Resource Allocation Techniques for IRS-Aided NOMA System

Zinian Dong

This dissertation is submitted for the degree of

Doctor of Philosophy

University of York

School of Physics, Engineering and Technology

September 2024

Abstract

To meet the demands of emerging technologies such as the Internet of Things (IoT), Internet of Everything (IoE), virtual reality (VR), and unmanned aerial vehicles (UAVs), future sixth-generation (6G) and beyond networks must deliver higher energy and spectral efficiency, ultra-low latency, massive connectivity, and improved capacity. Traditional orthogonal multiple access (OMA) schemes used in 5G may no longer suffice, necessitating more advanced solutions. Non-orthogonal multiple access (NOMA) has emerged as a promising multiple access (MA) technique due to its advantages in spectral and energy efficiency, user fairness, and flexibility. Moreover, NOMA is compatible with other technologies, such as multiple antenna systems, enabling further improvements in system performance and complexity management. Recently, intelligent reflecting surfaces (IRS) have gained attention for their ability to enhance coverage and energy efficiency. The integration of IRS and NOMA yields a synergistic effect, producing greater benefits than either technology alone. This thesis focuses on resource allocation strategies for IRS-aided NOMA networks. Specifically, it considers a multi-user downlink IRS-aided multiple-input-single-output (MISO) NOMA system where the base station (BS) has only imperfect channel state information (CSI). A robust beamforming design is proposed to minimize power consumption while satisfying user-specific quality-of-service (QoS) constraints based on outage probability. Additionally, a cluster-based IRS-aided NOMA network is investigated, where users within each cluster share the same frequency subcarrier. To further evaluate system performance, this work addresses age of information (AoI) minimization in uplink and sum throughput maximization in downlink ultra-reliable low-latency communication (URLLC) scenarios. To solve the resource allocation problems, both convex-optimization-based alternating optimization (AO) and deep reinforcement learning (DRL) algorithms are proposed.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. Some of the work presented in this thesis has been planned to be submitted to IEEE journals, which are listed as follows:

1. **Z.Dong**, X.Wei, A.Waraiet, K.Cumannan,"Robust Downlink Beamforming Design for an IRS-aided NOMA System", to be submitted to *IEEE Transactions on Vehicular Technology.*

2. **Z.Dong**, A.Waraiet, K.Cumannan,"AoI Minimization for Uplink IRS-aided NOMA Based IoT Network", to be submitted to *IEEE Transactions on Machine Learning in Communications and Networking.*

3. **Z.Dong**, A.Waraiet, K.Cumannan,"Sum-rate Maximization for IRS-aided URLLC-NOMA System", to be submitted to *IEEE Open Journal of Communication Society.*

Zinian Dong

September 2024

Acknowledgements

Firstly, I would like to express my deepest appreciation to my supervisor, Dr. Kanapathippillai Cumanan, for his unwavering support, insightful guidance, and patience throughout the development of this thesis. Without his encouragement and feedback, I would not complete my Ph.D. It is my distinct honour and privilege to have such a gracious supervisor. I also want to present my gratitude to Xinchen Wei, Abdulhamed Waraiet and Saeed Mohammad Zadeh for their advice and support in my Ph.D research.

I am also thankful to my family, whose constant love and encouragement have been a source of strength during my research journey.

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

Abbreviation	Definition
GPRS	General Packet Radio Service
CDMA	Code-Division Multiple Access
LTE	Long-Term Evolution
WiMAX	Worldwide Interoperability for Microwave Access
SCMA	Sparse Code Multiple Access
RSMA	Rate-Splitting Multiple Access
mmWave	Millimeter Wave
UDN	Ultra-dDense Networking
AR	Augmented Reality
VR	Virtual Reality
6G	Sixth Generation
UAV	Unmanned Aerial Vehicle
ITU-R	International Telecommunication Union-Radio
NOMA	Non-Orthogonal Multiple Access
IRS	Intelligent Reflecting Surface
THz	Terahertz
ISAC	Integrated Sensing and Communication

IoE	Internet of Everything
MA	Multiple Access
OMA	Orthogonal Multiple Access
FDMA	Frequency-Division Multiple Access
TDMA	Time-Division Multiple Access
SC	Superposition Coding
SIC	Successive Interference Cancellation
SISO	Single-Input Single-Output
MIMO	Multiple-Input Multiple-Output
BS	Base Station
AWGN	Additive White Gaussian Noise
SINR	Singal-to-Interference-plus-Noise Ratio
OFDMA	Orthogonal Frequency-Division Multiple Access
MISO	Multiple-Input Single-Output
QoS	Quality of Service
IoT	Internet of Thing
URLLC	Ultra Reliable Low Latency Communication
LoS	Line of Sight
CSI	Channel State Information
SDP	Semidefinite Programming
AoI	Age of Information
LP	Linear Programming
QP	Quadratic Programming

QCQP	Quadratic Constrained Quadratic Programming
SOCP	Second Order Cone Programming
ML	Machine Learning
RL	Reinforcement Learning
DRL	DDeep Reinforcement Learning
AI	Artificial Intelligence
MDP	Markov Decision Process
MAB	Multi-Arm Bandit
NN	Neural Network
DQN	Deep Q-Network
DDPG	Deep Deterministic Policy Gradient
TD3	Twin-Delayed Deep Deterministic Policy Gradient
1G	First-Generation
2G	Second-Generation
3G	Third-Generation
4G	Fourth-Generation
5G	Fifth-Generation
6G	Sixth-Generation
AO	Alternating Optimization
CSCG	Circularly Symmetric Complex Gaussian
SDR	Semidefinite Relaxation
NLUPA	Next Largest-difference User Pairing Algorithm
DNN	Deep Neural Network

List of Notations

Notation	Definition
$\mathbb{E}\{\cdot\}$	Statistical Expectation
$\mathcal{C}^N$	N-dimensional complex vectors
$\mathcal{R}^N$	N-dimensional real vectors
$\mathcal{S}^N$	N-dimensional semidefinite matrix
$\|x\|$	Absolute value of complex number x
$\|\cdot\|$	Euclidian vector norm
$\mathbf{x}$	Vector $\mathbf{x}$
$\mathbf{X}$	Matrix $\mathbf{X}$
$(\cdot)^T$	Transpose
$(\cdot)^H$	Hermitian Transpose
$\mathrm{Tr}\{\mathbf{X}\}$	Trace of matrix $\mathbf{X}$
$\mathbf{x} \succeq 0$	Each element in $\mathbf{x}$ is greater than zero
$\mathbf{X} \succeq 0$	Positive semi-definite matrix $\mathbf{X}$
$\min\{\cdot\}$	Minimum of function
$\max\{\cdot\}$	Maximum of function
$\mathcal{CN}(\mu, \sigma^2)$	Complex Gaussian random variable with mean μ and variance σ^2
$\mathbf{dom} f$	Domain of a function f
∇f	First derivative of function f
$\nabla^2 f$	Second derivative of function f
$[\mathbf{X}]_{n,m}$	The m row n column element of matrix $\mathbf{X}$
$\Re\{\cdot\}$	Real part of a complex number
$\Im\{\cdot\}$	Imaginary part of a complex number

Chapter 1

Introduction

1.1 Overview

In today's highly information-driven era, wireless communication has become crucial to modern society. As the standard and requirements of wireless communications have changed rapidly, they typically upgrade to a new generation every ten years. In particular, the evolution from 1G to 5G represents a significant advance in terms of speed, capacity, and functionality [1]. In addition, the development of multiple access (MA) techniques also played an important role in the allocation of network resources, enabling future applications in 6G and beyond.

The journey of wireless communication began with Heinrich Hertz in the late 19^{th} century, when he first generated and verified the existence of radio waves [2]. After mastering the fundamentals and applications of radio, the first wireless cellular communication was launched in Japan in 1979, leading to the announcement of the **1G!** (**1G!**) mobile system. In first-generation (1G), analog transmission supported voice service, in which frequency-division multiple access (FDMA) was applied [3]. Specifically, FDMA technique divides the available frequency bandwidth into multiple sub-frequency bands and assigns each of those to specific users [4]. However, due to the limitation of the available bandwidth, the number of devices that can be served by base station (BS) simultaneously was limited. In order to address this problem, a digital transmission system with more capacity, namely second-generation (2G), replaced the 1G network in 1991 [5].

In the 2G communication system, users were able to deliver text and picture messages by employing time-division multiple access (TDMA). Similarly, in TDMA, time resources

are divided into multiple time slots, within which users send their signals via the same frequency band. In addition, the 2G mobile system was integrated with general packet radio service (GPRS), such as 2.5G [3].

In the late 1980s, the third-generation (3G) cellular system was introduced to meet the needs of a larger number of serviced users, faster data transfer, and better quality of service (QoS). In 3G networks, code-division multiple access (CDMA) techniques were used, such as wideband CDMA and CDMA2000, which supported higher-data-rate access [6]. A CDMA network allocates a unique code to each transmitted data, allowing multiple signals to occupy the same frequency band without interfering with each other [7].

In 2008, the requirements of fourth-generation (4G) were specified, including peak speed at 100 Mbps for high-mobility communication and 1 Gbps for low-mobility communication [8]. With the two techniques of 4G, namely, Long-Term Evolution (LTE)-advanced and mobile Worldwide Interoperability for Microwave Access (WiMAX), devices were provided more rapid access to the Internet, which supported various services, such as video calls and file transmission. In the 4G network, orthogonal frequency-division multiple access (OFDMA) is adopted, which splits the whole available bandwidth into orthogonal multiple sub-carriers respectively serving each user simultaneously [3]. Benefiting from 4G, users were able to watch mobile streaming and pay on the Internet.

In the late 2010s, the fifth-generation (5G) was deployed worldwide, offering ultra-high data speeds and extremely low latency services [9]. In the 5G networks, OFDMA is still the primary MA technique, such as OFDMA, complemented by single carrier FDMA in the uplink and non-orthogonal multiple access (NOMA) for specific high-density scenarios. Furthermore, this MA technique are assisted by other methods, including sparse code multiple access (SCMA), rate-splitting multiple access (RSMA), and NOMA. Specifically, SCMA allows multiple users to share the same time-frequency resources by employing sparse codes [10]. RSMA splits the data of each user into common and private parts, which are encoded separately with different codes [11]. NOMA multiplexes signals intended for different users in the code and power domain at the transmitter to significantly enhance the spectral efficiency. In addition, 5G BS also employs other techniques, such as millimeter wave (mmWave), massive multiple-input multiple-output (MIMO), and ultra-ddense networking (UDN) [12]. With the assistance of these techniques, multiple

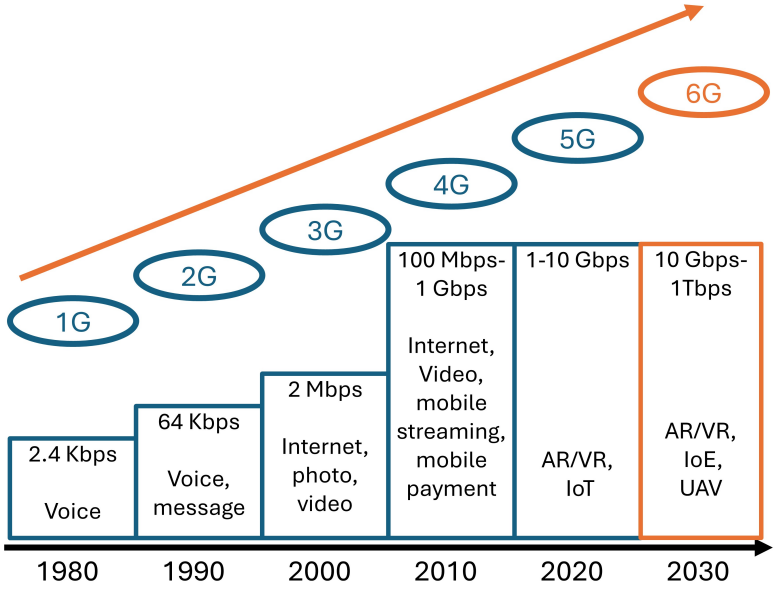


Figure 1.1: An illustration of wireless communication evolution.

applications can be employed, including augmented reality (AR), virtual reality (VR), and internet of thing (IoT).

1.2 Towards 6G and beyond

1.2.1 Requirements of 6G

In the previous section, the evolution of wireless communication generations is discussed, which is shown in Fig. 1.1. Since the wireless network generation is updated roughly every one decade, the research and development of sixth-generation (6G) have been undertaken by the global telecommunications industry. In particular, the International Telecommunication Union-Radio (ITU-R) consider eight parameter as performance indicators of 5G, which are also valid for 6G, as shown in Fig. 1.2. To further compare 5G and 6G, these key metrics are presented as follows [12, 13]:

- Peak Data Rate: Due to the development of a number of emerging technologies, such as AR andVR, higher data rate is a vital requirement for 6G communication system. This can be illustrated by peak data rate, which presents the fastest achievable data rate with perfect condition. 5G networks are supposed to provide services with 20 Gbps peak data rate, while the expected rate of 6G networks are up to 1 Tbps,

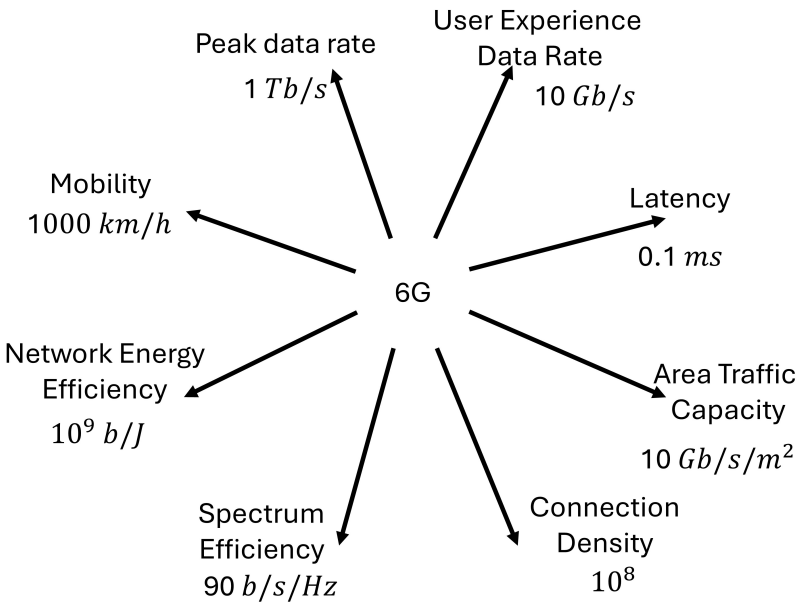


Figure 1.2: An illustration of key requirements of 6G networks.

which is more than 50 times over 5G [14].

- User Experience Data Rate: The user experience data rate represents the actual data throughput that can be ensured for the receivers with a probability of more than 95%. Specifically, 5G is able to support 100 Mbps, while 6G is expected to increase the rate 10 times, reaching 10 Gbps [13].

- Latency: Latency is a metric to measure the time delayed from the transmitters and receivers. The minimal latency of 5G networks is 1 ms, while 6G should be able to perform 10 times better, with 0.1 ms delay [12].

- Area Traffic Capacity: Area traffic capacity refers to the total data throughput that a network can serve within a specific geographic area. The minimal traffic capacity of 5G networks reaches 10 Mbps/m^2, which is expected to be 10 Gbps/m^2 in 6G.

- Connection density: Connection density represents the number of devices or connections that can be supported within a unit area, which is especially important for IoT applications. 5G networks have ability to support up to 10^6 devices per square kilometers, while 6G is supposed to support 10^8 devices per square kilometers [12].

- Spectrum Efficiency: Spectrum efficiency is a metric to evaluate how effectively the available frequency spectrum is utilized, which illustrate the highest data through-

put per unit bandwidth. For 5G, the spectrum efficiency is 30 bps/Hz. In contrast, the target spectrum efficiency of 6G is 3 times higher than that of 5G, reaching more than 90 bps/Hz [12].

- Network Energy Efficiency: Energy efficiency indicates the amount of data transmitted with unit power consumption. The 6G network should be able to perform 100 times better than 5G, transmitting 10^9 bit/J [13].

- Mobility: Mobility refers to the capability of a communication system to maintain continuous and reliable connectivity as users moving, which is measured by the maximal speed under QoS constraints. In 5G networks, the acceptable speed is up to 500 km/h. In 6G network, the transmission will be more stable, with maximal mobility at 1000km/h [13].

1.2.2 Enabling Techniques for 6G and Beyond

As previous subsection discussed, the existing techniques are not sufficient to meet the requirements of 6G or beyond wireless networks. Thus, various of techniques has recently been considered promising for future communication networks, including mmWave, Terahertz (THz) communication, massive MIMO, artificial intelligence (AI)-driven network, NOMA, intelligent reflecting surface (IRS), and integrated sensing and communication (ISAC) [12, 13]. It is worth mentioning that as the transition from 5G to 6G progresses, several key technologies introduced in 5G will continue to play a vital role but with enhanced capabilities. The basic concept of these techniques are briefly introduced as follows.

- mmWave: In order to extend bandwidth compared to conventional frequency band used in previous generation of wireless networks, mmWave communications operate in the frequency range of 30 GHz to 300 GHz and corresponding wavelength range of 1 mm to 10 mm [15]. This wide bandwidth provide more capacity, higher throughput, and lower latency. In addition, due to the short wavelength, mmWave communications enable large antenna arrays in physically small size antenna elements, leading to narrower beams. The constrict beams are able to improve signal strength and reduce interference [16]. The mmWave technology is one of the key components applied in 5G and still remain potential for 6G and beyond network.

- THz Communication: THz communication is considered as an upgrade of mmWave, which further enlarges the bandwidth in terahertz frequency range, such that from 0.1 THz to 10 THz [17]. Compared to mmWave, THz communication offers even broader bandwidth and the potential for ultra-high data rates up to several terabits per second. This makes THz ideal for applications requiring high speed data transfer in 6G, such as AR, VR, and unmanned aerial vehicle (UAV). Although mmWave technology has already established and integrated into current networks, THz technology still requires continued research to overcome its technical complexities and deployment challenges.

- Massive MIMO: Massive MIMO is an advanced multiple antennas technique designed to enhance wireless communication systems. In conventional MIMO networks, transmitters and receivers are equipped with a few antennas, while massive MIMO networks deploy a significantly large number of antennas, ranging around hundreds, at the BS [18]. This advanced MIMO network design beamformers to direct signals precisely towards multiple users simultaneously, providing stronger signals while minimizing interference [19]. Thus, massive MIMO is able to play a crucial role in meeting the growing demands for high-speed, high-capacity connectivity in 5G, 6G, and future wireless networks. In addition, there is another advanced technique, namely, cell-free massive MIMO, in which a large number of access points serve multiple users jointly and simultaneously across a wide geographic area [20, 21]. Unlike traditional cellular MIMO that relies on BSs within a defined area, cell-free massive MIMO eliminates the concept of cells.

- AI-Driven: An AI-driven network is an advanced communication system where AI is utilized to enhance the network performance, security and QoS. Through the analysis the real-time and previous data, the AI-driven network has ability to predict and address issues automatically, such as dynamically assigning the resource blocks and optimizing power consumption [22]. Besides, the AI is able to continuously learn from the varying environment, leading to seamless resources management. Therefore, driven by AI, complex allocation problems can be addressed, resulting in higher efficiency, improved throughput, reduced latency.

- NOMA: NOMA is an innovative MA technique designed to enhance the capac-

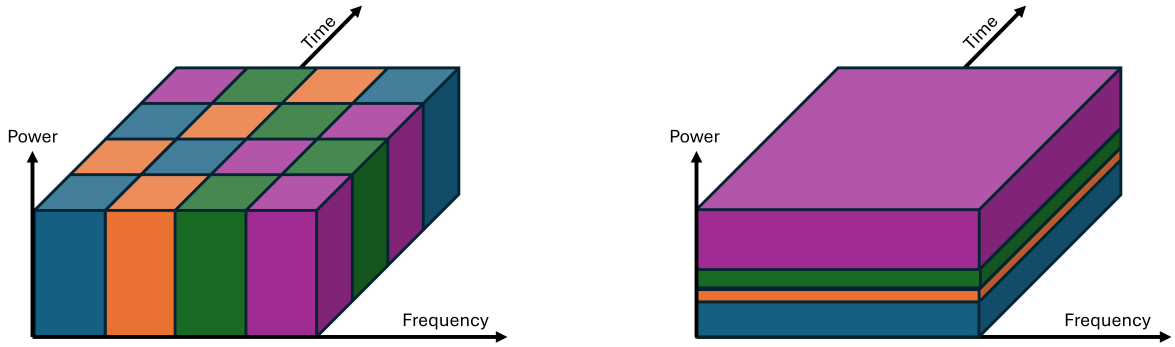
ity and efficiency of the network. Unlike conventional orthogonal multiple access (OMA) that assigns separate resources to each user, NOMA allows multiple users to share the same time or frequency resources simultaneously by applying superposition coding (SC) and successive interference cancellation (SIC) [9, 23]. Particularly, SC is exploited at the transmitters, which allocates different weights, namely, codes or power levels, based on users' channel condition, while SIC is exploited at the receivers, which allows users with stronger channels to encode signals intended for users with weaker channels to reduce interference [23,24]. Benefit from such spectrum sharing, the performance of NOMA network in terms of bandwidth efficiency, fairness and energy efficiency, can be significantly improved.

- IRS: IRS has been a promising technique in wireless communication, offering a novel way to enhance signal propagation and optimize network performance [25]. Specifically, IRS consists of numerous passive, controllable reflecting elements, each of which is able to reflect and steer signals by adding amplitude and phase shift to the incident signals [26, 27]. Therefore, IRS is able to regenerate stronger communication links, providing better performance.

- ISAC: ISAC is another emerging technology that combines communication and sensing ability within a single system, where the network is able to simultaneously transmit data sense the environment [28]. By using the same resource blocks for both sensing and communication functions, ISAC enhances network efficiency and reduces infrastructure costs. The dual functionality of ISAC supports various of applications, including UAV and Internet of everything (IoE), by providing immediate situational perception and connectivity [29].

1.3 Towards NOMA

As mentioned in the previous section, NOMA has been a promising technique for future wireless network [30]. The core principle of NOMA scheme is to enable multiple users to share the same resource blocks and transmit simultaneously. As for traditional OMA schemes, TDMA divides the whole time resources into multiple time slots, each of which is allocated to an unique user, while FDMA splits the available frequency resources into multiple subcarriers, which is occupied by only single user. On contrary,



(a) Resource allocation for OFDMA. (b) Resource allocation for NOMA.

Figure 1.3: Illustrations of resource allocation for OFDMA and NOMA scheme.

NOMA eliminates the orthogonality in previous MA techniques, considerably enhancing the spectrum efficiency. In general, there are two types of NOMA, code-domain NOMA and power-domain NOMA [9, 23]. Code-domain NOMA assigns multiple users unique non-orthogonal code sequences, enabling simultaneous communication [31]. On the other hand, power-domain NOMA distinguishes signals intended for different users based on their power levels. This thesis only considers power domain NOMA. The comparison of resource allocation methods of OFDMA and power-domain NOMA is presented in Fig. 1.3. As shown in Fig. 1.3a, OFDMA divides frequency and time orthogonally and assigns different users with same power, while in Fig. 1.3b, power-domain NOMA scheme allows different users to share all the time and frequency resources and allocates different power levels to different users. Particularly, the power allocation based on the channel strengths of users. NOMA allocates higher power to the users with poor channel strengths, superpose the signals for multiple users together at the transmitter by SC, and apply SIC at receivers to remove inter-user interference. Specifically, in downlink NOMA, the user with better channel condition decodes the signal of user with worse channel condition first, treating it as interference, and subtracts it from the received signal using SIC. On the other hand, in the uplink, the BS receives a superimposed signal and applies SIC to decode the signal with higher power. Then, the BS subtracts the decoded signal from the received mixed signal to obtain the weaker signal. In this manner, NOMA is able to improve user fairness and spectrum efficiency, which address the demand of massive connectivity for 6G and beyond networks.

In addition, NOMA can be integrated with other existing techniques, such as multiple-

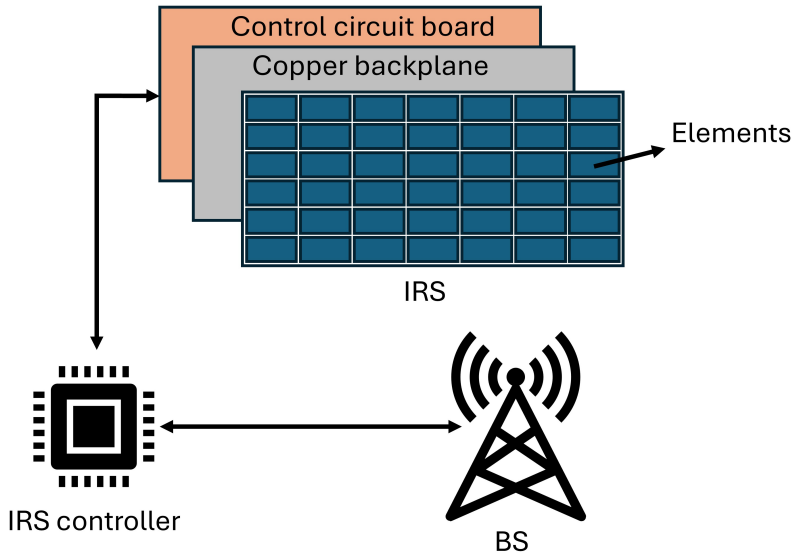


Figure 1.4: An illustration of IRS architecture

antenna technique and conventional OMA schemes [30]. Specifically, multiple-input single-output (MISO)-NOMA and MIMO-NOMA are explored in numerous literature [32, 33]. This hybrid approach aims to further enhance increase capacity, and improve energy efficiency by leveraging spatial domain resources in addition to power domain. OFDMA and TDMA are also combined with NOMA in [34,35], generating cluster-based NOMA, where clusters use resources divided by OMA and NOMA scheme is established in each cluster. Furthermore, NOMA can also be integrated with other emerging technologies, such as IRS, IoT, and ultra reliable low latency communication (URLLC).

1.4 Towards IRS

On the other hand, IRS is another enabling technology to achieve demands for future wireless communication [36]. Physically, as shown in 1.4, the architecture of IRS is illustrated, where IRS is composed of three layers and one controller [26, 37]. The first layer consists of multiple tunable elements, each of which is a dielectric material and able to manipulate the incident signals. The second cooper layer is exploited to reflect the signals by eliminating the power leakage. The third layer is the control circuit board, which can operate the elements to adjust the amplitude and/or phase shift dynamically. This operation is determined and controlled by IRS controller, which communicates with

a BS through wireless link.

By doing so, IRS can improve signal coverage, increase spectral efficiency, and reduce energy consumption, providing a cost-effective alternative to traditional methods, such as deploying additional BS [25, 26]. In addition, although another technique, namely, relay system can also achieve the same effect as IRS does, they do so through different mechanisms. Relays actively amplify and re-transmit signals, extending coverage and boosting signal strength but with higher power consumption and costs, while IRS, on the other hand, passively reflects and optimizes signals with minimal power usage and cost by offering dynamic adjustment of all elements [38].

Furthermore, due to numerous benefits of IRS, it can be integrated with roughly all the wireless networks, such as IRS-aided MISO [39] and IRS-aided MIMO [40]. By adjusting the phase and amplitude of incoming signals, IRS supports a great number of applications, from smart cities and industrial IoT to high-density urban environments, making it a crucial technology for realizing the full potential of future network [26].

1.5 Thesis Outline and Contributions

Nevertheless, NOMA and IRS still present several challenges despite of their significant advantages. The most critical challenge for NOMA is efficient resource allocation. Since power is assigned to users based on their channel strengths, dynamic adjustment is required to optimize performance and reduce interference across users with varying channel strength. Additionally, as NOMA applies SIC at the receivers to manage the interference, it is also crucial to determine the decoding order and user clustering. On the other hand, amplitude and phase shift of IRS elements requires dynamical optimization as well. Besides, the IRS phase shifts directly affect the channel gains, which further influence NOMA power allocation and SIC decoding order. This creates a highly non-convex and interdependent optimization problem. In conclusion, the coupling of power, beamforming, and phase shift optimization makes the problem significantly more complex. Therefore, compared to existing research on resource allocation problems in wireless communication, more iterative algorithms or machine learning (ML)-based solutions are required for efficient optimization. Thus, this thesis mainly focuses on different resources allocation and phase shift optimization techniques for IRS-aided NOMA networks. The

outline of this thesis is summarized as follows:

In Chapter 2, the basic fundamentals and literature review of both IRS and NOMA are presented. In the first section, the basic principle of NOMA are discussed, including SC and SIC. Then, a single-input single-output (SISO)-NOMA and MISO-NOMA scenarios are reviewed, respectively. In addition, fundamental principle of SIC is introduced in the following section, involving its signal model and advantages. To further illustrate system models, IRS-aided SISO-NOMA and MISO-NOMA scenarios are presented. Finally, detailed literature review related to resource allocation problems for IRS-aided NOMA networks is provided.

In Chapter 3, various resource allocation techniques to optimize different variables achieving higher performance of the system are introduced first. Secondly, to solve these problems, fundamental concept of convex optimization techniques and ML-based algorithms are discussed. In particular, basic of principle of convex set, function and problems are provided. Then, convex problems, including linear programming (LP), quadratic programming (QP), quadratic constrained quadratic programming (QCQP), second order cone programming (SOCP) and semidefinite programming (SDP) are presented in detail. On the other hand, the category of ML-based algorithm is introduced briefly, involving supervised learning, unsupervised learning, reinforcement learning (RL) and deep learning, while the training process of actor-critic dDeep reinforcement learning (DRL), especially deep deterministic policy gradient (DDPG) agent, is shown mathematically.

In Chapter 4, a robust downlink beamforming design to allocate resources for IRS-aided MISO NOMA network is proposed. In this design, it is assumed that the BS receives imperfect channel state information (CSI) of cascade channel through the IRS with uncertainty that follows circularly symmetric complex Gaussian (CSCG) distribution. Thus, the robust beamforming design is formulated as a total transmit power minimization problem subject to outage-probability-based singal-to-interference-plus-noise ratio (SINR) constraint for each user. To solve this non-convex problem, Bernstein-type inequality is applied to replace the constraints with outage probability with linear inequalities. Then, an alternating optimization (AO) algorithm is proposed, where the beamforming vectors and IRS phase shift are optimized separately and iteratively until the total power consumption converges. The simulation results shows effectiveness of the proposed algorithm and represents the benefit of NOMA and impact of IRS and

multiple-antenna technique to this network.

To further investigate cluster-based IRS-aided NOMA, an age of information (AoI) minimization problem for IoT-based network is considered in Chapter 5. In traditional NOMA, multiple users share the same time and frequency resources, allowing users with different channel conditions to be served simultaneously. In contrary, cluster-based NOMA, orthogonal frequency subcarriers are assigned to different clusters to avoid inter-cluster interference, while NOMA scheme is applied to allocate resource for users within a cluster. In this work, the problem aims to optimize devices clustering, power allocation and IRS phase shift to minimize AoI subject to total power consumption constraint for individual device. To solved this problem, an algorithm for device clustering according to channel gains is proposed first. Then, a sub-optimal convex-optimization-based algorithm is proposed, which optimizes IRS phase shift and power allocation in turn. In addition, after using the same device clustering methods, another DRL-based algorithm is presented to jointly optimize power allocation and phase shift. Then comparison between both algorithm is shown in the simulation results.

In Chapter 6, a sum throughput maximization problem for cluster-based IRS-aided URLLC NOMA network is formulated. Particularly, unlike conventional data rate provided by Shannon capacity theorem, in URLLC system, short packets communication and decoding error probability are both considered. Thus, the inverse Q function is involved in data rate formulas. To solve this non-convex problem, DRL-based algorithm is propose. In particular, users are clustered based next largest-difference user pairing algorithm (NLUPA). Then, the original problem is formulated into a DRL environment and then a DDPG agent is trained to optimize power allocation and IRS phase shift jointly. The performance of the proposed algorithm is presented in the simulation results.

Finally in Chapter 7, the conclusion of this thesis is presented first, focusing on the contributions and results of three works in Chapter 4, Chapter 5, and Chapter 6. Then, in the next section, the investigations, which should be extended from current works and future research are introduced. Particularly, the research works in Chapter 5 and Chapter 6 can be further investigated with multiple-antenna and imperfect CSI scenario. In addition, twin-delayed deep deterministic policy gradient (TD3) agent can be applied in DRL algorithm to solve power allocation and IRS phase shift optimization problems jointly. Finally, future works on resource allocation problems for ISAC-NOMA are discussed.

Chapter 2

Fundamental Concepts and Literature Review

In this chapter, the fundamental concepts of NOMA are introduced first, including two key techniques, namely SC and SIC. Secondly, scenarios of downlink transmission of SISO-NOMA and SISO-OMA are considered and comparison in terms of achievable data rate is also presented. Then, the advantages of NOMA scheme is reviewed. On the other hand, the basic principle and advantages of IRS are provided in the next section. Finally, the related literature review on NOMA, IRS and IRS-aided NOMA is respectively presented briefly.

2.1 NOMA Fundamentals

NOMA has received great interests due to its wide range of benefits and has been proposed as a promising MA technique for future wireless communication networks [9,23,24]. Compared with conventional OMA techniques, such as FDMA and TDMA, NOMA explores a new dimension that allows multiple users to transmit or receive signals simultaneously by sharing the same resource block, namely frequency and time resources. This improves the overall system capacity and enables more users to be served simultaneously.Therefore, NOMA is able to enhance spectral efficiency and support massive connectivity in next-generation wireless networks. In general, there are two types of NOMA schemes, power-domain and code-domain NOMA. In power-domain NOMA, different users are assigned with different power levels according to their channel strengths,

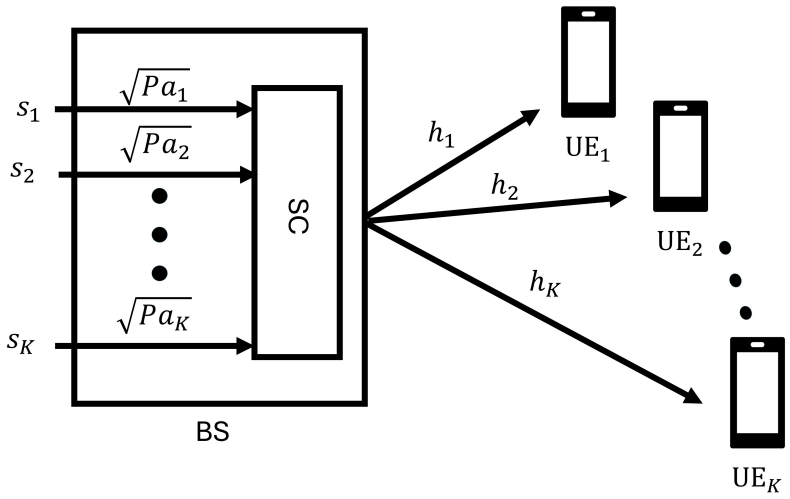


Figure 2.1: A multi-user SISO NOMA System with SC.

while in code-domain NOMA, different spreading codes are allocated to users over the whole available time-frequency resources. In this thesis, power-domain NOMA is mainly focused. The simultaneously sharing in power-domain NOMA is achieved through implementing SC and SIC [41, 42]. Particularly, signals of multiple users can be differentiated according to their unique power levels or code sequences. The basic principle of SC and SIC is introduced in the following subsection.

2.1.1 Superposition Coding and Successive Interference Cancellation

The implementation of NOMA depends on two key techniques: the SC at transmitters and the SIC at receivers [23]. SC is a technique for multi-user communication systems, which is firstly introduced in [43]. It enables a transmitter to send multiple signals for different users over the same resource block simultaneously by superimposing them with different priority weights [23, 44]. As shown in Fig. 2.1, a muti-user SISO NOMA scenario is considered, where there are K number of users equipped with single antenna served by single-antenna BS. In power-domain NOMA, K users are allocated with different power based on their channel strength. Then the signals with different power are superposed and sent to all the users.

