|^2}{|g_k h_u|^2}$, $\widetilde{\sigma}_k^2 = \frac{\sigma^2}{P_0|h_k|^2}$ and $k \in \{1,2\}$. The term $\sqrt{\alpha-\alpha^2}$ is concave when α is between 0 and 1 because its second-order derivative is negative. The objective function is a max-concave form, which gives convexity. However, constraints (5.14b), (5.14c) and (5.14d) are concave because a concave term located at the right-hand side of a greater sign gives concavity. As a result, problem P11 is non-convex and cannot be solved by CVX directly.

The main obstacle arises from the term $\sqrt{\alpha-\alpha^2}$, which exists in three constraints. One short-cut to handle the non-convex term is SCA, which approximates the non-convex term by its first-order Taylor series. The first-order Taylor series of $\sqrt{\alpha-\alpha^2}$ is expressed as

$$T(\alpha)=\sqrt{\alpha_0-\alpha_0^2}+\frac{1-2\alpha_0}{2\sqrt{\alpha_0-\alpha_0^2}}(\alpha-\alpha_0), \tag{5.15}$$

where α_0 is a fixed point initialised before solving the problem. By substituting $\sqrt{\alpha-\alpha^2}$ with $T(\alpha)$, a new sub-problem is obtained as follows:

$$\text{P12}: \max_{\{\alpha\}} \quad \sqrt{\alpha-\alpha^2} \tag{5.16a}$$

$$\text{s.t.} \quad (1-\alpha)\geq \epsilon_2(\alpha+\eta A_1 P_0(1+P_s T(\alpha))+\widetilde{\sigma}_1^2) \tag{5.16b}$$

$$(1-\alpha)\geq \epsilon_2(\alpha+\eta A_2 P_0(1+P_s T(\alpha))+\widetilde{\sigma}_2^2)) \tag{5.16c}$$

$$\alpha\geq \epsilon_1(\eta A_1 P_0(1+P_s T(\alpha)+\widetilde{\sigma}_1^2) \tag{5.16d}$$

$$0\leq\alpha\leq 1. \tag{5.16e}$$

Since $T(\alpha)$ is a linear term of α, all constraints in P12 are affine. As a result, problem P12 is a convex problem, which CVX can efficiently solve.

The next step is to optimise both sub-problems P10 and P12 iteratively. Assuming $\eta^{(t)*}$ and $\alpha^{(t)^*}$ are the optimal solution of P10 and P12 at the t-th iteration, respectively, $\eta^{(t)*}$ is the fixed value adopted by P12 in the current iteration, and $\alpha^{(t)^*}$ is the fixed value adopted by P10 in the next iteration. As for the fixed point α_0, it is also updated iteratively. The updating rule is given by

$$\alpha_0^{(t+1)}=\alpha^{(t)^*}. \tag{5.17}$$

Remark 1. *The alternating algorithm's performance is sensitive to the target data rate of two NOMA users because sub-problem P12 is sensitive to the target data rate of two NOMA users. Since the Taylor series $T(\alpha)$ substitutes the original term $\sqrt{\alpha-\alpha^2}$, the size of the feasible set greatly depends on $\epsilon_1 T(\alpha)$ and $\epsilon_2 T(\alpha)$. The approximate accuracy of SCA depends on the fixed point, specifically α_0 in P12. According to* (5.17)*, α_0 is equal to the optimal α in the previous iteration, which is related to ϵ_1 and ϵ_2; thus, P12 is sensitive to the target data rate of two NOMA users.*

Remark 2. *The initialisation is critical to the alternating algorithm. When constraints are* *stringent, randomly choosing initial points may make the optimisation problem infeasible. The details of finding feasible initial points were discussed in another work [88], and an algorithm was proposed.*

Algorithm 6 The alternating algorithm

Initialization $\alpha_0^{(0)}$, $t = 1$, $R_u^{(0)} = 0$ and t_d
while $R_u^{(t)} - R_u^{(t-1)} \geq t_d$ **do**
  Obtain $\eta^{(t)^*}$ by solving the sub-problem P10
  Update $\eta_0^{(t)} = \eta^{(t)^*}$
  Obtain $\alpha^{(t)^*}$ by solving the sub-problem P12
  Update $\alpha_0^{(t)} = \alpha^{(t)^*}$
  Calculate $R_u^{(t)}$ according to $\eta^{(t)^*}$ and $\alpha^{(t)^*}$
  Update $t = t + 1$
end while

The algorithm is summarised in Algorithm 6, where t_d is the threshold to determine when the loop stops and $\eta_0^{(t)}$ denotes the fixed value used to solve P12. Notably the initial $\alpha_0^{(0)}$ may not be the feasible choose of P10. In this case, more iterations are necessary until a feasible one is found.

5.4 Closed-form Derivation

In this section, the closed-form solution is derived. As discussed above, the original problem consists of two coupled optimisation variables. A common method to deal with a problem containing two variables is to fix one variable first and then find the optimal solution for another one. Note that the objective function and all constraints only contain the linear term of η, which indicates that an optimisation problem is a simple form with respect to η. For any given α, the optimisation problem only has one variable. We first fix α and find the optimal solution of η.

Once α is fixed, the objective function is a logarithmic function with the form of $\log_2(1 + x)$, which is a concave and monotonic function. We further notice that the objective function monotonously increases with η for any given α. Thus, the maximal value of the objective function is achieved when η reaches its upper bound. After some algebraic manipulations of constraints (5.10b) and (5.10c), there are three upper bounds of η as follows:

$$\eta \leq \frac{P_0|h_2|^2}{|g_2 h_u|^2|s_d|^2}\left(\frac{1-\alpha}{2^{R_t 2}-1} - \alpha - \frac{\sigma^2}{P_0|h_2|^2}\right) \triangleq B_1, \tag{5.18}$$

$$\eta \leq \frac{P_0|h_1|^2}{|g_1 h_u|^2|s_d|^2}\left(\frac{1-\alpha}{2^{R_{t2}}-1}-\alpha-\frac{\sigma^2}{P_0|h_1|^2}\right) \triangleq B_2, \tag{5.19}$$

and

$$\eta \leq \frac{P_0|h_1|^2}{|g_1 h_u|^2|s_d|^2}\left(\frac{\alpha}{2^{R_{t1}}-1}-\frac{\sigma^2}{P_0|h_1|^2}\right) \triangleq B_3. \tag{5.20}$$

It is noted that the largest value of η is the smallest upper bound among the inequalities above, which is related to R_{t1}, R_{t2}, α and channel gains. We first note that (5.18) and (5.19) have the same form, which indicates that only the channel gains, i.e., $|h_1|^2$, $|h_2|^2$, $|g_1|^2$ and $|g_2|^2$, affect the upper bound of η. For example, if $|h_1|^2$ is greater than $|h_2|^2$ and $|g_1|^2$ is smaller than $|g_2|^2$, than B_2 is larger than B_3. Thus, B_3 is the upper bound compared to B_2. Since the perfect CSI is well known by the BS, it is easy to decide the upper bound of η from (5.18) and (5.19). Without loss of generality, we first assume B_2 is smaller than B_3. Then, the next step is to compare B_2 and B_3. If $B_2 \leq B_3$, we have

$$\alpha \geq \frac{\epsilon_1}{\epsilon_1+\epsilon_1\epsilon_2+\epsilon_2}. \tag{5.21}$$

In this case, the upper bound of η is B_2 and the optimal solution

$$\eta^* = \min(B_2, 1). \tag{5.22}$$

Since the constraint (5.10e) limits the range of η, η^* also should be located in this range. To make the optimisation problem feasible, the constraint $B_2 \geq 0$ is necessary. After some algebraic manipulations, we have

$$\alpha \leq \frac{P_0|h_1|^2-\epsilon_2\sigma^2}{P_0|h_1|^2(1+\epsilon_2)}. \tag{5.23}$$

Remark 3. (5.23) *defines the upper bound of α whereas α is non-negative according to* (5.10d). *If $P_0|h_1|^2-\epsilon_2\sigma^2 \leq 0$, the optimisation problem is non-feasible. Note that the higher transmit power and the better channel quality will enlarge the feasible set. However, the higher noise power and the higher data rate requirement will narrow down the feasible set.*

If $\eta^* = B_2$, the original problem will be written as

$$\mathrm{P13}: \max_{\{\alpha\}} \quad \log_2\left(1 + \frac{|h_u|^4 A_1\left(\frac{1-\alpha}{\epsilon_2} - \alpha - \widetilde{\sigma}_1^2\right)}{\beta P_0 |h_{\mathrm{SI}}|^2 + \sigma_b^2}\right) \tag{5.24a}$$

$$\text{s.t.} \quad \frac{(1-\alpha)A_2}{\alpha A_2 + A_1\left(\frac{1-\alpha}{\epsilon_2} - \alpha - \widetilde{\sigma}_1^2\right) + \frac{\sigma^2}{|g_2 h_u|^2}} \geq \epsilon_2 \tag{5.24b}$$

$$\alpha \geq \epsilon_1\left(\frac{1-\alpha}{\epsilon_2} - \alpha\right) \tag{5.24c}$$

$$\frac{\epsilon_1}{\epsilon_1 + \epsilon_1\epsilon_2 + \epsilon_2} \leq \alpha \leq \frac{1 - \epsilon_2\widetilde{\sigma}_1^2}{1+\epsilon_2} \tag{5.24d}$$

$$B_2 \leq 1, \tag{5.24e}$$

by substituting $\eta^* = B2$ into the original problem.

Proposition 1. *The problem* P13 *can be equivalently transferred to* P14 *if* B_2 *is the smallest upper bound of* η*.*

$$\mathrm{P14}: \min_{\{\alpha\}} \quad \alpha \tag{5.25a}$$

$$\text{s.t.} \quad (5.24\text{d}) \quad (5.24\text{e})$$

Proof. We first define $f(\alpha) = \log_2\left(1 + C\left(\frac{1-\alpha}{\epsilon_2} - \alpha - \widetilde{\sigma}_1^2\right)\right)$, where $C = \frac{|h_u|^4 A_1}{\beta P_0 |h|_{\mathrm{SI}}^2 + \sigma_b^2}$. The first derivative of $f(\alpha)$ is expressed as

$$f'(\alpha) = \frac{-C - \epsilon_2}{\epsilon_2 + C(1 - \alpha - \epsilon_2\alpha - \epsilon_2\widetilde{\sigma}_1^2)} \tag{5.26}$$

According to constraint (5.24d) $\alpha \leq \frac{1-\epsilon_2\widetilde{\sigma}_1^2}{1+\epsilon_2}$, the minimum value of the liner term $1 - \alpha - \epsilon_2\alpha - \epsilon_2\widetilde{\sigma}_1^2$ is greater than 0. Therefore, $f'(\alpha) \leq 0$, which indicates $f(\alpha)$ is monotonously decreasing with α. The maximum value of $f(\alpha)$ is achieved when α arrives at its minimum value. Therefore, finding the minimum α is equivalent to finding the optimal solution of P13.

For constraint (5.24b), we first rewrite it as follows:

$$\frac{(1-\alpha)A_2}{\alpha A_2 + A_1\left(\frac{1-\alpha}{\epsilon_2} - \alpha - \widetilde{\sigma}_1^2\right) + A_2\widetilde{\sigma}_2^2} \geq \epsilon_2. \tag{5.27}$$

After some algebraic manipulations,

$$A_1(1-\alpha-\epsilon_2\alpha-\epsilon_2\widetilde{\sigma}_1^2) \leq A_2(1-\alpha-\epsilon_2\alpha-\epsilon_2\widetilde{\sigma}_2^2) \tag{5.28}$$

is obtained. Dividing the left-hand side and the right-hand side of (5.28) with $|s_d|^2$ gives the result that $B_2 \leq B_1$, which is the assumption of this case. As a result, the constraint is satisfied under the assumption that $B_2 \leq B_1$.

Constraint (5.24c) is the same as the left-hand side of constraint (5.24d); therefore, it can be removed. Finally, the proposition is proved. □

Constraint (5.24e) can be rewritten as

$$(A_1 + A_1\epsilon_2)\alpha + P_0 P_s \epsilon_2 \sqrt{\alpha - \alpha^2} \geq A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2. \tag{5.29}$$

Proposition 2. *The inequality* (5.29) *constructs a convex set.*

Proof. We first define $f(\alpha) = \sqrt{\alpha - \alpha^2}$. The second derivative of $f(\alpha)$ is expressed as follows:

$$f^{''}(\alpha) = -\frac{1}{4(\alpha-\alpha^2)\sqrt{\alpha-\alpha^2}}. \tag{5.30}$$

As $0 \leq \alpha \leq 1$, $f^{''}(\alpha) \leq 0$ and $P_0 P_s \epsilon_3 \geq 0$, the term $P_0 P_s \epsilon_2 \sqrt{\alpha - \alpha^2}$ is concave. $(A_1 + A_1\epsilon_2)\alpha$ is an affine term. Therefore, the linear combination of those two terms is concave. According to the convex optimisation, a concave term located at the left-hand side of a greater sign constructs convexity. Finally, the proposition is proved. □

The objective function of P14 and constraint (5.24d) are affine and constraint (5.24e) is convex. Therefore, problem P14 is a convex problem. Solving it using a convex optimisation toolbox, such as CVX in Matlab, is easy. Thus, the optimal solution for this case has been obtained. However, it is noted that the objective function is the optimisation variable itself. Therefore, P14 is a linear programming (LP) problem. The aim is to find the minimal value of the optimisation problem. Once the feasible set is settled, the minimum value can be obtained. The new target is to define the feasible set.