On the other hand, SIC is applied at the receivers, the basic concept of which is to

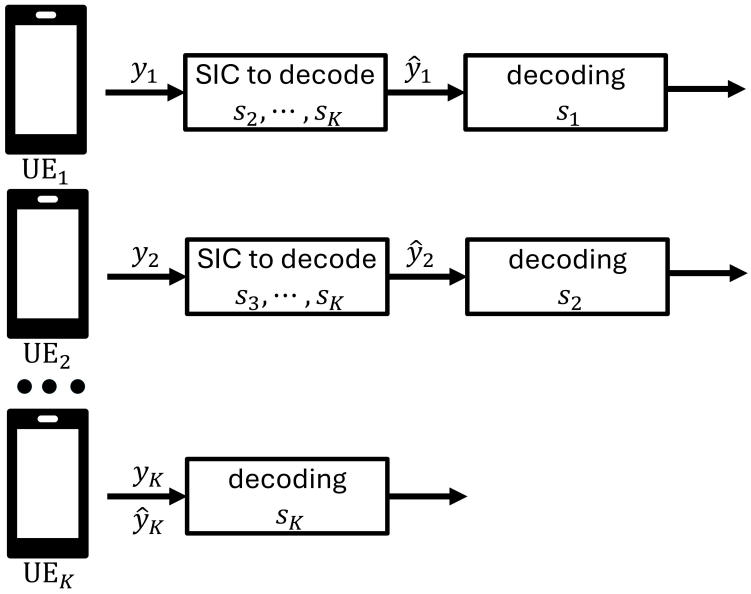


Figure 2.2: A downlink multi-user SISO-NOMA System with SIC.

successively decode the signals of different users and progressively remove the inter-user interference [9]. In downlink NOMA, the user with stronger channel condition is able to sequentially decode and subtract the signals intended for other users with weaker channel condition. Particularly, the signal for user with the weakest channel gain is decoded first and subtracted from the rest of received signal. Then, the second weakest user's signal is decoded. This iterative decoding process continues until the signal of the target user is decoded. Notably, SIC is applied at stronger users with lower error and the successful implementation of it demands higher power levels of weaker users' signal. As shown in Fig. 2.2, the channel strengths can be sorted as $h_1 \geq h_2 \geq \cdots \geq h_K$, where h_i indicated the channel strength of UE_i, $\quad i \in \{1, 2, \cdots, K\}$. UE_1 has the highest channel gains, therefore signals intended for other users can be decoded and subtracted. Thus, there is no interference from other users at UE_1. Similarly, UE_2 can decode other users' signals except for user with stronger channel gains, namely, UE_1 so signal for UE_1 is regarded as interference. This process continues until the last user decodes its own signal considering other users' signals as interference. On the other hand, in uplink NOMA, the BS receives a superimposed signal. In order to decode it, the BS decodes the signal with the highest power first. Then, the BS subtracts it from the received combined signal, which is repeated iteratively, decoding the next strongest signal at each step, until all signals are decoded. In particular, a SIC process of uplink multi-user SISO-NOMA system

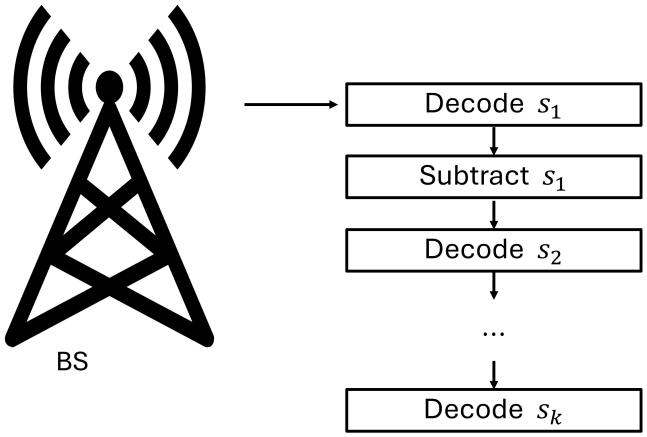


Figure 2.3: An uplink multi-user SISO-NOMA System with SIC.

is considered in Fig. 2.3, where s_i indicates the signal for UE_i, $i = 1, \cdots, k$. Assuming the strengths of received signal at BS are sorted as $|h_1 s_1|^2 \geq |h_2 s_2|^2 \geq \cdots \geq |h_k s_k|^2$, then the BS decodes s_1 first treating other signals as interference. Then the decoded s_1 is subtracted from the mixed signal. Then s_2 is decoded treating the s_3 to s_k as interference. The iterative process ends until the last signal s_k is decoded. The concept of SIC has also been exploited inMIMO system [45]. The SC and downlink SIC process are introduced mathematically in the next subsection.

2.1.2 A SISO-NOMA Scenario

To further illustrate the SC and SIC process in detail, a simpler two-user power-domain NOMA system is considered, assuming $K = 2$ shown in 2.1. The BS transmits two signals s_1 and s_2, intended for user 1 (UE_1) and user 2 (UE_2), respectively. Let p denote the total transmit power and $a_i, \quad i = 1, 2$ stand for the fraction of transmit power assigned for user i. Applying SC, the transmitted signal from the BS can be given by

$$x = \sqrt{pa_1} s_1 + \sqrt{pa_2} s_2, \tag{2.1}$$

where $a_1 + a_2 = 1$. Note that the power allocation, a_1 and a_2, depends on the channel gains of two users [9]. Notably, users with stronger channel gains are allocated with lower

power, while users with weaker channel gains are allocated with higher power. Denote the channel coefficient from UE_1 and UE_2 to the BS by h_1 and h_2, respectively and assume UE_1 has a stronger channel gain than that of UE_2 based on the distance between the BS, such as

$$|h_1|^2 \geq |h_2|^2. \tag{2.2}$$

Therefore, UE_1 is assigned with lower power, which is given by

$$a_1 \leq a_2. \tag{2.3}$$

At receivers, the received signals for $UE_i, \quad i = 1, 2$ are given by

$$y_i = h_i \left(\sqrt{pa_1}s_1 + \sqrt{pa_2}s_2\right) + n_i, \tag{2.4}$$

where n_i denotes the additive white Gaussian noise (AWGN) with zero mean and variance σ_i^2. As assumed in (2.2), UE_2 has worse channel condition than UE_1 and can only decode its own signal, while $UE_1$is able to decode signals of UE_2 and subtract it from the received signal y_1. As such, the interference from UE_2 can be removed from y_1. In contrast, UE_2 can only decode its own signal regarding the signal for UE_1 as interference. Thus, after SIC, the received signal for UE_1 and UE_2 can be rewritten as

$$\hat{y}_1 = h_1\sqrt{pa_1}s_1 + n1, \tag{2.5}$$

$$\hat{y}_2 = h_2 \left(\sqrt{pa_1}s_1 + \sqrt{pa_2}s_2\right) + n_2. \tag{2.6}$$

Note that both users share the same resource block, namely the frequency bandwidth. With the assumption that the bandwidth is set to be 1HZ, the SINRs for UE_1 and UE_2 can be represented as

$$\text{SINR}_1^1 = \frac{|h_1|^2 pa_1}{\sigma_1^2}, \tag{2.7}$$

$$\text{SINR}_2^1 = \frac{|h_1|^2 pa_2}{|h_1|^2 pa_1 + \sigma_1^2}, \tag{2.8}$$

$$\text{SINR}_2^2 = \frac{|h_2|^2 pa_2}{|h_2|^2 pa_1 + \sigma_2^2}, \tag{2.9}$$

where SINR_2^1 and SINR_2^2 denote the SINR of decoding the signal intended for UE_2 at UE_1 and at itself, respectively. Then the SINR of decoding signal for UE_2 is defined as [46]

$$\mathrm{SINR}_2 = \min\left\{\mathrm{SINR}_2^2, \mathrm{SINR}_2^1\right\} = \mathrm{SINR}_2^2, \tag{2.10}$$

where the equation always holds true when $\sigma_1^2 = \sigma_2^2$. According to the Shannon's capacity formulation [47], the achievable data rate for UE_1 and UE_2 is given by

$$R_1^{\mathrm{NOMA}} = \log_2(1 + \mathrm{SINR}_1^1) = \log_2\left(1 + \frac{|h_1|^2 p a_1}{\sigma_1^2}\right), \tag{2.11}$$

$$R_2^{\mathrm{NOMA}} = \log_2(1 + \mathrm{SINR}_2) = \log_2\left(1 + \frac{|h_2|^2 p a_2}{|h_2|^2 p a_1 + \sigma_2^2}\right). \tag{2.12}$$

In order to further validate the performance of NOMA in terms of achievable data rate improvement against conventional OMA, a two-user OFDMA scenario is considered. The available bandwidth is divided into two parts, W_1 for UE_1 and W_2 for UE_2, where $W_1 + W_2 = 1\mathrm{HZ}$. Therefore, the achievable rate for each user in the OFDMA system is expressed as

$$R_i^{\mathrm{OMA}} = W_i \log_2\left(1 + \frac{|h_i|^2 p a_i}{W_i \sigma_i^2}\right), \quad i = 1, 2. \tag{2.13}$$

Hence, the achievable sum rate for two-user NOMA and OFDMA can be written respectively as

$$R^{\mathrm{NOMA}} = \log_2\left(1 + \frac{|h_1|^2 p a_1}{\sigma_1^2}\right) + \log_2\left(1 + \frac{|h_2|^2 p a_2}{|h_2|^2 p a_1 + \sigma_2^2}\right), \tag{2.14}$$

$$R^{\mathrm{OMA}} = W_1 \log_2\left(1 + \frac{|h_1|^2 p a_1}{W_1 \sigma_1^2}\right) + W_2 \log_2\left(1 + \frac{|h_2|^2 p a_2}{W_2 \sigma_2^2}\right). \tag{2.15}$$

For the sake of simplicity, let assume the whole bandwidth is equally divided, such that $W_1 = W_2 = 0.5\mathrm{HZ}$. Moreover, the noise power is assumed to be same for both user, such that $\sigma_1^2 = \sigma_2^2 = 1$. Assuming $|h_1|^2 = 10$, $|h_2|^2 = 5$, and $p = 1\mathrm{W}$, the achievable sum rate for two scenarios are presented in Fig. 2.4. It is obviously shown that SISO-NOMA outperforms SISO-OMA in terms of achievable sum rate.

In addition, it is mathematically demonstrated in [48] that NOMA scheme is able to provide better performance in terms of achievable rate than OMA scheme.

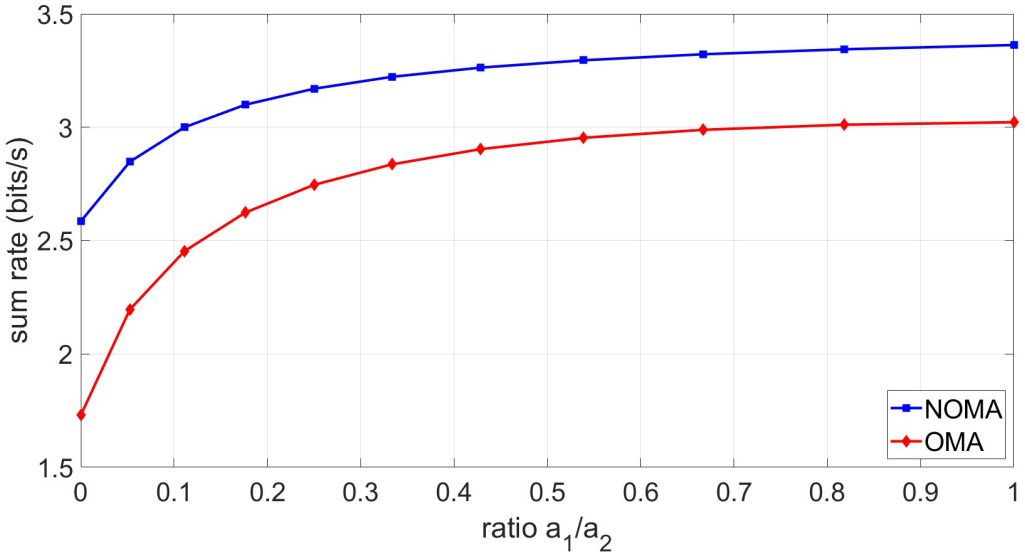


Figure 2.4: Achievable Sum Rate for SISO-NOMA and SISO-OMA.

2.1.3 A MISO-NOMA Scenario

To exploit spatial domain and enhance system performance, multiple antenna techniques have received great attention over the last two decades [49]. In particular, this is achieved by employing multiple antennas either at transmitters or receivers, such as MISO and MIMO network. Besides, considering the utilization of power domain by NOMA, more advantages can be obtained with the integration of NOMA and multiple antennas techniques [50–52]. When NOMA scheme is applied in a multiple antenna system, the capacity, spectrum efficiency, energy efficiency, and spatial density can be improved to a higher level. Note that as this thesis focuses on the SISO and MISO system with NOMA scheme, this subsection is narrowed to MISO-NOMA. As shown in Fig. 2.5, a multi-user MISO-NOMA system is considered, which consists of a BS equipped with N antennas and K single-antenna users. Let the k^{th} user denote by UE_k, $k \in \mathcal{K}$, where $\mathcal{K} = \{1, 2, \cdots, K\}$. The BS assignes beamforming vectors $\mathbf{w}_k \in \mathbb{C}^{N\times 1}$ to UE_k. Thus, the transmitted signal after SC at the BS is given by

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{w}_k s_k, \tag{2.16}$$

where s_k stands for the signal symbol for UE_k. Assuming that $s_k, \quad \forall k \in \mathcal{K}$ are independent and uncorrelated with each other, the signals have unit power, such as $E\left[|s_k|^2\right] = 1$.

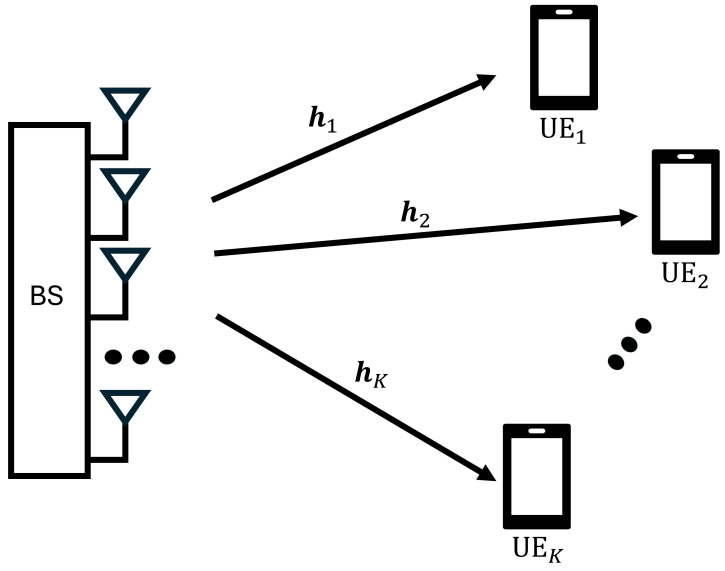


Figure 2.5: A Multi-user MISO-NOMA System.

Then, the received signal can be expressed as

$$y_k = \mathbf{h}_i^H \mathbf{x} + n_i = \mathbf{h}_k^H \left(\sum_{i=1}^{K} \mathbf{w}_i s_i \right) + n_i, \quad \forall k \in \mathcal{K}, \tag{2.17}$$

where $\mathbf{h}_k$ is the channel coefficients between the BS and UE_k and n_i is the zero-mean AWGN with variance σ_k^2 at UE_k.

In the previous subsection, SISO-NOMA allocates different power levels based on the users' channel conditions. Similarly, in MISO-NOMA, beamforming vectors are also designed based on channel conditions. To simplify the system, it is assumed that the channel strengths are ordered as

$$\|\mathbf{h}_1\|^2 \geq \|\mathbf{h}_2\|^2 \geq \cdots \geq \|\mathbf{h}_K\|^2. \tag{2.18}$$

Thus, as channel of UE_k is stronger than that of UE_{k+1} to UE_K, UE_k efficiently decode and subtract their signals. Consequently, the interference from UE_{k+1} to UE_K can be eliminated. Hence, the received signal after SIC at UE_k is represented by [53]

$$\hat{y}_k = \mathbf{h}_k^H \mathbf{w}_k s_k + \mathbf{h}_k^H \left(\sum_{i=1}^{k-1} \mathbf{w}_i s_i \right) + n_k, \quad \forall k \in \mathcal{K}. \tag{2.19}$$

Note that in (2.19), the first term is the signal for interest, the second term is the sum-

mation of interference from the stronger users, and the third term is the noise of UE_k. Therefore, the SINR and achievable rate are expressed respectively as

$$\mathrm{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{i=1}^{k-1} \mathbf{h}_k^H \mathbf{w}_i + \sigma_i^2}, \tag{2.20}$$

and

$$R_k = \log_2 (1 + \mathrm{SINR}_k) = \log_2 \left(1 + \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{i=1}^{k-1} \mathbf{h}_k^H \mathbf{w}_i + \sigma_i^2} \right), \quad \forall k \in \mathcal{K}. \tag{2.21}$$

2.1.4 Advantages of NOMA

NOMA represents a significant advancement in wireless communication technology, providing a wide range of advantages. The key benefits are listed as follows:

- Bandwidth efficiency: In OMA scheme, each user is allocated to a unique spectrum, such that in OFDMA, the available frequency resource blocks are divided orthogonally, each of which is assigned to a single user. With the increase in the number of users or the demands of QoS, underutilization is caused, leading to low spectrum efficiency and throughput of the overall system. Conversely, NOMA enables multiple users to share the same frequency and time resource by exploiting different power levels, which removes the concept of orthogonality in OMA. This simultaneously access scheme substantially improves the spectrum efficiency, as it allows for more efficient use of the available bandwidth. In additional, since the interference of weaker users can be eliminated by SIC at receivers, NOMA is able to enhance the throughput of the network. Therefore, NOMA leads to superior bandwidth efficiency and offers a higher data rate compared to conventional OMA.

- Fairness: In OMA, to optimize individual user experiences, users with higher channel conditions are typically granted with higher priority for accessing the spectrum [54], while users with poorer channel conditions are forced to wait. However, all the users in NOMA share the whole spectrum concurrently regardless of their channel conditions. Moreover, NOMA allocate more power to users with weaker channel gains. In this manner, NOMA achieves higher overall throughput while promoting fairness among users.

- Compatibility: NOMA is also regarded as an add-one technique [23], which can be integrated with current or future communication systems, including MISO and OMA system. Without additional modifications to the existing technique frameworks, NOMA can be combined with other networks and offer an extra degree of freedom [44].

- Massive connectivity: In OMA, each user is assigned with a private resource block, which cannot be shared by others. Hence, the number of connected users is considerably limited in practical implementation. However, since NOMA network is able to serve multiple users by allowing them to use the same resource block, it can offer massive connectivity, especially for IoT era [55].

- Latency: In OMA, particularly in TDMA, users with weaker channel gains are obliged to wait when spectrum is occupied by other users, leading to the dilemma to achieve the requirements in URLLC in future communication [12,56]. In contrast, multiple users can be simultaneously served in the same resource block in NOMA network. Thus, the latency can be significantly reduced.

2.2 Fundamentals of IRS

In this section, the basic concept, applications and advantages of IRS are introduced. IRS is a transformative technique to achieve intelligent and reconfigurable wireless communication [25–27, 57]. Generally, an IRS consists of a large number of elements, each of which is controlled to dynamically change the reflecting coefficients [58]. Through this adjustment, IRS is able to induce controllable amplitude and/or phase shift to incident signals. In other words, the wireless channels between transmitters and receivers in an IRS-aided communication system can be reconfigured to reach the desired QoS of users. Compared with traditional relay system, an IRS is a passive device, which consumes very little power and has lower complexity since it only requires passive reflective elements. On the other hand, traditional relay system actively receives, amplifies, and retransmits the signal, which consumes higher power and introduces more complex signal processing [38].

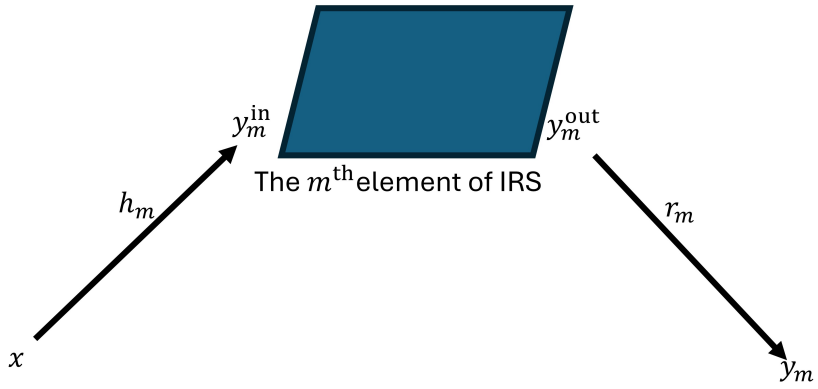


Figure 2.6: The Propagation of a Signal through the n^{th} of IRS Element.

2.2.1 IRS Signal Model

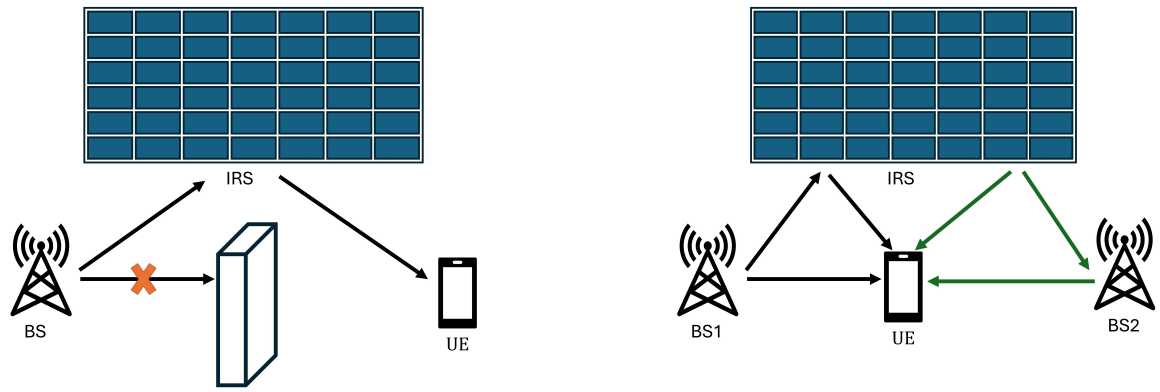
To present the effect of IRS mathematically, a simple point-to-point communication is considered, as shown in Fig. 2.6. To reduce complexity, it is assumed that the transmitter is equipped with N antennas, the receiver is equipped with a single antenna, and the IRS consists of M passive reflecting elements. Denote the equivalent complex-valued transmitted signal by $\mathbf{x} \in \mathbb{C}^{N\times 1}$, the complex coefficient for channel from the transmitter to the m^{th} IRS element by $\mathbf{h}_m \in \mathbb{C}^{N\times 1}$, where $m \in \mathcal{M}, \quad \mathcal{M} = \{1, 2, \cdots, M\}$. Thus, the incident signal to the m^{th} IRS elements is defines as

$$y_m^{\text{in}} = \mathbf{h}_m^H \mathbf{x}. \tag{2.22}$$

Since IRS adds amplitude and phase shift to incoming signals, the output signal from the m^{th} element is expressed as [26, 59]

$$y_m^{\text{out}} = y_m^{\text{in}} \beta_m e^{j\theta_m} = \mathbf{h}_m^H \beta_m e^{j\theta_m} \mathbf{x}, \tag{2.23}$$

where β_m and $\theta_m \in [0, 2\pi]$ stand for the amplitude and phase shift induced by element m, respectively. Hence, by ignoring the noise at the receiver, the received reflected signal



(a) Applying IRS to serve blocked user. (b) Applying IRS to serve cell-edge user.

Figure 2.7: Utilizing IRS to Assist Wireless Communication.

through the m^{th} element is given by

$$y_m = r_m y_m^{\text{out}} = \mathbf{h}_m^H \beta_m e^{j\theta_m} r_m^* \mathbf{x}. \tag{2.24}$$

As shown in (2.24), the corresponding signal consists of three terms, namely, channel from the transmitter to the IRS element, reflection coefficient, and channel from the IRS element to the receiver. Assuming all M IRS elements are able to reflect incident signals independently, the final received signal from the M elements IRS can be regarded as a superposition of their individual signal $y_m, \quad \forall m \in \mathcal{M}$, which is given by

$$y = \left(\sum_{m=1}^{M} \mathbf{h}_m^H \beta_m e^{j\theta_m} r_m^* \right) \mathbf{x} = \mathbf{r}^H \mathbf{\Theta} \mathbf{H} \mathbf{x}, \tag{2.25}$$

where $\mathbf{r} = [r_1, r_2, \cdots, r_M]^T$, $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_M]^H$, and $\mathbf{\Theta} = \text{diag}\left(\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, \cdots, \beta_M e^{j\theta_M}\right)$. It is worth mentioning that the effect of IRS can be considered as multiplexing a $M \times M$ matrix to incident signals. The matrix $\mathbf{\Theta}$ in (2.25) is diagonal due to the assumption that all M elements are independent to each other.

2.2.2 Advantages of IRS

IRS has drawn much attention as a groundbreaking technique due to its numerous advantages, which are listed as follows [57, 60]:

- Coverage: IRS can enhance signal strength and coverage significantly. Particularly, by coordinating the amplitude and phase shift of reflected signals, IRS is able to

regenerate a stronger and more reliable communication link. For instance, as shown in Fig. 2.7a, IRS can built a cascade channel to avoid the obstacle in the line of sight (LoS) link between the BS and users. For another scenario shown in Fig. 2.7b, by increasing the signal strength, IRS can greatly boost the QoS for users at the cell edge.

- Energy efficiency: By adjusting each element to optimize the propagation environment, wireless network with assistant of IRS allows users to consume less power than that without IRS to reach the same QoS level. In addition, IRS operates passively without active power resources, which reduce energy consumption of the whole system. Unlike conventional relay system, which decodes the received signals, regenerates and re-transmits them, without self interference, IRS is able to achieve the same effect in terms of improvement of signal strength with less operation and power costs [25, 38].

- Flexibility and compatibility: Not only benefiting the wireless system theoretically, but IRS also enjoys advantages in practical implementation. In particular, due to its light weight and low profile, IRS can be easily deployed on/ removed from environment, such as walls of buildings. Furthermore, IRS can be integrated straightly into other communication system.

2.3 IRS-aided NOMA

Given that both IRS and NOMA can improve QoS of a wireless network, the integration of these two techniques are considered in this section. IRS-aided NOMA has potential to tackle the inherent limitation of NOMA, particularly in interference management [61,62]. Specifically, although SIC can remove part of interference, as shown in (2.19), users still suffers from the interference of stronger users. This interference becomes a significant issue especially with the increase of user density, which substantially impact the overall system performance. However, assisted by IRS, there is opportunity to reconfigure the NOMA communication link, which can further increase the freedom and reduce the interference [63].

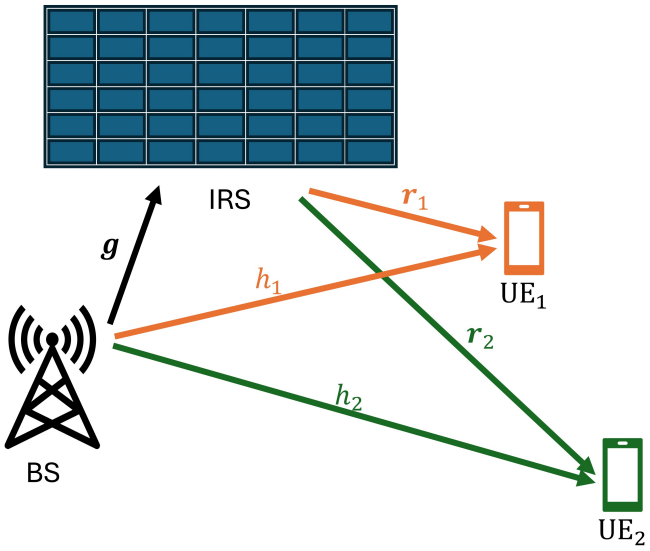


Figure 2.8: A Two-user IRS-aided NOMA System.

2.3.1 A IRS-aided SISO-NOMA Scenario

This subsection introduces an IRS-aided SISO-NOMA network to mathematically explain the its fundamental concepts. Fig. 2.8 shows IRS-aided NOMA system, where a single-antenna BS serves two single-antenna users, denoted by $\text{UE}_i, \quad i \in \{1, 2\}$. To exploit SC, the BS assigned different power, p_i, to two signals, $s_i, \quad i \in \{1, 2\}$. Then the superposed signal is given by

$$x = \sqrt{p_1} s_1 + \sqrt{p_2} s_2. \tag{2.26}$$

With the assistant of the M-element IRS, the wireless link between BS and users can be reconfigured by adding amplitude and phase shift matrix $\mathbf{\Theta} \in \mathbb{C}^{M \times M}$ similar to (2.25). Denote channel coefficient from the BS to UE_i by h_i and denote channel vector from BS to IRS and from IRS to UE_i respectively by $\mathbf{g} \in \mathbb{C}^{M \times 1}$ and $\mathbf{r}_i \in \mathbb{C}^{M \times 1}$. Then, the received signal is expressed as

$$y_i = \left(h_i + \mathbf{r}_i^H \mathbf{\Theta} \mathbf{g}\right) \left(\sqrt{p_1} s_1 + \sqrt{p_2} s_2\right) + n_i, \quad i \in \{1, 2\}, \tag{2.27}$$

where n_i is AWGN with zero mean and variance σ_i^2 at UE_i. Assuming that UE_1 has stronger channel condition, according to SIC, UE_1 can decode UE_2's signal and remove

the interference of it. Thus, the received signals of two users after SIC are given by

$$\hat{y}_1 = \left(h_1 + \mathbf{r}_1^H \mathbf{\Theta} \mathbf{g}\right) \sqrt{p_1} s_1) + n_1, \tag{2.28}$$

$$\hat{y}_2 = \left(h_2 + \mathbf{r}_2^H \mathbf{\Theta} \mathbf{g}\right) (\sqrt{p_1} s_1 + \sqrt{p_2} s_2) + n_2. \tag{2.29}$$

Thus, SINRs and achievable rates for UE_1 and UE_2 are defined respectively as

$$\mathrm{SINR}_1 = \frac{|h_1 + \mathbf{r}_1^H \mathbf{\Theta} \mathbf{g}|^2 p_1}{\sigma_1^2}, \tag{2.30}$$

$$\mathrm{SINR}_2 = \frac{|h_2 + \mathbf{r}_2^H \mathbf{\Theta} \mathbf{g}|^2 p_2}{|h_2 + \mathbf{r}_2^H \mathbf{\Theta} \mathbf{g}|^2 p_1 + \sigma_2^2}. \tag{2.31}$$

ans

$$R_1 = \log_2 \left(1 + \frac{|h_1 + \mathbf{r}_1^H \mathbf{\Theta} \mathbf{g}|^2 p_1}{\sigma_1^2}\right), \tag{2.32}$$

$$R_2 = \log_2 \left(1 + \frac{|h_2 + \mathbf{r}_2^H \mathbf{\Theta} \mathbf{g}|^2 p_2}{|h_2 + \mathbf{r}_2^H \mathbf{\Theta} \mathbf{g}|^2 p_1 + \sigma_2^2}\right). \tag{2.33}$$

It is worth mentioning that though interference has been removed by SIC in (2.30), UE_2 still suffers from the interference from UE_1, as shown in (2.31). Particularly, this interference consists of combined channel coefficient for UE_2 and power assigned for UE_1. Therefore, by jointly optimizing the amplitude, phase shift and power allocation, the interference can be optimized in order to achieve higher SINR, leading to higher throughput for overall system.

2.3.2 A IRS-aided MISO-NOMA Scenario

In this subsection, an IRS is further applied in a MISO-NOMA network. As shown in Fig. 2.9, with the assistant of a M elements IRS, the BS equipped with N antennas serves two single-antenna user, UE_1 and UE_2. Let $\mathbf{w}_1$ and $\mathbf{w}_2$ denote the beamforming vectors assigned by the BS to UE_1 and UE_2, respectively. Thus, the transmitted signal at the BS is given by

$$\mathbf{x} = \mathbf{w}_1 s_1 + \mathbf{w}_2 s_2, \tag{2.34}$$

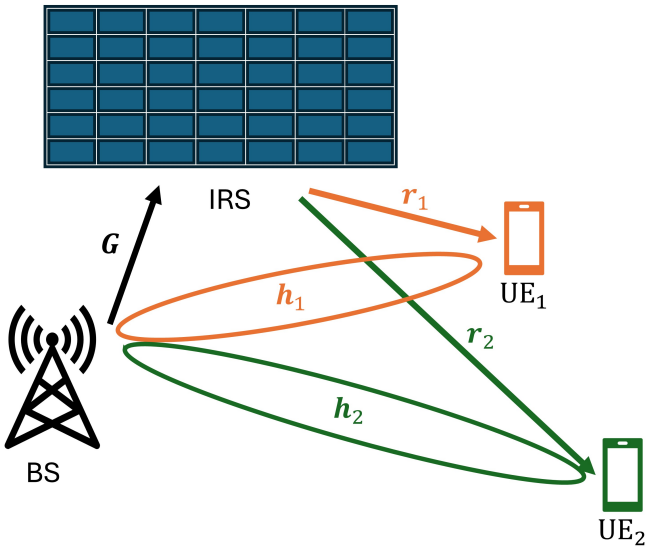


Figure 2.9: A Two-user IRS-aided MISO-NOMA System.

where s_1 and s_2 respectively indicate the signal for UE_1 and UE_2. Similarly, the assistant of IRS can be considered as adding amplitude and phase shift matrix $\Theta \in \mathbb{C}^{M\times M}$. Denote the channel coefficient vectors from the BS to UE_i by $\mathbf{h}_i \in \mathbb{C}^N$ and respectively denote channels from the BS to IRS and from IRS to UE_i by $\mathbf{G} \in \mathbb{C}^{M\times 1}$ and $\mathbf{r}_i \in \mathbb{C}^{M\times 1}$, where $i \in \{1,2\}$. Then, the received signal can be given by

$$y_i = \left(\mathbf{h}_i^H + \mathbf{r}_i^H \mathbf{\Theta} \mathbf{G}\right)(\mathbf{w}_1 s_1 + \mathbf{w}_2 s_2) + n_i, \quad i \in \{1,2\}, \tag{2.35}$$

where n_i is AWGN with zero mean and variance σ_i^2 at UE_i. It is worth mentioning that since NOMA decoding order is based on instantaneous random channel gains of multiple users, the channel estimation is crucial for IRS-aided NOMA system. In particular, the channel estimation for IRS is achieved through a process involving pilot signals, and efficient feedback mechanisms. Assuming UE_1 has much greater channel strength than UE_2, UE_1 is able to decode both signals intended for two users and eliminate the interference. Hence, after SIC, the received signal for UE_1 and UE_2 can be respectively represented as

$$\hat{y}_1 = \left(\mathbf{h}_1^H + \mathbf{r}_1^H \mathbf{\Theta} \mathbf{G}\right)\mathbf{w}_1 s_1 + n_1, \tag{2.36}$$

$$\hat{y}_2 = \left(\mathbf{h}_2^H + \mathbf{r}_2^H \mathbf{\Theta} \mathbf{G}\right)(\mathbf{w}_1 s_1 + \mathbf{w}_2 s_2) + n_2 \tag{2.37}$$

Thus, SINRs and achievable rates for UE_1 and UE_2 are respectively defined as

$$\mathrm{SINR}_1 = \frac{\left|\left(\mathbf{h}_1^H + \mathbf{r}_1^H \boldsymbol{\Theta} \mathbf{G}\right) \mathbf{w}_1\right|^2}{\sigma_1^2}, \tag{2.38}$$

$$\mathrm{SINR}_2 = \frac{\left|\left(\mathbf{h}_2^H + \mathbf{r}_2^H \boldsymbol{\Theta} \mathbf{G}\right) \mathbf{w}_2\right|^2}{\left|\left(\mathbf{h}_2^H + \mathbf{r}_2^H \boldsymbol{\Theta} \mathbf{G}\right) \mathbf{w}_1\right|^2 + \sigma_2^2}. \tag{2.39}$$

and

$$R_1 = \log_2(1 + \frac{\left|\left(\mathbf{h}_1^H + \mathbf{r}_1^H \boldsymbol{\Theta} \mathbf{G}\right) \mathbf{w}_1\right|^2}{\sigma_1^2}), \tag{2.40}$$

$$R_2 = \log_2(1 + \frac{\left|\left(\mathbf{h}_2^H + \mathbf{r}_2^H \boldsymbol{\Theta} \mathbf{G}\right) \mathbf{w}_2\right|^2}{\left|\left(\mathbf{h}_2^H + \mathbf{r}_2^H \boldsymbol{\Theta} \mathbf{G}\right) \mathbf{w}_1\right|^2 + \sigma_2^2}). \tag{2.41}$$

This integration of multiple antennas, NOMA and IRS provides a powerful solution for increasing spectral efficiency, improving signal coverage, mitigating interference, and reducing energy consumption, making it an ideal choice for addressing the demands of 6G and beyond.