We further notice that the intersection of two sets, i.e., (5.24d) and (5.24e), is the feasible set of P14. $\mathcal{S}_1$ and $\mathcal{S}_2$ denote the feasible set of (5.24d) and (5.24e), respectively. $\mathcal{S}_1$ is a line segment which has been decided. The next step is to analyse constraint (5.24e) and find $\mathcal{S}_2$. Since A_1 and ϵ_2 are greater than 0 and α is located in the range of (0,1), the left-hand side of (5.29) is greater than 0. Therefore, if $A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2 \leq 0$, constraint (5.29) holds on the domain of α. If $A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2 \geq 0$, constraint (5.29) can be rewritten as follows:

$$K_1\alpha^2 - K_2\alpha + K_3 \leq 0, \tag{5.31}$$

where $K_1 = (A_1 + A_1\epsilon_2)^2 + (P_0P_s\epsilon_2)^2$, $K_2 = (P_0P_s\epsilon_2)^2 + 2(A_1 + A_1\epsilon_2)(A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2)$ and $K_3 = (A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2)^2$. The function $f(\alpha) = K_1\alpha^2 - K_2\alpha + K_3$ is a quadratic function with u-curve because $K_1 \geq 0$. According to the knowledge of the quadratic function, if $K_2^2 - 4K_1K_3 \leq 0$, $f(\alpha) \geq 0$ holds under all circumstances, which means constraint (5.29) cannot be satisfied. Then, P14 is infeasible. If $K_2^2 - 4K_1K_3 \geq 0$, the equation $f(\alpha) = 0$ has two roots, which are

$$X_1 = \frac{K_2 - \sqrt{K_2^2 - 4K_1K_3}}{2K_1}$$

and

$$X_2 = \frac{K_2 + \sqrt{K_2^2 - 4K_1K_3}}{2K_1}.$$

As a result, α is located in the range of (X_1, X_2). Finally, the feasible set of constraint (5.29) is summarised as follows:

$$\mathcal{S}_2 = \begin{cases} [0,1] & \text{If } c_1 \text{ holds;} \\ [X_1, X_2] & \text{If } c_2 \text{ holds;} \\ \varnothing & \text{If } c_3 \text{ holds,} \end{cases} \tag{5.32}$$

where c_1 denotes $A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2 \leq 0$, c_2 denotes $A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2 \geq 0$ and $K_2^2 - 4K_1K_c \geq 0$ and c_3 denotes $A_1 - A_1\epsilon_2\widetilde{\sigma}_1^2 - P_0\epsilon_2 \geq 0$ and $K_2^2 - 4K_1K_c \leq 0$. The feasible set of P14 is defined as $\mathcal{S} \triangleq \mathcal{S}_1 \cap \mathcal{S}_2$, which is

$$\mathcal{S} = \begin{cases} \mathcal{S}_1 & \text{If } c_1 \text{ holds;} \\ \mathcal{S}_1 & \text{If } c_2 \text{ holds and } X_1 \leq b_1 \text{ and } X_2 \geq b_2; \\ [X_1, X_2] & \text{If } c_2 \text{ holds and } X_1 \geq b_1 \text{ and } X_2 \leq b_2; \\ [X_1, b_2] & \text{If } c_2 \text{ holds and } X_1 \geq b_1 \text{ and } X_2 \geq b_2; \\ [b_1, X_2] & \text{If } c_2 \text{ holds and } X_1 \leq b_1 \text{ and } X_2 \leq b_2; \\ \varnothing & \text{If } c_3 \text{ holds,} \end{cases} \tag{5.33}$$

where $b_1 = \frac{\epsilon_1}{\epsilon_1+\epsilon_1\epsilon_2+\epsilon_2}$ and $b_2 = \frac{1-\epsilon_2\widetilde{\sigma}_1^2}{1+\epsilon_2}$ denote the lower bound and the upper bound of (5.24d), respectively. Then, the optimal solution is $\alpha^* = \min(\mathcal{S})$, whose expression is given as follows:

$$\alpha^* = \begin{cases} b_1 & \text{If } \mathcal{S} = \mathcal{S}_1 \text{ or } \mathcal{S} = [b_1, X_2]; \\ X_1 & \text{If } \mathcal{S} = [X_1, X_2] \text{ or } \mathcal{S} = [X_1, b_2]; \\ null & \text{If } \mathcal{S} = \varnothing. \end{cases} \tag{5.34}$$

We further discuss the case that $\eta^* = 1$

Proposition 3. *The original problem can be equivalently transferred to* P15 *if* $\eta^* = 1$.

$$\text{P15} : \max_{\{\alpha\}} \quad \sqrt{\alpha - \alpha^2} \tag{5.35a}$$

$$\text{s.t.} \quad (5.24\text{d})$$

$$B_2 \geq 1 \tag{5.35b}$$

Proof. If $\eta^* = 1$, we have $1 \leq B_2 \leq \max(B_1, B_3)$. Note that R_1, R_2 and R_{12} are increasing with η decreasing. Therefore, R_1, R_2 and R_{12} are monotonically decreasing functions with respect to η. We further note that $R_{12}(\eta = B_2) = R_{t2}$, $R_2(\eta = B_1) = R_{t2}$ and $R_1(\eta = B_3) = R_{t1}$. As a result, $R_{12}(\eta = 1) \geq R_{t2}$, $R_2(\eta = 1) \geq R_{t2}$ and $R_1(\eta = 1) \geq R_{t1}$. Constraints (5.10b) and (5.10c) are satisfied. □

P15 is difficult to be solved due to the non-convexity of (5.35b) and the non-convex objective function. After applying the algebraic transformation to (5.35b), we have

$$K_1\alpha^2 - K_2\alpha + K_3 \geq 0, \tag{5.36}$$

Similar to analysis above, if $K_2^2 - 4K_1K_3 \leq 0$, the function $f(\alpha) \geq 0$ holds on the domain of α, which means constraint (5.35b) has already been satisfied. If $K_2^2 - 4K_1K_3 \geq 0$, X_1 and X_2 are the two roots of $f(\alpha) = 0$. As a result, (5.35b) can be equivalently transformed to $\alpha \leq X_1$ and $\alpha \geq X_2$ when $K_2^2 - 4K_1K_3 \geq 0$ holds. P15 is eventually transformed to

$$\text{P16} : \max_{\{\alpha\}} \quad \sqrt{\alpha - \alpha^2} \tag{5.37a}$$

$$\text{s.t.} \quad \text{(5.24d)}$$

$$\alpha \leq X_1 \tag{5.37b}$$

$$\alpha \geq X_2 \tag{5.37c}$$

Remark 4. *Constraints* (5.37b) *and* (5.37c) *only exist only when* $K_2^2 - 4K_1K_3 \geq 0$. *If* $K_2^2 - 4K_1K_3 \leq 0$*, the constraint* (5.35b) *is always satisfied, therefore, constraints* (5.37b) *and* (5.37c) *can be removed.*

It is noted that P16 is still a non-convex problem, but all linear constraints make P16 an LP problem. To address this problem, we first to find the feasible set of P16, which is given as follows:

$$\mathcal{S} = \begin{cases} \mathcal{S}_1 & \text{If } K_2^2 - 4K_1K_3 \leq 0 \\ & \text{or } X_2 \leq b_1 \text{ or } X_1 \geq b_2; \\ [b_1, X_1] \cup [X_2, b_2] & \text{If } X_1 \geq b_1 \text{ and } X_2 \leq b_2; \\ [b_1, X_1] & \text{If } b_1 \leq X_1 \leq b_2 \text{ and } X_2 \geq b_2; \\ [X_2, b_2] & \text{If } X_1 \leq b_1 \text{ and } b_1 \leq X_2 \leq b_2; \\ \varnothing & \text{If } X_1 \leq b_1 \text{ and } X_2 \geq b_2. \end{cases} \tag{5.38}$$

The next step is to find the optimal solution of α, which can minimise the objective function. From the proof of proposition 2, we know the objective function is concave. The optimal point is reached when its first-order derivative equals 0. By letting $f'(\alpha) = 0$, we have $\alpha^* = 0.5$ without any other constraints. If $0.5 \in \mathcal{S}$, the optimal solution of P16 is 0.5; otherwise, there are some cases.

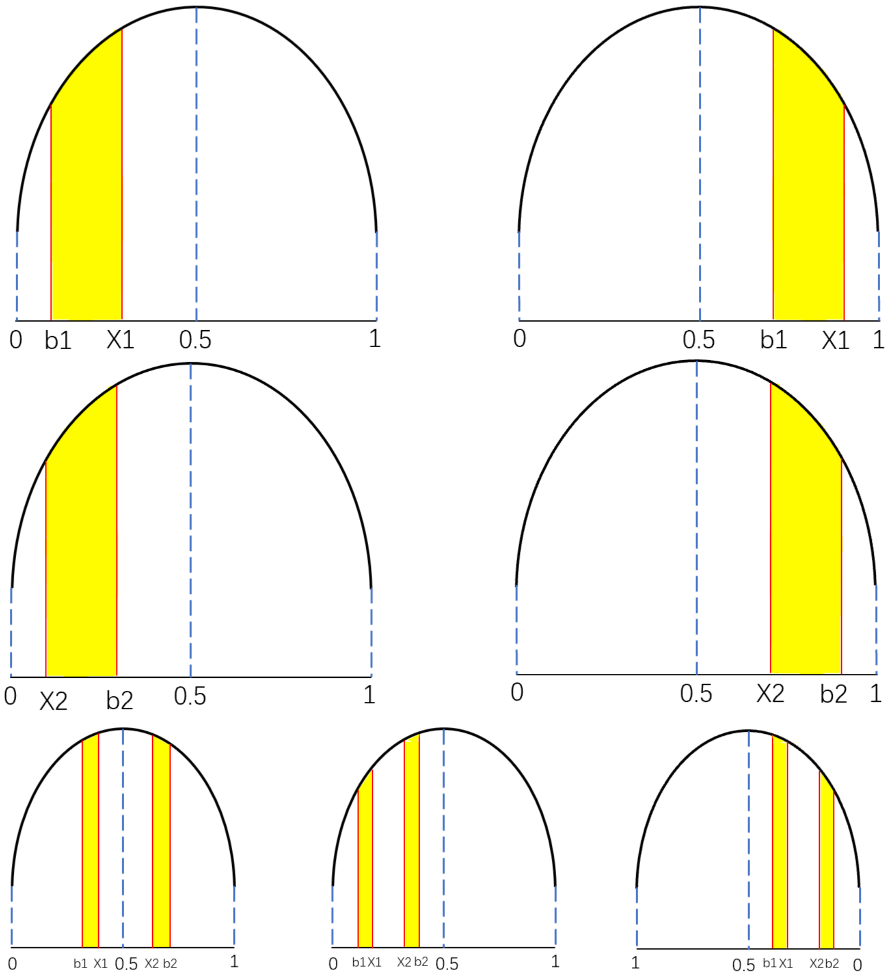

Figure 5.2. Several cases of the optimal solution.

Fig. 5.2 illustrates the optimal solution of α in different cases when $0.5 \notin \mathcal{S}$. Note that the optimal solution is always reached at the boundary of $\mathcal{S}$. Therefore, the boundary of $\mathcal{S}$ that can minimise the objective function is the optimal solution. The expression of the optimal solution of α is given by

If $0.5 \in \mathcal{S}$

$$\alpha^* = 0.5;$$

If $\mathcal{S} = \mathcal{S}_1 = [b_1, b_2]$

$$\alpha^* = \begin{cases} b_2 & \text{If } b_2 \leq 0.5 \\ b_1 & \text{If } b_1 \geq 0.5; \end{cases}$$

If $\mathcal{S} = [b_1, X_1] \cup [X_2, b_2]$

$$\alpha^* = \begin{cases} b_2 & \text{If } b_2 \leq 0.5 \\ b_1 & \text{If } b_1 \geq 0.5 \\ X_1 & \text{If } |X_1 - 0.5| \leq |X_2 - 0.5| \\ X_2 & \text{If } |X_1 - 0.5| \geq |X_2 - 0.5|; \end{cases}$$

If $\mathcal{S} = [b_1, X_1]$

$$\alpha^* = \begin{cases} X_1 & \text{If } X_1 \leq 0.5 \\ b_1 & \text{If } b_1 \geq 0.5; \end{cases}$$

If $\mathcal{S} = [X_2, b_2]$

$$\alpha^* = \begin{cases} b_2 & \text{If } b_2 \leq 0.5 \\ X_2 & \text{If } X_2 \geq 0.5. \end{cases}$$

We next assume B_3 is smaller than B_2, which indicates the upper bound of η is B_3 and the optimal solution

$$\eta^* = \min(B_3, 1). \tag{5.39}$$

Considering the same reason that η is positive, the constraint $B_3 \geq 0$ is necessary. If $\eta* = B3$, the original problem will be written as

$$\begin{aligned} \text{P17} : \max_{\{\alpha\}} \quad & \log_2 \left(1 + \frac{|h_u|^4 A_1 \left(\frac{\alpha}{\epsilon_2} - \widetilde{\sigma}_1^2 \right)}{\beta P_0 |h_{\text{SI}}|^2 + \sigma_b^2} \right) && \text{(5.40a)} \\ \text{s.t.} \quad & \frac{(1-\alpha) A_2}{\alpha A_2 + A_1 \left(\frac{\alpha}{\epsilon_1} - \widetilde{\sigma}_1^2 \right) + \frac{\sigma^2}{|g_2 h_u|^2}} \geq \epsilon_2 && \text{(5.40b)} \\ & \alpha \leq \epsilon_1 \left(\frac{1-\alpha}{\epsilon_2} - \alpha \right) && \text{(5.40c)} \\ & \epsilon_1 \widetilde{\sigma}_1^2 \leq \alpha \leq \frac{\epsilon_1}{\epsilon_1 + \epsilon_1 \epsilon_2 + \epsilon_2} && \text{(5.40d)} \\ & B_3 \leq 1 && \text{(5.40e)} \end{aligned}$$

Similar to the proof of Proposition 1, P17 can be equivalently transferred to P18, which is

$$\text{P18} : \max_{\{\alpha\}} \quad \alpha \tag{5.41a}$$

$$\text{s.t.} \qquad (5.40\text{d}) \quad (5.40\text{e}).$$

The constraint (5.40e) can be rewritten as

$$A_1\alpha - P_0P_s\epsilon_1\sqrt{\alpha - \alpha^2} \leq A_1\epsilon_1\widetilde{\sigma}_1^2 + \epsilon_2P_0. \tag{5.42}$$

According to the proof of Proposition 2, constraint (5.42) constructs a convex set. As a result, P18 is a convex problem that the convex optimisation toolbox can solve. Constraint (5.42) needs to be simplified to obtain the closed-form solution. To obtain the closed-form solution, constraint (5.42) needs to be simplified. It is noted that the right-hand side of (5.42) is greater than 0; however, it cannot determine the sign of the left side because of α. Since recklessly squaring both sides will change the direction of the inequality sign, it is important to decide on the sign on the left-hand side.