2.4 Literature Review

In this section, related literature on resource allocation techniques is discussed from three perspectives: NOMA network, IRS network, and IRS-aided NOMA network. This structured review allows for a comprehensive understanding of both individual technologies and their combined potential. In additional, this review also focuses on resources allocation problems and robust cases of this IRS-aided NOMA networks.

2.4.1 NOMA network

The authors of [9, 23, 24] propose the fundamental principle and present wide range of theoretical benefits of NOMA. The comparison between NOMA and classical OMA is provided in [64,65], revealing the superior performance of NOMA in system capacity and power consumption.

Furthermore, with high compatibility, NOMA has been integrated with a number of technologies. The concept of combination of NOMA and multiple antennas technique is first proposed in [66]. Then different works on beamforming-based MISO-

NOMA [32, 67–70] and MIMO-NOMA [33, 71, 72] schemes are proposed. In particular, the authors in [69] provides a beamforming design to minimize total transmit power under individual achievable rate constraint for a multi-user MISO-NOMA network. The power minimization problems are also addressed in [73]. Note that both these papers regard the assumption of perfect CSI as a premise. However, it is difficult for a BS to obtain perfect CSI, the problem is extended to a robust beamforming design with bounded uncertainty in [74], where the non-convex problem is transformed into a convex SDP problem. Besides power consumption, achievable sum rate maximization problem is considered in [75], where the authors proposed a algorithm to tackle the problem for downlink MISO-NOMA network.

In addition, the integrate of OMA and NOMA has also grabbed great interest, which is investigated in [76–79]. Particularly, in this OMA-NOMA system, the available resource block (i.e. time or frequency) is divided orthogonally into several sub-resource blocks, within each of which NOMA scheme is exploited [79–81]. For example, an energy efficiency maximization problem for a hybrid OFDMA-NOMA system is considered in [34], where the total bandwidth is divided into multiple subcarriers and users are categorized in to multiple clusters. Each cluster occupy a orthogonal subcarrier to avoid inter-cluster interference and users in a cluster transmit their signals based on NOMA to remove inter-user interference. Similarly, the authors in [35] address a minimal SINR maximization problem for a TDMA-NOMA network. In this network, the whole time resource is split into time slots, users are clustered, and only one cluster is allowed to transmit using NOMA at each time slot.

Moreover, the cluster-based MISO/MIMO-NOMA networks have been analyzed to release the complexity of SIC in [82–84]. For instance, the authors in [84] investigate an overall system capacity maximization problem for MIMO-NOMA system, where users are clustered based on their channel strength. A linear beamforming technique is proposed for antennas at the receivers to remove inter-cluster interference, and power allocation is provided for users to maximize the sum data rate. Another work about cluster-based MISO-NOMA in [82] proposes a clustering algorithm which two users with higher correlation and larger channel gain difference are grouped and each cluster is assigned with a subcarrier. Beamforming vectors are designed to maximize the sum capacity.

Besides, resource allocation plays a critical role in the performance of NOMA net-

works, since it directly impacts both spectral efficiency and user fairness [9, 23, 24, 85]. The authors in [86] explore the power allocation for multi-cell multi-carrier NOMA network, where a power minimization problem is solved under achievable rate constraints for individual devices. In [87], a resource allocation problem aiming to maximize spectral efficiency while to ensure the QoS of each user is formulated and solved. Generally, due to the non-convexity of the formulated resource allocation problems, most existed works tend to transport the original ones into convex problems or apply ML to solve it. For instance, a energy efficiency maximization problem is introduced in [88], where the original non-convex problem is transported and iteratively solved using approximated convex approach. On the other hand, authors in [89] proposes a deep learning-based approach to minimize power consumption for multi-carrier NOMA.

2.4.2 IRS-aided network

The basic concept of IRS is offered in [25, 26]. To illustrate the enhancement of IRS to a wireless network, the studies in [38, 90] compare IRS with decode-and-forward relay system in terms of spectral efficiency. Benefit from multiple advantages, IRS is investigated in several related works to assist different networks. For example, research papers on IRS-aided multi-user MISO system are provided in [39, 91–94]. Specifically, the authors in [39] propose a design to maximize spectral efficiency, which is able to optimize beamforming vectors and IRS phase shift jointly. The authors formulate a energy efficiency maximization problem for multi-user IRS-aided MISO network to reach the balance of achievable data rate and total power consumption, in [93], where the problem is divided into two sub-problems, namely, beamforming and phase shift optimization, and solved iteratively using AO method. Another resource allocation for multi-IRS assisted multi-cluster IoT network is formulated as a sum achievable data rate maximization problem in [94]. To address this problem, the quadratic transformation and the AO algorithm are applied to optimize optimal downlink and uplink IRS phase shifts. In stead of conventional optimization algorithm, such as convex optimization or AO, multiple works on resource allocation for IRS-aided networks are achieved by ML-based algorithm. For instance, the authors in [95] utilize deep learning to optimize amplitude and phase shift of IRS in an IRS-aided MIMO network. The work in [96] investigates a DRL-based algorithm to solve sum rate maximization problem for IRS-aided OFDMA-based system.

MIMO is also integrated with IRS in [40,97]. Specifically, IRS is combined with millimeter wave MIMO in [40], where a least mean square-based channel estimator is provided to markedly improve the performance of beamforming.

Additionally, since obtaining perfect CSI for the BS is challenging due to the passive nature of the IRS, which cannot transmit or receive pilot symbols. This makes CSI estimation in an IRS-aided system difficult. Therefore, several works in [98, 99] have introduced methods to estimate the cascaded channel, which is defined as the product of the BS-IRS and IRS-user channels. Based on this cascade channel estimation, robust resource allocation problems for IRS-aided network are investigated in [100–102]. Specifically, a solution to the robust outage-probability-based power minimization problem is proposed in [100], where the original problem is iteratively optimize beamforming vectors and phase shift matrix using convex optimization algorithm. In [101], the authors propose similar AO methods to jointly optimize time allocation, beamforming vectors and IRS reflecting coefficients achieving minimal power consumption for a robust case in IRS-aided wireless powered communication network while ensuring the QoS for individual user. In [102], a more efficient resource allocation method for IRS-aided MIMO communication is provided. In this work, the imperfect CSI is considered and a DRL-based sequential scheduling algorithm is proposed to maximize the sum data rate.

2.4.3 IRS-aided NOMA network

Except conventional OMA, several IRS-aided NOMA wireless networks have been considered in [61,62,103–107]. In particular, in [103], an IRS is implemented to assist NOMA network enhancing the cell-edge users' QoS, where only reflected communication link by IRS is considered. There are additional research works on resource allocation problems for IRS-aided NOMA networks. Specifically, the authors in [61] formulate a achievable sum rate maximization problem for an IRS-aided multi-subcarrier NOMA system, where power allocation, phase shifts design, and decoding order determination is proposed. In [107], a power minimization problem is formulated for an IRS-aided SISO-NOMA network, which is solved by the sequential rotation algorithm. Furthermore, in [108] a study investigates two-user communication scenario assisted by IRS with discrete phase shifts, which includes NOMA, FDMA, and TDMA. The simulation results indicate that TDMA is able to perform better than FDMA due to the lack of frequency-selective IRS coeffi-

cients. While comparing TDMA with NOMA, NOMA may perform worse than TDMA when both served users are close to IRS, namely lower channel gains difference. With increase of this difference, NOMA can yield superior performance.

Similarly, due to the passive elements of IRS, there are also a number of works considering that the BS attends imperfect CSI. For example, the authors in [109] formulate a problem to maximize the energy efficiency for IRS-aided NOMA, which is divided into two sub-problems. Then, two sub-problems are iteratively transformed into convex forms and solved by AO algorithm. The work in [105] presents similar convex optimization and AO method to handle a robust design for IRS-aided NOMA network to maximize secrecy rate. Though the convex-optimization-based algorithm mentioned above can achieve accurate solutions, DRL is able to jointly optimize power allocation and phase shift more efficiently [110]. In particular, the authors in [106] introduce a DRL-based algorithm to address robust beamforming design for IRS-aided MISO-NOMA system, where only cascade channels are considered. Moreover, in [111] a long-term robust sum-rate maximization problem is considered, which is further reformulated into a DRL environment and solved using TD3 agent.

2.5 Summary

In this chapter, the fundamental concepts of NOMA are introduced, especially SC and SIC. To further explain NOMA principle mathematically, SISO-NOMA and MISO-NOMA scenarios are provided. Then, the advantages of NOMA is listed, including bandwidth efficiency, fairness, compatibility, massive connectivity, and latency. In addition, IRS fundamentals are reviewed, where a signal model present the impact of it. Besides, the benefits of IRS is offered, including coverage, energy efficiency, flexibility and compatibility. Furthermore, A IRS-aided SISO-NOMA and MISO-NOMA scenarios are considered to demonstrate the integration of IRS and NOMA. Finally, the related literature on resource allocation for NOMA, IRS, and IRS-aided NOMA networks are reviewed, respectively.

Chapter 3

Mathematical Backgrounds

In this chapter, examples of different resource allocation techniques for this thesis are presented in the first section. Then, to efficiently solve the optimization problem, two methods are introduced. In the first one, the problem is able to be transformed to a convex optimization problem and can be solved using Matlab directly. On the other hand, when the problem is difficult to mathematically transformed, ML-based algorithm is considered.

3.1 Resource Allocation Techniques

Over decades, resource allocation technique plays an necessary role for transmitters and receivers to achieve the QoS requirements of higher-speed data communication system [112, 113]. In particular, the resource of a wireless network, including frequency slots, time slots and power levels, are properly assigned to achieve different requirements and capacities of different users. In another word, the overall system performance is able to be enhanced by assigning these resources based on communication environment, channel condition, and users' requirements. The resource allocation techniques considered in this thesis are introduced briefly in the follows.

3.1.1 Power Minimization Technique

Power minimization, also known as transmit power control, is a vital problem for wireless networks [69, 114]. This technique addressed the trade-off between life of battery and radio spectrum of the network [115]. Particularly, the power control aims to control the

transmit power while satisfying a QoS requirement for each user. For instance, a typical power minimization problem is defined as

$$\begin{aligned} \min_{p_k} \quad & \sum_k p_k \\ \text{s.t.} \quad & \text{SINR}_k \geq \overline{\text{SINR}}_k, \quad \forall k, \end{aligned} \tag{3.1}$$

where p_k, SINR_k, and $\overline{\text{SINR}}_k$ are the allocated transmit power, SINR, and SINR threshold for the k^{th} user. This problem is to minimize the total transmit power consumption under individual SINR constraint for each user.

3.1.2 Sum Rate Maximization Technique

In addition, sum rate maximization is also a central objective to improve the efficiency and capacity of 6G or beyond wireless communications [116, 117]. Specifically, by optimizing the resource (i.e. power, bandwidth, time slots) allocation among multiple users, the network can handle higher data rate connections with limited power consumption. For example, a problem of sum rate maximization through power control is defines as [118]

$$\begin{aligned} \max \quad & \sum_k R_k \\ \text{s.t.} \quad & \sum_k p_k \leq p_{\max}, \\ & R_k \geq \overline{R}_k, \quad \forall k, \end{aligned} \tag{3.2}$$

where R_k and $\overline{R}_k$ are the achievable rate and the data rate threshold for the k^{th} user,respectively, and $p_{\max}$ is the available total transmit power.

It is worth mentioning that this data rate in above technique is based on Shannon capacity with nearly zero decoding error probability [119]. However, URLLC system exploits short packets communication to achieve lower latency and has to suffer from decoding error. Hence, Shannon's capacity theorem is not appropriate for URLLC network. Thus, the throughput of the k^{th} user transmitting a L length packet in URLLC system can be given by [120]

$$R_k^{'} = \sum_l \log_2(1 + \text{SINR}_{k,l}) - Q^{-1}(\epsilon)\sqrt{a^2\left(1 - \frac{1}{(1+\text{SINR}_{k,l})^2}\right)}, \tag{3.3}$$

where ϵ is the decoding error probability, a is a invariant, $a = \log_2 e$, and Q^{-1} is the inverse function of Q function. $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{t^2}{2})dt$. Thus, a sum throughput maximization problem of a URLLC system is defines as

$$\begin{aligned} \max \quad & \sum_k R_k^{'} \\ \text{s.t.} \quad & \sum_k p_k \leq p_{\max}, \\ & R_k^{'} \geq \overline{R}_k, \quad \forall k, \end{aligned} \tag{3.4}$$

3.1.3 AoI Minimization Technique

Furthermore, AoI is another important metric used to measure the freshness of information of a wireless network [121, 122]. Unlike conventional metric focusing on throughput or energy consumption, AoI quantifies the time elapsed since the most recent data update was generated at the source and received at the destination. In order ensure the information freshness of all the users, AoI minimization techniques aim to allocate the optimal resources to achieve minimal average sum AoI of the communication system. Denote AoI of the k^{th} user at time slots $t \in \{1, 2, \cdots, T\}$ by $\alpha_k(t)$. Then, a standard AoI minimization problem is formulated as

$$\begin{aligned} \min \quad & \frac{1}{T} \sum_k \sum_t \alpha_k(t) \\ \text{s.t.} \quad & \sum_k p_k \leq p_{\max}, \quad \forall k, \end{aligned} \tag{3.5}$$

where the objective is to minimize average AoI among T time slots with power limitation by optimizing various network parameters, including scheduling and power allocation.

3.2 Convex Optimization

By applying the above resources allocation techniques, the objective can be formulated as an optimization problem, which is either convex or non-convex. A non-convex problem can be mathematically transformed into a convex form, while a convex problem is able to be solved directly by convex optimization tools or software [123, 124]. The convex fundamentals and standard convex problems are introduced in the following subsections.

(a) A convex set. (b) A non-convex set.

Figure 3.1: Illustrations of convex and non-convex sets.

The basic principle of convex is explained from three aspects: convex sets, convex cones, and convex functions.

3.2.1 Convex Sets

A set is convex if, for any two points in this set, it contains the line segments between them [124]. Let $\mathcal{C}$ denotes a set and $x_1 \in \mathcal{C}$, $x_2 \in \mathcal{C}$ denote two points in this set. Then, the description of convex set can be expressed as

$$\theta x_1 + (1-\theta)x_2 \in \mathcal{C}, \quad \forall x_1, x_2 \in \mathcal{C}, \quad \forall \theta \in [0,1]. \tag{3.6}$$

Geometrically, Fig. 3.1a show an example of convex set, which has a solid body without holes. A set with inward curve is always non-convex, as shown in Fig. 3.1b, where the yellow segments is not involved in the set. In order to define a convex set, it is sometimes difficult to use the definition to prove convexity. Hence, some operations preserving the convexity of a set is mentioned in the following:

1) Intersection: If sets $\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_n$ are convex, then their intersection set $\mathcal{C}$ is also convex, where

$$\mathcal{C} = \mathcal{C}_1 \bigcap \mathcal{C}_2 \bigcap \cdots \bigcap \mathcal{C}_n. \tag{3.7}$$

2) Affine Transform: If set $\mathcal{C} \in \mathbb{C}^{n\times 1}$ is convex, matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ and vector $\mathbf{b} \in \mathbb{C}^{m\times 1}$,

then the affine transform set $\mathbf{A}\mathcal{C}+\mathbf{b} \in \mathbb{C}^{m\times 1}$ remain convex, where

$$\mathbf{A}\mathcal{C}+\mathbf{b} = \{\mathbf{A}\mathbf{x}+\mathbf{b} | \mathbf{x} \in \mathcal{C}\}. \tag{3.8}$$

3) Perspective Transform: If set $\mathcal{C} \in \mathbb{R}^{n-1} \times \mathbb{R}_+$ is convex, then the perspective transform $f(\mathcal{C})$ is convex, where $f : \mathbb{R}^n \to \mathbb{R}^{n-1}$, such that

$$f(\mathcal{C}) = \left[x_1/x_n, x_2/x_n, \cdots, x_{n-1}/x_n\right]^T. \tag{3.9}$$

x_i stands for element of $\mathcal{C}$ and $\mathbb{R}_+$ denotes the set of positive numbers.

3.2.2 Convex Cones

Another fundamental concept in convex optimization is convex cone, which is a special case of convex set. Essentially a set $\mathcal{S}$ is a convex cone, then for any $x_1, x_2 \in \mathcal{C}$, for all non-negative factor $\theta_1, \theta_2 \geq 0$, $\theta_1 x_1 + \theta_2 x_2$ still lies in $\mathcal{C}$. Convex cones are integral in various optimization problems for wireless communication, which is briefly introduced as follows [124]:

1) Second-order Cone: A second order cone in $\mathbb{R}^{n+1}$ is defines as

$$\mathcal{C} = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} | \|\mathbf{x}\| \leq t\}. \tag{3.10}$$

2) Semidefinite Cone: In the space of $n \times n$ symmetric matrices, denoted by $\mathbb{S}^n$, a semidefinite cone $\mathcal{C}$ is defined as

$$\mathcal{C} = \{\mathbf{X} \in \mathbb{S}_+^n | \mathbf{X} \succeq 0\}, \tag{3.11}$$

where $\mathbf{X} \succeq$ means that $\mathbf{X}$ is positive semidefinite, i.e., for all vectors $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z}^T\mathbf{X}\mathbf{z} \geq 0$.

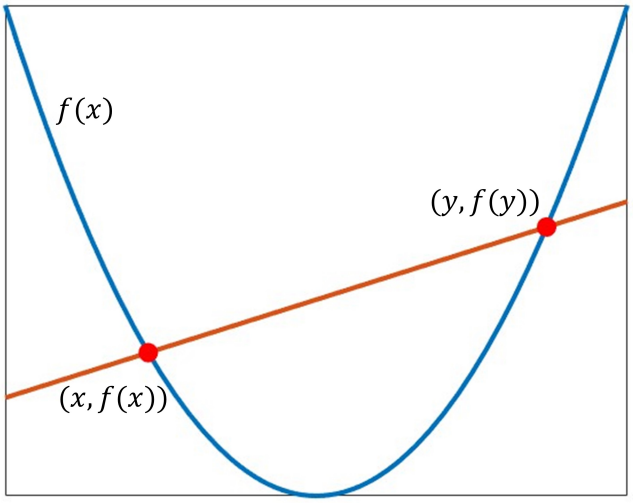


Figure 3.2: An illustration of convex function.

3.2.3 Convex Functions

Formally, a function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is convex if $\mathbf{dom}\, f$ is convex and if, $\forall \mathbf{x}, \mathbf{y} \in \mathbf{dom}\, f$ and any $\theta \in [0, 1]$, the following inequality holds [124]:

$$f\left(\theta \mathbf{x} + (1-\theta)\mathbf{y}\right) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}). \tag{3.12}$$

The right term of (3.12) stands for the line segment connecting $(\mathbf{x}, f(\mathbf{x}))$ and $(\mathbf{y}, f(\mathbf{y}))$, while the left term represents the function value when the variable sweeps from $\mathbf{x}$ to $\mathbf{y}$. Thus, geometrically, the figure of a convex function always lies under the line segment linking two points on it. The blue curve in Fig. 3.2 shows a convex function $f(x)$, which remains under the red line bridging two points on $f(x)$.

In addition, there are another two properties to check the convexity of function f with convex **dom**, which is listed as follows:

1) First Order Condition: If f is differentiable, then f is convex if and only if

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}). \tag{3.13}$$

2) Second Order Condition: If f is twice differentiable, then f is convex if and only if

$$\nabla^2 f(\mathbf{x}) \succeq 0. \tag{3.14}$$

3.2.4 Convex Optimization Problems

A convex optimization problem is shaped like following [124]:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \cdots, m, \\ & h_i(\mathbf{x}) = 0, \quad i = 1, 2, \cdots, p, \end{aligned} \tag{3.15}$$

where $\mathbf{x} \in \mathbb{R}^n$, $f, g_1, \cdots, g_m : \mathbb{R}^n \to \mathbb{R}$ are convex functions, $h_1, h_2, \cdots, h_p : \mathbb{R}^n \to \mathbb{R}$ are affine functions. $\mathbf{x}$ is called optimization variables, f is named as objective function, g_i, $i \in \{1, \cdots, m\}$ and h_i, $i \in \{1, \cdots, p\}$ are inequality and equality constraint functions, respectively. Normally, problem (3.15) seeks a optimal $\mathbf{x}$ to minimize $f(\mathbf{x})$ under all the inequality and equality constraints over a convex set, which is given by

$$\mathcal{D} = \mathbf{dom}\, f \bigcap_{i=1}^{m} \mathbf{dom}\, g_i \bigcap_{i=1}^{p} \mathbf{dom}\, h_i. \tag{3.16}$$

Furthermore, an optimality criterion of a differentiable objective function f is defined as follows [124]. Denote a feasible set by $\mathcal{X}$, where

$$\mathcal{X} = \{\mathbf{x} | g_i(\mathbf{x}) \leq 0, \, \forall i = 1, \cdots, m, \, h_i(\mathbf{x}) = 0, \, \forall i = 1, \cdots, p\}. \tag{3.17}$$

Then, $\mathbf{x}^*$ is optimal if and only if $\mathbf{x}^* \in \mathcal{X}$ and

$$\nabla f(\mathbf{x}^*)^T (\mathbf{y} - \mathbf{x}^*) \geq 0, \quad \forall \mathbf{y} \in \mathcal{X}. \tag{3.18}$$

According to above convex fundamentals, several standard convex optimization problems are introduced as follows:

1) Linear Programming:

In LP, objective function and constraint functions are all affine. Thus, a general LP

can be formulated as [124]

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} + d \\ \text{s.t.} \quad & \mathbf{G}\mathbf{x} \preceq \mathbf{h}, \\ & \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned} \tag{3.19}$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{G} \in \mathbb{R}^{m\times n}$, $\mathbf{A} \in \mathbb{R}^{p\times n}$, $\mathbf{h} \in \mathbb{R}^m$, and $\mathbf{b} \in \mathbb{R}^p$.

2) Quadratic Programming:

A convex optimization problem is called QP problem if its objective function is quadratic and the constraint functions are linear, which is given by [124]

$$\begin{aligned} \min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r \\ \text{s.t.} \quad & \mathbf{G}\mathbf{x} \preceq \mathbf{h}, \\ & \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned} \tag{3.20}$$

where $\mathbf{P} \in \mathbb{S}_+^n$ and $\mathbf{q} \in \mathbb{R}^n$. Note that LP is regarded as special case of QP when $\mathbf{P} = \mathbf{0}$.

Quadratic Constrained Quadratic Programming: By setting all the constraints and objective functions to be quadratic, the optimization problem is named QCQP problem, which is expressed as [124]

$$\begin{aligned} \min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{P}_0\mathbf{x} + \mathbf{q}_0^T\mathbf{x} + r_0 \\ \text{s.t.} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i^T\mathbf{x} + r_i \le 0, \quad i = 1, 2, \cdots, m, \\ & \mathbf{A}\mathbf{x} = b, \end{aligned} \tag{3.21}$$

where $\mathbf{P}_i \in \mathbb{S}_+^n, \quad i = 0, 1, \cdots, m$. By setting all $\mathbf{P}_i = \mathbf{0}$, QCQP can be formulated as a QP problem.

3) Second order Cone Programming:

A SOCP optimization problem is formulated as [124]

$$\min_{\mathbf{x}} \quad \mathbf{c}^T\mathbf{x}$$

$$\begin{aligned} \text{s.t.} \quad & \|\mathbf{A}_i\mathbf{x}+\mathbf{b}_i\| \leq \mathbf{c}_i^T\mathbf{x}+d_i, \quad i=1,\cdots,m, \\ & \mathbf{F}\mathbf{x}=\mathbf{g}, \end{aligned} \tag{3.22}$$

where $\mathbf{A}_i \in \mathbb{R}^{m\times n}$, $\mathbf{F} \in \mathbb{R}^{p\times n}$. The first constraint of problem in(3.22) is called second order cone constraint.

4) Semidefinite Programming:

With semidefinite matrices $\mathbf{G}, \mathbf{F}_1, \cdots, \mathbf{F}_n \in \mathbb{S}^n$, a SDP problem is written as [124]

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & \mathbf{F}_1x_1+\mathbf{F}_2x_2+\cdots+\mathbf{F}_kx_n+\mathbf{G} \preceq \mathbf{0}, \\ & \mathbf{A}\mathbf{x}=b, \end{aligned} \tag{3.23}$$

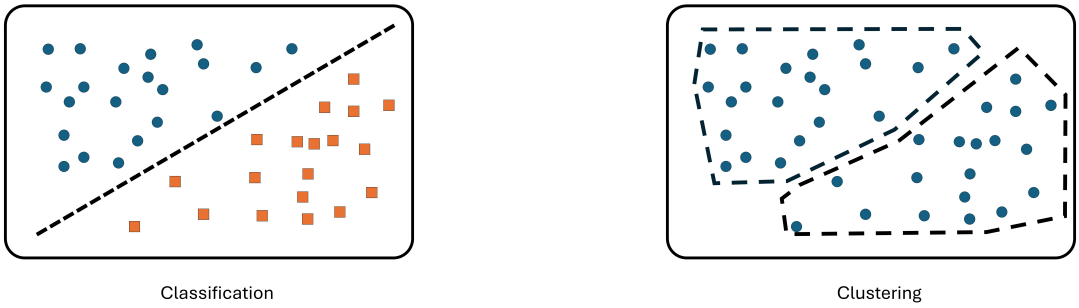
where $\mathbf{x}=\{x_1,x_2,\cdots,x_n\}^T$.

Another type of SDP can be formulated as linear equality constraints and a non-negativity constraint on the variable in the form of a matrix, such that [124],

$$\begin{aligned} \min_{\mathbf{X}} \quad & \mathbf{tr}(\mathbf{C}\mathbf{X}) \\ \text{s.t.} \quad & \mathbf{tr}(\mathbf{A}_i\mathbf{X})=b_i, \quad i=1,\cdots,m, \\ & \mathbf{X} \succeq \mathbf{0}, \end{aligned} \tag{3.24}$$

where $\mathbf{X}, \mathbf{C}, \mathbf{A}_1, \cdots \mathbf{A}_m \in \mathbb{S}^n$.

For a conventional resource allocation problem, after transforming the objective functions and constraints into convex functions and sets, the original one can be cast into a standard convex problem, such as SDP and SOCP. These type of problems can be efficiently solved using tool boxes, (i.e. CVX [125] and Yalmip [126]). Furthermore, it is guaranteed that every local optimal solution is also global optimal [124]. Thus, it is beneficial, efficient and low-complexity to address an optimization problem by transforming it into a convex form.



(a) An example of supervised learning. (b) An example of unsupervised learning.

Figure 3.3: An illustration of supervised learning and unsupervised learning.

3.3 Machine Learning

ML is an important subset of AI technique, focusing on the study of statistical algorithm which enables computers to learn from data and improve overtime performance without being explicitly programmed for each specific task [127–129]. On contrary, convex optimization techniques are well-suited for solving structured problems. Though convex-based algorithm benefits from its mathematical rigor, its still suffers from inflexible. For instance, in a convex optimization problem in (3.15), each time the parameters in constraint function g_i, h_i are changed, the problem has to be solved again. In particular, the SINR_k in the constraint in(3.1) is a function of allocated power levels and channel conditions of all the users, indicating that the problem needs to be solved multiple times for a time-varying channel model. However, a well trained ML algorithm is able to handle the optimization problem even when channel conditions differ with time, which significantly reduce the complexity. Furthermore, ML-based algorithm is also effective dealing with resource allocation problems (i.e., problems with discrete variables) that cannot or difficult to be rewritten as a convex form. Given the above benefits, ML will be the new trend to address the channel of future 6G and beyond networks [130].

ML can be broadly classified into three categories based on the nature of the learning process, namely, supervised learning, unsupervised learning, RL, DRL [131], which are introduced in the following subsections.

3.3.1 Supervised Learning

Supervised learning algorithm is trained on labeled dataset, where each input in the dataset is paired with a corresponding output dataset, providing the algorithm with correct behavior [128,129]. This framework is trained with labeled mapped input-output dataset and aims to predict correct output for future unseen input. Supervised learning is usually used in tasks of classification, where the machine identifies discrete objects, as shown in Fig. 3.3a. Supervised learning offers high accuracy and reliability when trained on sufficient and diverse labeled data. However, it is limited by well-labeled data, since it is usually complex and time-consuming to obtain.

3.3.2 Unsupervised Learning

On contrast, unsupervised learning trains algorithm with unlabeled dataset. In particular, by analyzing the inherent characteristics of the data, unsupervised learning seek a primary structure to identify the similarities and relationship without predefined output labels [128, 129]. Fig.3.3b shows a common task of unsupervised learning, namely clustering, where similar data points are grouped together based on their features. Without relying on labeled data, unsupervised learning is more data-efficient, while it offer less accuracy as lacking labeled data for guidance.

3.3.3 Reinforcement Learning

RL is a branch of ML concerned with how agents ought to take actions in an environment to maximize cumulative reward [132,133]. Particularly, unlike the above supervised learning and unsupervised learning, RL algorithm enables an agent to interact with the environment and receive feedback in the terms of rewards based on the actions taken. By analyzing the feedback, agent can discover an optimal policy which maximizes the accumulated reward over time. Therefore, RL is highly effective for dynamic and sequential decision-making tasks, such as robotics and game playing. ON the downside, RL requires significant computational resources and training time for the agent to explore and learn optimal strategies. The interaction cycle is described in Fig. 3.4. At time step (t), the agent takes an action $a^{(t)}$ according to a given state $s^{(t)}$. The environment is impacted by the taken action, and returns the reward $r^{(t)}$ and state for next time step $s^{(t+1)}$.

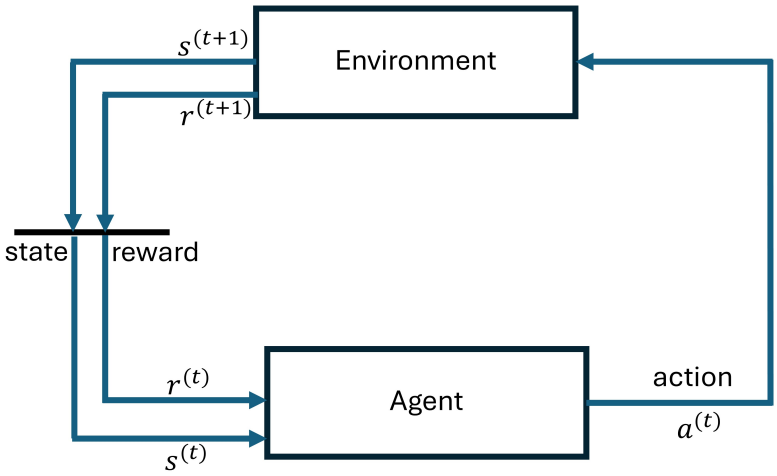


Figure 3.4: An illustration of RL cycle.

Generally, based on RL algorithm, an optimization problem can be formulated as a Markov decision process (MDP), which consists of a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P})$ [134]. Each component is defined as follows:

- $\mathcal{S}$: State space denotes by $\mathcal{S}$, representing all possible state within the environment. Each state illustrates the corresponding information for the agent to take an action.
- $\mathcal{A}$: Action space, denoted by $\mathcal{A}$, contains all possible action the agent can take, each of which can affect the environment.
- $\mathcal{R}$: The reward function, denoted by $\mathcal{R}$, defines the immediate reward value after taking a action. Particularly, the reward function maps state-action pair to a real number, indicating the desirability of the action in the state.
- $\mathcal{P}$: The transition probability defines the probability of transforming from one state to another given specific action.

To further illustrate the principle of RL, another simplified RL problem, named multi-arm bandit (MAB) is considered. The agent selects between multiple actions (arms), each of which leads to a random reward produced by a specific distribution, in order to maximize the summation of reward. In this framework, the goal can be achieved through the balance between exploration (trying out different arms to learn their reward distributions) and exploitation (choosing the best-known arm to maximize reward) [135].

Generally, there are two types of RL, namely, on-policy and off-policy RL. In particular, on-policy RL algorithms learn the value of the policy they are currently following. The agent's exploration and learning are based on the same policy, making the learning process directly influenced by the agent's current actions. On the other hand, off-policy RL algorithms learn from data generated by a different policy. In this manner, agent explores using one policy while learning the value of another.

In addition, RL can be classified based on learning strategies into policy-based and value-based RL. Particularly, policy-based RL agent directly learns a policy function that maps states to actions, which is beneficial for continuous action spaces and complex environments where directly learning the policy leads to more effective exploration. Value-based RL focuses on learning a value function, which estimates the long-term reward of taking specific actions from a given state. The value-based agent uses this value function to obtain an optimal policy by selecting the action with the highest estimated value.

One of the most powerful value-based RL algorithm to tackle optimization problems is called Q-learning, in which, the agent learn a optimal policy based on Q-values. In particular, for each time step (t) in an episode of infinite MDP, the agent observes the current state takes an action based on its policy π. Afterwards, the agent receives reward and moves to the next state. By keeping the trade-off between exploration and exploitation, the RL agent learns to maximize the long-term accumulative discounted reward, denoted by episode return $g^{(t)}$, which is defines as [132]

$$g^{(t)} = r^{(t+1)} + \gamma r^{(t+2)} + \gamma^2 r^{(t+3)} + \cdots = \sum_{k=0}^{\infty} \gamma^k r^{(t+k+1)}, \tag{3.25}$$

where $\gamma \in [0, 1]$ is a discount parameter indicating the priority of future reward. Formally, the agent takes actions according to the policy $\pi(a|s)$, which maps from states to probabilities of selecting each action. To validate a policy, the expected return the agent will achieve by taking action a in state s and following this policy π thereafter, called Q-value, which is defined as [132]

$$q_\pi(s^{(t)}, a^{(t)}) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k r^{(t+k+1)} | s = s^{(t)}, a = a^{(t)} \right]. \tag{3.26}$$

Thus, a well trained agent takes action with maximal return from each state. Using Q-values, the optimal policy can be expressed as selecting the action with the highest

	a_1	a_2
s_1	$Q(s_1, a_1)$	$Q(s_1, a_2)$
s_2	$Q(s_2, a_1)$	$Q(s_2, a_2)$
$\cdots$	$\cdots$	$\cdots$
s_N	$Q(s_N, a_1)$	$Q(s_N, a_2)$

Table 3.1: An example of Q-table

Q-value for each state, which is expressed as

$$\pi_*(s^{(t)}) = \arg\max_{a^{(t)} \in \mathcal{A}} Q(s^{(t)}, a^{(t)}). \tag{3.27}$$

In classical Q-learning, there is a q-table, which contains all state-action pairs. For instance, a q-table for agent in a model with N stats and two actions for each state is shown in Table. 3.1. Specifically, at a state, the agent will take action leading to a higher Q-value based on the table.

In the training process of a Q-value-based agent, the policy is iteratively improved through interactions with its environment. The process starts with initializing a Q-table. Then, within multiple training episodes, the agent takes actions based on ϵ-greedy policy. Particularly, the agent takes a random action with probability ϵ to explore, and takes an action according to Q-table with probability $1 - \epsilon$, where $\epsilon \in [0, 1]$. As the agent takes actions, it observes the resulting rewards and transitions to new states. The Q-values are updated using the Bellman equation, which is expressed as [132]

$$Q(s^{(t)}s^{(t)}) \leftarrow r^(t) + \gamma \max_{a^{(t)} \in \mathcal{A}} Q(s^{(t+1)}, a^{(t)}). \tag{3.28}$$

This iterative process continues until the Q-values converge to their true values, indicating that the agent has learn an optimal policy.

Note that the Q-learning mentioned is a category of model-free RL algorithms, which do not rely on a model of the environment's dynamics to learn optimal policies. Unlike model-based RL, which requires a representation of how actions affect the environment, model-free RL directly learns from the agent through its interactions with the environment. This makes model-free RL more flexible and applicable to a wider range of problems, especially resource allocation problems for wireless communication where the environment's dynamics are complex and difficult to model accurately. Key model-free

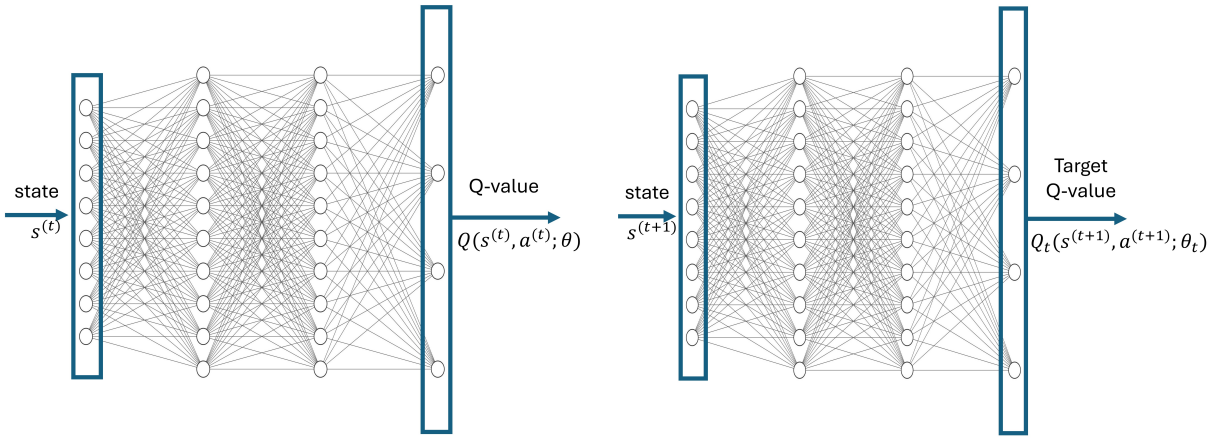
algorithms include Q-learning, deep Q-network (DQN), and Policy Gradient methods such as TD3 and DDPG, which are introduced in the next subsection.

3.3.4 Deep Learning

Deep learning is a subset of ML using deep neural network (DNN)s with many layers to model and solve complex problems. Inspired by the structure and function of the human brain, DNNs consist of more than one layers, which allows deep learning algorithm learns the relationship between the input and output layer. Deep learning algorithms are designed to automatically learn intricate patterns and relationships in data, making them particularly powerful for large-scale data and high-dimensional inputs [136]. The deep learning algorithm largely benefit from the availability of large datasets and the development of more efficient approximate function. Thus, motivated by the advantages of deep learning, RL can directly process complex inputs like images or sound. The integration of RL and deep learning is particularly beneficial for optimization problems including high-dimensional observations, where agents must interact with complex environment.

3.3.5 Deep Reinforcement Learning

DRL is an advanced approach that combines the principles of RL and neural network (NN) to enable agents to make decisions in more complex, higher-dimensional environments [134]. Unlike conventional RL, DRL applies neural networks to approximate the policy or Q-value functions, allowing the agent to generalize from past experiences and make decisions even in previously unseen states. For instance, a DRL framework, known as DQN, replaces the Q-table of Q-learning by NNs to generally estimate Q-values with similar states and actions. By doing this, DQN agent is able to handle optimization problems with continuous state space. However, DQN algorithm quickly becomes impractical since the action space is still discrete. Thus, another NN-based structure is integrated with DQN to extend the ability of agent to handle continuous action space, which is know as actor-critic DRL. Since in the most resource allocation problems, channel condition is considered as state and power allocation is considered as action, both state and action space are continuous. Thus, actor-critic DRL is more efficient for wireless communication. Both DQN and actor-critic DRL are introduced as follows:



(a) Main network of DQN. (b) Target Network of DQN.

Figure 3.5: An illustration of network structure of DQN agent, including Q-value and target networks.

- Deep Q-Network: The main structure of DQN agent are given in Fig. 3.5. DQN is firstly proposed in [137], where the authors introduce a replay buffer, a main network and a target network to traditional Q-learning. Both networks are used to estimate Q-values with θ and θ_t as the weights of NNs, respectively. During the training, the stats of interaction between the agent with the environment are recorded in the replay buffer, which can be described by a tuple: $\left[s^{(t)}, a^{(t)}, r^{(t)}, s^{(t+1)}\right]$. Then, the agent randomly select a mini batches with multiple tuples from the replay buffer and input them to the two NNs. In the main network, as shown in Fig. 3.5a, with current state $s^{(t)}$ inputted, the output is the estimated Q-values for all possible action, which is expressed by

$$Q(s^{(t)}, a^{(t)}; \theta). \tag{3.29}$$

To further validate the performance of the main value, the target network estimates the Q-value for given state for next step $s^{(t+1)}$, as shown in Fig. 3.5b, and formulates a target value, which is given by [134]

$$y^{(t)} = r^{(t)} + \gamma \arg\max_{a^{(t)} \in \mathcal{A}} Q(s^{(t+1)}, a^{(t+1)}; \theta_t). \tag{3.30}$$

To train the main network and update its weight, the agent minimize the loss between the actual Q-value generated by the main network and the target Q-value

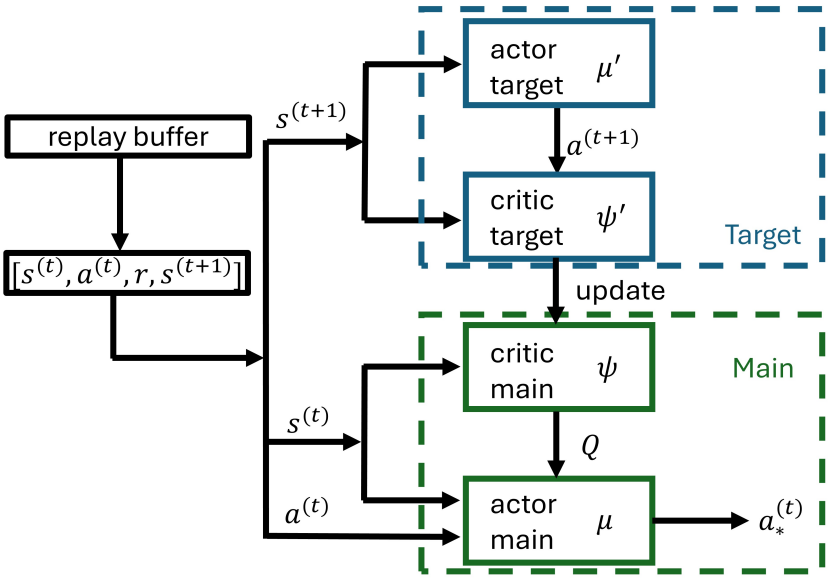


Figure 3.6: An illustration of DDPG-based DRL model.

generated by the target network. The loss function is expressed as [134]

$$L(\theta) = \left(y^{(t)} - Q(s^{(t)}, a^{(t)}; \theta)\right)^2. \tag{3.31}$$

Note that the weight of target network is a copy of that of the main network, but to make the target Q-value more stable, it is frozen for a number of steps, giving the main network opportunity to catch up.

- DDPG-based Algorithm: It is worth mentioning that both NN in DQN model return Q-values of all the possible actions. For instance, in the Fig. 3.5, there are four nodes in the output layer, indicating that agent can only take four actions at each state. Due to limitation of NN structure, it is impossible for DQN agent handle optimization problems with continuous action space. Thus, extra NNs are attached giving rise to a more powerful DRL algorithm, namely actor critic. Specifically, actor critic Q-value based DRL, such as DDPG or TD3, consists of one actor and at least one critic, each of which contains one main network [134, 138].

  In particular, DDPG-based DRL framework is considered and the training process is illustrated as follows. The actor NN, denoted by μ is responsible for determining the optimal actions in a given state based on the current policy and the critic NN, denoted by ψ, evaluates the actions taken by the actor by estimating the Q-value

function, which are expressed as

$$\mu(s^{(t)}, \theta_\mu) = a_*^t, \tag{3.32}$$

and

$$\psi(s^{(t)}, a^{(t)}; \theta_\psi) = Q_*, \tag{3.33}$$

where Q_* indicates the optimal Q-value and θ_i, $i = [\mu, \psi]$ stands for the parameters for the NN. In particular, the Fig. 3.6 shows the framework of a DDPG agent, where target networks as copies of both actor and critic networks generate the target values. Similar to DQN agent, a tuple is sampled from the replay buffer, such as $[s^{(t)}, a^{(t)}, r,^{(t)}, s^{(t+1)}]$. Let μ' and ψ' denote by target actor and critic, respectively and θ_i, $i = [\mu'\psi']$ denote the parameter for target network. Then, the target for a sampled tuple is given by [134]

$$y^{(t)} = r^{(t)} + \gamma\psi' \left(s^{(t+1)}, \mu'(s^{(t+1);\theta_{\mu'}}); \theta_{\psi'}\right). \tag{3.34}$$

Note that instead of calculating the Q-values for all the possible actions for the DQN agent in (3.30), actor network provides directly the optimal Q-values by target critic network. Hence, actor critic is able to deal continuous action space. The training of the main critic network is based on the following loss function [134]:

$$L(\theta_\psi) = \left(\psi(s^{(t)}, a^{(t)}; \theta_\psi) - y^{(t)}\right)^2. \tag{3.35}$$

Furthermore, the actor is trained by maximizing the Q-value provided by main critic network, which is expressed by

$$\max_{\theta_\mu} \quad \psi\left(s^{(t)}, \mu(s^{(t)}; \theta_\mu)\right). \tag{3.36}$$

As for the training for target networks, the parameters, $\theta_{\mu'}$ and $\theta_{\psi'}$, copy from learned actor and critic at each time step with lower learning rate, which can be

presented by [134]

$$\theta_i = \lambda\theta_i + (1-\lambda)\theta_i, \quad i = \mu', \psi', \tag{3.37}$$

where λ is the learning rate. To ensure the stability of training, target networks are trained with limited rate, such that $\lambda \ll 1$.

3.4 Summary

In this chapter, the required mathematical backgrounds for this thesis are presented. Firstly, the resource allocation technique are introduced, including power minimization, sum rate maximization and AoI minimization techniques, which can formulate optimization problems for resource allocations. Next, two algorithms, namely convex optimization and ML, are discussed. In convex optimization section, definitions and properties of convex set, cone and function are provided first. Then, multiple types of convex optimization problems are proposed. On the other hand, supervised learning, unsupervised learning, and RL are introduced in ML section. Subsequently, the Q-value based DRL is considered, focusing on DQN and actor critic DDPG.

Chapter 4

Robust Downlink Beamforming Design for an IRS-aided NOMA System

4.1 Introduction

Beamforming design plays a critical role in in enhancing the performance and efficiency of wireless networks. By focusing the transmission of signals in specific directions, beamforming allows for more efficient use of spectrum and power, improving the signal quality for intended users while reducing interference to others [139]. In particular, effective beamforming design not only enhances overall network capacity but also contributes to improved user experience and energy efficiency. Most of works on beamforming design for IRS-aided NOMA networks consider the perfect CSI. However, in practice, obtaining perfect CSI is usually unrealistic due to various limitations and challenges in wireless communication systems, especially in IRS-aided networks, since the IRS only reflect signals passively. Hence, motivated by this, we develop a robust beamforming design for MISO IRS-aided NOMA network. . In particular, channel uncertainties are taken into account in the beamforming design, while satisfying the data rate constraints on the system. This can be achieved by developing a robust optimization framework to design the transmit beamformers at the BS and the phase shift at the IRS. With the above background, this is the first work to investigate the robust beamforming design problem based on the imperfect cascaded channels through IRS. The main contibution can be sumarrized as

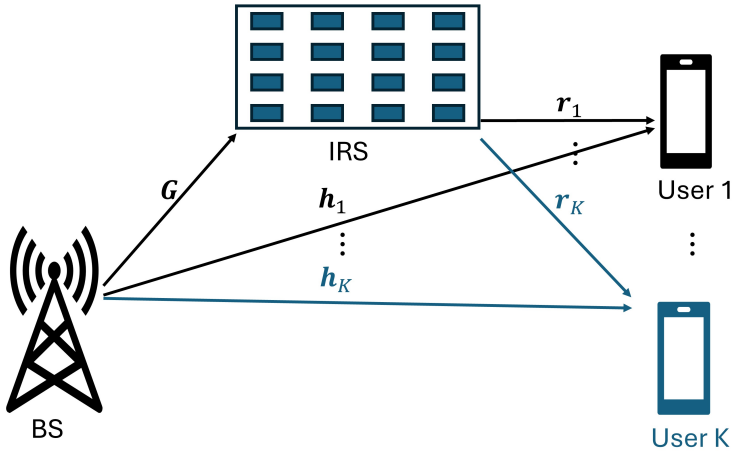


Figure 4.1: A multi-user MISO IRS-aided NOMA system.

follows:

- By taking into account the statistical channel uncertainty of the IRS-aided downlink NOMA system, we formulate a robust beamforming design, aiming to minimize the total transmit power under outage-probability-based data rate constraints, on the served users.

- Different mathematical matrix transformations are used to convert the original robust non-convex problem into a convex problem. In particular, a well-known, Bernstein-type inequality is exploited to deal with the outage probability constraints. Additionally, an AO-based iterative algorithm is developed to solve this problem efficiently.

- We provide several simulation results to demonstrate the performance of the proposed robust design, which reveal the superior performance of the proposed NOMA robust beamforming in terms of less power consumption and convergence speed, while meeting the required outage probability constraints. .

4.2 System Model

Generally, there are two types of robust beamforming design. The first one is worst-case robust beamforming design with bounded CSI error model, where the BS knows the radii of the uncertainty region. This design aims to optimize the beamformers such that it performs adequately under the worst possible channel realization within this bound. The second one is outage-constrained robust beamforming design with statistical CSI error model, where each CSI error vector is assumed to follow a specific distribution. Instead of focusing on the worst case, the design allows for a certain probability of failure, which is defined as outage-probability. In this work, we consider a downlink transmission of an IRS-aided NOMA communication system, where the BS with N_{tx} antennas transmits signals to K single-antenna users, $\mathcal{K} = \{1, 2, \cdots, K\}$, as shown in Fig.5.1. This transmission is assisted by the IRS that comprises M reflecting elements $\mathcal{M} = \{1, 2, \cdots, M\}$. We assume that users can receive signals via two links: direct links between BS and users and the links through IRS. Let the k^{th} user equipment denote by $\text{UE}_k, \quad \forall k \in \mathcal{K}$. Then, the channel coefficients between the BS and the UE_k are denoted by $\mathbf{h}_k \in \mathbb{C}^{N_{tx} \times 1}$, and the reflecting channel between the BS and IRS, the IRS and UE_k denoted $\mathbf{G} \in \mathbb{C}^{M \times N_{tx}}$ and $\mathbf{r}_k \in \mathbb{C}^{M \times 1}$, respectively. All the channel coefficients are modelled with large-scale and small-scale fading components. Taking $\mathbf{h}_k$ as an example, the channel coefficient can be expressed as $\mathbf{h}_k = \sqrt{\text{PL}_0 d_k^{-\alpha}} \tilde{\mathbf{h}}_k$ [140]. PL_0 is the path loss at a unit distance, d_k stands for the distance of the communication links from the BS to the user k, α denotes the path loss exponent, and $\tilde{\mathbf{h}}$ follows as Rayleigh fading distribution with zero mean and unit variance as $\tilde{\mathbf{h}} \in \mathcal{CN}(0, \mathbf{I})$, representing zero-mean CSCG distribution with unit variance, where $\mathbf{I}$ is an identity matrix. The phase shift matrix of IRS is defined as $\mathbf{\Theta} = \text{diag}(\beta_1 e_1, \beta_2 e_2, \cdots, \beta_M e_M)$, where β_m and e_m denote the amplitude and phase shift coefficient for the m^{th} element of IRS, respectively. Without loss of generality, we assume that the amplitude phase shift fixed, such that $\beta_m = 1, \forall m$. Therefore, the received signal at the UE_k is expressed as

$$y_k = (\mathbf{h}_k^H + \mathbf{r}_k^H \mathbf{\Theta} \mathbf{G}) \sum_{i=1}^{K} \mathbf{w}_i s_i + n_i, \forall k \in \mathcal{K}, \tag{4.1}$$

where $\mathbf{w}_k$ represents the beamforming vector, s_k denotes the transmit data symbol, and $n_k \sim \mathcal{CN}(0, \sigma^2)$ represents the zero mean circularly symmetric complex AWGN with variance σ^2. Assuming the power of data symbol for UE_k is $\mathbb{E}\{|s_k|^2\} = 1$, the total transmit power can be defined as

$$p_{\text{tx}} = \sum_{k=1}^{K} \|\mathbf{w}_k\|^2, \forall k \in \mathcal{K}. \tag{4.2}$$

For the sake of brevity we define $\mathbf{H}_k = \text{diag}(\mathbf{r}_k)\mathbf{G}$ and set $\mathbf{e} = [e_1, e_2, \cdots, e_M]^H$ to be the phase shift vector of reflecting coefficients where $e_m = e^{j\theta_m}, \forall m$ and θ_m is the phase shift of the m^{th} reflecting coefficient. Then, the received signal in (4.1) can be rewritten as follows

$$y_k = (\mathbf{h}_k^H + \mathbf{e}^H \mathbf{H}_k) \sum_{k=1}^{K} \mathbf{w}_k s_k + n_k, \quad \forall k \in \mathcal{K}. \tag{4.3}$$

We assume the direct channel between the BS and the users is perfect. However, we are considering a challenging scenario that involves uncertainty in the channel (imperfect CSI) due to the passive IRS elements.(i.e. The reflecting channel $\mathbf{H}_k$ is imperfect). Thus, the cascade channel can be expressed as follows

$$\mathbf{H}_k = \hat{\mathbf{H}}_k + \Delta\mathbf{H}_k, \qquad \forall k \in \mathcal{K}, \tag{4.4}$$

where $\hat{\mathbf{H}}_k$ is the estimated channel, and $\Delta\mathbf{H}_k$ is the channel uncertainties following CSCG distribution as

$$\text{vec}(\Delta\mathbf{H}_k) \sim \mathcal{CN}(0, \mathbf{\Sigma}_k), \quad \mathbf{\Sigma}_k \succeq 0, \forall k \in \mathcal{K}, \tag{4.5}$$

where $\mathbf{\Sigma}_k$ denotes the positive semidefinite covariance matrices of channel uncertainties [141]. Using (4.4), the received signal at UE_k can be reformulated as

$$y_k = (\mathbf{h}_k^H + \mathbf{e}^H(\hat{\mathbf{H}}_k + \Delta\mathbf{H}_k)) \sum_{k=1}^{K} \mathbf{w}_k s_k + n_k, \ \forall k \in \mathcal{K}. \tag{4.6}$$

Since SIC is employed at the receiver ends of a NOMA transmission, the decoding order is a crucial element of the overall system performance. SIC first decodes the stronger user's signal, subtracts it from the received signal, and then decodes the weaker user's

signal. This decoding order is essential because it impacts the interference levels experienced by each user and, consequently, their achievable data rates and overall system performance. Correctly determining the optimal decoding order can significantly enhance the efficiency and fairness of the NOMA system. It is assumed that the UEs are ordered based on their channel strengths, such that $\|\mathbf{h}_1\| \geq \|\mathbf{h}_2\| \geq \cdots \geq \|\mathbf{h}_K\|$. The channel strengths mentioned above only consider the direction link form BS to UE, since the channel strengths of cascade channels impact significantly less than direct communication links. For instance, the UE_k decodes the signals intended for UEs from UE_{k+1} to UE_K using SIC before decoding its signal and the signals intended for UEs from UE_1 to UE_{k-1} are regarded as interference. On the other hand, the signals intended for $\mathrm{UE}_{k+1}, \cdots, \mathrm{UE}_K$ cannot be removed completely by UE_k due to the channel uncertainties. Therefore, based on SIC, intended signal of the UE_k at the UE_l can be given as

$$\begin{aligned} y_{k,l} =& (\mathbf{h}_l^H + \mathbf{e}^H \mathbf{H}_l)\mathbf{w}_k s_k + \sum_{i=1}^{k-1} (\mathbf{h}_l^H + \mathbf{e}^H \mathbf{H}_l)\mathbf{w}_i s_i + \sum_{i=k+1}^{K} \mathbf{e}^H \Delta\mathbf{H}_l \mathbf{w}_i s_i + n_l, \\ & \forall k \in \mathcal{K}, \forall l \in \{1, \cdots, k\}, \end{aligned} \tag{4.7}$$

where $\mathbf{H}_l = \hat{\mathbf{H}}_l + \Delta\mathbf{H}_l$ and the first term denotes the desired signal for the UE_k, the second term is the interference introduced by signals for $\mathrm{UE}_1 \cdots \mathrm{UE}_{k-1}$, and the third term is the interference due to the channel uncertainty. Hence, the SINR of the signal intended for the UE_k at the UE_l is defined as [74]

$$\begin{aligned} \Gamma_{k,l} =& \frac{|(\mathbf{h}_l^H + \mathbf{e}^H \mathbf{H}_l)\mathbf{w}_k|^2}{\sum\limits_{i=1}^{k-1} |(\mathbf{h}_l^H + \mathbf{e}^H \mathbf{H}_l)\mathbf{w}_i|^2 + \sum\limits_{i=k+1}^{K} |\mathbf{e}^H \Delta\mathbf{H}_l \mathbf{w}_i|^2 + \sigma_l^2}, \\ & \forall k \in \mathcal{K}, \forall l \in \{1, \cdots, k\}, \end{aligned} \tag{4.8}$$

and the achievable rate of UE_k can be written as

$$R_k = \log_2(1 + \min\{\Gamma_{k,1}, \Gamma_{k,2}, \cdots, \Gamma_{k,k}\}), \quad \forall k \in \mathcal{K}. \tag{4.9}$$

4.3 Problem Formulation

We consider an outage probability-based robust design for the MISO IRS-aided NOMA system defined above. In particular, a robust power minimization problem is formulated under the phase shift unit-modulus and outage probability based rate constraints as follows:

$$\mathbf{P1} \qquad \min_{\mathbf{w}_k, \mathbf{e}} \quad \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \tag{4.10}$$

$$\text{s.t.} \quad \Pr\{R_k \geq \mu_k\} \geq 1 - \rho_k, \quad \forall k \in \mathcal{K}, \tag{4.11}$$

$$|[\mathbf{e}]_m| = 1, \quad \forall m \in \mathcal{M}, \tag{4.12}$$

where μ_k denotes the required data target rate for the UE_k user. Due to the uncertainty of CSI, it is impossible to determine a practical joint design of beamforming vectors and IRS phase shift to meet the required QoS of users. Therefore, (4.11) guarantees that UEs will meet their data rate goals with an outage probability of ρ_k. Additionally, the constraint in (4.12) ensures that the IRS elements only modify phase of the signal without changing its strength or attenuating its amplitude, thereby preserving the overall integrity of the signal. However, it can be seen that the outage probability constraints for the SINR of users and the unit-modulus constraints for IRS elements are non-convex. This means that problem **P1** cannot be directly solved using convex optimization techniques. To solve the original problem **P1**, it must be transformed into a convex form. In the following sections, we present approach to solve the original non-convex problem.

4.4 Proposed Methodology and Algorithm

In this section, we first use approximation techniques to transform the non-convex outage probability constraints into convex ones. Next, we propose an alternative iterative optimization algorithm to determine the beamforming vectors and phase shifts of the formulated robust power minimization problem in **P1** while satisfying the outage probability and unit-modulus phase shift constraints.

4.4.1 Problem Transformation

Due to the non-convex outage probability constraint in (4.11), the robust problem cannot be solved directly using standard convex optimization techniques. Hence, we exploit Bernstein-type inequality to overcome this non-convexity issue [100].

Lemma 1: Berstein-Type Inequality: Assume $f(\mathbf{x}) = \mathbf{x}^H\mathbf{U}\mathbf{x} + 2\Re\{\mathbf{u}^H\mathbf{x}\} + u$, where $\mathbf{U} \in \mathbb{H}^{n\times n}$, $\mathbf{u} \in \mathbb{C}^{n\times 1}$, $u \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{C}^{n\times 1} \sim \mathcal{CN}(0, \mathbf{I})$, where $\mathbf{I}$ is an identity matrix. Then, for any $\rho \in [0, 1]$, the following approximations hold:

$$\Pr\{\mathbf{x}^H\mathbf{U}\mathbf{x} + 2\Re\{\mathbf{u}^H\mathbf{x}\} + u \geq 0\} \geq 1 - \rho \tag{4.13}$$

$$\Rightarrow \mathrm{Tr}\{\mathbf{U}\} - \sqrt{2\ln(1/\rho)}x + \ln(\rho)\lambda^{+}_{\max}(-\mathbf{U}) + u \geq 0 \tag{4.14}$$

$$\Rightarrow \begin{cases} \mathrm{Tr}\{\mathbf{U}\} - \sqrt{2\ln(1/\rho)}\nu + \ln(\rho)\psi + u \geq 0 \\ \sqrt{\|\mathbf{U}\|_F^2 + 2\|\mathbf{u}\|^2} \leq \nu \\ \psi\mathbf{I} + \mathbf{U} \succeq \mathbf{0}, \quad \psi \geq 0, \end{cases} \tag{4.15}$$

where ν and ψ are slack variables and $\lambda^{+}_{max}(-\mathbf{U}) = \max\{\lambda_{\max}(-\mathbf{U}), 0\}$. The proof of Lemma 1 is given in [100].

Based on Lemma 1, the outage probability constraint of UE_k in (4.11) can be rewritten as

$$\Pr\{R_k \geq \mu_k\} \geq 1 - \rho_k, \forall k \in \mathcal{K}$$
$$\Rightarrow \begin{cases} \Pr\{\log_2(1 + \Gamma_{k,1}) \geq \mu_k\} \geq 1 - \rho_k, \\ \Pr\{\log_2(1 + \Gamma_{k,2}) \geq \mu_k\} \geq 1 - \rho_k, \\ \qquad\qquad \cdots \\ \Pr\{\log_2(1 + \Gamma_{k,k}) \geq \mu_k\} \geq 1 - \rho_k, \end{cases} \quad \forall k \in \mathcal{K} \tag{4.16}$$

by substituting $\Gamma_{k,l}$ in (4.8) in (4.16), the outage probability of use UE_k can be reformulated as (4.17). In (a), the absolute value of the product for two vectors is expanded by exploiting the following transform, $|\mathbf{z}_1^H\mathbf{z}_2| = \mathbf{z}_1^H\mathbf{z}_2\mathbf{z}_2^H\mathbf{z}_1$, where $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{C}^{N_{tx}\times 1}$. Additionally, we define $\mathbf{A}_k = \dfrac{\mathbf{w}_k\mathbf{w}_k^H}{2^{\mu_k} - 1} - \sum_{i=1, i\neq k}^{K}(\mathbf{w}_i\mathbf{w}_i^H)$ and $\mathbf{B}_k = \dfrac{\mathbf{w}_k\mathbf{w}_k^H}{2^{\mu_k} - 1} - \sum_{i=1}^{k-1}(\mathbf{w}_i\mathbf{w}_i^H)$. Finally, in (b), we substitute $\mathbf{A}_k$, $\mathbf{B}_k$ and apply the transform $\mathbf{z}_1^H\mathbf{z}_3\mathbf{z}_3^H\mathbf{z}_2 + \mathbf{z}_2^H\mathbf{z}_3\mathbf{z}_3^H\mathbf{z}_1 = 2\Re\{\mathbf{z}_1^H\mathbf{z}_3\mathbf{z}_3^H\mathbf{z}_2\}$, where $\mathbf{z}_3 \in \mathbb{C}^{N_{tx}\times 1}$.

For sack of simplicity, we define $\mathbf{E} = \mathbf{e}\mathbf{e}^H$ and applying mathematical matrices transfor-

$$
\begin{aligned}
&\Pr\left\{\log_2\left(1+\Gamma_{k,l}\right)\geq\mu_k\right\}\\
&=\Pr\left\{\frac{|(\mathbf{h}_l^H+\mathbf{e}^H\mathbf{H}_l)\mathbf{w}_k|^2}{\sum\limits_{i=1}^{k-1}|(\mathbf{h}_l^H+\mathbf{e}^H\mathbf{H}_l)\mathbf{w}_i|^2+\sum\limits_{i=l+1}^{K}|\mathbf{e}^H\Delta\mathbf{H}_l\mathbf{w}_i|^2\sigma_l^2}\geq 2^{\mu_k}-1\right\}\\
&\overset{(a)}{=}\Pr\Bigg\{(\mathbf{h}_l^H+\mathbf{e}^H\mathbf{H}_l)\left(\frac{\mathbf{w}_k\mathbf{w}_k^H}{2^{\mu_k}-1}-\sum_{i=1}^{k-1}\mathbf{w}_i\mathbf{w}_i^H\right)(\mathbf{h}_l+\mathbf{H}_l^H\mathbf{e})\\
&\qquad-\sum_{i=k+1}^{K}(\mathbf{e}^H\Delta\mathbf{H}_l\mathbf{w}_i\mathbf{w}_i^H\Delta\mathbf{H}_l^H\mathbf{e})-\sigma_l^2\geq 0\Bigg\}\\
&\overset{(b)}{=}\Pr\Bigg\{\mathbf{e}^H\Delta\mathbf{H}_l\mathbf{A}_k\Delta\mathbf{H}_l^H\mathbf{e}+2\Re\left\{\mathbf{e}^H\Delta\mathbf{H}_l\mathbf{B}_k(\mathbf{h}_l+\widehat{\mathbf{H}}_l^H\mathbf{e})\right\}\\
&\qquad+(\mathbf{h}_l^H+\mathbf{e}^H\widehat{\mathbf{H}}_l^H)\mathbf{B}_k(\mathbf{h}_l+\widehat{\mathbf{H}}_l^H\mathbf{e})-\sigma_l^2\geq 0\Bigg\} \qquad (4.17)\\
&\geq 1-\rho_k,
\end{aligned}
$$

---

mations $\mathrm{Tr}(\mathbf{Z_1Z_2Z_3Z_4})=\mathrm{vec}^T(\mathbf{Z}_4^T)(\mathbf{Z}_3^T\otimes\mathbf{Z}_1)\mathrm{vec}(\mathbf{Z}_2)$ and $\mathrm{Tr}(\mathbf{Z}_1^H\mathbf{Z}_2)=\mathrm{vec}^H(\mathbf{Z}_1)\mathrm{vec}(\mathbf{Z}_2)$, where $\mathbf{Z}_i\in\mathbb{C}$. Therefore, (4.17) can be rewritten as follows:

$$
\begin{aligned}
&\Pr\Bigg\{\mathrm{vec}^H(\Delta\mathbf{H}_l)(\mathbf{A}_k^H\otimes\mathbf{E})\mathrm{vec}(\Delta\mathbf{H}_l)\\
&+2\Re\left\{\mathrm{vec}^H\left((\mathbf{e}\mathbf{h}_l^H+\mathbf{E}\widehat{\mathbf{H}}_l)\mathbf{B}_k\right)\mathrm{vec}(\Delta\mathbf{H}_l)\right\}\\
&+(\mathbf{h}_l^H+\mathbf{e}^H\widehat{\mathbf{H}}_l)\mathbf{B}_k(\mathbf{h}_l+\widehat{\mathbf{H}}_l^H\mathbf{e})-\sigma_l^2\geq 0\Bigg\}\geq 1-\rho_k. \qquad (4.18)
\end{aligned}
$$

For the convenience of derivation, we rewrite the channel uncertainty in (4.4) into $\mathrm{vec}(\Delta\mathbf{H}_k)=\epsilon_k\mathbf{a}_k$ where $\epsilon_k^2\mathbf{I}=\mathbf{\Sigma}_k$ and $\mathbf{a}_k\sim\mathcal{CN}(0,\mathbf{I})$. Therefore, (4.18) can be represented as

$$
\Pr\left\{\mathbf{a}_l^H\mathbf{U}_{k,l}\mathbf{a}_l+2\Re\{\mathbf{u}_{k,l}^H\mathbf{a}_1\}+u_{k,l}\geq 0\right\}\geq 1-\rho_k, \qquad (4.19)
$$

where $\mathbf{U}_{k,l}=\epsilon_l^2\mathbf{A}_k\otimes\mathbf{E}$, $\mathbf{u}_{k,l}=\epsilon_l\mathrm{vec}\left((\mathbf{e}\mathbf{h}_l^H+\mathbf{E}\widehat{\mathbf{H}}_l^H)\mathbf{B}_k\right)$, and

$$
u_{k,l}=(\mathbf{h}_l^H+\mathbf{e}\widehat{\mathbf{H}}_l)\mathbf{B}_k(\mathbf{h}_l+\widehat{\mathbf{H}}_l^H\mathbf{e})-\sigma_l^2. \qquad (4.20)
$$

Then, the Bernstein-Type Inequality is utilized to approximate (4.19) as

$$\mathrm{Tr}\{\mathbf{U}_{k,l}\} - \sqrt{2\ln(1/\rho_k)}\nu_{k,l} + \ln(\rho_k)\psi_{k,l} + u_{k,l} \geq 0, \tag{4.21}$$

$$\sqrt{\|\mathbf{U}_{k,l}\|_F^2 + 2\|\mathbf{u}_{k,l}\|^2} \leq \nu_{k,l}, \tag{4.22}$$

$$\psi_{k,l}\mathbf{I} + \mathbf{U}_{k,l} \succeq 0, \qquad \psi_{k,l} \geq 0. \tag{4.23}$$

The constraints in (4.21) can be further simplified by applying the following transformations

$$\begin{aligned}\mathrm{Tr}\{\mathbf{U}_{k,l}\} &= \epsilon_l^2 \mathrm{Tr}\{\mathbf{A}_k \otimes \mathbf{E}\} = \epsilon_l^2 \mathrm{Tr}\{\mathbf{A}_k\}\mathrm{Tr}\{\mathbf{E}\} \\ &= \epsilon_l^2 M \mathrm{Tr}\{\mathbf{A}_k\}, \end{aligned} \tag{4.24}$$

$$\begin{aligned}\|\mathbf{U}_{k,l}\|_F^2 &= \epsilon_l^4 \|\mathbf{A}_k \otimes \mathbf{E}\|_F^2 = \epsilon_l^4 \|\mathbf{A}_k\|_F^2 \|\mathbf{E}\|_F^2 \\ &= \epsilon_l^4 M^2 \|\mathbf{A}_k\|_F^2, \end{aligned} \tag{4.25}$$

$$\begin{aligned}\|\mathbf{u}_{k,l}\|^2 &= \epsilon_l^2 \|\mathrm{vec}\left((\mathbf{e}\mathbf{h}_l^H + \mathbf{E}\widehat{\mathbf{H}}_l)\mathbf{B}_k\right)\|^2 \\ &= \epsilon_l^2 M \|\mathrm{vec}\left((\mathbf{h}_l^H + \mathbf{e}^H\widehat{\mathbf{H}}_l)\mathbf{B}_k\right)\|^2, \end{aligned} \tag{4.26}$$

$$\begin{aligned}\lambda_{\max}(\mathbf{U}_{k,l}) &= \epsilon_l^2 \lambda_{\max}(\mathbf{A}_k \otimes \mathbf{E}) = \epsilon_l^2 \lambda_{\max}(\mathbf{A}_k)\lambda_{\max}(\mathbf{E}) \\ &= \epsilon_1^2 M \lambda_{\max}(\mathbf{A}_k), \end{aligned} \tag{4.27}$$

With these approximated convex constraints, the original non-convex optimization problem in (4.10) can be reformulated as the following convex problem:

$$\mathbf{P2} \qquad \min_{\mathbf{w}_k, \mathbf{e}, \nu_{k,l}, \psi_{k,l}} \quad \sum_{k=1}^{K} \|\mathbf{w}_k\|^2 \tag{4.28}$$

$$\text{s.t.} \quad \epsilon_l^2 M \mathrm{Tr}\{\mathbf{A}_k\} - \sqrt{2\ln(1/\rho_k)}\nu_{k,l} - \ln(1/\rho_k)\psi_{k,l} + u_{k,l} \geq 0, \tag{4.29}$$

$$\left\| \begin{array}{c} \epsilon_l^2 M \mathrm{vec}(\mathbf{A}_k) \\ \sqrt{2M}\epsilon_l \mathbf{B}_k(\mathbf{h}_l + \widehat{\mathbf{H}}_l^H \mathbf{e}) \end{array} \right\| \leq \nu_{k,l}, \tag{4.30}$$

$$\psi_{k,l}\mathbf{I} + \epsilon_l^2 M \mathbf{A}_k \succeq \mathbf{0}, \tag{4.31}$$

$$\psi_{k,l} \geq 0, \quad \forall k \in \mathcal{K}, \forall l \in \{1, \cdots, k\}, \tag{4.32}$$

$$(4.12).$$

However, it is evident that the joint design of beamforming vectors and phase shift to

all users in the system adds complexity, making it more challenging to solve. Therefore, an alternative optimization method will be used to find the solution in the next subsection.