Given the fact that the right-hand side of (5.42) is positive, the inequality holds under all circumstances if the right-hand side is negative. By letting $A_1\alpha - P_0P_s\epsilon_1\sqrt{\alpha - \alpha^2} \leq 0$, the set

$$0 \leq \alpha \leq \frac{(P_0P_s\epsilon_1)^2}{A_1^2 + (P_0P_s\epsilon_1)^2} \tag{5.43}$$

is obtained. For the simple notation, b_3 and b_4 are used to denote $\frac{(P_0P_s\epsilon_1)^2}{A_1^2+(P_0P_s\epsilon_1)^2}$ and $\epsilon_1\widetilde{\sigma}_1^2$, respectively. If α is from this set, constraint (5.42) is always satisfied. P18 can be split into two sub-problems based on (5.43), which are as follows:

$$\text{P18}_1 : \max_{\{\alpha\}} \quad \alpha \tag{5.44a}$$
$$\text{s.t.} \qquad (5.40\text{d}) \quad (5.43),$$

and

$$\text{P18}_2 : \max_{\{\alpha\}} \quad \alpha \tag{5.45a}$$
$$\text{s.t.} \qquad (5.40\text{d}) \quad (5.42)$$
$$\alpha \geq b_3. \tag{5.45b}$$

All constraints in P18_1 are linear, and the optimal solution is the maximal value of the feasible set. Therefore, the optimal solution α_1^* of P18$_1$ are

$$\alpha_1^* = \begin{cases} \min(b_1, b_3) & \text{If } b_4 \le b_3; \\ null & \text{If } b_4 \ge b_3. \end{cases} \tag{5.46}$$

For the P18$_2$, the constraint (5.42) can be rewritten as

$$K_4\alpha^2 - K_5\alpha + K_6 \le 0, \tag{5.47}$$

where $K_4 = A_1^2 + (P_0P_s\epsilon_1)^2$, $K_5 = 2A_1(A_1\epsilon_1\widetilde{\sigma}_1^2 + \epsilon_2 P_0) + (P_0P_s\epsilon_1)^2$ and $K_6 = (A_1\epsilon_1\widetilde{\sigma}_1^2 + \epsilon_2 P_0)^2$, by squaring both sides of (5.42) because of the restriction that forces both sides of (5.42) to be positive from constraint (5.45b). The function $g(\alpha) = K_4\alpha^2 - K_5\alpha + K_6$ is a quadratic function with u-curve because $K_4 \ge 0$. If $K_5^2 - 4K_4K_6 \ge 0$, the equation $g(\alpha) = 0$ has two roots, which are

$$X_3 = \frac{K_5 - \sqrt{K_5^2 - 4K_4K_6}}{2K_4}$$

and

$$X_4 = \frac{K_5 + \sqrt{K_5^2 - 4K_4K_6}}{2K_4}.$$

The following analysis is similar to the above. We only provide the result due to space limitations.

$$\alpha_2^* = \begin{cases} X_4 & \text{If } b_4 \le b_3 \le b_1 \text{ and } b_4 \le X_4 \le b_3; \\ b_3 & \text{If } b_4 \le b_3 \le b_1 \text{ and } X_3 \le b_3 \le X_4; \\ X_4 & \text{If } b_3 \le b_4 \text{ and } b_4 \le X_4 \le b_1; \\ b_1 & \text{If } b_3 \le b_4 \text{ and } X_3 \le b_1 \le X_4; \\ null & \text{others.} \end{cases} \tag{5.48}$$

Note that α_2^* only exists when $K_5^2 - 4K_4K_6 \ge 0$, otherwise α_2^* is $null$. Therefore, the optimal solution of P18 is

$$\alpha^* = \max(\alpha_1^*, \alpha_2^*). \tag{5.49}$$

When $\eta^* = 1$, the original problem can be equivalently transferred to

$$\text{P19} : \max_{\{\alpha\}} \quad \sqrt{\alpha - \alpha^2} \tag{5.50a}$$

$$\text{s.t.} \quad (5.40\text{d})$$

$$B_3 \geq 1. \tag{5.50b}$$

Constraint (5.50b) is non-convex, which can be rewritten as

$$A_1\alpha - P_0P_s\epsilon_1\sqrt{\alpha - \alpha^2} \geq A_1\epsilon_1\widetilde{\sigma}_1^2 + \epsilon_2P_0. \tag{5.51}$$

Note that if $A_1\alpha - P_0P_s\epsilon_1\sqrt{\alpha - \alpha^2}$ is negative, the problem is infeasible. Therefore, P19 can be transformed to

$$\text{P20} : \max_{\{\alpha\}} \quad \sqrt{\alpha - \alpha^2} \tag{5.52a}$$

$$\text{s.t.} \quad (5.40\text{d})$$

$$\alpha \geq b_3 \tag{5.52b}$$

$$K_4\alpha^2 - K_5\alpha + K_6 \geq 0. \tag{5.52c}$$

Two liner constraints (5.40d) and (5.52b) construct a set which is

$$\mathcal{S}_3 = \begin{cases} [b4, b1] & \text{If } b_3 \leq b_4 \\ [b3, b1] & \text{If } b_4 \leq b_3 \leq b_1 \\ \varnothing & \text{If } b_3 \geq b_1. \end{cases} \tag{5.53}$$

Following the strategy solving P16, the feasible set of P20 is given by

If $\mathcal{S}_3 = [b_4, b_1]$,

$$\mathcal{S} = \begin{cases} \mathcal{S}_3 & \text{If } K_5^2 - 4K_4K_6 \leq 0 \\ & \text{or } X_4 \leq b_4 \text{ or } X_3 \geq b_1 \\ [b_4, X_3] \cup [X_4, b_1] & \text{If } X_3 \geq b_4 \text{ and } X_4 \leq b_1 \\ [b_4, X_3] & \text{If } b_4 \leq X_3 \leq b_1 \text{ and } X_4 \geq b_1 \\ [X_4, b_1] & \text{If } X_3 \leq b_4 \text{ and } b_4 \leq X_4 \leq b_1 \\ \varnothing & \text{If } X_3 \leq b_4 \text{ and } X_4 \geq b_1; \end{cases} \tag{5.54}$$

If $\mathcal{S}_3 = [b_3, b_1]$,

$$\mathcal{S} = \begin{cases} \mathcal{S}_3 & \text{If } K_5^2 - 4K_4K_6 \leq 0 \\ & \text{or } X_4 \leq b_3 \text{ or } X_3 \geq b_1 \\ [b_3, X_3] \cup [X_4, b_1] & \text{If } X_3 \geq b_3 \text{ and } X_4 \leq b_1 \\ [b_3, X_3] & \text{If } b_3 \leq X_3 \leq b_1 \text{ and } X_4 \geq b_1 \\ [X_4, b_1] & \text{If } X_3 \leq b_3 \text{ and } b_3 \leq X_4 \leq b_1 \\ \varnothing & \text{If } X_3 \leq b_3 \text{ and } X_4 \geq b_1. \end{cases} \tag{5.55}$$

Then, the strategy described in Fig. 5.2 is applied to obtain the optimal solution of α. The optimal solution of α^* can be summarised as follows:

If $0.5 \in \mathcal{S}$

$$\alpha^* = 0.5;$$

If $\mathcal{S} = \mathcal{S}_3 = [b_4, b1]$

$$\alpha^* = \begin{cases} b_1 & \text{If } b_1 \leq 0.5 \\ b_4 & \text{If } b_4 \geq 0.5; \end{cases}$$

If $\mathcal{S} = [b_4, X_3] \cup [X_4, b_1]$

$$\alpha^* = \begin{cases} b_1 & \text{If } b_1 \leq 0.5 \\ b_4 & \text{If } b_4 \geq 0.5 \\ X_3 & \text{If } |X_3 - 0.5| \leq |X_4 - 0.5| \\ X_4 & \text{If } |X_3 - 0.5| \geq |X_4 - 0.5|; \end{cases}$$

If $\mathcal{S} = [b_4, X_3]$

$$\alpha^* = \begin{cases} X_3 & \text{If } b_3 \leq 0.5 \\ b_4 & \text{If } b_4 \geq 0.5; \end{cases}$$

If $\mathcal{S} = [X_4, b_1]$

$$\alpha^* = \begin{cases} b_1 & \text{If } b_1 \leq 0.5 \\ X_4 & \text{If } X_4 \geq 0.5; \end{cases}$$

If $\mathcal{S} = \mathcal{S}_3 = [b_3, b_1]$

$$\alpha^* = \begin{cases} b_1 & \text{If } b_1 \leq 0.5 \\ b_3 & \text{If } b_3 \geq 0.5; \end{cases}$$

If $\mathcal{S} = [b_3, X_3] \cup [X_4, b_1]$

$$\alpha^* = \begin{cases} b_1 & \text{If } b_1 \leq 0.5 \\ b_3 & \text{If } b_3 \geq 0.5 \\ X_3 & \text{If } |X_3 - 0.5| \leq |X_4 - 0.5| \\ X_4 & \text{If } |X_3 - 0.5| \geq |X_4 - 0.5|; \end{cases}$$

If $\mathcal{S} = [b_3, X_3]$

$$\alpha^* = \begin{cases} X_3 & \text{If } X_3 \leq 0.5 \\ b_3 & \text{If } b_3 \geq 0.5. \end{cases}$$

Finally, the closed-form of the optimal solution is obtained.

Remark 5. *If a network contains multiple users, it is impossible to apply the NOMA technique to all users in practice because the complexity of SIC is very high, especially for the last decoded user. One efficient strategy is to divide users into different clusters and NOMA is only applied within a cluster. By assuming one cluster only contains two users, the closed-form solution above is valid in this cluster.*

5.5 Simulation Results

In this section, simulation results are provided for the alternating algorithm, the closed-form solution and the random selection strategy. Monte Carlo method is adopted to ensure the fairness of the simulation. The total number of experiments is set as 50. The theoretical and experimental results of the alternating algorithm and random selection strategy are calculated by sharing the same CSI in each experiment. All curves are generated by calculating the mean of all theoretical results and experimental results of the alternating algorithm and the random selection strategy. In the simulation, all the channels between every two nodes follow the Rician fading distribution, which can be modelled as follows:

$$\mathbf{f} = \sqrt{\frac{\kappa}{1+\kappa}}\mathbf{f}^{\mathrm{LoS}} + \sqrt{\frac{1}{1+\kappa}}\mathbf{f}^{\mathrm{nLoS}}, \tag{5.56}$$

where κ is the Rician factor, $\mathbf{f}^{\mathrm{LoS}}$ denotes the line-of-sight (LoS) component and $\mathbf{f}^{\mathrm{nLoS}}$ denotes the non-LoS (nLoS) component. In our simulation, $\mathbf{f}^{\mathrm{LoS}}$ is a fixed constant 1, $\mathbf{f}^{\mathrm{nLoS}}$ is generated from a Rayleigh fading distribution and κ is set as 1. The impact of distance on the channel gain is also considered. The channel gains are generated by

$$h_k = \frac{h_k^*}{\sqrt{d_k^{\delta_k}}}, k \in \{1, 2, u\}, \tag{5.57}$$

where h_k^* follows the distribution described by (5.56), d_k denotes the distance between every two nodes and δ_k denotes is the path loss exponent. In the simulation, the path loss exponents of each link are set to be fixed, which is $\delta_1 = \delta_2 = 1.6$ and $\delta_u = 1.3$. The noise power is set as -80 dBm.

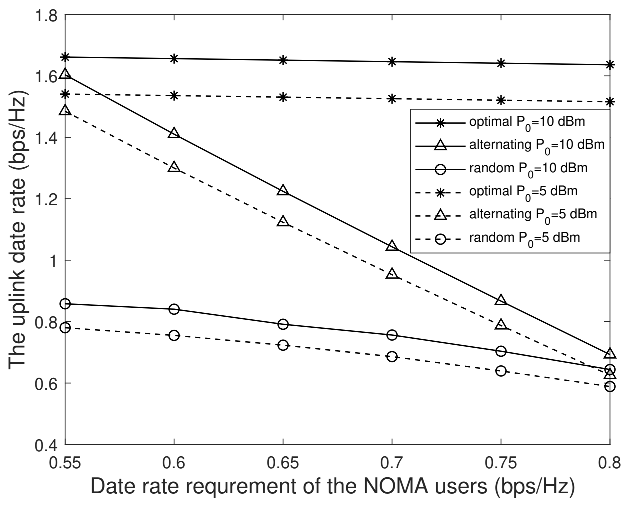

Figure 5.3. The uplink data rate as a function of the minimal data rate requirement of the NOMA users.

Fig. 5.3 indicates the uplink data rate as a function of the minimal data rate requirement of two NOMA users. The transmit power P_0 is set as 5 dBm and 10 dBm, respectively. The minimal data rate requirements of two NOMA users are assumed to be identical. From the figure, we notice that the optimal solution is not sensitive to the minimal data rate requirement, which means improving the QoS of two NOMA users will not scarify the uplink data rate. However, the alternating algorithm and the random benchmark are more sensitive to the minimal data rate requirement of the NOMA users. When the QoS requirement increases, the uplink data rate will decrease accordingly.

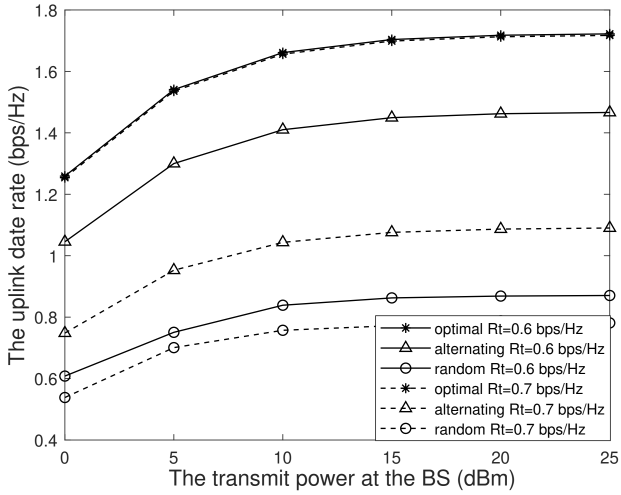

Figure 5.4. The uplink data rate as a function of the transmit power at the BS.

Fig. 6.2 shows the uplink data rate as a function of the transmit power at the BS. The target data rate of two NOMA users' R_t is set as 0.6 bps/Hz and 0.7 bps/Hz, respectively. The uplink data rate increases with increasing the transmit power at the BS in all schemes. However, it is noted that the uplink data rate will enter a plateau when the transmit power reaches a certain level. This figure shows that the benefit of high transmit power may be marginal, resulting in energy waste. In practice, it is critical to choose a transmit power wisely which can guarantee the best performance without causing energy waste.

Fig. 5.5 shows the uplink data rate as a function of the distance between the uplink device and the BS. The transmit power P_0 is set as 5 dBm and 10 dBm, respectively. In this simulation, we assume that the distance between the uplink device and two NOMA users is the same and only the distance between the BS and the uplink device changes. The results show that when the uplink device is closer to the BS, the uplink data rate improves. When the uplink device is too far from the BS, the data rate is nearly 0. Therefore, the uplink device should be deployed to a location close to the BS. The results also show that the optimal solution is better than all benchmarks.

Fig. 5.6 shows the uplink data rate as a function of the distance between the BS and two

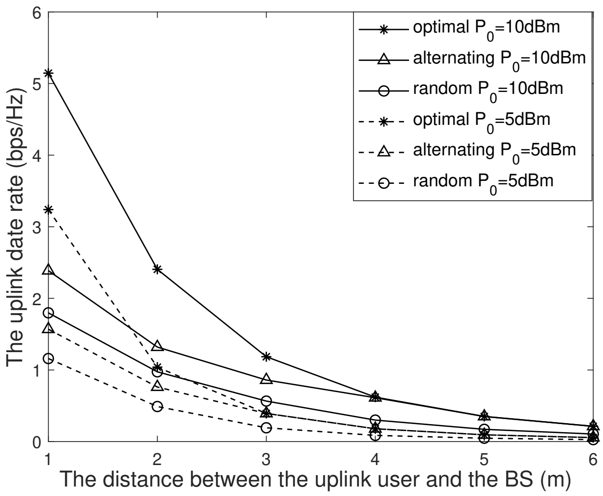

Figure 5.5. The uplink data rate as a function of the distance between the uplink device and the BS.