4.4.2 Altering Optimization-based Iterative Algorithm

In this subsection, we develop an AO framework to efficiently solve problem **P2**. In particular, we divide the problem into beamforming design and phase shift optimization. Both subproblems are solved alternately until the required convergence accuracy is achieved.

Beamforming Design

For a given set of phase shifts, $\mathbf{e}$, problem **P2** can be reformulated equivalently as the following SDP problem

$$
\begin{aligned}
\mathbf{P3} \qquad & \min_{\mathbf{W}_k, \nu_{k,l}, \psi_{k,l}} && \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_k\} && (4.33)\\
& \text{s.t.} && (4.29) \quad (4.30) \quad (4.31) \quad (4.32) \\
& && \mathbf{W}_k \succeq 0, \mathbf{W}_k = \mathbf{W}_k^H, && (4.34)\\
& && \mathrm{rank}(\mathbf{W}_k) = 1, \forall k \in \mathcal{K}, \forall l \in \{1, \cdots, k\}, && (4.35)
\end{aligned}
$$

where $\mathbf{W}_k = \mathbf{w}_k \mathbf{w}_k^H$. It is shown in [142] that there is always at least one rank rank-one solution for an SDP problem. Thus, in problem **P3**, we can relax the non-convex rank-one constraint in (4.35) by adopting semidefinite relaxation (SDR). In this way, the constraints' objective function is turned into convex in $\mathbf{W}_k$. Then, the problem **P3** can be solved efficiently and conveniently by multiple software packages, such as the CVX toolbox in Matlab [143]. By multiplexing square root of the maximal eigenvalue of the optimal matrix with the corresponding eigenvector, the beamforming vectors $\mathbf{w}_k$s can be extracted from the matrix $\mathbf{W}_k$.

Phase Shift Optimization

Next, we consider the optimization of phase shift of the IRS for a given set of beamformers. Since the transmit power only depends on the beamformers, the IRS phase shift vector

$\mathbf{e}$ does not affect the power value. Thus, the original optimization problem **P1** can be reformulated into a feasibility checking problem for a given phase shift vector. Then, for a given set of beamformers, problem **P1** is reformulated as

$$\begin{aligned} \mathbf{P4} \qquad & \underset{\mathbf{e},\nu_{k,l}}{\text{Find}} && \mathbf{e} \\ & \text{s.t.} && (4.18) \quad (4.12), \end{aligned} \tag{4.36}$$

where the objective is to find a feasible solution for $\mathbf{e}$ that satisfies all the constraint in **P1**.

In order to modify the problem **P4** into a convex problem, we introduce a slack variables $\vartheta_{k,l}, \quad \forall k \in \mathcal{K}, \quad \forall l \in \{1, \cdots, k\}$. Then, constraint in (4.18) can be rewritten as

$$\begin{aligned} \Pr\Big\{ & \text{vec}^H(\Delta\mathbf{H}_l)(\mathbf{A}_k^H \otimes \mathbf{E})\text{vec}(\Delta\mathbf{H}_l) + 2\Re\left\{\text{vec}^H\left((\mathbf{e}\mathbf{h}_l^H + \mathbf{E}\widehat{\mathbf{H}}_l)\mathbf{B}_k\right)\text{vec}(\Delta\mathbf{H}_l)\right\} \\ & + (\mathbf{h}_l^H + \mathbf{e}^H\widehat{\mathbf{H}}_l)\mathbf{B}_k(\mathbf{h}_l + \widehat{\mathbf{H}}_l^H\mathbf{e}) - \sigma_l^2 - \vartheta_{k,l} \geq 0 \Big\} \geq 1 - \rho_k. \end{aligned} \tag{4.37}$$

In the following, we utilize the result of Lemma 1 and (4.20) and define $u_{k,l}^{\mathbf{e}}$ as

$$\begin{aligned} u_{k,l}^{\mathbf{e}} &= (\mathbf{h}_l^H + \mathbf{e}^H\widehat{\mathbf{H}}_l^H)\mathbf{A}_k(\mathbf{h}_l + \widehat{\mathbf{H}}_l^H\mathbf{e}) - \vartheta_{k,l} - \sigma_l^2 \\ &= \mathbf{h}_l^H\mathbf{A}_k\mathbf{h}_l + \mathbf{e}^H\widehat{\mathbf{H}}_l^H\mathbf{A}_k\widehat{\mathbf{H}}_l^H\mathbf{e} + \mathbf{e}\widehat{\mathbf{H}}_l^H\mathbf{A}_k\mathbf{h}_l + \mathbf{h}_l^H\mathbf{A}_k\widehat{\mathbf{H}}_l^H\mathbf{e} - \vartheta_{k,l} - \sigma_l^2 \\ &= \mathbf{h}_l^H\mathbf{A}_k\mathbf{h}_l + \text{Tr}\{\mathbf{C}_{k,l}\mathbf{E}_{\text{ext}}\} - \vartheta_{k,l} - \sigma_l^2, \end{aligned} \tag{4.38}$$

where

$$\mathbf{E}_{\text{ext}} = \begin{bmatrix} \mathbf{e} \\ 1 \end{bmatrix}\begin{bmatrix} \mathbf{e}^H & 1 \end{bmatrix}, \quad \mathbf{C}_{k,l} = \begin{bmatrix} \widehat{\mathbf{H}}_l^H\mathbf{A}_k\widehat{\mathbf{H}}_l^H & \widehat{\mathbf{H}}_l\mathbf{A}_k\mathbf{h}_l \\ \mathbf{h}_l^H\mathbf{A}_k\widehat{\mathbf{H}}_l^H & 0 \end{bmatrix}. \tag{4.39}$$

Using the definition in (4.26), the norm of (4.38) can be written as

$$\begin{aligned} \|\mathbf{u}_{k,l}^{\mathbf{e}}\|^2 &= \epsilon_l^2 M\|\text{vec}\left((\mathbf{h}_l^H + \mathbf{e}^H\widehat{\mathbf{H}}_l)\mathbf{B}_k\right)\|^2 \\ &= \epsilon_l^2 M\left(\mathbf{h}_l^H\mathbf{B}_k\mathbf{B}_k^H\mathbf{h}_l + \text{Tr}\{\mathbf{D}_{k,l}\mathbf{E}_{\text{ext}}\}\right), \end{aligned} \tag{4.40}$$

where

$$\mathbf{D}_{k,l} = \begin{bmatrix} \widehat{\mathbf{H}}_l^H \mathbf{B}_k \mathbf{B}_k^H \widehat{\mathbf{H}}_l^H & \widehat{\mathbf{H}}_l \mathbf{B}_k \mathbf{B}_k^H \mathbf{h}_l \\ \mathbf{h}_l^H \mathbf{B}_k \mathbf{B}_k^H \widehat{\mathbf{H}}_l^H & 0 \end{bmatrix}. \tag{4.41}$$

Since in this subsection, we assume that the beamforming vector is fixed, and the variable $\psi_{k,l}$ is only related to beamformers. Thus the variables $\psi_{k,l}$ in problem **P2** can be reformulated as

$$\psi_{k,l} = \max(0, \lambda_{\max}(-\epsilon_l^2 M \mathbf{A}_k)), \tag{4.42}$$

and, constraint in (4.30) can be given by

$$\epsilon_l^4 M^2 \|\mathbf{A}_k\|_F^2 + 2M\epsilon_l^2(\mathbf{h}_l^H \mathbf{A}_k \mathbf{A}_k^H \mathbf{h}_l + \mathrm{Tr}\{\mathbf{D}_{k,l}\mathbf{E}_{\text{ext}}\}) \leq 2\Re\{\nu_{k,l}^{(n)} \nu_{k,l}\} - \nu_{k,l}^{(n)2}, \tag{4.43}$$

where $\nu^{(n)}$ is the optimal value in the n^{th} iteration. Therefore, based on the problem **P4**, the feasibility checking problem for phase shift is reformulated as

$$\begin{aligned}
\mathbf{P5} \quad \max_{\mathbf{E}_{\text{ext}}, \nu_{k,l}, \vartheta_{k,l}} \quad & \sum_{k=1}^{K} \sum_{l=1}^{k-1} \vartheta_{k,l} && (4.44)\\
\text{s.t.} \quad & \epsilon_l^2 M \mathrm{Tr}\{\mathbf{A}_k\} - \sqrt{2\ln(1/\rho_k)}\nu_{k,l} - \ln(1/\rho_k)\psi_{k,l} + u_{k,l}^{\mathbf{e}} \geq 0 && (4.45)\\
& \epsilon_l^4 M^2 \|\mathbf{A}_k\|_F^2 + 2M\epsilon_l^2(\mathbf{h}_l^H \mathbf{A}_k \mathbf{A}_k^H \mathbf{h}_l + \mathrm{Tr}\{\mathbf{D}_{k,l}\mathbf{E}_{\text{ext}}\}) \\
& \quad \leq 2\Re\{\nu_{k,l}^{(n)} \nu_{k,l}\} - \nu_{k,l}^{(n)2}, \quad \forall k \in \mathcal{K}, \forall l \in \{1, \cdots, k\}, && (4.46)\\
& \mathbf{E}_{\text{ext}} \succeq \mathbf{0}, \quad \mathbf{E}_{\text{ext}} = \mathbf{E}_{\text{ext}}^H, && (4.47)\\
& [\mathbf{E}_{\text{ext}}]_{m,m} = 1, \quad \forall m \in \mathcal{M}, && (4.48)\\
& \mathrm{rank}(\mathbf{E}_{\text{ext}}) = 1, && (4.49)
\end{aligned}$$

where constraint in (4.47) ensures that matrix $\mathbf{E}_{\text{ext}}$ is a Hermitian semidefinite matrix. The constraint (4.48) is transformed from (4.12). Similar to the approach in problem **P3**, constraint (4.49) can be relaxed by applying SDR, and, therefore, the problem can be solved efficiently using CVX. In conclusion, as shown in problem **P5**, by maximizing the variable $\vartheta_{k,l}$, we can find a feasible phase shift matrix $\mathbf{E}_{\text{ext}}$ which satisfies both constraints for QoS (4.11) and IRS phase shift (4.12).

So far, we have converted the original problem **P1** into two SDP sub-problems: optimization problem **P3** for beamformers and feasibility checking problem **P5** for IRS phase shift. Then, the proposed AO method is summarized in Algorithm 1, where problem **P3** and problem **P5** are alternately solved in each iteration until meeting the required accuracy, which is defined by a threshold τ. Moreover, if the problem **P3** and **P5** are infeasible, we introduce a new variable ς to ensure the iteration number cannot exceed a certain value.

Algorithm 1 Robust Power Minimization Problem using AO

Initialization: set random vector $\mathbf{e}^{(0)}$, iteration number $n = 0$, initial transmit power $p^{(0)} = 0$,
Input: channel coefficients $\mathbf{h}_k$, $\hat{\mathbf{H}}_k$, $\forall k \in \mathcal{K}$, the uncertainty δ,
while $n <= \varsigma$ **do**
    $n \leftarrow n + 1$,
    Solve the problem **P3** in (4.33) based on $\mathbf{e}^{(n-1)}$,
    Update the $\mathbf{W}_k^{(n)}$, $\nu_{k,l}^{(n)}$ matrix and extract $\mathbf{w}_k^{(n)}$,
    $p_{\text{tx}}^{(n)} \leftarrow \sum_{k=1}^{K} \|\mathbf{w}_k^{(n)}\|^2$,
    $\mathbf{A}_k^{(n)} \leftarrow \frac{\mathbf{w}_k^{(n)}\mathbf{w}_k^{(n)H}}{2^{\mu_k}-1} - \sum_{i=1,i\neq k}^{K} \left(\mathbf{w}_i^{(n)}\mathbf{w}_i^{(n)H}\right)$,
    $\mathbf{B}_k^{(n)} \leftarrow \frac{\mathbf{w}_k^{(n)}\mathbf{w}_k^{(n)H}}{2^{\mu_k}-1} - \sum_{i=1}^{k-1} \left(\mathbf{w}_i^{(n)}\mathbf{w}_i^{(n)H}\right)$,
    $\psi_{k,l}^{(n)} \leftarrow \max(0, \lambda_{\max}(-\epsilon_l^2 M \mathbf{A}_k^{(n)}))$,
    Solve the problem **P5** in (4.44) based on $\mathbf{A}_k^{(n)}$, $\mathbf{B}_k^{(n)}$, and $\psi_{k,l}^{(n)}$,
    Update $\mathbf{E}_{\text{ext}}^{(n)}$ and extract $\mathbf{e}^{(n)}$,
    if $|p_{\text{tx}}^{(n)} - p_{\text{tx}}^{(n-1)}| \leq \tau$ **then**
        break
    end if
end while
Output: optimal beamforming vectors $\mathbf{w}_k^*$, $\forall k \in \mathcal{K}$ and optimal phase shift vector $\mathbf{e}^*$.

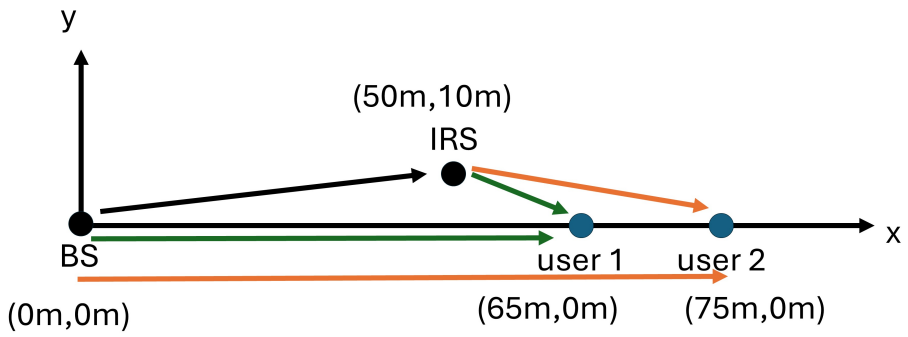


Figure 4.2: The system network setup.

4.5 Complexity Analysis

As summarized in Algorithm 1, the two sub-problems **P3** and **P4** are solved iteratively until the transmit power reaches convergence. Each problem is a SDP problem and can be solved directly with SDR using convex optimization tools such as CVX in Matlab. However, using SDR cannot guarantee that the rank of the optimal matrix is always one. Thus, if a not-rank-one solution is obtained, the Gaussian randomization technique can be utilized to obtain the optimal solution. The first sub-problem **P3** aims to optimize beamformers with fixed IRS phase shift, whose complexity can be given by $\mathcal{O}(N_{tx}^6)$ per iteration. The second sub-problem **P4** optimizes the IRS phase shifts for a given set of beamformers, whose complexity is given by $\mathcal{O}(M^6)$ per iteration.

4.6 Simulation Results

In this section, numerical results are provided to evaluate the performance of the proposed robust algorithm. As shown in Fig. 4.2,it is assumed that the BS is located at point (0 m, 0 m) in the coordinate system, and IRS is placed at point (50 m, 10 m). We assume that two UEs are placed at (65 m, 0 m) and (75 m, 0 m) to guarantee the decoding

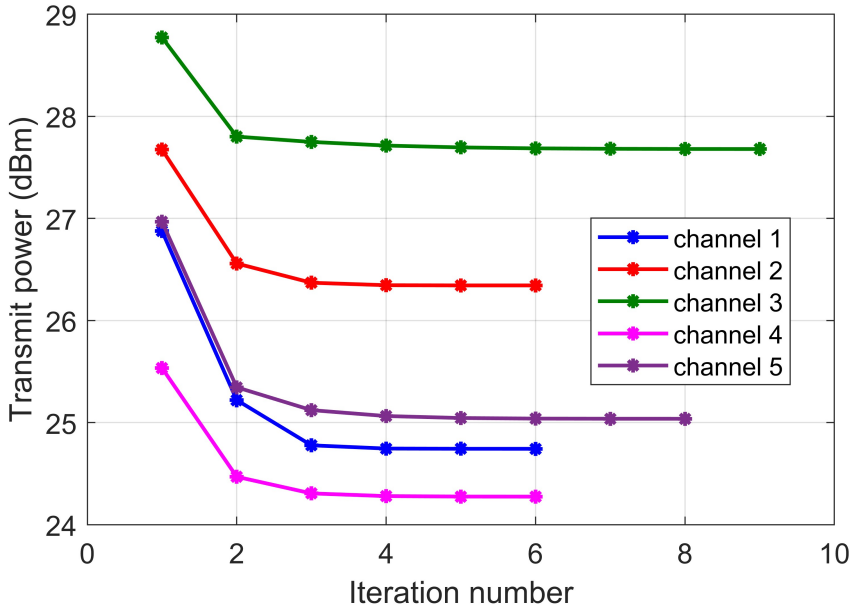


Figure 4.3: Convergence of the proposed algorithm for different sets of channels when $N_{tx} = M = 3$, $K = 2$.

order. Both large-scale and small-scale fading coefficients model the channel coefficients. The large-scale fading can be expressed with a path loss component, which is represented as $\sqrt{\mathrm{PL}_0 d^{-\alpha}}$, where PL_0 is the path loss at one meter, d is the link distance, and α is the path loss exponent. The values of α are set to $\alpha_{BU} = 4$, $\alpha_{BI} = 2.2$ and $\alpha_{IU} = 2$, which stand for the path loss exponent of the BS-user link, BS-IRS link and IRS-UE link, respectively. The small-scale fading component in $\mathbf{h}_k, \mathbf{H}_k, \forall k$ are generated with a Rayleigh distribution. For the uncertainty model, the variance of $\mathrm{vec}(\mathbf{H}_k)$ is defined as $\epsilon_k^2 = \delta^2 \|\mathrm{vec}(\mathbf{H}_k)\|^2, \forall k$, where $\delta \in [0, 1)$ stands for the relative amount of CSI uncertainty. The noise variance is defined as $\sigma^2 = -80\mathrm{dBm}$. The convergence tolerance τ in Algorithm 1 is set to be 10^{-4} [144].

Fig. 4.3 shows the convergence of optimal total transmit power obtained by the proposed algorithm by randomly generating five sets of channels. In this simulation, the target data rate μ_k is set to 2 bits/s/Hz, the outage probability ρ_k is set to 0.05, and the relative error variance δ is set to 0.01. As shown in Fig. 4.3, the algorithm converges within a few numbers of iterations, but different channel realizations need different optimal transmit powers to meet the QoS requirements.

Fig. 4.4 presents the minimum transmit power with different target data rates for three different transmission schemes using a specific set of channel coefficients. The

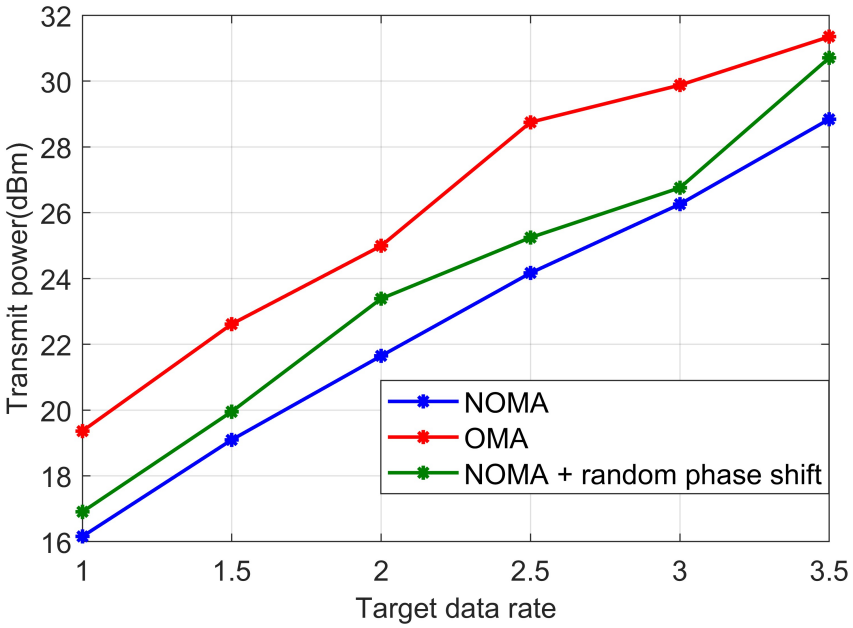


Figure 4.4: Total transmit power for different target data rates and different sets of channels when $N_{tx} = M = 6$, $K = 2$.

number of antennas and IRS elements are set as $N_{tx} = M = 6$. As shown in Fig. 4.4 for all three schemes, the minimum transmit power increases with the target data rate. Comparing OMA and NOMA schemes, NOMA consumes less power than OMA does with the same target rate constraint. In addition, using random phase shift consumes more power than NOMA, which illustrates the competitive performance of the proposed algorithm. Note that random phase shift performs better than traditional OMA, since NOMA has more impact on energy efficiency than IRS.

Fig. 4.5 depicts probability density functions of achieved data rates of the user with lower channel strength when $M = N_{tx} = 3$, $\delta = 0.01$, and $\rho = 0.05$. In the simulation, 100,000 channel error matrices are generated, and the corresponding achieved data rates are calculated. Compared with the target rate of 2, these results in the blue solid curve confirm that the proposed robust design satisfies the target data rate requirements with a predefined outage probability. However, if the beamformers and IRS phase shift of the non-robust optimization problem are used, the red dash curve shows that the true rates cannot achieve the outage probability. In particular, due to the error in the estimation of cascade channels, the BS should allocate more power to all the users than non-robust scenarios to achieve the same SINR constraint. Thus, with the proposed algorithm, optimal power is assigned to users, which guarantees that the achievable data rates are

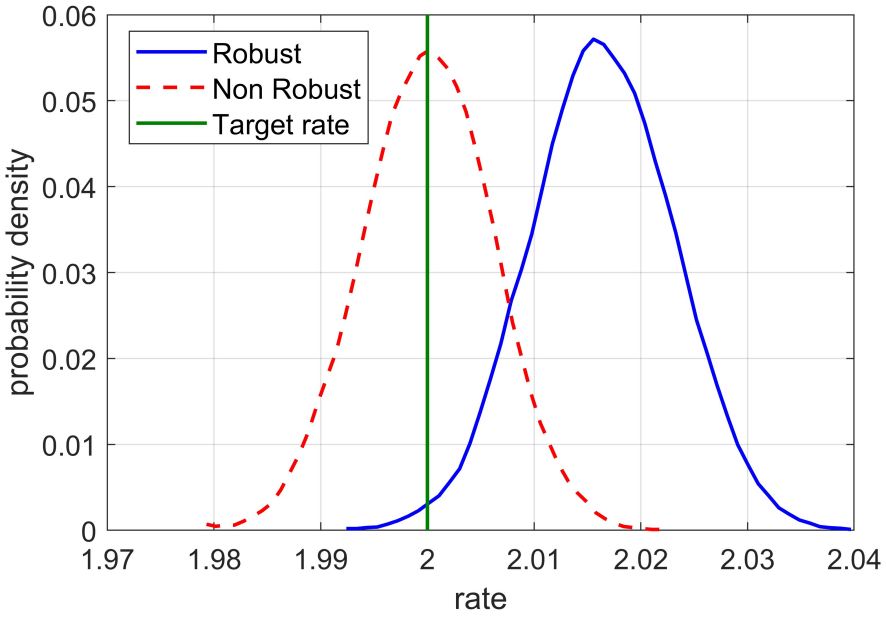


Figure 4.5: probability density of data rate of weak user when $N_{tx} = M = 3$, $K = 2$.

higher than the target with the outage probability despite the channel uncertainty.

Fig. 4.6 shows the required minimum transmit power against different numbers of antennas N_{tx} or IRS elements M. We first investigate the case with the fixed $N_{tx} = 6$. As shown by curve plots with markers diamond in Fig 4.6, the required minimum transmit power decreases with the increase of M because the IRS can improve the system performance by enhancing the channel strength. Meanwhile, when the relative channel uncertainty changes from $\delta = 0.01$ to $\delta = 0.05$, the required minimum transmit power also increases. This is because the BS requires more transmit power to satisfy the constraints with larger channel uncertainties. On the other hand, curve plots with markers square reveal that with fixed IRS elements number $M = 6$, the required minimum transmit power decreases with the increase of antenna number N_{tx}. In addition, we can observe that the gradients of red curves are much less than those of blue curves. This performance difference is justified through the following argument. The increase of element number M also leads to more channel uncertainties in the cascaded channel, which means that the BS requires more transmit power to meet the requirements of QoS.

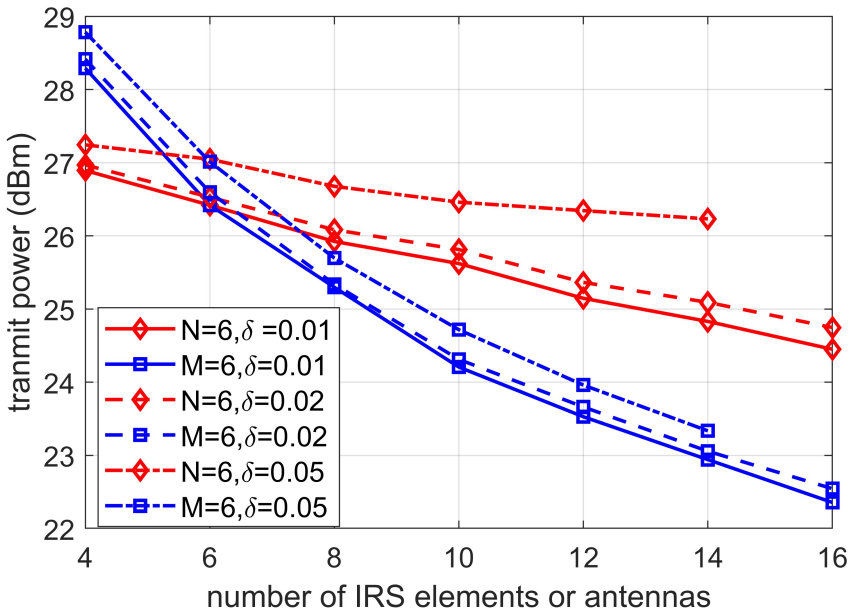


Figure 4.6: Total transmit power with different transmit antennas or IRS elements number. For the x-axis, the number of IRS elements M is considered for the line plot with markers diamond while the number of antennas N_{tx} is considered for the line plot with markers square.

4.7 Summary

In this work, we proposed a robust downlink beamforming design for a MISO IRS-aided NOMA wireless network In particular, the original robust design problem was formulated into a robust power minimization problem with outage probability based rate and unit-modulus phase shift constraints. The inevitable channel uncertainties were take into account in the robust design. However, the formulated power minimization problem was non-convex. To address the problem, we converted the original non-convex problem into convex form based on the Berstein Type Inequality and SDR. Then, the AO method was developed to solve the original problem iteratively. The simulation results showed that the proposed iterative algorithm can successfully solve the power minimization problem, while satisfying the QoS requirements at users. In addition, the increase in the number of IRS elements and antennas at BS can decrease the required minimum transmit power at the BS.

Chapter 5

AoI Minimization for Uplink IRS-aided NOMA Based IoT Network

5.1 Introduction

Besides power control problems, data freshness also plays a crucial role in wireless communication. Outdated information in wireless networks can significantly reduce system efficiency, decision-making accuracy, and the overall QoS. In order to measure it, AoI offers a thorough approach to quantify the information freshness especially for 6G and beyond networks, which required notably require timely transmission [145]. The most existing works on AoI problem of IRS-aided NOMA system divide the resource allocation problem into two sub-problems and optimize power allocation and IRS phase shift separately. For instance, in [122], the authors introduce a AoI minimization problem for IRS-aided NOMA network. Then, the authors optimize IRS phase shift using convex optimization, address power allocation using a sub-optimal solution, and apply exhaustive search for clustering. However, both convex optimization and exhaustive search remarkably increase the computational complexity, which are not applicable in use-cases and services with stringent delay constraints. Thus, in this chapter, we aim to propose a dynamic devices pairing algorithm according to their channel gains. Then, two approaches are proposed to tackle power allocation and phase shift.

The contribution of this chapter can be listed as follows:

- We consider a resource allocation problem for uplink IRS-aided NOMA IoT network. In particular, the problem aims to obtain optimal power allocation, IRS phase shift and devices clustering to minimize average sum AoI under unit-modulus constraints for IRS.
- The original AoI minimization problem is non-convex and contains different coupled design variables. Hence, we propose a dynamic clustering method. Then, a closed-form optimization and a DRL-based approach are proposed to optimize both power allocation and phase shift of IRS.
- Various simulation results of both methods show their benefits in terms of power consumption and complexity comparing with that of the similar work in the literature. In addition, though closed-form optimization can obtain more accurate solution than DRL-based approach, it still suffers from higher computational complexity.

5.2 System Model and SINR Analysis

AoI management is generally a more serious issue in uplink compared to downlink communication, due to limited transmission resources at transmitting devices, especially in massive IoT networks [146]. Thus, we consider a uplink transmission of an IRS-aided NOMA communication system, which consist of a BS equipped with single antenna, an IRS with M elements, and $2I$ single-antenna devices, as shown in Fig. 5.1. The time is divided into time slots $t \in \mathcal{T}$, where $\mathcal{T} = \{1, 2, \cdots, T\}$. In addition, let index $i \in \mathcal{I}$ denote the devices set, where $\mathcal{I} = \{1, 2, \cdots, 2I\}$, and $m \in \mathcal{M}$ denote the elements of the IRS, where $\mathcal{M} = \{1, 2, \cdots, M\}$. The devices are separated into two groups based on their channel strengths: strong group $\mathcal{S} = \{1, 2, \cdots, I\}$, where devices have stronger channel gains, and weak group $\mathcal{W} = \{1, 2, \cdots, I\}$, where devices have weaker channel gains. Let $h_s(t) \in \mathbb{C}$ and $h_w(t) \in \mathbb{C}$ denote the channel from the user $s \in \mathcal{S}$ to the BS and from the user $w \in \mathcal{W}$ to the BS at time slot t, respectively. Then, the following formulation holds $|h_s(t)| \geq |h_w(t)|, \forall s \in \mathcal{S}, \forall w \in \mathcal{W}, \forall t \in \mathcal{T}$. The channel from user $w \in \mathcal{W}$ to the IRS and the channel from IRS to the BS can be represented by $\mathbf{r}_w(t) \in \mathbb{C}^{M \times 1}$ and $\mathbf{g}(t) \in \mathbb{C}^{M \times 1}$. All the channels consist of both large-scale and small-sale fading, which can be expressed as $h(t) = \sqrt{\mathrm{PL}_0 d(t)^{-\beta}} \hat{h}(t)$ for $h \in \{h_s(t), h_w(t), \mathbf{r}_w(t), \mathbf{g}(t)\}$, where PL_0

indicates the path loss at unit distance, $d(t)$ is the distance of the communication link at time slot t, and β is the path loss exponent. The small-scale fading components, $\hat{h}(t)$ is modelled as Rayleigh fading with zero means and unit variance, $\hat{h}(t) \in \mathcal{CN}(0, \mathbf{I})$, where $\mathbf{I}$ refers to an identity matrix. All $2I$ devices are grouped into I clusters, by paring one user from $\mathcal{S}$ and one user from $\mathcal{W}$. Uplink NOMA scheme is applied in each cluster. In addition, different cluster communicate with the BS via orthogonal sub-carriers, so that inter-cluster interference can be removed. Moreover, in order to enhance the performance of users in $\mathcal{W}$, an IRS is applied. The phase shift matrix of IRS at time slot t is $\mathbf{\Theta}(t) = \text{diag}(\beta_1(t)e_1(t), \beta_2(t)e_2(t), \cdots, \beta_M(t)e_M(t))$, where $\beta_m(t)$ and $e_m(t)$, $\forall m \in \mathcal{M}$, stand for the amplitude and phase shift coefficient of the m^{th} element of IRS, respectively. It is assumed that the phase shift of each IRS element has unit amplitude shift, $\beta_m(t) = 1, \quad \forall m \in \mathcal{M}, \quad \forall t \in \mathcal{T}$.

In each cluster, since SIC is employed, decoding order is crucial important. Due to the grouping strategy mentioned above, $|h_s(t)| \geq |h_w(t)|$, the BS can decode the signal of user s before decoding signal of user w. Therefore, the interference from user s can be removed. On the other hand, for user s, the signal for user w is regarded as an interference at the BS. Thus, the SINR at the BS of user s and w can be respectively expressed as

$$\Gamma_s(t) = \frac{p_s(t)|h_s(t)|^2}{p_w(t)|h_w(t) + \mathbf{r}_w^H(t)\mathbf{\Theta}(t)\mathbf{g}(t)|^2 + \sigma_s^2} \tag{5.1}$$

$$\Gamma_w(t) = \frac{p_w(t)|h_w(t) + \mathbf{r}_w^H(t)\mathbf{\Theta}(t)\mathbf{g}(t)|^2}{\sigma_w^2}, \tag{5.2}$$

$$\forall s \in \mathcal{S}, \quad \forall w \in \mathcal{W},$$

where $p_i(t)$ and σ_i^2 refer to power allocation and noise power at time slot t of zero-mean AWGN for user i, $i \in \{\mathcal{S}, \mathcal{W}\}$, respectively.

5.3 AoI Analysis

The AoI indicate the freshness of transmitted information, which can be modelled as the duration time of the intended signal from the transmission to received at the BS. In this chapter, it is assumed that the signal is transmitted and received at the same time slot. Denote the AoI of the devices i at the t^{th} time slot by $\alpha_i(t)$. If the signal is transmitted from the i^{th} device and received at the BS successfully at time slot t, the

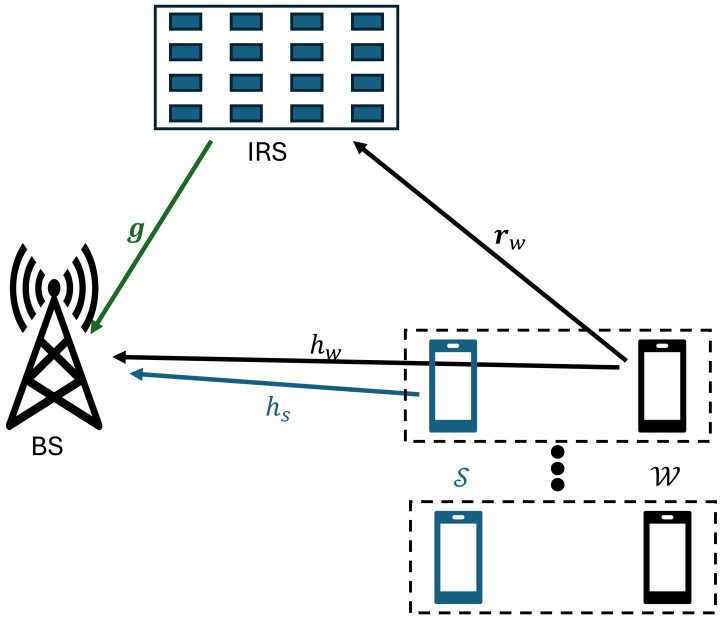


Figure 5.1: IRS-aided uplink NOMA communication system.