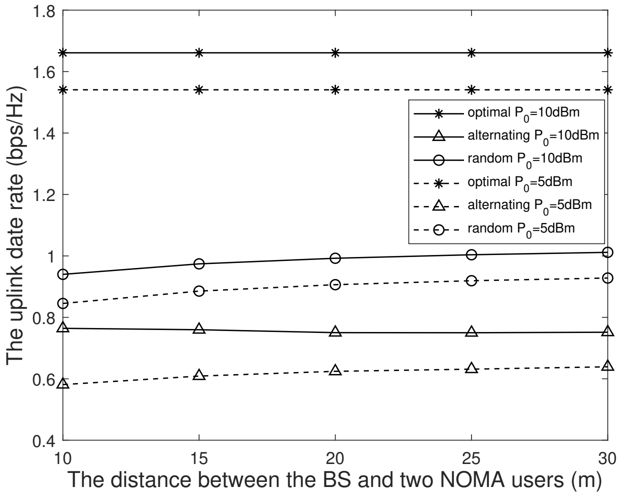

Figure 5.6. The uplink data rate as a function of the distance between the BS and two NOMA users.

NOMA users. The transmit power P_0 is set as 5 dBm and 10 dBm, respectively. In this simulation, the distances between the BS and two NOMA users are assumed to be the same. The target data rate of two NOMA users is set as 0.7 bps/Hz. The results show that this distance will not affect the uplink data rate in all schemes. The reason is that the backscattering coefficient η can always reach the maximum value of 1 under the current QoS condition of NOMA users. In other words, the QoS condition of two NOMA users is still satisfied when the uplink device fully backscatters its signal. Since the objective function (5.3) is only related to η and η is always equals to 1, the uplink data rate remains the same. Although the uplink data rate will not be affected by this distance, the longer distance between two NOMA users and the BS will bring difficulty in satisfy the QoS requirement. Hence, this distance cannot be too large in practice.

Fig. 5.7 compares the grid search result and the optimal result to verify the authenticity of

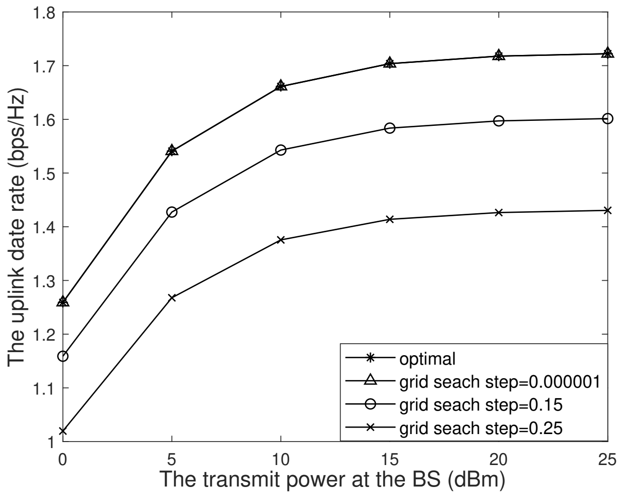

Figure 5.7. The optimal solution versus grid search solutions.

the optimal solution. We first choose many different values of α from 0 to 1 and calculate the corresponding value of η. Then, these α-η pairs are used to calculate the corresponding uplink data rate. The desired outcome is the maximum uplink data rate. In this simulation, three step sizes are adopted. When the step size is set to 0.000001, the grid search is nearly upgraded to an exhaustive search, whose performance nearly coincides with the optimal solution. If the step size increases, the performance will decrease; however, the time complexity will also decrease.

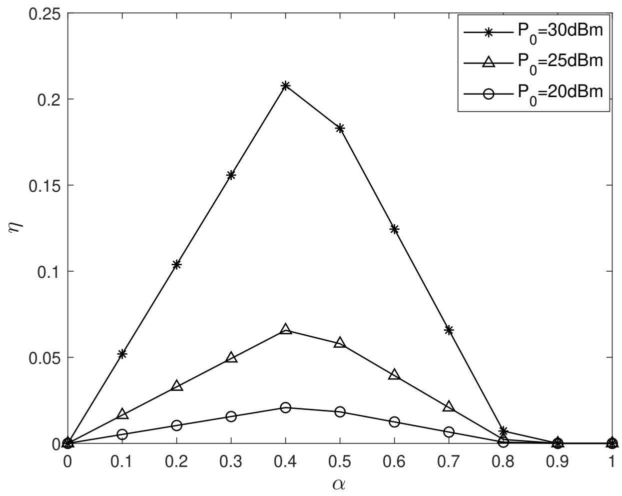

Figure 5.8. The relationship between α and η.

Fig. 5.8 shows the relationship between α and η under different transmit power P_0. The figure shows that the large value of α will cause η to become 0, which means some QoS constraints are violated. Three curves represent the $\alpha - \eta$ relationship under different transmit power levels. A large transmit power will cause a larger η.

The optimal result significantly outperforms the alternating algorithm. One reason is that SCA scarifies optimality because the Taylor series substitutes the original non-convex term. Another reason is that the initialisation of the fixed point introduced by SCA also affects the performance of the alternating algorithm. A wiser strategy of initialising the fixed point can generate better performance. The coefficients in the random strategy are not optimised; thus, its performance is the worst.

5.6 Conclusion

This study investigated a backscatter-assisted NOMA network, consisting of two downlink NOMA users and one uplink device. This chapter aimed to maximise the data rate of the uplink device while also ensuring the QoS requirements of two NOMA users. The alternating algorithm was proposed to solve this problem efficiently, and it can be extended to solve a scenario containing multiple downlink NOMA users. Then, the closed-form solution of a two-user case was derived. The study has practical significance because it allows an extra IoT device to be added to an existing NOMA network, significantly improving the legacy system's capacity.

In order to solve optimisation problems in a backsttering-NOMA network efficiently, the deep learning will be introduced in the next chapter. The comparison of the deep learning algorithm and the traditional convex algorithm is investigated as well.

Chapter 6

BAC-NOMA for Secondary Transmission

6.1 Introduction

Ultra-massive machine type communications (umMTC) as a key technique of the envisioned sixth-generation (6G) communication system is the focus of research [89]. The main feature of umMTC is that massive low-power devices, e.g., energy-constrained Internet of Things (IoT) sensors, are connected. It is challenging to serve a huge number of devices simultaneously due to the spectrum constraint.

To tackle this spectrum challenge, non-orthogonal multiple access (NOMA) has been proposed as a promising technique to promote spectrum cooperation between wireless users and devices [66]. NOMA allows different users to share the spectrum resource simultaneously, where successive interference cancellation (SIC) is adopted to partially remove co-channel interference. The energy constraint is another challenge in umMTC networks, since it is difficult to equip small IoT devices with batteries. One way to achieve energy cooperation among energy-constrained devices is backscatter communication (BackCom), where a backscatter (BAC) circuit of a BAC device can be excited by the signal from another device[90].

The combination of NOMA and BAC can further improve resource efficiency. Some existing works have investigated NOMA-BAC scenarios [84], [91]–[93]. The system models considered in [91]–[93] contain two NOMA users, one backscatter device (BD) and a single-antenna base station (BS). The work [84] investigated the system model which contains a single-antenna BS and two NOMA users equipped with backscatter circuit. Unlike the existing works above, the BS in this work is equipped with multiple antennas and the result shows that the a multiple-antenna BS can help to improve the performance. This chapter also has a different focus compared with the aforementioned works. The aim of this work is to find the optimal beam vectors and backscattering coefficient for maximising the sum rate of two NOMA users while aforementioned works focused on performance analysis. First, a sum rate of two NOMA users maximisation problem is formulated. A closed form expression of the optimal backscattering coefficient is provided, and then a learning based algorithm and a semi-definite relaxation (SDR) based algorithm are proposed to design beamforming vectors. The computer simulation results show both algorithms have their own advantages and disadvantages, which provides flexible strategy options.

6.2 System Model and Problem Formulation

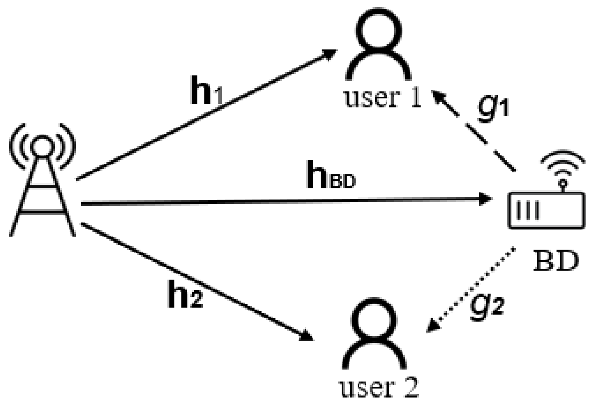

Figure 6.1. The system model.

The system model shown in Fig. 6.1 consists of one BS equipped with M antennas, one signal-antenna passive BD and two signal-antenna NOMA users. The BD only communicates with user 1 by backscattering its signal, while the reflected signal also interferes with user 2. Since the BD is a passive device, all energy consumed to backscatter signals is from the received signal from the BS. The reflecting coefficient of the BD, denoted by η, can be optimized to balance the achievable data rate and the interference. The channels of user 1, user 2 and the BD are denoted by $\mathbf{h}_1 \in \mathbb{C}^{M\times 1}$, $\mathbf{h}_2 \in \mathbb{C}^{M\times 1}$ and $\mathbf{h}_{\mathrm{BD}} \in \mathbb{C}^{M\times 1}$, respectively. The channels between the BD and two users are denoted by g_1 and g_2, respectively. In this paper, the Rician fading model is considered for all channels. The BS transmits the superposed signal, as shown in follows

$$\mathbf{x} = \mathbf{w}_1 s_1 + \mathbf{w}_2 s_2, \tag{6.1}$$

where $\mathbf{w}_1 \in \mathbb{C}^{M\times 1}$ and $\mathbf{w}_2 \in \mathbb{C}^{M\times 1}$ denote the beamforming vectors for two users and s_1 and s_2 denote the signals sent to two users, satisfying $\mathrm{E}\{|s_1|^2\} = \mathrm{E}\{|s_2|^2\} = 1$, where $\mathrm{E}\{\cdot\}$ is the expectation operation. As a result, the received signal at two users can be expressed as follows:

$$y_i = \mathbf{h}_i^H \mathbf{x} + \sqrt{\eta} g_i \mathbf{h}_{\mathrm{BD}}^H \mathbf{x} c + w_i, i \in \{1,2\}, \tag{6.2}$$

where ^{H} is the Hermitian transpose, c denotes the device-to-device (D2D) signal from the BD to user 1 and w denotes the additive white Gaussian noise (AWGN). It is worth pointing out that the meaning of the second term in (6.2) is different for user 1 and user 2. Specifically, it is the desired signal for user 1 and the interference signal for user 2.

The SIC technique is performed at the user with the strong channel gain to remove extra interference. [1] In this paper, it is assumed that user 1 first decodes user 2's signal followed by its own signal and then BD's signal. The data rate for decoding user 2's signal at user 1 is given by

$$R_{1,2} = \log_2\left(1 + \frac{|\mathbf{h}_1^H \mathbf{w}_2|^2}{|\mathbf{h}_1^H \mathbf{w}_1|^2 + \sum_{k=1}^{2} \eta |g_1|^2 |\mathbf{h}_{\mathrm{BD}}^H \mathbf{w}_k|^2 + \sigma_1^2}\right), \tag{6.3}$$

where σ_1^2 denotes the noise power at user 1. The signal for user 2 can be removed after successfully decoding it. The achievable data rate of user 1 for decoding s_1 is given by

$$R_1 = \log_2\left(1 + \frac{|\mathbf{h}_1^H \mathbf{w}_1|^2}{\sum_{k=1}^{2} \eta |g_1|^2 |\mathbf{h}_{\mathrm{BD}}^H \mathbf{w}_k|^2 + \sigma_1^2}\right). \tag{6.4}$$

At this stage, both s_1 and s_2 are known by user 1, which enables the decoding of c. When decoding c, the superposed signal $\mathbf{x}$ plays the role of fast-varying channel components [71], making the fast channel $\hat{h}_{\mathrm{BD}} = \mathbf{h}_{\mathrm{BD}}^H \mathbf{x}$. The achievable data rate of c under the fast fading is given by

$$R_{1,c} = \log_2\left(1 + \frac{\eta |g_1|^2 |\hat{h}_{\mathrm{BD}}|^2}{\sigma_1^2}\right). \tag{6.5}$$

$\hat{h}_{\mathrm{BD}}$ varies by the two symbols, including s_1 and s_2, and thus the average rate for decoding c can be written as

$$\bar{R}_{1,c} = \mathrm{E}_{|\hat{h}_{\mathrm{BD}}|^2}\left[\log_2\left(1 + \frac{\eta |g_1|^2 |\hat{h}_{\mathrm{BD}}|^2}{\sigma_1^2}\right)\right]. \tag{6.6}$$

For user 2, it decodes the desired signal directly by treating other terms as interference. The achievable data rate of user 2 is given by

$$R_2 = \log_2\left(1 + \frac{|\mathbf{h}_2^H \mathbf{w}_2|^2}{|\mathbf{h}_2^H \mathbf{w}_1|^2 + \sum_{k=1}^{2} \eta |g_2|^2 |\mathbf{h}_{\mathrm{BD}}^H \mathbf{w}_k|^2 + \sigma_2^2}\right), \tag{6.7}$$

where σ_2^2 is the noise power at user 2.

The aim of this paper is to maximise the sum rate of two users while guaranteeing the quality of service (QoS) for BD. Therefore, a sum rate maximisation problem can be formulated as follows:

[1] A passive device will cause time synchronisation error to a NOMA user during SIC [94]. It is assumed synchronisation error is ignored and this work can be considered as an upper bound of the practical result.

$$
\begin{aligned}
\text{P21} \max_{\{\mathbf{w}_1,\mathbf{w}_2,\eta\}} \quad & R_1 + \min\{R_{1,2}, R_2\} && (6.8a)\\
\text{s.t.} \quad & \bar{R}_{1,c} \geq R_{tc}, && (6.8b)\\
& ||\mathbf{w}_1||^2 + ||\mathbf{w}_2||^2 \leq P_0, && (6.8c)\\
& 0 \leq \eta \leq 1, && (6.8d)
\end{aligned}
$$

where R_{tc} denotes the target data rate of BD. It is worth pointing out that the achievable data rate for decoding s_2 is determined by both $R_{1,2}$ and R_2 because s_2 needs to be decoded twice in user 1 and user 2 in SIC. Thus, the achievable data rate for decoding s_2 is denoted by $\min\{R_{1,2}, R_2\}$. The formulated problem P21 is non-convex and the non-convexity exists in both the objective function and constraints, which is difficult to be solved directly.