AoI can be refreshed to be one, i.e., $\alpha_i(t) = 1$. In contrast, if the BS fails to decode the signal, the AoI will increase by one, i.e., $\alpha_i(t) = \alpha_i(t-1) + 1$. Furthermore, the success of transmission can be demonstrated by whether the SINR of the i^{th} user greater than a given threshold Γ_{th}. Thus, the AoI value of the i^{th} user can be given by [122]

$$\alpha_i(t) = \begin{cases} 1, & \text{if } \Gamma_i \geq \Gamma_{\text{th}}, \\ \alpha_i(t-1) + 1, & \text{otherwise}, \end{cases} \quad \forall i \in \{\mathcal{S}, \mathcal{W}\} \tag{5.3}$$

5.4 Problem Formulation

We consider a power allocation, devices pairing, and IRS phase shift optimization problem, focusing on minimizing the sum average AoI under power consumption and IRS elements unit-modulus constraints, which can be formulated as an optimization problem as follows:

$$\mathbf{P1} \quad \min_{\mathbf{B},\mathbf{p}_s,\mathbf{p}_w,\boldsymbol{\Theta}} \quad \frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2I}\sum_{s=1}^{I}\sum_{w=1}^{I}\left(\alpha_s(t)+\alpha_w(t)\right)b_{s,w}(t)\right] \tag{5.4}$$

$$\text{s.t.} \quad b_{s,w}(t) \in \{0,1\}, \tag{5.5}$$

$$\sum_{s=1}^{I} b_{s,w}(t) \leq 1, \tag{5.6}$$

$$\sum_{w=1}^{I} b_{s,w}(t) \leq 1, \tag{5.7}$$

$$p_s(t) \leq p_{\max}, \tag{5.8}$$

$$p_w(t) \leq p_{\max}, \quad \forall s \in \mathcal{S}, \quad \forall w \in W, \tag{5.9}$$

$$|[\mathbf{\Theta}(t)]_{m,m}| = 1, \quad \forall m \in \mathcal{M}, \quad \forall t \in \mathcal{T}, \tag{5.10}$$

where $p_{\max}$ is the maximum power for each devices and $b_{s,w}$ is a binary variable which stands for the user cluster. For instance, if devices $s \in \mathcal{S}$ and $w \in \mathcal{W}$ are paired in to a cluster, $b_{s,w} = 1$, otherwise, $b_{s,w} = 0$. $\mathbf{B} = \{b_{s,w} | s \in \mathcal{S}, w \in \mathcal{W}\}$, $\mathbf{p}_s = \{p_s | s \in \mathcal{S}\}$, and $\mathbf{p}_w = \{p_w | w \in \mathcal{W}\}$. In the original problem **P1**, constraints (5.6) and (5.7) guarantee that device $s \in \mathcal{S}$ is paired with at most one device $w \in \mathcal{W}$. Constraints (5.8) and (5.9) ensure that the transmit power for both groups do not exceed the maximum power budget. In addition, the constraint ((5.10)) ensures that all the elements in the main diagonal of the IRS phase shift matrix $\mathbf{\Theta}(t)$ have unit absolute value, which illustrates that IRS elements will not effect the signal's strength but only change its phase which has been assumed in the literature.

5.5 Proposed Methodology

Due to the fact that the original problem **P1** is apparently a non-convex problem and contains multiple coupled design variables, including continuous variables $\mathbf{p}_s$, $\mathbf{p}_w$, $\mathbf{\Theta}$, and binary variable $\mathbf{B}$, it is challenging to solve the original problem jointly. Thus, we propose two algorithms: closed-form optimization and DRL-based approach, to design the clustering strategy, IRS phase shift, and power allocation.

5.5.1 Closed-Form Optimization

In this optimization algorithm, the original problem **P1** is divided into three sub-problems to solve devices pairing, IRS phase shift optimization, and power allocation separately. To be more specific, we first develop a user pairing strategy to pair devices into clusters. Secondly, we exploit SDP to optimize the phase shift of the IRS. Then, the optimal power allocation can be found by mathematical transforms.

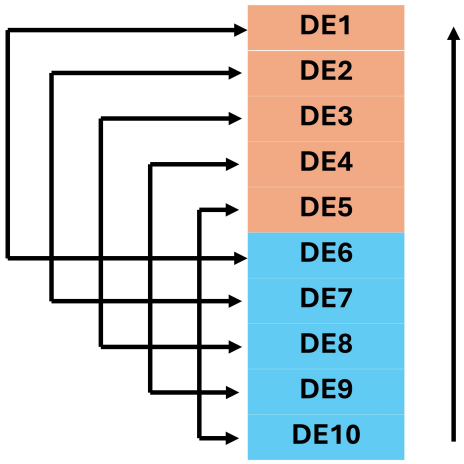


Figure 5.2: Illustration of user pairing strategy.

User Clustering Strategy

For devices clustering, the downlink NOMA prefers NLUPA, which aims to guarantee the maximal channel gains of stronger user and weaker user in a cluster [147]. Therefore, the sum rate performance can be achieved in each cluster. Particularly, in NLUPA, the device with strongest channel gain is paired with the device with weakest channel gains in a cluster. However, for uplink NOMA, NLUPA cannot perform the same because of the distinction of SINR definitions [83]. In downlink NOMA, the user with strongest channel is able to decode other users' signals, so their signals are not regarded as interference to the strongest user, while in uplink NOMA, user with strongest channel gain suffers from interference of all the other users as shown in (5.1).

The proposed suboptimal user grouping strategy for uplink NOMA is shown in Fig.5.2, where the device denotes by DEi, $\forall i \in \{1, \cdots, 10\}$. All the DE$i$s are sorted based on their channel gains in a descending order, which is describe by $|h_{\text{DE1}}| \geq |h_{\text{DE2}}| \geq \cdots |h_{\text{DE10}}|$. Then, the DE$i$s are divided into two groups, shown by two different colors, orange and blue, respectively, in Fig.6.2. In NOMA scheme, due to the SIC, the cluster performs better as the channel gains difference increase. Thus, in this grouping strategy, we pair the first device of orange group and the first one in blue group in a cluster. In this way, the average difference between channel gains of two devices in a cluster can remain almost the same which leads to a better over-all performance. Then, the user clustering

strategy for given $2I$ channel coefficients is summarized as in Algorithm 2.

Algorithm 2 Devices pairing algorithm

1: Sort $2I$ devices based on their channel gains: $|h_1| \geq |h_2| \geq \cdots \geq |h_i| \geq \cdots \geq |h_{2I}|, \quad \forall i \in \mathcal{I}$
2: Divide devices in two groups $\mathcal{S}$ and $\mathcal{W}$:
3: **if** $i \leq I$ **then**
4: $\quad i \in \mathcal{S}$
5: **else if** $i \geq I+1$ **then**
6: $\quad i \in \mathcal{W}$
7: **end if**
8: Group $2I$ devices into clusters:
9: 1^{st} cluster $= \{h_1, h_{I+1}\}$
10: $\cdots$
11: i^{th} cluster $= \{h_i, h_{I+i}\}$
12: $\cdots$
13: I^{th} cluster $= \{h_I, h_{2I}\}, \quad \forall i \in \{1, \cdots I\}$

Based on user pairing proposed strategy in Algorithm 2, the user clustering problem is addressed. Then, assuming the device $s \in \mathcal{S}$ and the device $w \in \mathcal{W}$ are paired in a cluster, the original problem **P1** can be reduced to the following optimization problem

$$\mathbf{P2} \qquad \min_{\mathbf{p}_s, \mathbf{p}_w, \boldsymbol{\Theta}} \qquad \frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2I}\sum_{s=1}^{I}\sum_{w=1}^{I}(\alpha_s(t)+\alpha_w(t))\right] \tag{5.11}$$

$$\text{s.t.} \qquad (5.8), (5.9), (5.10), \tag{5.12}$$

$$\forall s \in \mathcal{S}, \quad \forall w \in \mathcal{W}, \quad \forall t \in \mathcal{T}.$$

IRS Phase Shift Optimization Approach

In this sub-problem, we aim to optimize the IRS phase shift for a given fixed devices cluster. Since the IRS focuses on enhancing the channel strengths between devices $w \in \mathcal{W}$ and the BS, the optimal phase shift can be determined by maximizing the minimum channel strengths among all w devices, which can be expressed as follows:

$$\mathbf{P3} \quad \max_{\boldsymbol{\Theta}(t)} \quad \min_{w \in \mathcal{W}} |h_w(t) + \mathbf{r}_w^H(t)\boldsymbol{\Theta}(t)\mathbf{g}(t)|^2 \tag{5.13}$$

$$\text{s.t.} \quad |[\boldsymbol{\Theta}(t)]_{m,m}| = 1, \quad \forall m \in \mathcal{M}, \quad \forall t \in \mathcal{T}. \tag{5.14}$$

It can be observed from the problem **P3** that it is challenging to optimize the phase shift matrix due to its non-convexity structure. Based on [122,148], we introduce another variable ν to address the non-convexity issue. Then, the problem **P3** can be reformulated as

$$\mathbf{P4} \quad \max_{\alpha, \boldsymbol{\Theta}(t)} \quad \nu \tag{5.15}$$

$$\text{s.t.} \quad |h_w(t) + \mathbf{r}_w^H(t)\boldsymbol{\Theta}(t)\mathbf{g}(t)|^2 \geq \nu, \quad \forall w \in \mathcal{W} \tag{5.16}$$

$$\nu \geq 0, \tag{5.17}$$

$$|[\boldsymbol{\Theta}(t)]_{m,m}| = 1, \quad \forall t \in \mathcal{T}. \tag{5.18}$$

Note that the constraint in (5.18) is non-convex and cannot be solved directly. We define $\boldsymbol{\Phi}_w(t) = \text{diag}(\mathbf{r}_w(t))\mathbf{g}(t)$ as the cascade channel coefficient for device w and transform the phase shift matrix $\boldsymbol{\Theta}(t)$ into a vector $\mathbf{e}(t) = [[\boldsymbol{\Theta}(t)]_{1,1}, [\boldsymbol{\Theta}(t)]_{2,2}, \cdots, [\boldsymbol{\Theta}(t)]_{M,M}]^H$, and therefore, the following transformation can be applied, $\mathbf{r}^H(t)\boldsymbol{\Theta}(t)\mathbf{g}(t) = \mathbf{e}^H(t)\boldsymbol{\Phi}(t)$. Thus, defining $\mathbf{E}_{ext}(t) = \begin{bmatrix} \mathbf{e}(t) \\ 1 \end{bmatrix} [\mathbf{e}^H(t), 1]$, the problem **P4** can be rewritten as

$$\mathbf{P5} \quad \max_{\alpha, \mathbf{e}(t)} \quad \nu \tag{5.19}$$

$$\text{s.t.} \quad \text{Tr}\{\mathbf{E}_{ext}\mathbf{B}(t)\} + |h_w|^2 \geq \nu, \tag{5.20}$$

$$\mathbf{E}_{ext}(t) \succeq \mathbf{0}, \tag{5.21}$$

$$[\mathbf{E}_{ext}(t)]_{m,m} = 1, \tag{5.22}$$

$$\text{rank}(\mathbf{E}_{ext}(t)) = 1, \quad \forall m \in \mathcal{M}, \quad \forall t \in \mathcal{T}, \tag{5.23}$$

$$(5.17),$$

where

$$\mathbf{B}(t) = \begin{bmatrix} \mathbf{\Phi}(t)\mathbf{\Phi}^H(t) & \mathbf{\Phi}(t)h_w(t) \\ h_w(t)\mathbf{\Phi}^H(t) & 0 \end{bmatrix}. \tag{5.24}$$

By applying SDR, the constraint in (5.23) can be dropped [142, 143]. Then, the problem **P5** can be efficiently solved using CVX and the optimal phase shift $\mathbf{e}(t)$ can be extracted from $\mathbf{E}_{ext}(t)$ by using eigenvalue decomposition.

Power Allocation Approach

For a given set of devices cluster and the optimal IRS phase shifts, the constraints in (5.5), (5.7), (5.6) and (5.10) can be removed. The original problem **P1** is transformed into a power allocation problem to minimize the sum average AoI. In order to find the minimal AoI, optimal power should be allocated to s and w devices to enable successful signal transmission, in another words, to make the SINRs greater than the threshold. Thus, assuming that the devices $s \in \mathcal{S}$ and $w \in \mathcal{W}$ are paired in a cluster, the problem can be expressed as follows:

$$\begin{aligned}
\mathbf{P6} \quad & \min_{\mathbf{p}_s, \mathbf{p}_w} && \sum_{s=1}^{I} p_s(t) + \sum_{w=1}^{I} p_w(t) && (5.25)\\
& \text{s.t.} && \Gamma_s(t) \geq \Gamma_{th}, && (5.26)\\
& && \Gamma_w(t) \geq \Gamma_{th}, && (5.27)\\
& && p_s(t), p_w(t) \leq p_{\max}, && (5.28)\\
& && \forall s \in \mathcal{S}, \quad \forall w \in \mathcal{W}, \quad \forall t \in \mathcal{T},
\end{aligned}$$

where the objective of problem **P6** in (5.25) indicates the total transmit power is minimized to achieve the SINR threshold for all the users, so that AoI can be refreshed. Through thematic derivation, closed-form results can be obtained as follows,

$$p_w^*(t) = \min\left\{p_{\max}, \frac{\Gamma_{th}\sigma_w^2}{|h_w(t) + \mathbf{r}_w^H(t)\mathbf{\Theta}(t)\mathbf{g}(t)|^2}\right\}, \tag{5.29}$$

$$p_s^*(t) = \min\left\{p_{\max}, \frac{\Gamma_{th}(|h_w(t) + \mathbf{r}_w^H(t)\mathbf{\Theta}(t)\mathbf{g}(t)|^2 p_w^*(t) + \sigma_s^2)}{|h_s(t)|^2}\right\}, \tag{5.30}$$

$$\forall w \in \mathcal{W}, \quad \forall s \in \mathcal{S},$$

where $p_i^*(t), \forall i \in \{\mathcal{S}, \mathcal{W}\}$ stands for the optimal power allocated for the device i at time slot t. Particularly, if the problem **P6** is feasible, i.e., $p_i^*(t) \leq p_{\max}$, AoI is refreshed to be one, such that $\alpha_i = 1$.

5.5.2 DRL-based Approach

As explained in the above section, the original problem **P2** is divided into three sub-problems based on variables and, therefore, the IRS phase shift and power allocation are optimized separately.To jointly solve the problem P2 in terms of $\mathbf{p}_s$, $\mathbf{p}_w$, and $\mathbf{\Theta}$, we develop another DRL-based approach. Firstly, the problem **P2** is reformulated into a DRL environment, the details of which is given in the following subsections.

Problem Reformulation

In the following section, the problem is modelled as a MDP. To address the optimization problem in MDP, we further apply a DRL framework, which consist of an agent and an environment. At time step n, the agent takes action $\mathbf{a}^n$ according to the state $\mathbf{s}^n$ provided by the environment. Based on the taken action, a new state for next time step $\mathbf{s}^{n+1}$ and reward R^n are given by the environment. Then, through multiple interactions with the environment, an optimal policy $\pi^*(\mathbf{s}, \mathbf{a})$ can be learnt by the agent, which maps any state to the optimal action leading to the highest reward value. Thus, DRL framework can convert the optimization problem **P2** into a sequence of actions taken by agent to achieve maximum rewards. In particular, we define the action vector $\mathbf{a}^n$, the state vector $\mathbf{s}^n$ and the reward function R^n as follows:

- The action space $\mathbf{a}$: Since the objective of the optimization problem **P2** is a function in term of power allocation and IRS phase shift, such as $\mathbf{p}_s$, $\mathbf{p}_w$, and $\mathbf{\Theta}$, they are defined as elements in the action space vector. Particularly, the action space vector at step n is expressed by follows,

$$\mathbf{a}^n = \left[(\mathbf{p}_w^n)^T, (\mathbf{p}_s^n)^T, (\tilde{\mathbf{e}}^n)^T\right]^T. \tag{5.31}$$

  Note that a DNN framework is used and it only accept real numbers in both action

and sate space vectors. Thus, we introduce a new vector $\tilde{\mathbf{e}}$, which consists of both real and imaginary part of IRS phase shift vector at the n^{th} step $\mathbf{e}^n$. To be more specific, $\tilde{\mathbf{e}}^n = \left[\Re\{\mathbf{e}^n\}^T, \Im\{\mathbf{e}^n\}^T\right]^T$. Consequently, the action space vector for the n^{th} time step can be expressed as $\mathbf{a}^n \in \mathbb{R}^{2I+2M}$ with only real numbers.

- The state space $\mathbf{s}$: In order to help the agent be well trained through the interaction with the environment, we consider action vector for previous time slot in the state space. Moreover, because the SINR for devices s and w are affected by the channel gains as shown in (5.1) and (6.4), the channel sets of direct channel from devices to the BS, $\mathbf{h}^n = \{h_i^n | i \in \{\mathcal{S}, \mathcal{W}\}\}$, and the cascade channels $\mathbf{\Phi}_w^n, \quad \forall w \in \mathcal{W}$, should be contained in the state space as well. Accordingly, sate vector at step n is defined as

$$\mathbf{s}^n = \left[|(\mathbf{h}^n)|^T, \|\mathbf{\Phi}_1^n\|, \|\mathbf{\Phi}_2^n\|, \cdots, \|\mathbf{\Phi}_I^n\|, (\mathbf{a}^{n-1})^T\right]^T, \tag{5.32}$$

where $\mathbf{s}^n \in \mathbb{R}^{5I+2M}$. Similarly, as the state vector should only include real numbers, we take the absolute value of channel coefficients vectors and the norm value of channel coefficient matrices. Note that including previous actions in the state vector improves the agent's decision-making. Since the current state alone may not always capture the full context, adding past actions helps the agent retain memory of recent decisions, making it better equipped to handle temporal dependencies. This is particularly beneficial in continuous action spaces, where maintaining smooth and consistent policies is essential.

- The reward function r^n: As shown in the problem **P2**, the optimization problem aims to minimize the sum average AoI of all the devices while consuming minimum amount of power. Hence, both α_i and p_i, $i \in \{\mathcal{S}, \mathcal{W}\}$, are included in the reward function, which is expressed as follows:

$$R^n = -\sum_{s=1}^{I}\sum_{w=1}^{I}\left[\alpha_s^n + \alpha_w^n\right] + 2Ip_{\max} - \sum_{s=1}^{I}\sum_{w=1}^{I}\left[p_s^n + p_w^n\right], \tag{5.33}$$

where the first term is the negative value of sum AoI and the second term is the difference between maximum power and power consumption of all the devices at time step t. By maximizing this function in (6.18), the agent provides the optimal

IRS phase shift, and achieve the minimum sum average AoI over $\mathcal{T}$ consuming minimum amount of power.

To ensure that the agent can select the power allocation and IRS phase shift within a feasible region, the action vector should be scaled and normalized. Thus, the feasible power allocation for the i^{th} device at step n is expressed as

$$\bar{\mathbf{p}}^n = \mathbf{p}^n p_{\max}, \tag{5.34}$$

where $\bar{\mathbf{p}}^n = \{\bar{p}_i^n | i \in \mathcal{S}, \mathcal{W}\}$ is the scaled power allocation vector, and $\mathbf{p}^n = \{p_i^n | i \in \mathcal{S}, \mathcal{W}\}$ is the power allocation coefficient vector provided by the agent, where $p_i^n \in [0, 1]$. On the other hand, the feasible value of each element in the IRS phase shift $\mathbf{e}^n$ is given by

$$[\bar{\mathbf{e}}^n]_m = \frac{[\mathbf{e}^n]_m}{|[\mathbf{e}^n]_m|}, \quad \forall m \in \mathcal{M}, \tag{5.35}$$

where $\bar{\mathbf{e}}^n$ is the scaled phase shift vector. By applying this, the constraint in (5.10) can be always satisfied.

DDPG agent

In this section, basic the fundamental concept and framework of a DDPG agent are introduced. The structure of DDPG agent is shown in Fig. 5.3.

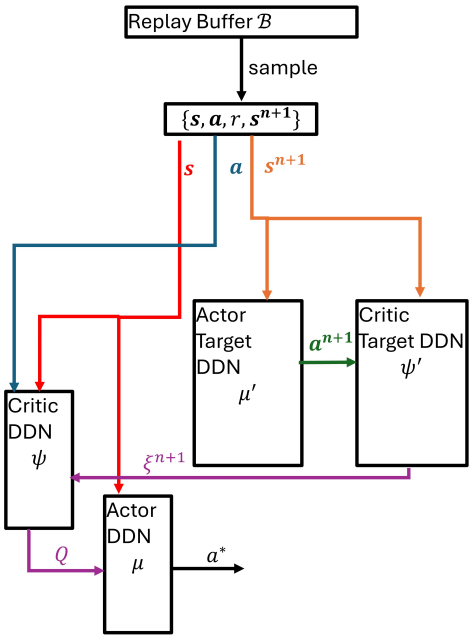


Figure 5.3: The actor-critic framework for DDPG agent.

In a standard DRL session, the interaction between an agent and an environment is described by a tuple $\{\mathbf{s}, \mathbf{a}, r, \mathbf{s}^{n+1}\}$. Through this process, the agent is trained to learn a policy $\pi(\mathbf{s}, \mathbf{a})$ to maximize a long term reward. However, the DRL agent is limited to discrete action and state spaces. Thus, in this work, the power allocation and IRS phase shift optimization is based on DDPG agent.

Compared with traditional RL agent, DDPG agent introduces parts: actor DNN and critic DNN, to exploit the RL algorithm with continuous sate and action spaces. In particular, the actor network provides optimal action for a given sate, which is formulated as

$$\mu(\mathbf{s}) = \mathbf{a}^*, \tag{5.36}$$

where μ refers to the actor network, $\mathbf{s}$ and $\mathbf{a}^*$ denote the input state and output optimal action vectors, respectively. On the other hand, the critic network is designed to evaluate the performance of actor network by calculating the Q-value of the optimal action provided by the actor. Let ψ denote the critic network, then, the expression for the critic is presented as follows,

$$\psi(\mathbf{s}, \mathbf{a}) = Q^*, \tag{5.37}$$

where Q^* indicates the optimal Q-value. In addition, DDPG agent have a copies of both actor and critic networks to generate training target, named as target networks. After the interaction with the environment, a tuple of experience, such as $\{\mathbf{s}^n, \mathbf{a}^n, r, \mathbf{s}^{n+1}\}$, is saved in the replay buffer, denote by $\mathcal{B}$. Then, denote actor and critic target networks by $\mu^{'}$ and $\psi^{'}$, respectively, the target for a sampled tuple is given by

$$\xi(r, \mathbf{s}) = r^n + \delta\psi^{'}\left(\mathbf{s}^{n+1}, \mu^{'}(\mathbf{s}^{n+1})\right), \tag{5.38}$$

where $\delta \in (0, 1]$ is a discount factor determining the value for the reward of next time step. As shown in (5.38), the target is calculated using Q-value obtained from the target critic network with the optimal action obtained from the target action network. Then,

the training of the critic network is based on the following loss function:

$$L(\psi) = \mathbb{E}_{\{\mathbf{s},\mathbf{a},r,\mathbf{s}^{n+1}\}\sim\mathcal{B}}\left[\left(\psi(\mathbf{s},\mathbf{a}) - \xi(r,\mathbf{s})\right)^2\right]. \tag{5.39}$$

Notably, expectation operation of the difference between the Q-value and the target is performed as a batch of tuples sampled from the replay buffer $\mathbf{B}$. The critic network is trained by minimizing the mean squared error in (5.39). Furthermore, the actor can be trained according to the Q-value supplied by the trained critic DNN. For example, the actor network is optimized by maximizing the Q-value given by the critic network, which is given by

$$\max_{\mu} \quad \mathbb{E}_{\mathbf{s}\sim\mathcal{B}}\left[\psi(\mathbf{s},\mu(\mathbf{s}))\right]. \tag{5.40}$$

Moreover, the target actor and critic network for DDPG copy from the learned actor and critic DNN in each time step, but with much lower learning rate. Particularly, the update coefficients of both target DNNs are formulated as

$$\mu^{'} = \epsilon\mu + (1-\epsilon)\mu^{'} \tag{5.41}$$

$$\psi^{'} = \epsilon\psi^{'} + (1-\epsilon^{'}), \tag{5.42}$$

where $\epsilon \ll 1$ is the learning rate. It means that the target value is learning with a limited speed, which ensure the stability of the training.

Considering that at the beginning, all the coefficients of all DNN are randomly generated, the action selected by the agent can be regarded as a random process. Thus, to yield this problem and help the agent to explore, a random noise vector is added to the action taken by the agent. Considering the range of action $[\mathbf{a}]_i \in [a_{\min}, a_{\max}]$, $\forall i \in \{1,\cdots,2I+2M\}$, each element in the real action vector is formulated as

$$[\mathbf{a}]_i = U\left([\mu(\mathbf{s})]_i + [\mathbf{n}]_i\right), \quad \forall i \in \{1,\cdots,2I+2M\} \tag{5.43}$$

where $U(x)$ is a step function for range of action, i.e.,

$$U(x) = \begin{cases} 1, & \text{if} \quad x \in [a_{\min}, a_{\max}], \\ 0, & \text{otherwise.} \end{cases} \tag{5.44}$$

The noise vector $\mathbf{n} \sim \mathcal{N}(0, \sigma \mathbf{I})$ follows a normally distribution with zero means and σ variance, which is gradually decrease as the training proceed.

DDPG-based Algorithm

The proposed DDPG-based algorithm is summarized in Algorithm 3.

Algorithm 3 DDPG-based power allocation and phase shift optimization algorithm

Initialize: DDPG agent's hyperparameters for actor DDN μ, critic DDN ψ, action noise vecotr $\mathbf{n}$, replay buffer $\mathcal{B}$, step number $n = 0$
Update hyperparameters for target DDNs: $\mu^{'} \leftarrow \mu$ and $\psi^{'} \leftarrow \psi$
while $\text{no.}_{\text{epi}} \leq \text{No.}_{\text{epi}}$ **do**
    Obtain the channel coefficients $h_i(t)$, $\mathbf{r}_w(t)$, $\forall s \in \mathcal{S}$, $\forall w \in \mathcal{W}$, $\mathbf{g}(t)$ from current sate $\mathbf{s}^n$
    Cluster the devices based on NLUPA
    Set random initial action vector $\mathbf{a}^0$ and calculate state vector based on (6.17)
    while $n \leq N_{\text{step}}$ **do**
        Get action vector $\mathbf{a}^n$ based on (5.43)
        Normalize the action vector to calculate power allocation $\bar{\mathbf{p}}^n$ and phase shift $\bar{\mathbf{e}}^n$ according to (5.34) and (6.20), respectively
        Calculate the SINR for each device according to (5.1) and (6.4)
        Calculate reward r based on the function in (6.18)
        Obtain the next-step state vector $\mathbf{s}^{n+1}$ from the environment
        Save the tuple $\{\mathbf{s}^n, \mathbf{a}^n, r^n, \mathbf{s}^{n+1}\}$ to $\mathcal{B}$
        Sample a example from $\mathcal{B}$
        Calculate target ξ based on (5.38)
        Train critic DDN using (5.39)
        Update the target actor and critic DDNs
        Update the target DDNs using (5.41)
        Step number $n = n + 1$
    end while
    $\text{no.}_{\text{epi}} = \text{no.}_{\text{epi}} + 1$
end while
Obtain the optimal power allocation $\bar{\mathbf{p}}^*$ and IRS phase shift $\bar{\mathbf{e}}^*$

Note that the well-trained DDPG agent is able to provide optimal power and phase shift, that can let all the devices achieve higher SINR, for given channel gains at every single time slot t. By doing this, the agent can ensure success transmissions for all the feasible channels. In other words, the average sum AoI is minimized over $\mathcal{T}$.

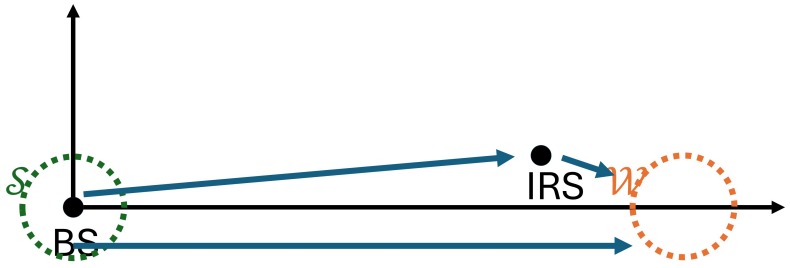


Figure 5.4: The system network setup.

5.6 Training and Simulation Results

The original problem **P1** is divided into two sub-problems: user pairing problem and IRS phase shift optimization, power allocation problem in **P2**. To address **P2**, we propose closed-form optimization and DDPG-based algorithm in this work. Thus, in this section, we provide simulation results to evaluate both methods for the sum average AoI of IRS-aided NOMA uplink system. In the simulation, the system setup of the network is shown in Fig.5.4, where the BS and the IRS are placed at (0m,0m) and (90m,20m) in the coordinate axis, respectively. The device $s \in \mathcal{S}$ is randomly and uniformly located in a circle area centred at (0m,0m) with radius 20m, as the green circle shown in Fig.5.4. Similarly, the circle area for device $w \in \mathcal{W}$ is centred at (110m,0m) with the same radius, as the orange circle shown in Fig. 5.4.

In addition, other simulation parameter are listed in Table 5.1. The channel coefficient are modelled with both large-scale and small-scale fading components. The large-scale fading model is defined as $\sqrt{\mathrm{PL}_0 d^{\beta}}$, where PL_0 is the path loss at unit distance, d stands for the distance of the communication link, and β is the path loss exponent. Particularly, β_{direct} refers to the path loss exponent of the channel from devices to the BS; β_{cascade} denote the exponent of channels from device $w \in \mathcal{W}$ to the IRS and channel from the IRS to the BS. The small-scale fading is assumed to be Rayleigh fading.

Table 5.1: System Parameters

Parameter	Value
Number of devices $2I$	16
Number of IRS elements M	10
Noise power σ_i^2	-90 dBm
Maximum transmit power $p_{\max}$	27 dBm
SINR threshold Γ_{th}	20 dB
Path loss at unit distance PL_0	10^{-2}
Path loss exponent (UE_i to BS) β_{direct}	-3.5
Path loss exponent (UE_w to IRS, IRS to BS) β_{cascade}	-2

5.6.1 Simulation Results

The original problem **P1** is divided into user clustering, phase shift optimization, power allocation problems and solved separately. Then two methods are proposed to address the non-convex problem, namely, close-form optimization and DDPG-based algorithm. In order to evaluate the performance of both approaches, we simulate the sum average AoI of the system using two approaches separately and then compare closed-form optimization and DRL-based approach. In addition, we also consider the following baseline algorithm as benchmarks:

- Baseline 1: as the author of [122] proposed, the IRS phase shift is optimized via convex optimization first, and users in $\mathcal{S}$ are allocated with maximum power. Then, the power allocated to users in $\mathcal{W}$ is random values that satisfy inequality in (5.27). Finally, exhaustive search is applied to obtain optimal user pairing solution.

- Baseline 2: similar to baseline 1, user clustering and IRS phase shift are optimized using same algorithm, but users in $\mathcal{S}$ are allocated with $p_{\max}$ and the power for users in $\mathcal{W}$ are optimized based on (5.26) [122].

- Baseline 3: in this benchmark, random IRS phase shift vector is generated, which replaces the result of convex optimization or the result in action space of DRL-based algorithm.

- Baseline 4: no IRS is exploited in this system. The power allocation is optimized using the same methods as the close-form optimization does.

- Baseline 5: OMA scheme is applied as this benchmark. Particularly, devices transmit with BS using OFDMA, meaning that only single user can transmit in a time

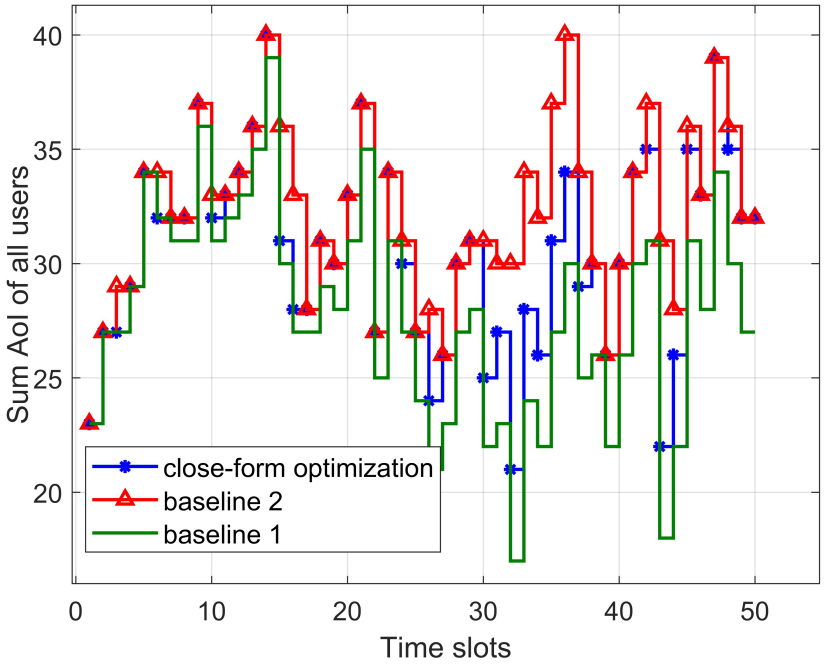


Figure 5.5: Sum AoI of all users at each time slot when Γ_{th} =20dB.

slot within a subcarrier.

Fig. 5.5 shows how AoI changes from time slot $t = 0$ to $t = 50$. In particular, when simulated individual SINR is greater than the threshold, 15 dB, which means the transmission is successful, the AoI of this user will be updated to be one. In another word, the stair plot will go lower. If the transmission fails, the stair plot will go upward. It can be observed from this figure, baseline 2 performs the worst, because when users in $\mathcal{W}$ are allocated with maximum power, users in $\mathcal{S}$ suffer from higher interference due to SIC, as shown in (5.1). In addition, the proposed closed-form optimization algorithm performs slightly worse than than of baseline 1, since the proposed user clustering method is a sub-optimal algorithm. Despite the loss of accuracy, comparing the proposed devices pairing strategy and exhaustive search in baseline 1, complexity can be significantly reduced. As the main goal of this work is to minimize AoI with lower power consumption, total power consumption should also be considered. Fig. 5.6 presents the changing of power consumption from time slot $t = 0$ to $t = 50$. Since in baseline 1, maximum power is allocated to users in $\mathcal{S}$, total power consumption is apparently higher than the proposed closed-form optimization.

To further compare the proposed algorithm with baseline 1, we investigate the average power consumption and average AoI. We simulate three schemes, namely, closed-form

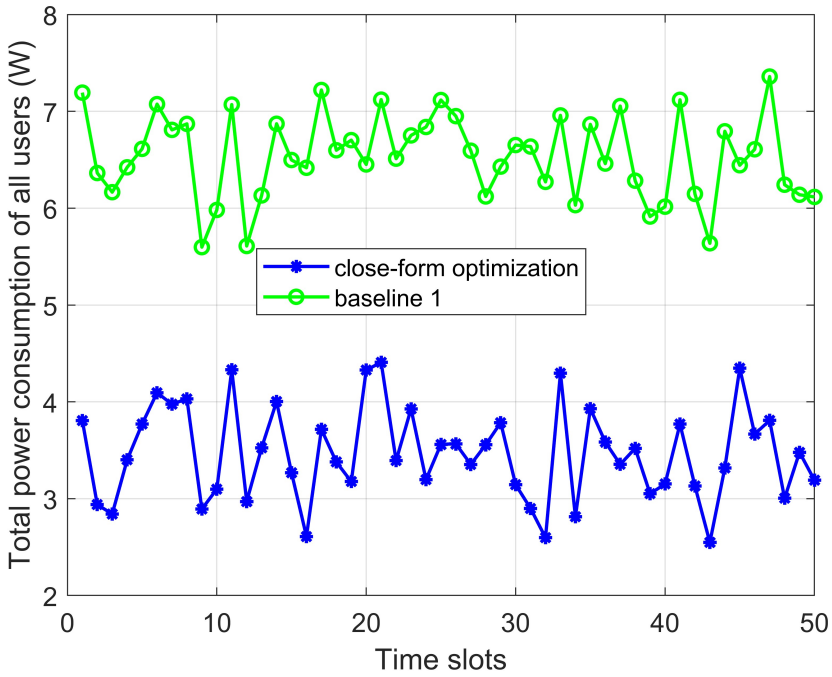


Figure 5.6: Average sum AoI when Γ_{th} =20dB.

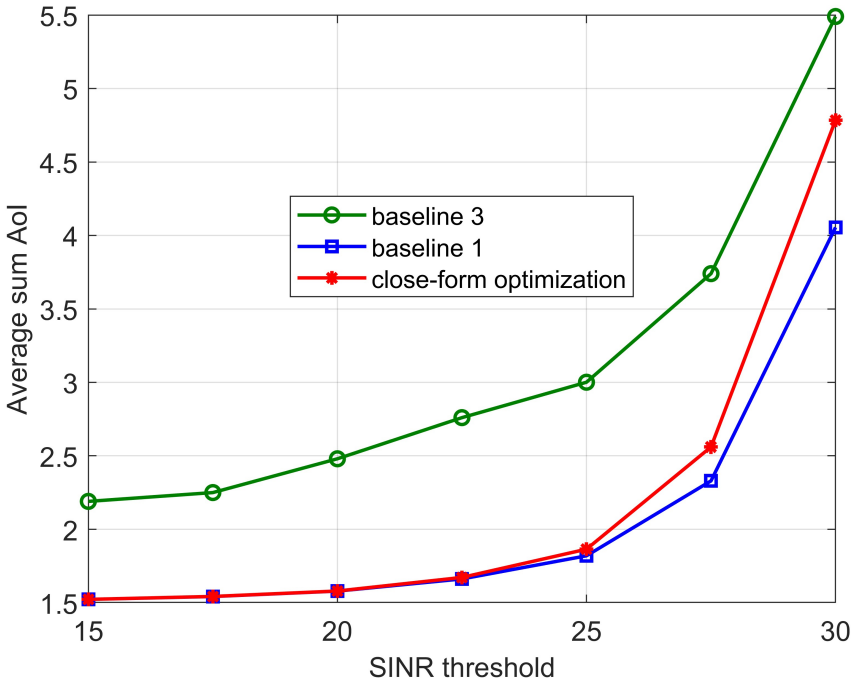


Figure 5.7: Average sum AoI with different SINR threshold Γ_{th}.

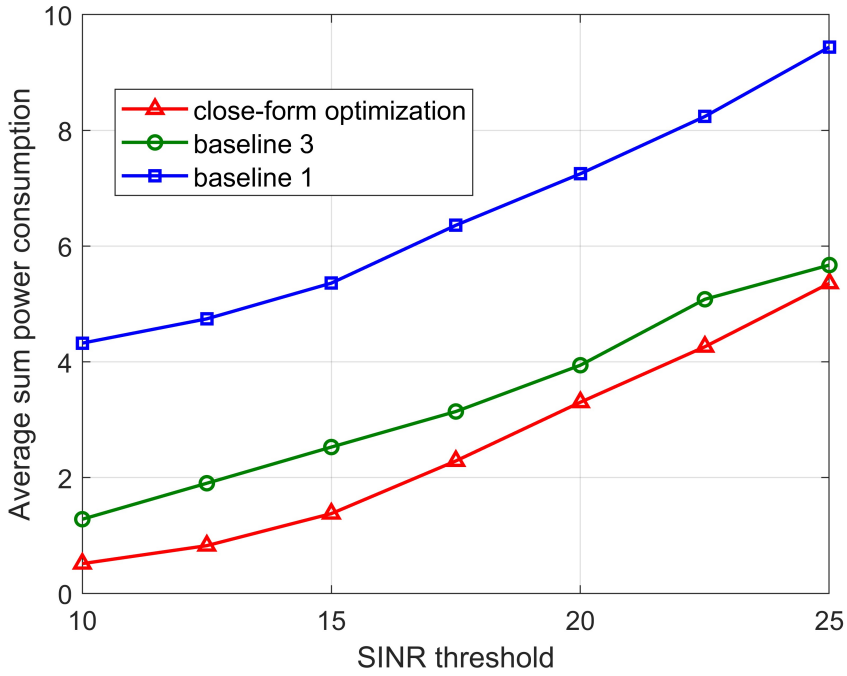


Figure 5.8: Average total power consumption with different SINR threshold Γ_{th}.

optimization, baseline 1 and baseline 3, over 500 time slots. Fig. 5.7 shows average sum AoI when SINR threshold increase. It can be observed that when SINR threshold increases, average AoIs of all schemes rise. Moreover, the closed-form optimization algorithm performs better than random devices pairing in baseline 3, but worse than exhaustive search in baseline 1 because of the same reason mentioned above. As for average power consumption, Fig. 5.8 illustrates that the proposed algorithm consumes the least power among these three schemes.

In addition, to demonstrate the impact of IRS, we investigate the average AoI with different numbers of IRS elements. Particularly, as Fig. 5.9 shown, IRS is able to significantly improve the performance of the system in terms of AoI minimization, comparing baseline 4 to the proposed closed-form optimization. On the other hand, exhaustive search still can obtain the optimal AoI compared to the proposed approach.

As for the training and simulation of DDPG agent in the proposed DRL-based algorithm, the hyperparameters are listed in the Table. 5.2.

Fig. 5.10 illustrates the convergence of average reward in the training process of DDPG agent in the formulated environment. The SINR thresholds for both scenarios are set to be 25 dB and 15 dB, respectively. As observed, agents in both scenarios converge at the end of training, but due to the varying channel condition, it takes much more

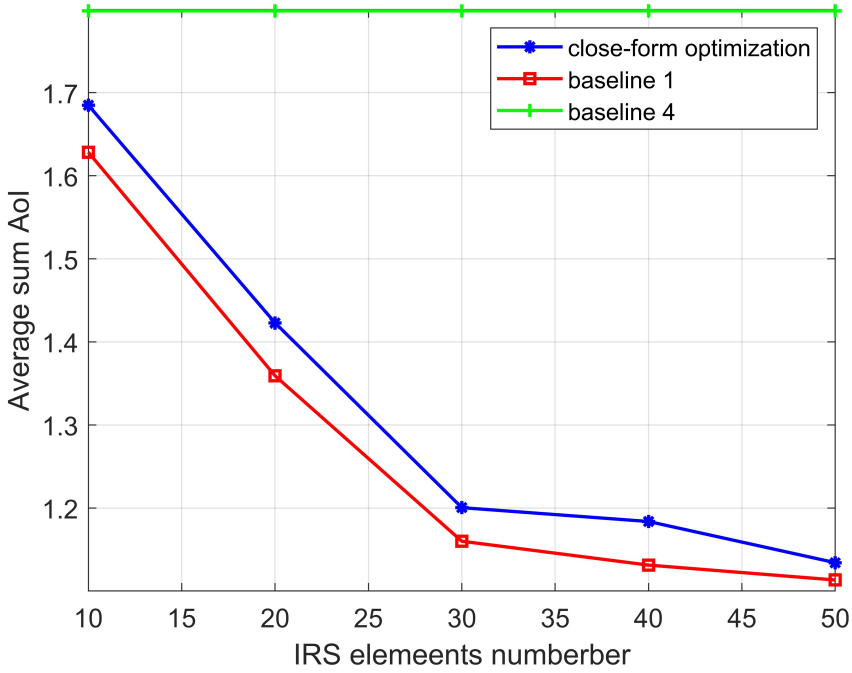


Figure 5.9: Average AoI with different IRS elements number.

Table 5.2: Hyperparameters for DDPG agent

Hyperparameter	**Value**
Actor learning rate (fixed/dynamic channels)	0.0007/0.0007
Critic learning rate (fixed/dynamic channels)	0.001/0.0008
Discount factor	0.99
Batch size	128
Replay buffer size ($\mathcal{B}$)	100000
Target smooth factor (fixed/dynamic channels)	0.0005/0.0005
Target update frequency	1
Number of episodes (fixed/dynamic channels)	800/1800
Number of steps per episode (fixed/dynamic channels)	300/500

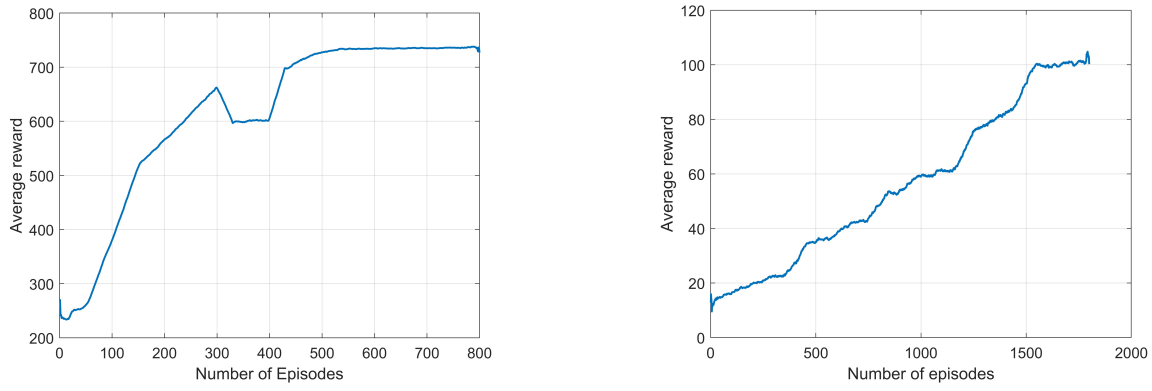


Figure 5.10: Convergence of DDPG agent with fixed (left) and dynamic (right) channel scenarios.

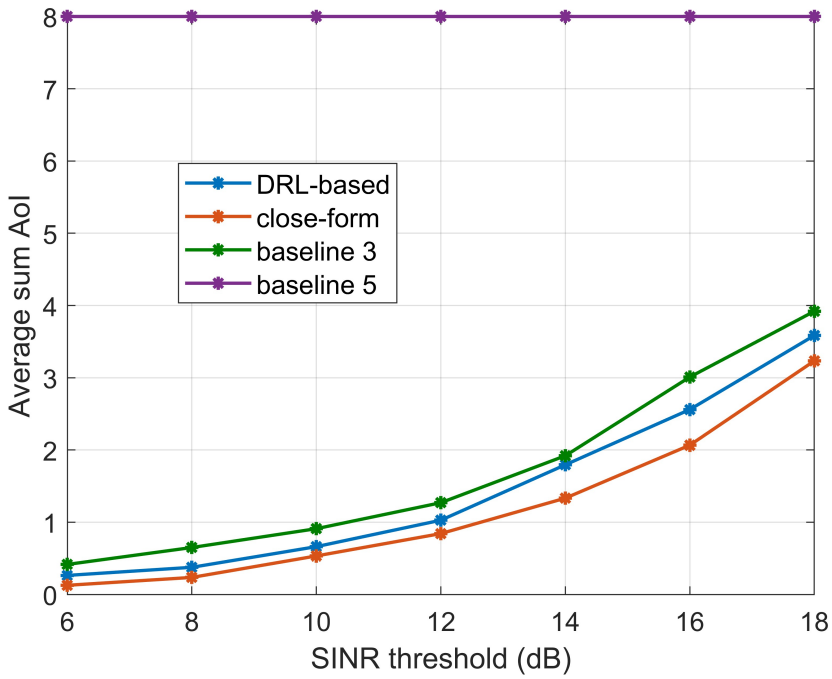


Figure 5.11: Average sum AoI with different SINR threshold.

episodes for the agent to learn the optimal policy.

To further evaluate the performance of trained DDPG agent, we simulate the agent trained with varying channel using 1000 different channel sets. The average sum AoI is shown in Fig. 5.11. It can be observed that the trained DDPG can obtain lower AoI than random IRS phase shift in baseline 3, but it is slightly worse than closed-form optimization method. Note that although in OMA scheme, there is no interference between devices, only single device is able to send its signal in a time slot within a subcarrier. Thus, at least half of devices cannot transmit successfully. It is also worth mentioning that despite the training target SINR is 15 dB, DDPG agent can still handle the resource allocation problem with different SINR. As for power consumption, Fig. 5.12 illustrates the average power consumption comparison between DRL-based algorithm and closed-form optimization. Closed-form optimization method consumes less power. As can be observed, closed-form optimization outperforms the DRL-based algorithm in terms of both AoI and power consumption. The reason of this is the closed-form algorithm uses convex optimization approach, which is much more accurate than DRL-based agents. However, DRL-based algorithm is significantly more efficient than conventional convex optimization based approaches. Specifically, once the agent is well trained, it can deal with the similar problem without any computational complexity. Besides, since the orig-

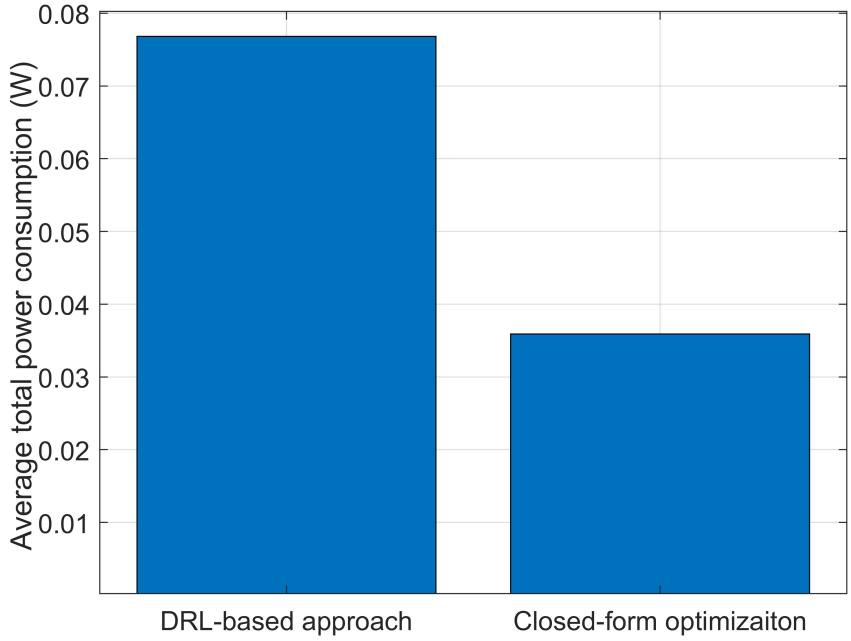


Figure 5.12: Average total power consumption.

inal problem cannot be solved directly, accurate power allocation and IRS phase shift cannot be obtained. Therefore, the data-based training is not considered.

5.7 Summary

In this work, a resource allocation problem for uplink IRS-aided NOMA network was considered in this chapter. Particularly, the problem minimized average sum AoI consuming less power and clustering devices subject to unit-modulus constraint for IRS elements. Due to the non-convexity of the original problem, we proposed two approaches to solve it. First, devices were paired according to their channel gains. Then, to obtain optimal phase shift and power allocation, we apply a close-form optimization, wherewith the IRS phase shift is optimized using convex optimization and power allocation is obtained with mathematical transform. Another algorithm is based on DRL, in which DDPG agent is trained to provide optimal phase shift and power jointly. Numerous simulation results confirmed that both algorithm only provided higher AoI than that of existing literature, but are able to significantly reduce powerw consumption. Comparing closed-form optimization and DRL-based algorithm, close-form optimization performs better in terms of AoI and power. On the other hand, DRL-based algorithm considerably decrease the

complexity while achieving a similar performance.

Chapter 6

Sum-rate Maximization for IRS-aided URLLC-NOMA System

6.1 Introduction

In recent years, the performance of short packets communication has attracted considerable interest due to delay intolerant applications. In particular, URLLC is one of the key techniques of 5G and beyond networks, which is designed to support applications that require extremely reliable communication with minimal delay, such as autonomous vehicles and remote surgery [149]. The most existing literature is on URLLC using OFDMA scheme [150], where the resource allocation problem is divided into sub-problems and solved iteratively. However, there are seldom works integrating NOMA and URLLC. The integration of NOMA and URLLC can combine the advantaged of both techniques, including high spectrum efficiency, lower latency and also improving the reliability of URLLC users. Thus, in this work, we study the resource allocation DRL-based algorithm for downlink NOMA-URLLC systems, where both BS and user are equipped with single antenna. The contribution of this work can be summarized as follows:

- We consider a IRS-aided NOMA-URLLC system and formulate a resource allocation problem optimizing NOMA user clustering, power allocation and IRS phase shift to maximize sum throughput of the system subject to QoS constraint for individual user and IRS unit-modulus constraint.

- The original formulated problem is non-convex and difficult to be solved. Thus, we

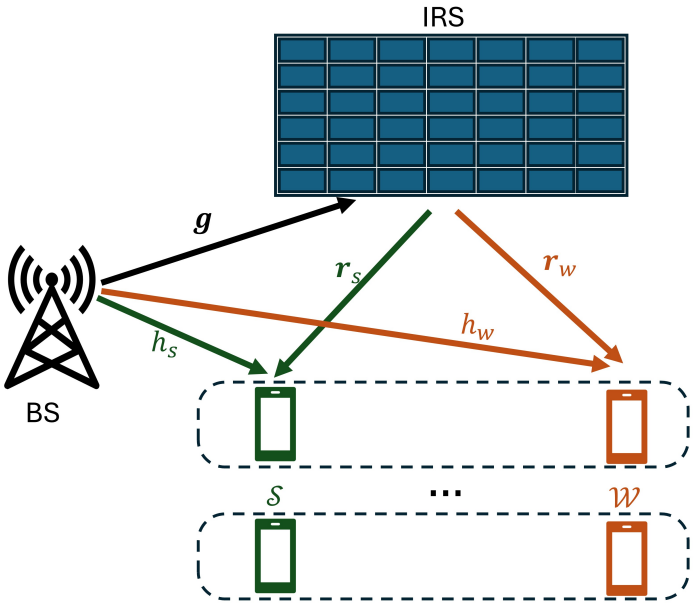


Figure 6.1: A cluster-based multi-user SISO IRS-aided NOMA system.

first cluster the users based on NLUPA, then propose a DRL-based algorithm to obtain the optimal power allocation and IRS phase shift much more efficiently.

- Numerous simulation results for both fixed and dynamic channel show the performance of the proposed algorithm. Furthermore, the results illustrate the DDPG-based design outperforms conventional sub-optimal benchmark schemes.

6.2 System Model

A downlink transmission of IRS-aided NOMA system as shown in Fig. 6.1 is considered, where the BS and all the user equipments are equipped with single antenna and the IRS is equipped with M reflecting elements. The IRS elements are denoted by index $m \in \mathcal{M}$, where $\mathcal{M} = \{1, 2, \cdots, M\}$, which assists the downlink transmission from the BS to users of number of $2I$. The users are equally separated into two groups according to their channel strengths. Particularly, assuming the users are sorted in a descending order according to their channel gains, such that $|h_{\mathrm{UE}_1}| \geq \cdots \geq |h_{\mathrm{UE}_I}| \geq \cdots \geq h_{\mathrm{UE}_{2I}}$, users from UE_1 to UE_I are grouped into a strong group, which is denoted by $\mathcal{S}$, while users from UE_{I+1} to UE_{2I} are grouped into a weak group, which is denoted by $\mathcal{W}$. Let the coefficients of channel from the BS to UE_s and from the BS to UE_w denote by $h_s \in \mathbb{C}$ and $h_w \in \mathbb{C}$, respectively, where $s \in \mathcal{S}$ and $w \in \mathcal{W}$. According to this grouping strategy,

the following inequality holds: $|h_s| \geq |h_w|$, $\forall s \in \mathcal{S}, \quad w \in \mathcal{W}$. The channel coefficients from the BS to the IRS, from the IRS to UE_s and from the IRS to UE_w are defined by $\mathbf{g} \in \mathbb{C}^M$, $\mathbf{r}_s \in \mathbb{C}^M$ and $\mathbf{r}_w \in \mathbb{C}^M$, respectively. These channel coefficients consists of both large-scale and small-scale fading, where the large-scale fading is defined by path loss and the small-scale fading is defined as Rayleigh fading. In particular, the channel coefficients can be expressed by $h = \sqrt{\text{PL}_0 d^{-\alpha}}\hat{h}$, where $h \in \{h_s, h_w, \mathbf{g}, \mathbf{r}_s, \mathbf{r}_w\}$ represents all the channels in this network. The square root part indicates the large scale fading components, where PL_0 denotes the path loss at unit distance, d represents the distance of the relative wireless communication link, and α is the path loss exponent. The small-scale fading part is modelled as Rayleigh fading with zero means and unit variance, $\hat{h} \in \mathcal{CN}(0, \mathbf{I})$, where $\mathbf{I}$ refers to identity matrix. In addition, to enhance the performance of the whole system, an IRS with m elements deployed. The phase shift matrix of IRS is defined as $\mathbf{\Theta} = \text{diag}(\beta_1 e_1, \beta_2 e_2, \cdots, \beta_M e_M)$, where β_m and e_m, $\quad \forall m \in \mathcal{M}$ are amplitude and phase shift of the m^{th} element of IRS, respectively. Without loss of generality, we assume the fixed amplitude shift, such that $\beta_i = 1, \quad \forall i \in \mathcal{M}$.

All the $2I$ users are divided into I clusters, pairing one user from the strong group $\mathcal{S}$ and one user from the weak group $\mathcal{W}$. Each cluster is assigned with a specific orthogonal subcarrier to avoid inter-cluster interference. In each cluster, NOMA scheme is applied to mitigate the interference between those two users, UE_s and UE_w. Let assume that UE_s and UE_w are paired in a cluster. Then, based on SC, the transmitted signal from the BS in a particular subcarrier to those two users in the same cluster, by allocating different power levels,

$$x = \sqrt{p_s} s_s + \sqrt{p_w} s_w, \tag{6.1}$$

where s_s and s_w represent the signal intended for UE_s and UE_w, respectively. The power of both signals are defined to be one, which is expressed by $\mathbb{E}\{s_i\} = 1, \quad \forall i \in \{s, w\}$. Thus, the received signal at UE_i is given by

$$y_i = \left(h_i + \mathbf{r}_i^H \mathbf{\Theta} \mathbf{g}\right) \left(\sqrt{p_s} s_s + \sqrt{p_w} s_w\right) + n_i, \quad \forall i \in \{s, w\}. \tag{6.2}$$

where $n_i \in \mathcal{CN}(0, \sigma^2)$ is the AWGN with zero mean and variance σ^2. Based on the SIC at the receiver ends, UE_s is able to decode the signal intended for UE_w and subtract it from

the received signal before decoding its own signal. Hence, the interference from UE_w can be eliminated. As for UE_w, due to its poorer channel strength, it can only decode its own signal by considering the signal intended for UE_s as interference. Therefore, the SINRs for paired UE_s and UE_w can be expressed respectively as

$$\Gamma_s = \frac{p_s|h_s + \mathbf{r}_s^H \boldsymbol{\Theta} \mathbf{g}|^2}{\sigma^2}, \tag{6.3}$$

$$\Gamma_w = \frac{p_w|h_w + \mathbf{r}_w^H \boldsymbol{\Theta} \mathbf{g}|^2}{p_s|h_w + \mathbf{r}_w^H \boldsymbol{\Theta} \mathbf{g}|^2 + \sigma^2}. \tag{6.4}$$

In general, the achievable rate for any typical/conventional resource allocation problem is defined based on well-known information theoretic Shannon's capacity theorem, where the packet length is approximately infinity and decoding error probability approaches to zero [119]. However, URLLC system transmits short packets to achieve lower latency and but it suffers from decoding error. Hence, Shannon's capacity theorem is not an appropriate performance metric for any type of URLLC network. To evaluate the performance of short packets communication, a performance framework is introduced in [120]. In particular, the maximum number of bits Ψ conveyed in a packet with L symbols can be approximately expressed as [151]:

$$\Psi = \sum_{i=1}^{L} \log 2(1 + \Gamma[i]) - Q^{-1}(\epsilon) \sqrt{\sum_{i=1}^{L} a^2 \left(1 - \frac{1}{(1+\Gamma[i])^2}\right)}, \tag{6.5}$$

where ϵ is the decoding error probability, Q^{-1} is the inverse Q-function, and $\Gamma[i]$ is the SINR for the i^{th} symbol.

In this work, we consider a URLLC system with IRS-aided NOMA. Unlike conventional OMA schemes, such as OFDMA and TDMA, the short packets are divided and transmitted via different resource blocks, namely time or frequency. By applying the cluster-based NOMA to the considered URLLC system, UEs are allowed to transmit through whole time slots and two users in each cluster can share the assigned subcarrier. Thus, the achievable data rate for UE_i in this URLLC system with IRS-aided NOMA URLLC network is given by

$$R_i = \log_2(1 + \Gamma_i) - Q^{-1}(\epsilon) \sqrt{a^2 \left(1 - \frac{1}{(1+\Gamma_i)^2}\right)}, \quad \forall i \in \{\mathcal{S}, \mathcal{W}\}. \tag{6.6}$$

6.3 Problem Formulation

In the considered URLLC system with IRS-aided NOMA, we consider a resource allocation problem to maximize the sum rate with a given total power individual data rate constraints. Particularly, we are aiming to determine the optimal power allocation at the BS, IRS phase shift and user clustering to achieve maximal sum data rate under phase shift unit-modulus and individual rate constraints, which is given by

$$\mathbf{P1} \quad \max_{\mathbf{p},\mathbf{B},\mathbf{\Theta}} \quad \sum_{s=1}^{I}\sum_{w=1}^{I}(R_s + R_w)b_{s,w} \tag{6.7}$$

$$\text{s.t.} \quad \sum_{s=1}^{I}\sum_{w=1}^{I}(p_s + p_w) \leq p_{\max}, \tag{6.8}$$

$$R_s \geq R_{\text{th}}, R_s \geq R_{\text{th}}, \quad \forall s \in \mathcal{S}, \forall w \in \mathcal{W}, \tag{6.9}$$

$$|[\mathbf{\Theta}]_{m,m}| = 1, \forall m \in \mathcal{M}, \tag{6.10}$$

where the design variables for power allocation is represented by a vector $\mathbf{p} = \{p_i | i \in \{\mathcal{S}, \mathcal{W}\}\}$ and user clustering matrix $\mathbf{B} = \{b_{s,w} | s \in \mathcal{S}, w \in \mathcal{W}\}$. $b_{s,w}$ is defined as a binary factor. When an UE_s from $\mathcal{S}$ and an UE_w from $\mathcal{W}$ are paired in a cluster, $b_{s,w} = 1$; otherwise, $b_{s,w} = 0$. The constraint in (6.8) ensures that the total power consumption is lower than the total available power $p_{\max}$. Constraint (6.9) guarantees that all the users can be served with transmission data rate higher than a predefined threshold R_{th}. In addition, constraint (6.10) confirms that IRS can only passively reflect signal without changing its strength or amplitude. The above power allocation problem is non-convex due to coupled nature of design variables and non-convex constraint, which makes the problem challenging to realise a feasible solution. Therefore, we utilize a DRL-based approach, the details of which are provided in the next subsection.

6.4 DRL-based Solution Approach

As mentioned above, the original problem **P1** is challenging to solve jointly in terms of power allocation coefficients, phase shifts of IRS and user clustering. Therefore, we divide the original problem **P1** into two sub-problems: 1) user clustering problem and 2) power allocation and phase shift optimization problem. Firstly, the user clustering problem is

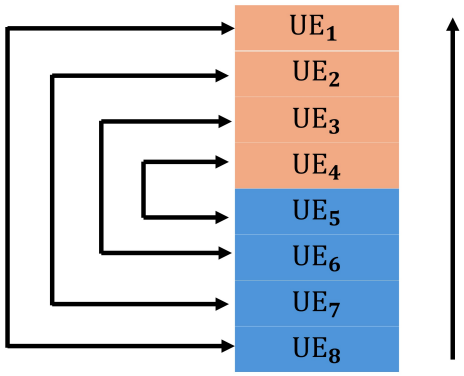


Figure 6.2: Illustration of user pairing scheme based on NLUPA.

solved using an appropriate user pairing strategy, namely, NLUPA. Then, for a given user cluster, we propose a DRL-based approach to jointly optimize optimal power allocation and IRS phase shift. The following subsections provide the details of user clustering approach as well as DRL based power allocation and phase shift optimization approach.

6.4.1 User Clustering

For clustering UEs based on the channel strengths, the downlink NOMA prefers NLUPA [147], which aims to guarantee the minimum channel gains of stronger user and weaker user in a cluster. Therefore, the sum rate performance can be achieved in each cluster . In particular, as shown in Fig. 6.2,NLUPA pairs the UE with strongest channel gain with the UE with weakest channel gains in a cluster to maintain a higher channel gain difference between the stronger UEs. In the downlink NOMA, based on SIC, the users with strongest channel strength is able to decode and subtract the signals from the weaker users, as shown in (6.3). Therefore, after SIC, user with highest channel gain does not suffer from interference of others and its throughput only depends on its channel gain and allocated power. In order to contribute more to the sum throughput of the considered system, it is necessary to assigned a UE with high channel gain into different clusters. As for the data rate of UEs with weak channel gain, they still struggle with interference from others, as shown in (6.4). Hence, to mitigate the interference more efficiently, it

is beneficial to pair a UE with weak channel gain with a UE with strong channel gain. Considering that UEs with high channel strength are able to achieve high throughput even with lower power allocation, the interference suffered from the UEs with weak channel gain can be eliminated. As result, the proposed user grouping for dowonlink NOMA network is shown in Fig.6.2, where the user equipments denotes by $\mathrm{UE}_i, \forall i \in \{1, \cdots, 8\}$. All the UEs are sorted in a descending order based on their channel gains, which is described by $|h_{\mathrm{UE}_1}| \geq |h_{\mathrm{UE}_2}| \geq \cdots |h_{\mathrm{UE}_8}|$. Then, the UE$i$s are divided into two groups $\mathcal{S}$ and $\mathcal{W}$, shown by two different colors, orange and blue in Fig. 6.2, respectively. In this grouping strategy, we pair the first strongest UE of group $\mathcal{S}$ and the first weakest UE of group $\mathcal{W}$ is assigned to the first cluster, while the second strongest UE of group $\mathcal{S}$ and the second weakest UE in group $\mathcal{W}$ in the second cluster, and so on. In this way, it is guaranteed that UEs with strong channel gains can contribute more to the system sum throughput, meanwhile introducing less interference to UEs with weak channel gain. Then, the user clustering scheme for given $2I$ users with their channel coefficients is summarized in Algorithm 4.

Algorithm 4 User pairing algorithm

Input: channel coefficients $h_i, \quad i \in \{1, \cdots, I\}$
Sort $2I$ users, UE_i, based on their channel gains: $|h_1| \geq |h_2| \geq \cdots \geq |h_i| \geq \cdots \geq |h_{2I}|, \quad \forall i \in \mathcal{I}$, where h_i denotes the channel gains for UE_i.
if $i \leq I$ **then**
    $\mathrm{UE}_i \in \mathcal{S}$
else
    $\mathrm{UE}_i \in \mathcal{W}$
end if
1^{st} cluster $= \{\mathrm{UE}_1, \mathrm{UE}_{I+1}\}$
$\cdots$
i^{th} cluster $= \{\mathrm{UE}_i, \mathrm{UE}_{I+i}\}, \quad \forall i \in \{1, \cdots I\}$
$\cdots$
I^{th} cluster $= \{\mathrm{UE}_I, \mathrm{UE}_{2I}\}$
Output: clustering binary factor $b_{s,w}, \quad \forall s \in \mathcal{S}, \quad \forall w \in \mathcal{W}$

6.4.2 DRL Environment

In the previous subsection, the clustering strategy based on NLUPA is presented which means that the design parameter, the binary variable $b_{s,w}$ in the original problem is determined. In other words, the $2I$ users are equally divided into groups $\mathcal{S}$ and $\mathcal{W}$ based

on their channel gains, and then they are paired into I clusters. With the given user pairs, the binary variables $b_{s,w}$ can be removed without loss of generality, and the original problem **P1** can be further reformulated into the following optimization problem:

$$\mathbf{P2} \quad \max_{\mathbf{p},\boldsymbol{\Theta}} \quad \sum_{s=1}^{I}\sum_{w=1}^{I}(R_s + R_w) \tag{6.11}$$

$$\text{s.t.} \quad \sum_{i=1}^{2I} p_i \leq p_{\max}, \tag{6.12}$$

$$R_i \geq R_{\text{th}} \quad \forall i \in \{\mathcal{S}, \mathcal{W}\}, \tag{6.13}$$

$$(6.10)$$

It is challenging to solve this optimization problem **P2**, due to the multiple coupled optimization variables as well as due to the non-convexity of the objective function. Thus, we propose a DRL-based approach to realise a low-complex feasible solution to **P2**, which is described in detail as follows.

The optimization problem **P2** is formulated as a MDP, where state space, action space, as well as an appropriate reward function are defined. To solve this MDP, we utilize a DRL framework, which comprises an agent and an environment. Specifically, at each time step n, the agent selects an action $\mathbf{a}^n$ based on the current state $\mathbf{s}^n$ provided by the environment. The environment, in turn, generates a new state $\mathbf{s}^{n+1}$ for the next time step and a corresponding reward R^n. Through repeated interactions with the environment, the agent learns an optimal policy $\pi^*(\mathbf{s}, \mathbf{a})$, which is basically able to map each state to an optimal action that maximizes the cumulative reward. Consequently, the DRL framework transforms the optimization problem **P2** into a sequence of actions that the agent takes to maximize the objective. In particular, we define the action vector (action space) $\mathbf{a}^n$, the state vector (state space) $\mathbf{s}^n$, and the reward function R^n as follows:

- The action space $\mathbf{a}$: To make the optimization problem simpler for the agent, the power allocation p_w for users $\text{UE}_s, \forall s \in \mathcal{S}$, and p_w for $\text{UE}_w, \forall w \in \mathcal{W}$, in a same cluster are expressed in terms of new variables, such as

$$p_w = p_{\text{sub}_i} a_w, \tag{6.14}$$

$$p_s = p_{\text{sub}_i}(1 - a_w), \quad \forall i \in \{1, \cdots, I\}, \tag{6.15}$$

where p_{sub_i} is the power allocated to the whole subcarrier, a_w stands for the ratio of power for user UE_w and p_s, p_w are the power allocated to UE_s and UE_w, respectively. Since the variables of the optimization problem **P2** consist of both power allocation and IRS phase shift, both of them are considered in the action space vector. Then, the action space vector at step n is defined as

$$\mathbf{a}^n = \left[(\mathbf{p}_{\text{sub}}^n)^T, (\mathbf{a}^n)^T, (\tilde{\mathbf{e}}^n)^T\right]^T, \tag{6.16}$$

where $\mathbf{p}_{\text{sub}}^n = \left\{p_{\text{sub}_i}^n | i \in \{1, \cdots, I\}\right\}$ and $\mathbf{a}^n = \{a_w^n | w \in \mathcal{W}\}$. It is worth mentioning that DNN frameworks are applied in this DRL-based approach and only real numbers are accepted in both action and sate space vectors. Define that the phase shift vector $\mathbf{e} \in \mathbb{C}^M$, which satisfies $\mathbf{\Theta} = \text{diag}(\mathbf{e})$. Then, we introduce a new vector $\tilde{\mathbf{e}}$, which consists of both real and imaginary part of IRS phase shift vector at the n^{th} step $\mathbf{e}^n$, such that $\tilde{\mathbf{e}}^n = \left[\Re\{\mathbf{e}^n\}^T, \Im\{\mathbf{e}^n\}^T\right]^T$. As result, the action space vector at the n^{th} time step can be expressed as $\mathbf{a}^n \in \mathbb{R}^{2I+2M}$ with only real numbers.

- The state space $\mathbf{s}$: Since the original problem **P1** aims to maximize the sum throughput of the considered system for given CSI, the channel coefficients h_is, $\mathbf{r}_i$s and $\mathbf{g}$, $\forall i \in \{\mathcal{S}, \mathcal{W}\}$, should be considered in the state vector. In addition, in order to help agent learn better the problem and reach the optimal policy faster, action vector for previous time step is also included. In particular, to further simplify the size of state space vector, the phase shift vector for time step n is extracted from the previous action space and substituted into coefficients for the communication link through IRS, leading to total channel gains f_i for all the users at time step n, which is given by

$$f_i^n = |h_i + \mathbf{r}_i^H \mathbf{\Theta}^n \mathbf{g}| = |h_i + \mathbf{r}^H \text{diag}(\mathbf{e}^n)\mathbf{g}|, \forall i \in \{\mathcal{S}, \mathcal{W}\}.$$

Note that to avoid complex number in state space vector, absolute value of channel coefficients are taken. Considering f_i already includes previous phase shift vector, only power allocation for subcarriers and power ratio for users in $\mathcal{W}$ should be added. Hence, the sate vector at step n is defined as

$$\mathbf{s}^n = \left[(\mathbf{f}^n)^T, (\mathbf{p}_{\text{sub}}^n)^T, (\mathbf{a}^n)^T\right]^T, \tag{6.17}$$

where $\mathbf{f}_i^n = \{f_i^n | i \in \{1, \cdots, 2I\}\}$, and $\mathbf{s}^n \in \mathbb{R}^{4I}$.

- The reward function r^n: As shown in the problem **P2**, the optimization problem aims to maximize the achievable sum rate for the considered URLLC system. Hence, the reward function is expressed as follows:

$$r^n = \begin{cases} \sum_{s=1}^{I} \sum_{w=1}^{I} \left(R_s^n + R_w^n\right), & \text{if} \quad R_s, R_w \geq R_{th} \\ \sum_{s=1}^{I} \sum_{w=1}^{I} \left(R_s^n + R_w^n\right) - 2R_{th}, & \text{otherwise,} \end{cases} \tag{6.18}$$

  where the conditional function shows the award and punishment to the agent. In particular, if the QoS of all the users are guaranteed, such that their achievable data rates are higher than required target data rates, the reward function is expressed as the sum data rate. On the other hand, if the throughput of one of users is lower, the reward function is represented as the negative value of the difference between the achieved sum data rates and the sum target rates.