6.3 The Proposed Algorithms

In this section, the non-convex problem is mathematically transformed to a simple version. A SDR-based algorithm from convex optimisation and a learning based algorithm from data science are proposed to solve this problem.

The first step is to deal with constraint (6.8b) which currently is a form with expectation

$$
\mathrm{E}_{|\hat{h}_{\mathrm{BD}}|^2}\left[\log_2\left(1+\frac{\eta|g_1|^2|\hat{h}_{\mathrm{BD}}|^2}{\sigma_1^2}\right)\right] \geq R_{tc}. \quad (6.9)
$$

By assuming that $s_1(n)$ and $s_2(n)$ are complex Gaussian distributed with zero mean and unit variance, $\hat{h}_{\mathrm{BD}}$ follows complex Gaussian distribution $\mathcal{CN}(0,\lambda)$, where $\lambda = |\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_1|^2 + |\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_2|^2$.

Lemma 2. *The squared envelope of $\hat{h}_{\mathrm{BD}}$ follows the exponential distribution and the probability density function (PDF) is $f(x) = \frac{1}{\lambda}e^{-\lambda x}$.*

Proof. First, we have:

$$
\hat{h}_{\mathrm{BD}} = \mathbf{h}_{\mathrm{BD}}^H\mathbf{x} = \mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_1 s_1 + \mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_2 s_2. \quad (6.10)
$$

Let us assume $\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_1 = a_1 + jb_1$, $\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_2 = a_2 + jb_2$, $s_1 = x_1 + jy_1$ and $s_2 = x_2 + jy_2$.

Therefore, $\hat{h_{\mathrm{BD}}}$ can be expressed as follows:

$$\hat{h}_{\mathrm{BD}} = (a_1x_1 - b_1y_1 + a_2x_2 - b_2y_2) + j(a_1y_1 + b_1x_1 + a_2y_2 + b_2x_2). \tag{6.11}$$

As the assumption that s_1 and s_2 follow the complex Gaussian distribution with zero mean and unit variance, i.e. $\mathcal{CN}(0,1)$, and hence, the real random variables x_1, y_1, x_2 and y_2 are distributed with $\mathcal{N}(0, \frac{1}{2})$. $\Re(\hat{h}_{\mathrm{BD}})$ and $\Im(\hat{h}_{\mathrm{BD}})$ are the linear combination of four Gaussian distributed random variables. Therefore, $\Re(\hat{h}_{\mathrm{BD}})$ and $\Im(\hat{h}_{\mathrm{BD}})$ are also Gaussian distribution random variables and have the same mean and variance. The mean and the variance of $\Re(\hat{h}_{\mathrm{BD}})$ and $\Im(\hat{h}_{\mathrm{BD}})$ are 0 and $\frac{1}{2}(a_1^2+b_1^2+a_2^2+b_2^2)$. Therefore, $\hat{h}_{\mathrm{BD}}$ follows a complex Gaussian distribution $\mathcal{CN}(0,\lambda)$, where $\lambda = |\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_1|^2 + |\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_2|^2 = a_1^2 + b_1^2 + a_2^2 + b_2^2$. Then, we can have the conclusion that the square envelope of $\hat{h}_{\mathrm{BD}}$ follows an exponential distribution with the parameter $\lambda = |\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_1|^2 + |\mathbf{h}_{\mathrm{BD}}^H\mathbf{w}_2|^2$. The lemma is proofed. $\square$

Thus, the average rate of $c(n)$ can be rewritten as follows:

$$\bar{R}_{1,c} = \int_0^\infty \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \log_2(1+\beta x)\, dx, \tag{6.12}$$

where $\beta = \frac{\eta|g_1|^2}{\sigma_1^2}$. With some algebraic manipulation, $\bar{R}_{1,c}$ can be expressed as follows:

$$\bar{R}_{1,c} = -\log_2(e) e^{\frac{1}{\beta\lambda}} E_i(-\frac{1}{\beta\lambda}), \tag{6.13}$$

where $E_i(x) \triangleq \int_{-\infty}^{x} (1/u)e^u\, du, x \leq 0$ denotes the exponential integral function. Although the expectation is omitted, $E_i(\cdot)$ introduces new challenging to the problem. To tackle the obstacle, a new function $f(x) \triangleq -e^{\frac{1}{x}} E_i(-\frac{1}{x})$ is constructed. Referring to the proof in [71], $f(x)$ is a monotonically increasing function of x. It is assumed that there exists a $\hat{y}$ satisfying that $f(\hat{y}) = \frac{R_{tc}}{\log_2(e)}$, then the constraint (6.8b) can be expressed as follows:

$$f(\beta\lambda) \geq f(\hat{y}). \tag{6.14}$$

$f(x)$ is a monotonically increasing function of x, and thus the constraint (6.8b) can be reformulated into a compact way, i.e.,

$$\beta\lambda \geq \hat{y}. \tag{6.15}$$

As an efficient algorithm to solve multi-variable optimisation problems, the alternating algorithm is adopted in this paper. Specifically, η is optimized with given beamforming vectors. Consider a function $g_i(x) = \log_2(1 + \frac{A_i}{B_i x + C_i})$, it can be easily proved by the first-order derivative that $g_i(x)$ is monotonically decreasing with x if A_i, B_i and C_i are non-negative. Therefore, the objective function, as a sum of two decreasing functions, is also monotonically decreasing. Thus, (6.8a) can be equivalently transformed to find the lower bound of η. The range of η can be found in constrains (6.8d) and (6.15), as shown in follows:

$$\frac{\hat{y}\sigma_1^2}{|g_1|^2 \sum_{k=1}^{2} |\mathbf{h}_{\mathrm{BD}}^H \mathbf{w}_k|^2} \leq \eta \leq 1. \tag{6.16}$$

According to the monotonicity of the objective function, the optimal solution of η is its lower bound

$$\eta^* = \frac{\hat{y}\sigma_1^2}{|g_1|^2 \sum_{k=1}^{2} |\mathbf{h}_{\mathrm{BD}}^H \mathbf{w}_k|^2}. \tag{6.17}$$

By substituting η in P21 with η^*, P21 is transformed as follows:

$$\begin{aligned}
\text{P22}: \max_{\{\mathbf{w}_1,\mathbf{w}_2\}} \; & \log_2\left(1 + \frac{|\mathbf{h}_1^H \mathbf{w}_1|^2}{\hat{y}\sigma_1^2 + \sigma_1^2}\right) \\
& + \min\{\log_2(1 + A_1), \log_2(1 + A_2)\} && (6.18\text{a}) \\
\text{s.t.} \quad & \frac{|\mathbf{h}_1^H \mathbf{w}_2|^2}{|\mathbf{h}_1^H \mathbf{w}_1|^2 + \hat{y}\sigma_1^2 + \sigma_1^2} \geq \epsilon_2, && (6.18\text{b}) \\
& \frac{\hat{y}\sigma_1^2}{|g_1|^2 \sum_{k=1}^{2} |\mathbf{h}_{\mathrm{BD}}^H \mathbf{w}_k|^2} \leq 1, && (6.18\text{c}) \\
& (6.8\text{c}),
\end{aligned}$$

where $A_1 = \frac{|\mathbf{h}_2^H \mathbf{w}_2|^2}{|\mathbf{h}_2^H \mathbf{w}_1|^2 + \frac{|g_2|^2}{|g_1|^2}\hat{y}\sigma_1^2 + \sigma_2^2}$, $A_2 = \frac{|\mathbf{h}_1^H \mathbf{w}_2|^2}{|\mathbf{h}_1^H \mathbf{w}_1|^2 + \hat{y}\sigma_1^2 + \sigma_1^2}$ and $\epsilon_2 = 2^{R_{tc}} - 1$.

Constraint (6.18b) is obtained from (6.8b) and constraint (6.18c) is obtained from (6.16), which guarantees that the BD can receive enough power to transmit its signal. It is worth mentioning that constraint (6.8b) is always satisfied by η^*

Note that problem P22 is difficult to be solved because there exists non-convexity in the objective function and the constraints. Two algorithms are provided in the following to solve P22 by conventional convex optimisation and deep learning, respectively.

6.3.1 SDR-based Algorithm

In this case, auxiliary variables can be utilized to transform P22 into a convex problem. By introducing the auxiliary variable t, the original problem is expressed as follows:

$$\begin{aligned}
\text{P23}: \quad & \max_{\{\mathbf{w}_1,\mathbf{w}_2,t\}} \log_2\left(1+\frac{|\mathbf{h}_1^H\mathbf{w}_1|^2}{\hat{y}\sigma_1^2+\sigma_1^2}\right)+\log_2(1+t) && \text{(6.19a)}\\
& \text{s.t.} \quad A_1 \geq t, && \text{(6.19b)}\\
& \quad\quad\;\; A_2 \geq t, && \text{(6.19c)}\\
& \quad\quad\;\; \text{(6.8d)}, \text{(6.18b)}, \text{(6.18c)}.
\end{aligned}$$

The objective function is still non-convex due to R_1. Another auxiliary variable α is introduced for further transformation. By substituting α to the first logarithmic function, P23 is equivalently transformed to

$$\begin{aligned}
\text{P24}: \quad & \max_{\{\mathbf{w}_1,\mathbf{w}_2,\alpha,t\}} \log_2(1+\alpha)+\log_2(1+t) && \text{(6.20a)}\\
& \text{s.t.} \quad \frac{|\mathbf{h}_1^H\mathbf{w}_1|^2}{\hat{y}\sigma_1^2+\sigma_1^2} \geq \alpha, && \text{(6.20b)}\\
& \quad\quad\;\; \text{(6.8d)}, \text{(6.18b)}, \text{(6.18c)}, \text{(6.19b)}, \text{(6.19c)}.
\end{aligned}$$

However, P24 is still a non-convex problem because of the non-convex constraints (6.18b), (6.18c), (6.19b), (6.19c) and (6.20b). A square norm can be equivalently transformed into a trace form as shown below

$$|\mathbf{h}_i^H\mathbf{w}_j|^2 = \text{Tr}(\mathbf{H}_i\mathbf{W}_j), i,j \in \{1,2,\text{BD}\}, \tag{6.21}$$

where $\mathbf{H}_i = \mathbf{h}_i\mathbf{h}_i^H$ and $\mathbf{W}_j = \mathbf{w}_j\mathbf{w}_j^H$. It is indicated that beamforming matrix $\mathbf{W}_j$ is positive semi-definite and the rank is 1. By applying this transformation, P24 can be equivalently transformed into

$$\begin{aligned}
\text{P25}: \quad & \max_{\{\mathbf{W}_1,\mathbf{W}_2,\alpha,t\}} \log_2(1+\alpha)+\log_2(1+t) && \text{(6.22a)}\\
& \text{s.t.} \quad \text{Tr}(\mathbf{H}_1\mathbf{W}_1) \geq \alpha c_1, && \text{(6.22b)}
\end{aligned}$$

$$\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_2) \geq t\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_1) + c_2 t, \tag{6.22c}$$

$$\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_2) \geq t\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1) + c_1 t, \tag{6.22d}$$

$$\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_2) - \epsilon_2\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1) \geq c_1\epsilon_2, \tag{6.22e}$$

$$\mathrm{Tr}(\mathbf{H}_{\mathrm{BD}}\mathbf{W}_1) + \mathrm{Tr}(\mathbf{H}_{\mathrm{BD}}\mathbf{W}_2) \geq \frac{\hat{y}\sigma_1^2}{|g_1|^2}, \tag{6.22f}$$

$$\mathrm{Tr}(\mathbf{W}_1) + \mathrm{Tr}(\mathbf{W}_2) \leq P_0, \tag{6.22g}$$

$$\mathbf{W}_i \succeq 0, i \in \{1,2\}, \tag{6.22h}$$

$$\mathrm{rank}(\mathbf{W}_i) = 1, i \in \{1,2\}, \tag{6.22i}$$

where $c_1 = \hat{y}\sigma_1^2 + \sigma_1^2$ and $c_2 = \frac{|g_2|^2}{|g_1|^2}\hat{y}\sigma_1^2 + \sigma_2^2$. All constraints except (6.22c), (6.22d) and (6.22i) construct a convex set. To deal with the non-convexity of constraints (6.22c) and (6.22d), a basic inequality is considered below

$$\frac{a^2+b^2}{2} \geq ab. \tag{6.23}$$

By adopting (6.23), the upper bound of $t\mathrm{Tr}(\mathbf{H}_i\mathbf{W}_j)$ is given by $\frac{t^2+\mathrm{Tr}(\mathbf{H}_i\mathbf{W}_j)^2}{2}$. By introducing this relaxation, the constraints (6.22c) and (6.22d) are respectively transformed into

$$\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_2) \geq \frac{t^2+\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_1)^2}{2} + c_2 t \tag{6.24}$$

and

$$\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_2) \geq \frac{t^2+\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)^2}{2} + c_1 t. \tag{6.25}$$

It is noted that when inequalities (6.24) and (6.25) are satisfied, constraints (6.22c) and (6.22d) must be satisfied. The last obstacle is the non-convex rank-one constraint (6.22i). A common method, namely, SDR, is adopted, where the rank one constraint is temporally ignored and the problem becomes a SDP problem. By dropping the rank-one constraint, P25 can be rewritten as follows:

$$\mathrm{P26}: \max_{\{\mathbf{W}_1,\mathbf{W}_2,\alpha,t\}} \log_2(1+\alpha) + \log_2(1+t) \tag{6.26a}$$

$$\text{s.t. } \mathrm{Tr}(\mathbf{H}_2\mathbf{W}_2) \geq \frac{t^2+\mathrm{Tr}(\mathbf{H}_2\mathbf{W}_1)^2}{2} + c_2 t \tag{6.26b}$$

$$\mathrm{Tr}(\mathbf{H}_1\mathbf{W}_2) \geq \frac{t^2 + \mathrm{Tr}(\mathbf{H}_1\mathbf{W}_1)^2}{2} + c_1 t \tag{6.26c}$$

$$(6.22\mathrm{b}), (6.22\mathrm{e}) - (6.22\mathrm{h}).$$

P26 is a convex problem , which can be solved by convex optimisation tool boxes such as CVX. If the rank of the obtained solution $\mathbf{W}_1^*$ and $\mathbf{W}_2^*$ is not 1, Gaussian randomisation stated in [12] is required to regenerate feasible suboptimal solution of $\mathbf{w}_1^*$ and $\mathbf{w}_2^*$ from $\mathbf{W}_1^*$ and $\mathbf{W}_2^*$.

6.3.2 Learning based Algorithm

For the learning based algorithm, a five-layer fully connected neural network is adopted to directly solve P22. It is worth to point out that the aim of this algorithm is to use an existing network for solving a constrained optimisation problem and to compare with the conventional SDR-based algorithm.

Dataset Generation

The training dataset is constructed with different channel sets. Denote

$$s^{(i)} = \{\mathbf{h}_1^{(i)}, \mathbf{h}_2^{(i)}, \mathbf{h}_{\text{bd}}^{(i)}, g_1^{(i)}, g_2^{(i)}\} \tag{6.27}$$

as a channel sample, which describes the channel condition at moment i. The training dataset should collect as much channel samples as possible. These collected channel samples are per-stored in a memory buffer and then grouped into several mini-batches for training. In this chapter, the size of the training dataset is 50000.

Neural Network Design

A five-layer full connected neural network, consisting of one input layer, two hidden layers, one output layer and an additional normalisation layer, is constructed. According to the channel sample s^i, the dimension of the input layer is $6M+2$. Two hidden layers have 256 neurons and 512 neurons, respectively. The output layer's dimension is decided by the dimension of two beams, i.e., $4M$. The Relu function is inserted between every two adjacent layers to enhance the non-linearity of the network and efficiently avoid overfitting. The normalisation layer normalizes the result based on

$$\mathbf{w}_j^{*(i)} = \sqrt{\frac{P_0}{||\mathbf{w}_1^{(i)}||^2 + ||\mathbf{w}_2^{(i)}||^2}}\mathbf{w}_j^{(i)}, j \in \{1, 2\}, \tag{6.28}$$

where $\mathbf{w}_1^{(i)}$ and $\mathbf{w}_2^{(i)}$ denote the output of the output layer and $\mathbf{w}_1^{*(i)}$ and $\mathbf{w}_2^{*(i)}$ denote the output of the normalisation layer corresponding to $s^{(i)}$. Hence, the output after normalisation always satisfies constraint (6.8c).

Loss Function Design

A wise design of the loss function can locate the output of the neural network within the feasible set of P22. In other words, the output of the network cannot violate the constraints of P22. Constraint (6.8c) has been handled by the normalisation layer. We introduce two activation functions

$$\mathcal{C}_1\left(\mathbf{w}_1^{*(i)}, \mathbf{w}_2^{*(i)}\right) = \mathbb{I}\left(\frac{|\mathbf{h}_1^{(i)H}\mathbf{w}_2^{*(i)}|^2}{|\mathbf{h}_1^{(i)H}\mathbf{w}_1^{*(i)}|^2 + \hat{y}\sigma_1^2 + \sigma_1^2} - \epsilon_2\right) \tag{6.29}$$

and

$$\mathcal{C}_2\left(\mathbf{w}_1^{*(i)}, \mathbf{w}_2^{*(i)}\right) = \mathbb{I}\left(1 - \frac{\hat{y}\sigma_1^2}{|g_1|^{(i)2}\sum_{k=1}^{2}|\mathbf{h}_{\mathrm{BD}}^{(i)H}\mathbf{w}_k^{*(i)}|^2}\right), \tag{6.30}$$

where an indicator function is denoted by

$$\mathbb{I}(x) = \begin{cases} 0, & \text{if } x \geq 0, \\ |x|, & \text{if } x < 0, \end{cases} \tag{6.31}$$

to handle constraint (6.18b) and constraint (6.18c), respectively. The third activation function denoted by $\mathcal{F}$ corresponds to the objective function, which is expressed as follows:

$$\begin{aligned}\mathcal{F}\left(\mathbf{w}_1^{*(i)}, \mathbf{w}_2^{*(i)}\right) = \log_2\left(1 + \frac{|\mathbf{h}_1^{(i)H}\mathbf{w}_1^{*(i)}|^2}{\hat{y}\sigma_1^2 + \sigma_1^2}\right) \\ + \log_2\left(1 + \min(A_1^{*(i)}, A_2^{*(i)})\right),\end{aligned} \tag{6.32}$$

where $A_1^{*(i)}$ denotes $\frac{|\mathbf{h}_2^{(i)H}\mathbf{w}_2^{*(i)}|^2}{|\mathbf{h}_2^{(i)H}\mathbf{w}_1^{*(i)}|^2+\frac{|g_2|^{(i)2}}{|g_1|^{(i)2}}\hat{y}\sigma_1^2+\sigma_2^2}$ and $A_2^{*(i)}$ denotes $\frac{|\mathbf{h}_1^{(i)H}\mathbf{w}_2^{*(i)}|^2}{|\mathbf{h}_1^{(i)H}\mathbf{w}_1^{*(i)}|^2+\hat{y}\sigma_1^2+\sigma_2^2}$. The customized loss function is given by

$$\mathcal{L}\left(\mathbf{w}_1^{*(i)}, \mathbf{w}_2^{*(i)}\right) = -\mathcal{F} + \gamma_1 \mathcal{C}_1 + \gamma_2 \mathcal{C}_2, \tag{6.33}$$

where γ_1 and γ_2 are punishment coefficients. Note that when the output from the network satisfies constraints (6.18b) and (6.18c), $\mathcal{C}_1 = \mathcal{C}_2 = 0$ and the minimal value of $\mathcal{C}_1$ and $\mathcal{C}_2$ is 0 because of (6.31). Because the loss function is minimized, the network will try to make $\mathcal{C}_1$ and $\mathcal{C}_2$ equal to 0. In this way, the loss function can lead the network to output a feasible solution. A mini-batch usually serves as the training data in each epoch, thus the network is trying to minimize the mean of the loss functions over the mini-batch. The average loss function is given by

$$\tilde{\mathcal{L}} = \frac{1}{|\mathcal{P}|} \sum_{i \in \mathcal{P}} \mathcal{L}\left(\mathbf{w}_1^{*(i)}, \mathbf{w}_2^{*(i)}\right), \tag{6.34}$$

where $\mathcal{P}$ denotes the mini-batch and $|\mathcal{P}|$ denotes the batch size. The network will avoid infeasible outputs violating any constraint by minimizing (6.34).

Note that it is challenging to define an activation function for each constraints when an optimisation problem contains a lot of constraints. We can simplify the original optimisation problem with the help from the traditional optimisation method. In this paper, the closed-form expression of η is derived and the original optimisation problem is simplified, which is possible to be solved by deep learning algorithms.

6.3.3 Complexity Analysis

Referring to [12], the worst-case complexity of solving a general SDR problem by CVX can be expressed as follows:

$$\mathcal{O}\left(n^{4.5}\log(\frac{1}{\epsilon})\right), \tag{6.35}$$

where n denotes the problem size and ϵ denotes the solution accuracy. In particular, n is related to the number of antennas and users in wireless communication field.

For the learning based algorithm, we only consider the forward propagation since the model is pre-trained. The forward propagation consists of matrix multiplication and activation functions. By assuming the i-th layer has i neurons and the j-th layer has j neurons, from the i-th layer to the j-th layer, the calculation of this operation is

$$\mathbf{S}_j^{j\times t} = \mathbf{W}^{j\times i} * \mathbf{Z}_i^{i\times t}, \tag{6.36}$$

where t denotes the number of training samples and $\mathbf{Z}_i^{i\times t}$, $\mathbf{W}^{j\times i}$ and $\mathbf{S}_j^{j\times t}$denote the input matrix, the weight matrix and the output matrix of the i-th layer, respectively. The operation (6.36) has $\mathcal{O}(jit)$. Then, the activation function

$$\mathbf{Z}_j^{j\times t} = f(\mathbf{S}_j^{j\times t}) \tag{6.37}$$

is applied, where $\mathbf{Z}_j^{j\times t}$ denotes the input matrix of the j-th layer and the time complexity of this operation is $\mathcal{O}(jt)$. In total, the time complexity of forward propagation for a 4-layer network is

$$\mathcal{O}(t(ij + jk + kl)), \tag{6.38}$$

where i, j, k and l are the numbers of neurons for the input layer, two hidden layers and the output layer, respectively.

6.4 Simulation Results

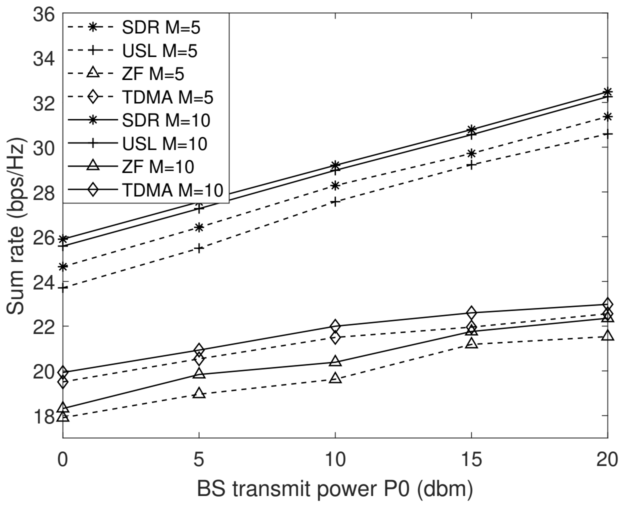

Figure 6.2. Sum rate versus transmit power.

In this section, simulation results of two proposed algorithms and benchmarks are provided. Rician fading model and Rayleigh fading model are utilized. The path loss exponent

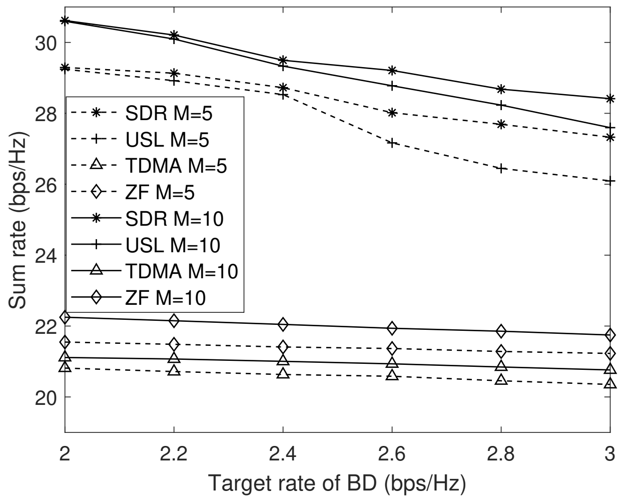

Figure 6.3. Sum rate versus the target rate of BD.

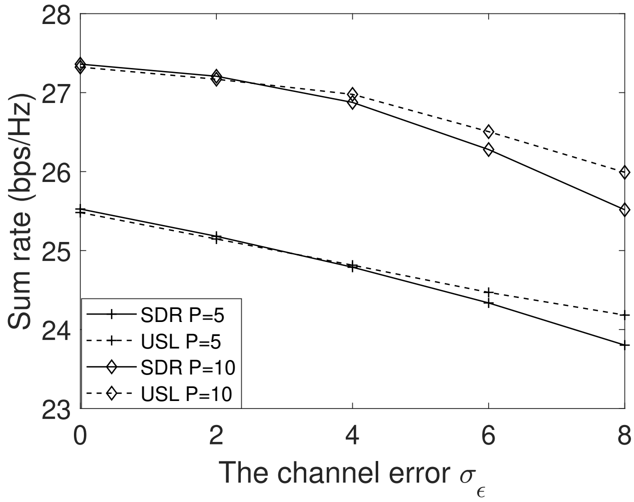

Figure 6.4. Sum rate versus channel error.

is set as 2.5. The distance between the BS and users is set as 2 m, and the noise power at two users is set to -96 dBm. Time-division multiple access (TDMA) and zero-forcing (ZF) are included as the benchmarks.

Fig. 6.2 and Fig. 6.3 show how the transmit power at the BS and the target rate of BD impact the sum rate of two NOMA users. We set the target rate of BD as 2 bps/Hz in Fig. 6.2 and the transmit power at the BS as 10 dbm in Fig. 6.3, respectively. The sum rate increases with the increasing of the transmit power at the BS, however, it decreases when the target rate of BD increases. It is noted that the performance gap between two algorithms is smaller when the BS is equipped with more antennas and the BD requires a smaller data rate. Two simulation sets with different numbers of antennas, $M = 5$ and $M = 10$, are conducted in both figures. The number of antenna can help improve the performance.

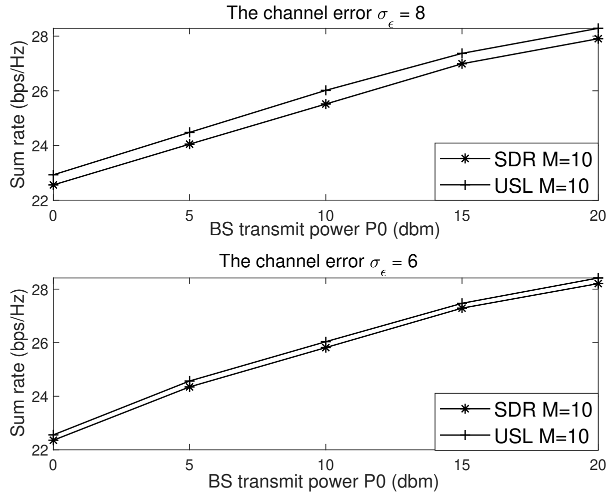

Figure 6.5. The performance under different channel error.

According to experimental experiences, (6.38) is usually great smaller than (6.35). When a well-trained network is operated in the evaluation mode, t in (6.38) equals to 1 because the real-CIS is just one sample. Therefore, the network can instantly predict beam vectors, and hence beam vectors can be frequently updated in a time slot. However, the SDR-based algorithm is time-consuming so it is impossible to optimize beam vectors every time even the real-time CSI is available. As a result, beam vectors calculated by the SDR-based algorithm are fixed in a time slot. Fig. 6.4 shows the performance of two proposed algorithms when the CSI is changing. σ_ϵ is a channel error unit to describe how fast the channel changes. The target rate of BD is set as 2 bps/Hz and the transmit power at the BS is set as 10 dBm. This figure shows the learning based algorithm is more robust than the SDR-based algorithm under a large channel error. Fig. 6.5 shows the learning based algorithm can achieve a better performance than the SDR-based algorithm if a channel error exists and the performance gap increases when the channel error becomes larger. The reason is that beam vectors calculated by the SDR-based algorithm is outdated but the learning based algorithm can timely update beam vectors when the channel changes fast.

6.5 Conclusion

In this chapter, an add-on BAC-NOMA network was investigated to comprehensively improve the spectral efficiency. The sum rate of two NOMA users was maximised while guaranteeing the QoS of the BD. Both algorithms based on convex optimisation and deep learning can achieve competitive performance compared to the benchmarks. The SDR-based algorithm is able to outperform the learning based algorithm when the channel is unchanged or changes slowly. When the channel changes fast, the learning based algorithm, as a low-complexity algorithm, is more robust than the SDR-based algorithm. Based on the proposed algorithms, a dynamic strategy selection mechanism can be constructed. Since only a fully connected neural network is adopted, the performance of the learning based algorithm can by further improved by particularly designer the network structure for this optimisation problem. How to design a neural network is a promising research direction for the future work.

Chapter 7

Conclusions and Future Works

7.1 Conclusions

This thesis investigated the optimisation problems on how to allocate resource in different NOMA scenarios, including RIS-NOMA and BAC-NOMA. Convex optimisation, alternating algorithm, SDR, SCA and machine learning are applied to solve aforementioned optimisation problems to improve the efficiency and the sum rate of the system. The simulation results demonstrated that the optimised network with the optimal resource allocation scheme can achieve the best performance. The specific conclusions of each chapter are summarised as follows.

In chapter 3, the joint optimisation of beamforiming, power allocation and RIS phase shift in a NOMA-RIS assisted multi-cluster network was investigated. A transmit power minimisation problem was formulated. The primal problem was split into two subproblems and these two subproblems were transformed to convex problems by inequality approximation, SDR and SCA. An alternating algorithm was proposed to minimise the transmit power by iteratively solving beamforming optimisation and phase shifting feasibility until the algorithm converges. Additionally, a low-complexity solution is provided for this scenario based on a partial exhaustive search approach. The simulation results demonstrated that the alternating algorithm outperforms the partial exhaustive search algorithm but exhibits higher complexity.

In chapter 4, a sum rate maximization problem in a RIS-assisted NOMA downlink network was investigated. A DDPG based algorithm was proposed to jointly optimise beamforming and phase shift. The proposed DDPG algorithm not only achieves competitive performance but also adapts to varying channel scenarios. However, conventional convex optimisation is primarily suitable for fixed channel scenarios. More specifically, machine learning provides a new solution for wireless communication networks and also can be applied for more complicated scenarios, which will be a powerful tool for developing 6G network.

In chapter 5, a backscatter-assisted NOMA network, consisting of two downlink NOMA users and one uplink device, was investigated. This work achieved the maximum data rate of the uplink device and meanwhile QoS requirements of two NOMA users are met. The proposed alternating algorithm efficiently addressed this problem and can be extended to handle scenarios involving multiple downlink NOMA users. Moreover, a closed-form solution for a two-user case was derived. This study holds practical significance as it enables the addition of an extra IoT device to a legacy NOMA network, resulting in a substantial capacity enhancement for the legacy system.

In chapter 6, an add-on BAC-NOMA network was investigated to comprehensively improve the spectral efficiency. The sum rate of two NOMA users was maximised and meanwhile the QoS of the backscatter device was guaranteed. Two algorithms based on convex optimisation and unsupervised learning were proposed to solve the primal problem, respectively. The simulation results show that the SDR-based algorithm outperforms the learning-based algorithm when the channel remains unchanged or changes slowly. However, in scenarios with fast-changing channels, the learning-based algorithm, as a low-complexity approach, exhibits greater robustness compared to the SDR-based algorithm. This study provided a dynamic strategy selection mechanism, which can select different algorithms based on the environment.

In summary, NOMA enhances spectrum efficiency in wireless networks, RIS offers advanced signal manipulation for improved wireless communication, and backscattering is a low-power communication technique often used in IoT applications. These technologies represent important advancements in the wireless communication field, each with its unique benefits and applications. With the combination of NOMA with RIS and backscattering becoming more and more mature, it has become a protential solution for the next generation communication system.

7.2 Future Works

7.2.1 STAR-RIS and Active RIS

All the works related to RIS in this thesis is about the conventional passive RIS, which can only reflect the incident signal to the destination. Recently, simultaneously transmitting and reflecting (STAR) RIS and active RIS were proposed to introduce more benefits. The STAR-RIS can simultaneously transmit and reflect the incident signals. In particular, two phases named the transmission phase and the reflection phase are introduced [95]. The STAR-RIS allows users with different requirements to be served simultaneously. The STAR-RIS assisted NOMA network has been investigated in some works [96]–[98]. However, the topic on multi-user scenario in a STAR-RIS network is still a valuable research direction in the future.

Active RISs introduce a new feature compared with the conventional RIS. The active RIS can amplify the incident signal while reflecting the signal [99], [100]. Both the amplitude coefficient and the phase shift of each reflecting unit can be adjusted. Therefore, more algorithms can be properly designed to find the optimal amplitude coefficients and phase shifts.

7.2.2 Optimisations in THz-NOMA Networks

Due to the severe congestion observed in the sub-6 GHz bands, utilizing the Terahertz (THz) band has emerged as a promising solution to meet the demands of emerging applications that necessitate super-fast broadband speeds and ultra-low latencies. THz-NOMA can further improve the spectrum efficiency in a THz transmission environment [101]–[103]. Some advantages of THz-NOMA are summarised as follows.

- High data rate: Since the super high frequency of the THz channel, the data rate for each device can be extremely high. Numerous bandwidth-intensive applications, such as live streaming and 8K videos, can greatly benefit from the implementation of THz-NOMA networks.
- Massive connectivity: The THz-NOMA network can accommodate massive devices because of its extremely large capabilities. This caters for the requirement of IoT and Internet of Everything (IoE) networks.

However, there are still many challenges in THz-NOMA networks. For example, the channel model in a THz network is completely different to conventional channel model. A new mathematical model is necessary to describe the channel in THz-NOMA networks. Moreover, high frequency requires complex hardware.

7.2.3 Federated Learning

Federated learning is a distributed machine learning approach that enables multiple distributed devices collaboratively train a shared model [104], [105]. In federated learning, each device trains its model locally and transmits the model parameters to a central point. The central point then collects the parameters from all devices and updates the shared model accordingly. Federated learning offer a lot of benefits.

- Data Privacy: Privacy is the most important advantage of federated learning since data is only trained locally and there is no data sharing between different devices.
- Low Communication Costs: Only the updated model parameters need to be transmitted between the central point and distributed devices, which is suitable for low-bandwidth or unstable network environments.
- Decentralisation: The decentralised stature enables data can be trained locally, which protects the privacy of the user who cannot or do not want to share their data.

The research of federated learning from the wireless communication perspective focuses on how the quality of wireless communication impacts the model's performance, given that wireless channels are the link between the central point and all the devices [106], [107]. There are several aspects the wireless communication should consider.

- Bandwidth efficiency: The wireless communication channel has a bandwidth limit, which means it may not have the capability to transmit parameters of a large size model. Therefore, some wireless communication techniques are necessary to be utilised to improve the bandwidth efficiency.
- Low latency: Updating models timely is crucial for efficient federated learning. Low-latency wireless communication is important for the real-time coordination between

distributed devices and the central point. Therefore, how to reduce the latency is an important topic in federated learning.

- Reliability: Bit errors will occur when signals are being transmitted by wireless communication channels. Therefore, error correction mechanisms and robust data transmission techniques need to be considered to guarantee a reliable transmission.

7.2.4 Age of Information

Age of information (AoI) is a concept to describe the timeliness or freshness of information at the destination [108]. Some works have considered to use NOMA to decrease AoI [109], [110]. The most interesting question is how to use NOMA to reduce AoI. In this topic, we can investigate the resource allocation and beamforming design. It is expected to provide valuable insights into the advantages and challenges of using NOMA techniques to optimise information freshness in future communication networks. Introducing NOMA into AoI has the potential to significantly enhance the performance and efficiency of real-time applications in various scenarios, which can be a perfect research direction in the future.

References

[1] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, 2021 (cited on pp. 16, 17).

[2] C.-X. Wang, X. You, X. Gao, *et al.*, "On the road to 6G: Visions, requirements, key technologies and testbeds," *arXiv preprint arXiv:2302.14536*, 2023 (cited on pp. 16, 17).

[3] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz, and J. M. Lopez-Soler, "A survey on 5G usage scenarios and traffic models," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 905–929, 2020 (cited on p. 17).

[4] B. Artmann, *Euclid—the creation of mathematics*. Springer Science & Business Media, 2012 (cited on p. 17).

[5] S. Kiranyaz, T. Ince, and M. Gabbouj, *Multidimensional particle swarm optimization for machine learning and pattern recognition*. Springer, 2014 (cited on p. 18).

[6] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple access techniques for 5G wireless networks and beyond*. Springer, 2019 (cited on pp. 19, 53).

[7] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *2013 IEEE 77th vehicular technology conference (VTC Spring)*, IEEE, 2013, pp. 1–5 (cited on p. 27).

[8] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun Mag*, 2019 (cited on pp. 30, 54).

[9] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019 (cited on p. 30).

[10] C. Boyer and S. Roy, “—invited paper—backscatter communication and RFID: Coding, energy, and MIMO analysis,” *IEEE Transactions on Communications*, vol. 62, no. 3, pp. 770–785, 2013 (cited on p. 33).

[11] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004 (cited on pp. 35, 67).

[12] Z. Luo, W. Ma, A. M. So, Y. Ye, and S. Zhang, “Semidefinite relaxation of quadratic optimization problems,” *IEEE Signal Process Mag*, vol. 27, no. 3, pp. 20–34, 2010 (cited on pp. 42, 43, 68, 71, 72, 136, 138).

[13] M. Razaviyayn, “Successive convex approximation: Analysis and applications,” Ph.D. dissertation, University of Minnesota, 2014 (cited on p. 43).

[14] Z.-H. Zhou, *Machine learning*. Springer Nature, 2021 (cited on p. 43).

[15] J. Heaton, “Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The MIT press, 2016, 800 pp, isbn: 0262035618,” *Genetic programming and evolvable machines*, vol. 19, no. 1-2, pp. 305–307, 2018 (cited on pp. 43, 45).

[16] S. Russell and P. Norvig, “Artificial intelligence: A modern approach,” 2002 (cited on p. 47).

[17] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning.[sl]*, 2012 (cited on p. 47).

[18] G. E. Hinton, T. J. Sejnowski, T. A. Poggio, *et al.*, *Unsupervised learning: foundations of neural computation*. MIT press, 1999 (cited on p. 47).

[19] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987 (cited on p. 47).

[20] W. H. L. Pinaya, S. Vieira, R. Garcia-Dias, and A. Mechelli, “Autoencoders,” in *Machine learning*, Elsevier, 2020, pp. 193–208 (cited on p. 47).

[21] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, “Generative adversarial networks: An overview,” *IEEE signal processing magazine*, vol. 35, no. 1, pp. 53–65, 2018 (cited on p. 47).

[22] J. A. Hartigan and M. A. Wong, “Algorithm AS 136: A k-means clustering algorithm,” *Journal of the royal statistical society. series c (applied statistics)*, vol. 28, no. 1, pp. 100–108, 1979 (cited on p. 48).

[23] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996 (cited on p. 48).

[24] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, “An introduction to deep reinforcement learning,” *arXiv preprint arXiv:1811.12560*, 2018 (cited on pp. 48, 50, 88).

[25] K. Feng, Q. Wang, X. Li, and C.-K. Wen, “Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems,” *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 745–749, 2020 (cited on p. 50).

[26] Z. Zhang, Y. Xiao, Z. Ma, *et al.*, “6G wireless networks: Vision, requirements, architecture, and key technologies,” *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019 (cited on p. 53).

[27] 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, 2019 (cited on p. 53).

[28] F. Tariq, M. Khandaker, K.-K. Wong, M. Imran, M. Bennis, and M. Debbah, “A speculative study on 6G,” *arXiv preprint arXiv:1902.06700*, 2019 (cited on p. 53).

[29] V. Jamali, A. Tulino, G. Fischer, R. Müller, and R. Schober, "Intelligent reflecting and transmitting surface aided millimeter wave massive MIMO," *arXiv e-prints*, 2019 (cited on p. 53).

[30] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun Mag*, vol. 52, no. 2, pp. 186–195, 2014 (cited on p. 53).

[31] Y. Guo, *Method and apparatus of small cell enhancement in a wireless communication system*, US Patent 9,295,077, Mar. 2016 (cited on p. 53).

[32] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *EEE Trans. Wirel. Commun.*, vol. 15, no. 1, pp. 537–552, 2015 (cited on pp. 53, 82).

[33] K. Saito, A. Benjebbour, Y. Kishiyama, Y. Okumura, and T. Nakamura, "Performance and design of SIC receiver for downlink NOMA with open-loop SU-MIMO," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, IEEE, 2015, pp. 1161–1165 (cited on pp. 53, 82, 98).

[34] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *2018 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2018, pp. 1–6 (cited on p. 53).

[35] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless communications: A tutorial," *arXiv preprint arXiv:2007.02759*, 2020 (cited on pp. 53, 58).

[36] X. Sun, N. Yang, S. Yan, *et al.*, "Joint beamforming and power allocation in downlink NOMA multiuser MIMO networks," *EEE Trans. Wirel. Commun.*, vol. 17, no. 8, pp. 5367–5381, 2018 (cited on p. 54).

[37] F. Fang, H. Zhang, J. Cheng, S. Roy, and V. C. Leung, "Joint user scheduling and power allocation optimization for energy-efficient NOMA systems with imperfect

CSI," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2874–2885, 2017 (cited on p. 54).

[38] G. Surabhi, R. M. Augustine, and A. Chockalingam, "On the diversity of uncoded OTFS modulation in doubly-dispersive channels," *EEE Trans. Wirel. Commun.*, vol. 18, no. 6, pp. 3049–3063, 2019 (cited on p. 54).

[39] W. Tang, M. Z. Chen, J. Y. Dai, *et al.*, "Wireless communications with programmable metasurface: New paradigms, opportunities, and challenges on transceiver design," *arXiv preprint arXiv:1907.01956*, 2019 (cited on p. 54).

[40] F. Fang, H. Zhang, J. Cheng, and V. C. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access network," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3722–3732, 2016 (cited on p. 54).

[41] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J Wirel. Commun. Netw.*, vol. 2019, no. 1, pp. 1–20, 2019 (cited on p. 54).

[42] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019. DOI: 10.1109/TWC.2019.2936025 (cited on pp. 54, 82, 84).

[43] W. Tang, M. Z. Chen, X. Chen, *et al.*, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *arXiv preprint arXiv:1911.05326*, 2019 (cited on p. 54).

[44] J. Zhu, Y. Huang, J. Wang, K. Navaie, and Z. Ding, "Power efficient IRS-assisted NOMA," *arXiv preprint arXiv:1912.11768*, 2019 (cited on p. 54).

[45] Z. Ding, R. Schober, and H. V. Poor, "On the impact of phase shifting designs on IRS-NOMA," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1596–1600, 2020 (cited on p. 54).

[46] M. Fu, Y. Zhou, and Y. Shi, "Reconfigurable intelligent surface empowered downlink non-orthogonal multiple access," *arXiv preprint arXiv:1910.07361*, 2019 (cited on pp. 54, 59, 62).

[47] X. Liu, Y. Liu, Y. Chen, and H. V. Poor, "RIS enhanced massive non-orthogonal multiple access networks: Deployment and passive beamforming design," *arXiv preprint arXiv:2001.10363*, 2020 (cited on pp. 54, 60).

[48] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, "Exploiting intelligent reflecting surfaces in multi-antenna aided NOMA systems," *arXiv preprint arXiv:1910.13636*, 2019 (cited on p. 54).

[49] J. Zuo, Y. Liu, Z. Qin, and N. Al-Dhahir, "Resource allocation in intelligent reflecting surface assisted NOMA systems," *arXiv preprint arXiv:2002.01765*, 2020 (cited on p. 54).

[50] M. Zeng, X. Li, G. Li, W. Hao, and O. Dobre, "Sum rate maximization for IRS-assisted uplink NOMA," *arXiv preprint arXiv:2004.10791*, 2020 (cited on p. 54).

[51] Y. Li, M. Jiang, Q. Zhang, and J. Qin, "Joint beamforming design in multi-cluster MISO NOMA intelligent reflecting surface-aided downlink communication networks," *arXiv preprint arXiv:1909.06972*, 2019 (cited on pp. 55, 56, 61, 62).

[52] W. Ni, X. Liu, Y. Liu, H. Tian, and Y. Chen, "Resource allocation for multi-cell irs-aided NOMA networks," *arXiv preprint arXiv:2006.11811*, 2020 (cited on pp. 55, 56).

[53] Z. Ding and H. V. Poor, "A simple design of IRS-NOMA transmission," *IEEE Communications Letters*, vol. 24, no. 5, pp. 1119–1123, 2020 (cited on p. 58).

[54] S. Yan, X. Zhou, D. W. K. Ng, J. Yuan, and N. Al-Dhahir, "Intelligent reflecting surface for wireless communication security and privacy," *arXiv preprint arXiv:2103.16696*, 2021 (cited on p. 59).

[55] X. Zhou, S. Yan, Q. Wu, F. Shu, and D. W. K. Ng, “Intelligent reflecting surface (IRS)-aided covert wireless communication with delay constraint,” *arXiv preprint arXiv:2011.03726*, 2020 (cited on p. 60).

[56] X. You, C. Wang, J. Huang, *et al.*, “Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts,” *Science China Information Sciences*, vol. 64, no. 1, pp. 1–74, 2021 (cited on p. 82).

[57] C. Huang, R. Mo, and C. Yuen, “Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1839–1850, 2020 (cited on p. 82).

[58] J. Gao, C. Zhong, X. Chen, H. Lin, and Z. Zhang, “Unsupervised learning for passive beamforming,” *IEEE Communications Letters*, vol. 24, no. 5, pp. 1052–1056, 2020 (cited on p. 82).

[59] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, “Deep reinforcement learning: A brief survey,” *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017 (cited on p. 82).

[60] B. Lyu, Z. Yang, H. Guo, F. Tian, and G. Gui, “Relay cooperation enhanced backscatter communication for Internet-of-Things,” *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2860–2871, 2018 (cited on p. 98).

[61] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi, and M. Mustaqim, “Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios,” *Ieee Access*, vol. 8, pp. 23 022–23 040, 2020 (cited on p. 98).

[62] L. Chen, S. Thombre, K. Järvinen, *et al.*, “Robustness, security and privacy in location-based services for future IoT: A survey,” *Ieee Access*, vol. 5, pp. 8956–8977, 2017 (cited on p. 98).

[63] T. O. Timoudas, R. Du, and C. Fischione, “Enabling massive IoT in ambient backscatter communication systems,” in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, IEEE, 2020, pp. 1–6 (cited on p. 98).

[64] H. Hejazi, H. Rajab, T. Cinkler, and L. Lengyel, “Survey of platforms for massive IoT,” in *2018 IEEE international conference on future IoT technologies (future IoT)*, IEEE, 2018, pp. 1–8 (cited on p. 98).

[65] D. A. L. Galappaththige, F. Rezaei, C. Tellambura, and S. Herath, “Link budget analysis for backscatter-based passive IoT,” *IEEE Access*, vol. 10, pp. 128 890–128 922, 2022 (cited on p. 98).

[66] Z. Ding, Y. Liu, J. Choi, *et al.*, “Application of non-orthogonal multiple access in LTE and 5G networks,” *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, 2017 (cited on pp. 98, 128).

[67] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, “A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, 2017 (cited on p. 98).

[68] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, “System-level performance of downlink non-orthogonal multiple access (NOMA) under various environments,” in *2015 IEEE 81st vehicular technology conference (VTC Spring)*, IEEE, 2015, pp. 1–5 (cited on p. 98).

[69] A. S. Parihar, P. Swami, V. Bhatia, and Z. Ding, “Performance analysis of SWIPT enabled cooperative-NOMA in heterogeneous networks using carrier sensing,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10 646–10 656, 2021 (cited on p. 98).

[70] E. M. Shaheen and M. R. Soleymani, “Performance analyses of SWIPT-NOMA enabled IoT relay networks,” in *2022 International Symposium on Networks, Computers and Communications (ISNCC)*, IEEE, 2022, pp. 1–6 (cited on p. 98).

[71] Z. Ding and H. V. Poor, “On the application of BAC-NOMA to 6G umMTC,” *IEEE Communications Letters*, 2021 (cited on pp. 98–100, 102, 130, 132).

[72] Z. Ding, “Harvesting devices’ heterogeneous energy profiles and QoS requirements in IoT: WPT-NOMA vs BAC-NOMA,” *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 2837–2850, 2021 (cited on p. 98).

[73] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, “Symbiotic radio: A new communication paradigm for passive Internet of Things,” *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1350–1363, 2019 (cited on p. 98).

[74] C. Yao, Y. Liu, X. Wei, G. Wang, and F. Gao, “Backscatter technologies and the future of Internet of Things: Challenges and opportunities,” *Intelligent and Converged Networks*, vol. 1, no. 2, pp. 170–180, 2020 (cited on p. 98).

[75] Y. You, C. Qin, and Y. Gong, “Resource allocation for a full-duplex base station aided OFDMA system,” in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, 2017, pp. 1–4 (cited on p. 98).

[76] C. Zhong and Z. Zhang, “Non-orthogonal multiple access with cooperative full-duplex relaying,” *IEEE Communications Letters*, vol. 20, no. 12, pp. 2478–2481, 2016 (cited on p. 98).

[77] Z. Zhang, Z. Ma, M. Xiao, Z. Ding, and P. Fan, “Full-duplex device-to-device-aided cooperative nonorthogonal multiple access,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4467–4471, 2016 (cited on p. 98).

[78] C. Zheng, W. Cheng, and H. Zhang, “Optimal resource allocation for two-user and single-DF-relay network with ambient backscatter,” *IEEE Access*, vol. 7, pp. 91 375–91 389, 2019 (cited on p. 99).

[79] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, “Ambient backscatter: Wireless communication out of thin air,” *ACM SIGCOMM computer communication review*, vol. 43, no. 4, pp. 39–50, 2013 (cited on p. 99).

[80] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: A contemporary survey," *IEEE Communications surveys & tutorials*, vol. 20, no. 4, pp. 2889–2922, 2018 (cited on p. 99).

[81] S. Gong, X. Huang, J. Xu, W. Liu, P. Wang, and D. Niyato, "Backscatter relay communications powered by wireless energy beamforming," *IEEE Transactions on Communications*, vol. 66, no. 7, pp. 3187–3200, 2018 (cited on p. 99).

[82] J. Xu, J. Li, S. Gong, K. Zhu, and D. Niyato, "Passive relaying game for wireless powered Internet of Things in backscatter-aided hybrid radio networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8933–8944, 2019 (cited on p. 99).

[83] G. Yang, X. Xu, and Y.-C. Liang, "Resource allocation in NOMA-enhanced backscatter communication networks for wireless powered IoT," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 117–120, 2019 (cited on pp. 99, 100).

[84] W. Chen, H. Ding, S. Wang, D. B. da Costa, F. Gong, and P. H. J. Nardelli, "Backscatter cooperation in NOMA communications systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3458–3474, 2021 (cited on pp. 100, 128).

[85] W. U. Khan, X. Li, M. Zeng, and O. A. Dobre, "Backscatter-enabled noma for future 6G systems: A new optimization framework under imperfect SIC," *IEEE Communications Letters*, vol. 25, no. 5, pp. 1669–1672, 2021 (cited on pp. 100, 103).

[86] D. W. K. Ng, Y. Wu, and R. Schober, "Power efficient resource allocation for full-duplex radio distributed antenna networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2896–2911, 2016 (cited on p. 102).

[87] W. U. Khan, M. A. Jamshed, E. Lagunas, S. Chatzinotas, X. Li, and B. Ottersten, "Energy efficiency optimization for backscatter enhanced NOMA cooperative V2X communications under imperfect CSI," *IEEE Transactions on Intelligent Transportation Systems*, 2022 (cited on p. 103).

[88] X. Xie, F. Fang, and Z. Ding, "Joint optimization of beamforming, phase-shifting and power allocation in a multi-cluster IRS-NOMA network," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7705–7717, 2021 (cited on p. 108).

[89] G. Kirubasri, S. Sankar, D. Pandey, B. K. Pandey, H. Singh, and R. Anand, "A recent survey on 6G vehicular technology, applications and challenges," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, IEEE, 2021, pp. 1–5 (cited on p. 128).

[90] A. Khazali, D. Tarchi, M. G. Shayesteh, H. Kalbkhani, and A. Bozorgchenani, "Energy efficient uplink transmission in cooperative mmwave NOMA networks with wireless power transfer," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 391–405, 2021 (cited on p. 128).

[91] Q. Zhang, L. Zhang, Y.-C. Liang, and P.-Y. Kam, "Backscatter-noma: A symbiotic system of cellular and Internet-of-Things networks," *IEEE Access*, vol. 7, pp. 20 000–20 013, 2019 (cited on p. 128).

[92] X. Li, H. Liu, G. Li, Y. Liu, M. Zeng, and Z. Ding, "Effective capacity analysis of AmBC-NOMA communication systems," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 10, pp. 11 257–11 261, 2022 (cited on p. 128).

[93] X. Li, Q. Wang, M. Zeng, *et al.*, "Physical-layer authentication for ambient backscatter-aided NOMA symbiotic systems," *IEEE Transactions on Communications*, vol. 71, no. 4, pp. 2288–2303, 2023 (cited on p. 128).

[94] M. W. Akhtar, S. A. Hassan, S. Saleem, and H. Jung, "STBC-aided cooperative NOMA with timing offsets, imperfect successive interference cancellation, and imperfect channel state information," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 712–11 727, 2020 (cited on p. 130).

[95] Y. Liu, X. Mu, J. Xu, *et al.*, "STAR: Simultaneous transmission and reflection for 360° coverage by intelligent surfaces," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 102–109, 2021 (cited on p. 145).

[96] Q. Gao, Y. Liu, X. Mu, M. Jia, D. Li, and L. Hanzo, "Joint location and beamforming design for STAR-RIS assisted NOMA systems," *IEEE Transactions on Communications*, vol. 71, no. 4, pp. 2532–2546, 2023. DOI: 10.1109/TCOMM.2023.3247753 (cited on p. 145).

[97] P. Wang, H. Wang, and Y. Fu, "Average rate maximization for mobile STAR-RIS-Aided NOMA system," *IEEE Communications Letters*, vol. 27, no. 5, pp. 1362–1366, 2023. DOI: 10.1109/LCOMM.2023.3261442 (cited on p. 145).

[98] F. Fang, B. Wu, S. Fu, Z. Ding, and X. Wang, "Energy-efficient design of STAR-RIS aided mimo-noma networks," *IEEE Transactions on Communications*, vol. 71, no. 1, pp. 498–511, 2023. DOI: 10.1109/TCOMM.2022.3223706 (cited on p. 145).

[99] Z. Zhang, L. Dai, X. Chen, *et al.*, "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Transactions on Communications*, vol. 71, no. 3, pp. 1707–1725, 2022 (cited on p. 145).

[100] 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 Communications Letters*, vol. 26, no. 5, pp. 1150–1154, 2022. DOI: 10.1109/LCOMM.2022.3159525 (cited on p. 145).

[101] Z. Ding and H. V. Poor, "Design of THz-NOMA in the presence of beam misalignment," *IEEE Communications Letters*, vol. 26, no. 7, pp. 1678–1682, 2022 (cited on p. 145).

[102] O. Maraqa, S. Al-Ahmadi, A. S. Rajasekaran, H. U. Sokun, H. Yanikomeroglu, and S. M. Sait, "Energy-efficient optimization of multi-user NOMA-assisted cooperative THz-SIMO MEC systems," *IEEE Transactions on Communications*, vol. 71, no. 6, pp. 3763–3779, 2023. DOI: 10.1109/TCOMM.2023.3265123 (cited on p. 145).

[103] H. Zhang, Y. Duan, K. Long, and V. C. M. Leung, "Energy efficient resource allocation in Terahertz downlink NOMA systems," *IEEE Transactions on Communications*,

vol. 69, no. 2, pp. 1375–1384, 2021. DOI: 10.1109/TCOMM.2020.3037597 (cited on p. 145).

[104] J. Konečnỳ, B. McMahan, and D. Ramage, "Federated optimization: Distributed optimization beyond the datacenter," *arXiv preprint arXiv:1511.03575*, 2015 (cited on p. 146).

[105] P. Singh, M. K. Singh, R. Singh, and N. Singh, "Federated learning: Challenges, methods, and future directions," in *Federated Learning for IoT Applications*, Springer, 2022, pp. 199–214 (cited on p. 146).

[106] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, and D. Niyato, "Federated learning for 6G communications: Challenges, methods, and future directions," *China Communications*, vol. 17, no. 9, pp. 105–118, 2020. DOI: 10.23919/JCC.2020.09.009 (cited on p. 146).

[107] P. Behmandpoor, P. Patrinos, and M. Moonen, "Federated learning based resource allocation for wireless communication networks," in *2022 30th European Signal Processing Conference (EUSIPCO)*, 2022, pp. 1656–1660. DOI: 10.23919/EUSIPCO55093.2022.9909708 (cited on p. 146).

[108] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021 (cited on p. 147).

[109] Z. Ding, O. A. Dobre, P. Fan, and H. Vincent Poor, "A new design of CR-NOMA and its application to AoI reduction," *IEEE Communications Letters*, pp. 1–1, 2023. DOI: 10.1109/LCOMM.2023.3292859 (cited on p. 147).

[110] S. Wu, Z. Deng, A. Li, J. Jiao, N. Zhang, and Q. Zhang, "Minimizing Age-of-Information in HARQ-CC aided NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 2, pp. 1072–1086, 2023. DOI: 10.1109/TWC.2022.3201115 (cited on p. 147).