To ensure that constraints in (6.12) and (6.10) are satisfied, the power allocated to clusters and the IRS phase shift vector are normalized as follows:

$$\bar{p}_{\text{sub}_i}^n = p_{\max} \frac{p_{\text{sub}_i^n}}{\sum_{i=1}^{I} p_{\text{sub}_i}^n}, \quad i \in \{1, \cdots, I\}, \tag{6.19}$$

$$[\bar{\mathbf{e}}^n]_m = \frac{[\mathbf{e}^n]_m}{|[\mathbf{e}^n]_m|}, \quad \forall m \in \mathcal{M}, \tag{6.20}$$

where $\bar{p}_{\text{sub}_i}^n$ and $\bar{\mathbf{e}}^n$ are the scaled power allocation for i^{th} cluster and the m^{th} element of phase shift vector, respectively.

6.4.3 DDPG-based Approach

In the previous subsection, the problem in **P2** has been reformulated into a RL environment, and based on the fundamental framework Of DDPG agent introduced in Chapter 3, we propose a DDPG-based approach to solve the sub-problem P2 and the developed algorithm is summarized in Algorithm 5.

Algorithm 5 DDPG-based power allocation and phase shift optimization algorithm for IRS-aided NOMA URLLC network

Initialize: DDPG agent's hyperparameters for actor DDN μ, critic DDN ψ, action noise vecotr $\mathbf{n}$, replay buffer $\mathcal{B}$, step number $n = 0$
Update hyperparameters for target DDNs: $\mu^{'} \leftarrow \mu$ and $\psi^{'} \leftarrow \psi$
while $\text{no.}_{\text{epi}} \leq \text{No.}_{\text{epi}}$ **do**
    Obtain the channel coefficients $h_i(t)$, $\mathbf{r}_i(t)$, $\forall i \in \{\mathcal{S}, \mathcal{W}\}$, $\mathbf{g}(t)$ from current sate $\mathbf{s}^n$
    Cluster the devices based on algorithm 4
    Set random initial action vector $\mathbf{a}^0$ and calculate state vector based on (6.17)
    while $n \leq N_{\text{step}}$ **do**
        Get action vector $\mathbf{a}^n$ based on (6.16)
        Normalize the action vector to calculate power allocation $\bar{\mathbf{p}}_{\text{sub}}^n$ and phase shift $\bar{\mathbf{e}}^n$ according to (6.19) and (6.20), respectively
        Calculate the SINR for each device according to (5.1) and (6.4)
        Calculate reward r based on the function in (6.18)
        Obtain the next-step state vector $\mathbf{s}^{n+1}$ from the environment
        Save the tuple $\{\mathbf{s}^n, \mathbf{a}^n, r^n, \mathbf{s}^{n+1}\}$ to $\mathcal{B}$
        Sample a example from $\mathcal{B}$
        Calculate target ξ based on (5.38)
        Train critic DDN using (5.39)
        Update the target actor and critic DDNs
        Update the target DDNs using (5.41)
        Step number $n = n + 1$
    end while
    $\text{no.}_{\text{epi}} = \text{no.}_{\text{epi}} + 1$
end while
Obtain the optimal power allocation $\bar{\mathbf{p}}^*$ and IRS phase shift $\bar{\mathbf{e}}^*$

6.5 Simulation Results

In this section, the parameters of DDPG agent and the considered URLLC system with IRS-aided NOMA are provided. Then, the training data and simulation results for both

Table 6.2: Parameters for IRS-aided NOMA URLLC network

system parameter	**value**
Cell radius	120 m
Number of IRS elements	10
Number of users $2I$	8
Number of Clusters I	4
Total transmit power	40 dBm
noise power	-110 dBm
Decoding error probability (ϵ)	0.1
Path loss (BS to users) α_1	4
Path loss (BS to IRS, IRS to users) α_2	2.2
Ttarget data rate (R_{th}) (fixed channels)	2.0 Bits/s/Hz
Ttarget data rate (R_{th}) (dynamic channels)	1.0 Bits/s/Hz

fixed channel and dynamic channel scenarios are presented.

6.5.1 Agent and System Parameters

The Hyperparameters for DDPG agent are provided in Table 6.1. Note that the default structure of the DDPG agent is defined in Matlab, where the number of neurons in each hidden layer is set to be 400.

Table 6.1: Hyperparameters for DDPG agent

Hyperparameter	**Value**
Actor learning rate (fixed/dynamic channels)	0.0005/0.0005
Critic learning rate (fixed/dynamic channels)	0.0007/0.0006
Discount factor	0.99
Batch size	128
Replay buffer size ($\mathcal{B}$)	100000
Target smooth factor (fixed/dynamic channels)	0.0005/0.0006
Target update frequency	1
Number of episodes (fixed/dynamic channels)	800/3000
Number of steps per episode (fixed/dynamic channels)	300/800

As for the system model, we consider a downlink transmission of a URLLC system with IRS-aided NOMA, which is defined using the following parameters shown in Table 6.2.

As mentioned above, the channel model of all the communication links are considered with large-scale fading and small-scale fading components. Specifically, the large-scale fading is expressed with path loss, where the path loss exponent of channel from BS directly to users is given by $\alpha_1 = 4$ and the path loss exponent of channels from BS to IRS as well as from IRS to users are given by $\alpha_2 = 2.2$.

In addition, to validate the performance of proposed algorithm, we use a benchmark scheme, the details of which are provided below:

- Baseline: In this benchmark scheme, the same user clustering algorithm, NLUPA, is applied. Then, the problem **P2** is divided into two sub-problems to optimize IRS phase shift and power allocation individually. Firstly, the phase shift is optimized via a closed-form solution approach, which is formulated as follows

$$\begin{aligned} \mathbf{P3} \quad & \max_{\boldsymbol{\Theta}(t)} \quad \min_{i \in \{1,\cdots,2I\}} |h_i + \mathbf{r}_i^H \boldsymbol{\Theta} \mathbf{g}|^2 \\ & \text{s.t.} \quad (6.10), \end{aligned} \tag{6.21}$$

  which can be converted into a SDP problem similar to **P3** in Chapter 5 and can be solved by utilizing SDR. The complexity of this SDP problem is given by $\mathcal{O}(M^6)$. For power allocation, the total transmit power is equally allocated to each cluster, while in all clusters, 80% of p_{sub_i} is allocated to user with weaker channel gains, which can be mathematically expressed as

$$p_{\text{sub}_i} = p_{\max}/I, \quad \forall i \in \{1, \cdots, I\} \tag{6.22}$$

$$a_w = 0.8, \quad \forall w \in \mathcal{W} \tag{6.23}$$

6.5.2 Fixed Channel Scenario

Firstly, we consider a simulation scenario in which the channel strengths for all the users are fixed throughout all the episodes. The convergence of the proposed agent is shown in Fig. 6.3, where the DDPG agents are trained with different target data rate. The change of the DDPG agent's policy can be observed, indicating that the agent is able to reach sufficiently high reward after a number of training episodes. The achievable throughput of the whole URLLC system is shown in Fig. 6.4. It is obvious that the proposed DDPG

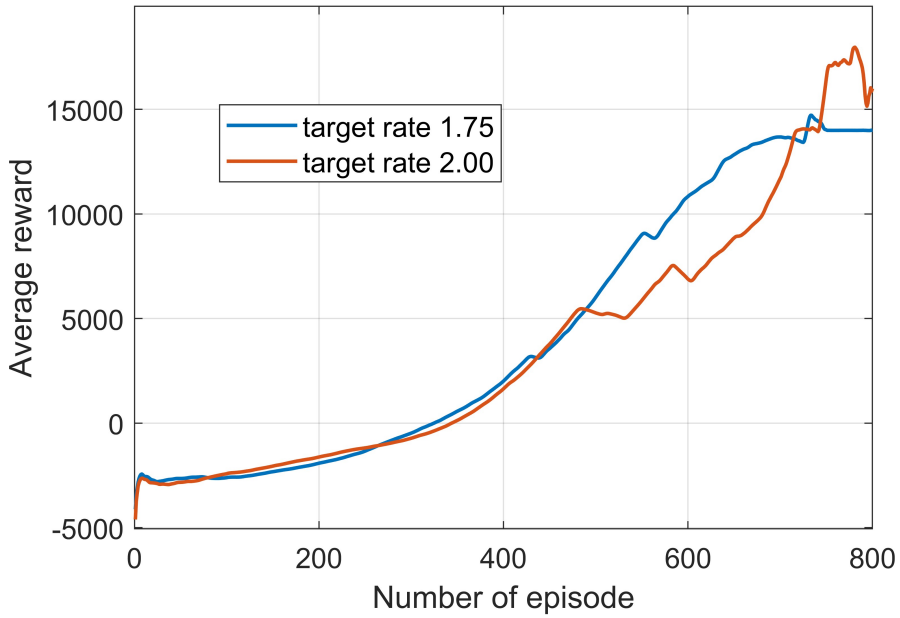


Figure 6.3: Convergence of the proposed DDPG agent for the fixed-channels scenario.

based approach achieve a better system sum throughput than that of the the baseline methods.

6.5.3 Dynamic Channel Scenario

However, the simulation results for fixed channel scenario can only show partial performance of the proposed algorithm and is not practical in scenario in other dynamic environments, which is the case in the most of the wireless communication systems. The channel is frequently varying in practice and creates dynamic environments, especially when users are moving. Hence, we consider dynamic channel condition case to further investigate the proposed DDPG approach. In this dynamic channel case, the agents are trained with a set of 15 channels. Furthermore, unlike the fixed-channel scenario, the UEs are randomly distributed in the cell radius. Therefore, an optimal policy, which is generalized to deal with never-seen-before channels, can be developed by the agent.

Fig. 6.5 shows the convergence of three DDPG agents trained with different target rates, 0.75 bits/s/Hz, 1.00 bits/s/Hz, 1.25 bits/s/Hz, respectively. After a certain number of episodes, all the agents are able to settle down to a high-reward policy. However, with the increase of training target rate, the required number of episodes number is also increasing, indicating agent requires more time to reach to an optimal policy. It is

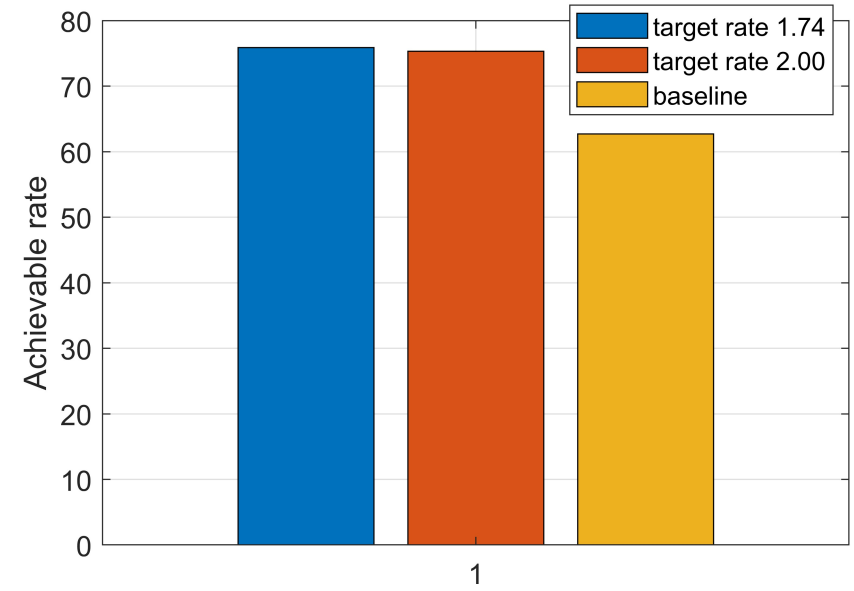


Figure 6.4: Bar reward plot of proposed DDPG agent for the fixed-channels scenario.

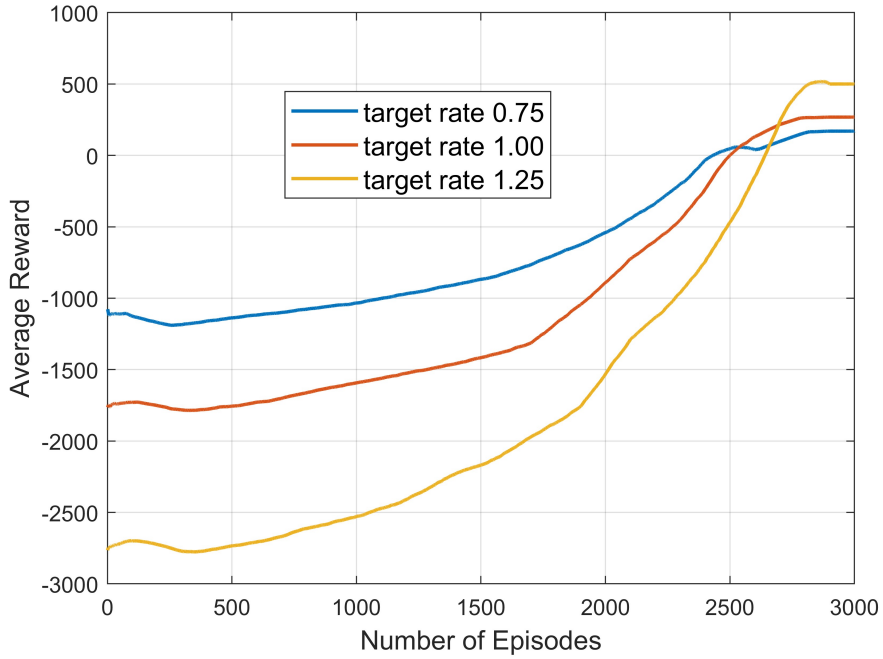


Figure 6.5: Convergence of the proposed DDPG agent for the dynamic-channels scenario.

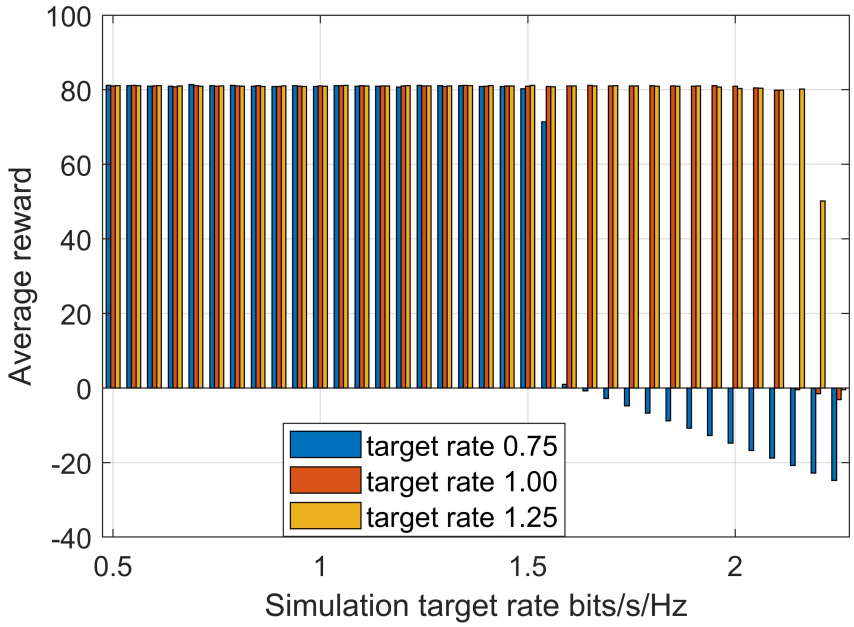


Figure 6.6: The simulation reward of three trained agent tested against varying target rates.

also worth mentioning that the average reward of agent trained with higher target rate convergences to a higher value. The reason is that the agent has potential to achieve high data rate, and setting higher target can efficiently enhance the agent's performance.

To further evaluate the performance of the agents, we simulate them using different sets of channels. Fig. 6.6 illustrates the average rewards with different target rates. Note that these three agents are trained with fixed target rates, but they are tested against varying target rates. It can be observed that all the agents can achieve positive and high reward when the target rate is lower or slightly higher than the training target. However, when the the target rate is relatively high, the policies for these agents turn to provide negative rewards. In addition, the outage probabilities of agents providing positive rewards are shown in Fig. 6.7. It can be observed from these simulation results that even though the agents are trained with fixed target rate, they are still able to deal with dynamic scenario with lower or higher target rate in simulation. Particularly, the agents trained with proposed algorithm can manage to offer high reward even with target rate double higher than the trained target rate. Note that this is impossible to realize with conventional optimization approaches, which highlight the benefits of deploying DRL agent to solve this type of resource allocation problems.

Fig. 6.8 presents the impact of total transmit power against the achievable sum data

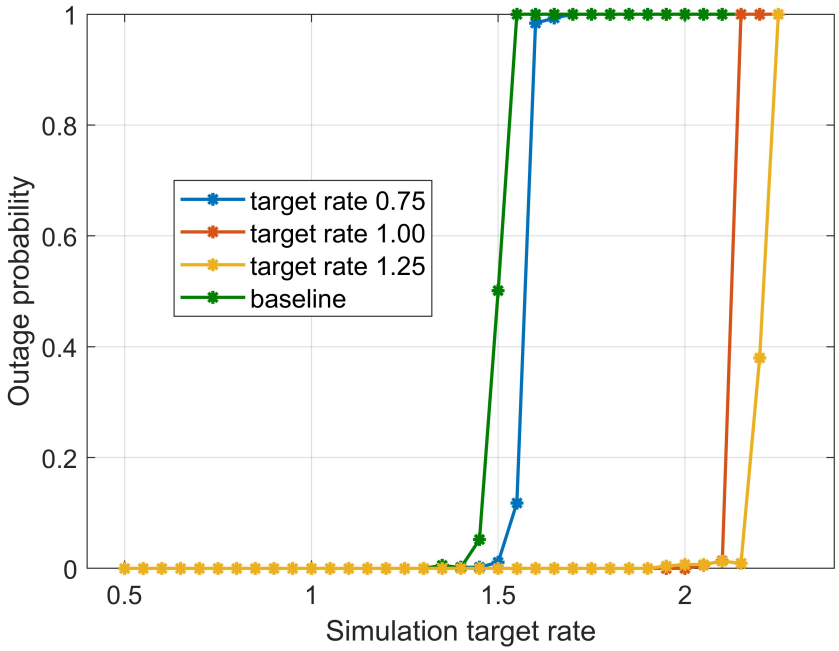


Figure 6.7: The outage probability of three trained agent tested against varying target rates.

rate. In this simulation, we define the throughput of each user as

$$R_i = \begin{cases} R_i, & \text{if } \mathbf{P2} \text{ is feasible} \\ 0, & \text{otherwise} \end{cases}, \quad \forall i \in \{\mathcal{S}, \mathcal{W}\} \tag{6.24}$$

The three trained agents are able to achieve similar performance when the available transmit power increase. As can be observed, below a certain value of total available transmit power, the achievable sum throughput is almost zero. The reason is that the transmit power is not sufficient to support the required target rates of individual user and all the users also suffer from the decoding error probability in the considered URLLC network. Furthermore, the agents keep providing near-zero throughput until the transmit power reaches about 29 dBm, while the baseline algorithm offers positive throughput when transmit power is beyond which is around 32 dBm. Thus, the proposed DRL-based approach outperform the benchmark scheme as well as it adopts easily to different dynamic environments with different target rates.

Fig. 6.9 and 6.10 show the influence of decoding error probability ϵ. As shown in (6.6), ϵ directly impact the throughput for each user. In particular, in both plots, with the increase of ϵ, the both reward and sum throughput increase. The reason of that

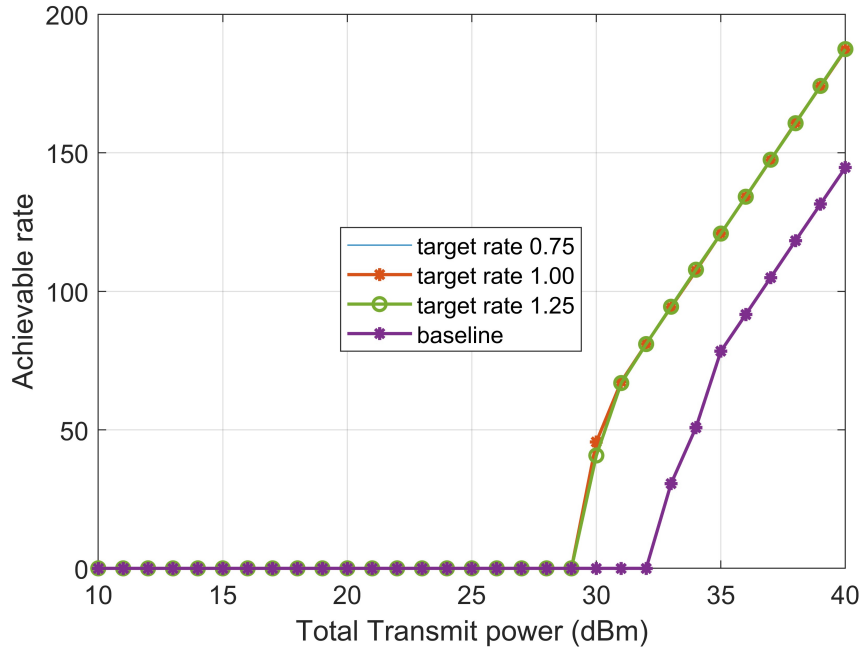


Figure 6.8: The sum rate with different available total transmit power at the BS.

is when ϵ increase, the value of the inverse function, Q^{-1}, decrease leading to higher individual data rate. Note that agents are trained with $\epsilon = 0.1$. Therefore, the agents can even handle the scenario with lower error probability. Furthermore, it is the same case for baseline benchmark, but the trained agent with DDPG approach by proposed algorithm can always yield a higher system sum rate.

6.6 Summary

A resource allocation problem for a URLLC system with IRS-aided NOMA is considered in this chapter. In particular, a sum throughput maximization problem under unit-modulus constraint for IRS elements and individual rate requirement constraint for each user is formulated. Since the original problem is not convex and contains multiple design variables, it is challenging to be solved directly. Hence, the original problem is separated into user clustering and power phase shift optimization problems. We first propose NLUPA to address the user clustering. Then, to jointly obtain optimal power allocation and IRS phase shift, conventional convex optimization method is unable to solve this issue. Thus, we formulate the original sum rate maximization problem as a DRL environment and use DDPG agent to solve the reformulated problem. Number of simulation

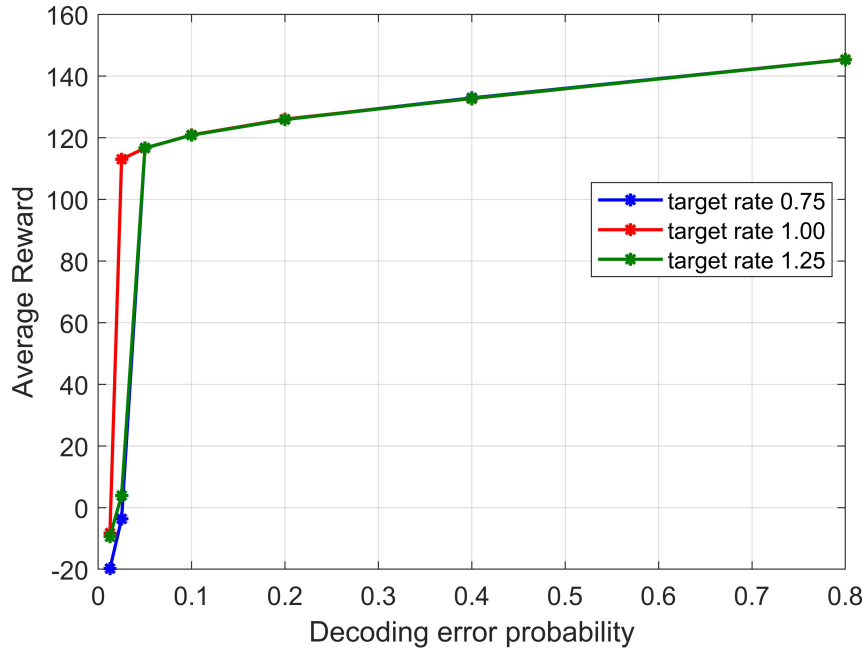


Figure 6.9: The average reward with different decoding error probability.

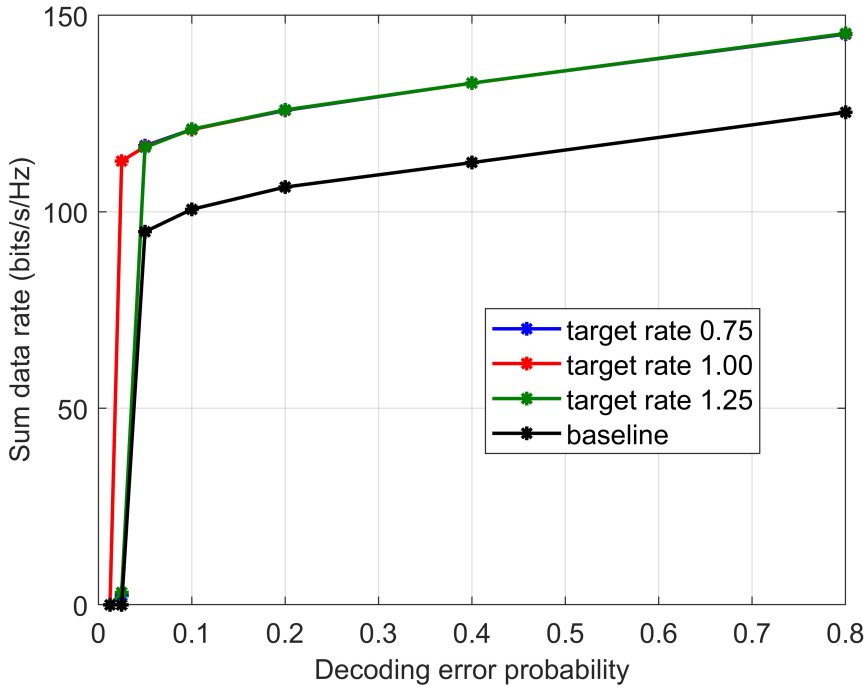


Figure 6.10: The sum rate with different decoding error probability.

results illustrate the agents are well trained, reaching higher rewards and outperform the benchmark sub-optimal algorithm. Specifically, the proposed DDPG based approach can achieve higher throughput with the same power consumption, or higher throughput with same decoding error probability. In addition, even the DDPG agents are trained with fixed target rate and decoding error, they still are able to provide high-reward policy when these parameters shift in a small range in other words, in dynamic environments.

Chapter 7

Future Work and Conclusion

7.1 Conclusion

NOMA has been considered a promising MA technique for 6G or beyond networks because of its numerous benefits, including higher spectral and energy efficiencies, fairness, and compatibility with other techniques, such as OMA, IRS. Unlike conventional OMA, such as TDMA and OFDMA, in which only a single user occupies a single time or frequency resource block, NOMA enables multiple users to share the whole resource. Particularly, multiple users are allowed to transmit their signals via the same frequency band simultaneously. In power-domain NOMA, by applying SC at the transmitters, signals of multiple users can be multiplexed together with different power levels according to their channel strengths, while at the receivers, SIC is exploited at stronger channel gain users to decode the signals intended for weak channel users to remove their interference. Due to the compatibility of NOMA, it can be integrated with multiple other techniques, such as OMA schemes and multiple antenna techniques. In addition, IRS is another revolutionary technique capable of signal reflection. Adding additional amplitude or phase shift to the intended signals allows reconfiguring the communication link in an IRS-aided network. Combining NOMA and IRS in a network, the interference suffered by weak channel users can be further eliminated. Therefore, energy efficiency and spectrum efficiency can be significantly improved. In order to design IRS-aided NOMA networks, resource allocation is a driving force. In particular, the allocation of resources is formulated into an optimization problem, which obtains optimal power allocation, IRS phase shift, and user clustering to minimize/maximize the objective function, such as power consumption or

sum throughput. Hence, this thesis focuses on different resource allocation problems for IRS-aided NOMA networks.

In conclusion, this thesis presents a comprehensive investigation into resource allocation techniques for IRS-aided NOMA networks, addressing key challenges in power efficiency, information freshness, and throughput maximization. By integrating IRS with NOMA, the proposed methods effectively enhance spectral and energy efficiency while reducing interference, making them suitable for 6G and beyond networks. The research works formulate and solve multiple optimization problems for different system, including downlink MISO-NOMA, uplink NOMA-based IoT, and URLLC-NOMA networks. Particularly, in the downlink MISO-NOMA system, the proposed AO algorithm efficiently minimizes total power consumption while satisfying outage-probability-based QoS constraints. For the uplink NOMA-based IoT network, the thesis introduces a sum AoI minimization framework, where IRS improves the transmission efficiency of weaker channel users. To solve the complex resource allocation problem, both convex optimization and a DDPG-based DRL algorithm is employed, demonstrating the effectiveness of DRL in reducing complexity while achieving near-optimal performance. Finally, the URLLC-NOMA system addresses sum-rate maximization through DRL-based solution, achieving higher reliability and lower latency. The simulation results consistently illustrated that IRS-aided NOMA significantly enhances system performance compared to conventional NOMA and OMA systems. These contributions offer valuable insights and practical algorithms for next-generation wireless networks, leading the way for more adaptive resource allocation strategies.

7.2 Future Works

This thesis has contributed considerably to resource allocation techniques for IRS-aided NOMA networks. However, several investigations remain to extend current research and future works, presented in the following.

7.2.1 Multiple Antennas Technique and Imperfect CSI

The current works in Chapter 5 and Chapter 6 in this thesis assume the BS are equipped with single antennas and are able to obtain perfect CSI. However, in the practical wire-

less communication system for 6G or beyond, the BS is usually equipped with multiple antennas. In addition, due to the existence of IRS, the signals are reflected passively. Therefore, errors are highly likely to occur in the channel estimations and quantization, which can significantly reduce the system's overall performance. Specifically, in NOMA networks, since both SC and SIC depend mainly on CSI, channel uncertainty can considerably increase the design for NOMA networks. Furthermore, integrating multiple antenna techniques and IRS-aided NOMA networks also increases the complexity, reforming the power allocation problem into a beamforming design problem. In summary, further investigation of robust beamforming designs for IRS-aided MISO-NOMA networks to maximize sum URLLC throughput and to minimize sum average AoI are considered as future directions to extend the current works.

7.2.2 TD3-Agent-Based Resource Allocation Techniques

The algorithm applied in works in Chapter 5 and Chapter 6 mainly depends on DDPG agent. To further extend the current works, a more effective agent, namely TD3 agent, can be exploited to address resource allocation problems. Though DDPG agents are able to tackle the joint optimization of power allocation and IRS phase shift, they still tend to overestimate bias and instability during training due to the use of a single critic network for estimating Q-values. TD3 agents address this issue by applying three improvements. Firstly, TD3 agents employ additional critic DNN, which estimate the targets separately to reduce bias. Secondly, delay policy update is introduced, where the actor-network is updated less frequently than the critic DNNs, leading to more stable updates and improving convergence. Finally, target policy smoothing is applied by adding noise to target actions, which reduces errors in the Q-value estimation and further improves learning stability. Thus, with the help of TD3 agents, the user clustering, power allocation, and IRS phase shift are able to be optimized jointly to significantly raise the accuracy of the solution for resource allocation problems in current works.

7.2.3 Resource Allocation for ISAC-NOMA networks

ISAC indicates an emerging technique that combines sensing and communication functionalities. Instead of considering both sensing and communication as separate entities. This technique enables the simultaneous sharing of resources for both functions, which

enhances overall efficiency. By optimizing the resource allocation for ISAC systems, the integration improves spectrum utilization and allows communication systems to benefit from enhanced sensing data. On the other hand, NOMA allows multiple users to transmit simultaneously by sharing the whole resource, enhancing connectivity and improving resource efficiency. Particularly, in power-domain NOMA, SC multiplexes signals for multiple users at transmitters and SIC are applied at receivers to eliminate interference. Motivated by this discussion, a downlink IRS-aided NOMA-ISAC system is considered an interesting research direction.

7.2.4 User Clustering Strategy for IRS-NOMA networks

In NOMA, users with different channel conditions share the same frequency or time resources, and their signals are decoded by applying SIC. However, efficiently user clustering strategy is crucial for maximizing energy efficiency, throughput and fairness. Traditional clustering methods often rely on fixed or heuristic approaches, which is suboptimal in dynamic environments. By applying ML algorithms, such as DRL, the system can automatically and adaptively group users based on real-time channel conditions and QoS requirements. This AI-driven approach enables the network to dynamically optimize user clustering, reduce interference, and optimize overall resource allocation jointly, making NOMA systems in practical deployments.

References

[1] A. Goldsmith, *Wireless Communications*. Cambridge university press, 2005.

[2] W. Brown, "The History of Power Transmission by Radio Waves," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, pp. 1230–1242, Sep. 1984.

[3] P. Sharma, "Evolution of mobile wireless communication networks-1G to 5G as well as future prospective of next generation communication network," *International Journal of Computer Science and Mobile Computing*, vol. 2, no. 8, pp. 47–53, 2013.

[4] C. Ciochina and H. Sari, "A Review of OFDMA and Single-carrier FDMA," in *2010 European Wireless Conference (EW)*, pp. 706–710, April 2010.

[5] L. J. Vora, "Evolution of Mobile Generation Technology: 1G to 5G and Review of Upcoming Tireless Technology 5G," *International journal of modern trends in engineering and research*, vol. 2, no. 10, pp. 281–290, 2015.

[6] H. Honkasalo, K. Pehkonen, M. Niemi, and A. Leino, "WCDMA and WLAN for 3G and Beyond," *IEEE Wireless Communications*, vol. 9, pp. 14–18, April 2002.

[7] L. L. Hanzo, L.-L. Yang, E.-L. Kuan, and K. Yen, *CDMA Overview*, pp. 35–80. IEEE, 2004.

[8] Y. Kim and R. Prasad, "4G Roadmap and Emerging Communication Technologies (Universal Personal Communications), Artech House," *Inc., Norwood, MA, USA*, 2006.

[9] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-s. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," *IEEE Communications Surveys Tutorials*, vol. 19, pp. 721–742, Secondquarter 2017.

[10] M. Rebhi, K. Hassan, K. Raoof, and P. Chargé, "Sparse Code Multiple Access: Potentials and Challenges," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1205–1238, 2021.

[11] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-Splitting Multiple Access: Fundamentals, Survey, and Future Research Trends," *IEEE Communications Surveys Tutorials*, vol. 24, pp. 2073–2126, Fourthquarter 2022.

[12] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, "On the Road to 6G: Visions, Requirements, Key Technologies, and Testbeds," *IEEE Communications Surveys Tutorials*, vol. 25, pp. 905–974, Secondquarter 2023.

[13] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies," *IEEE Vehicular Technology Magazine*, vol. 14, pp. 28–41, Sep. 2019.

[14] M. Series, "Minimum requirements related to technical performance for IMT-2020 radio interface (s)," *Report*, vol. 2410, pp. 2410–2017, 2017.

[15] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter Wave Communication: A Comprehensive Survey," *IEEE Communications Surveys Tutorials*, vol. 20, pp. 1616–1653, thirdquarter 2018.

[16] B. K. J. Al-Shammari, I. Hburi, H. R. Idan, and H. F. Khazaal, "An Overview of mmWave Communications for 5G," in *2021 International Conference on Communication Information Technology (ICICT)*, pp. 133–139, June 2021.

[17] N. U. Saqib, M. S. Haroon, H. Y. Lee, K. Park, H.-G. Song, and S.-W. Jeon, "THz Communications: A Key Enabler for Future Cellular Networks," *IEEE Access*, vol. 11, pp. 117474–117493, 2023.

[18] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 742–758, Oct 2014.

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

[20] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the Uplink Max–Min SINR of Cell-Free Massive MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 2021–2036, April 2019.

[21] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "Energy Efficiency of the Cell-Free Massive MIMO Uplink With Optimal Uniform Quantization," *IEEE Transactions on Green Communications and Networking*, vol. 3, pp. 971–987, Dec 2019.

[22] A. S. Elgamal, O. Z. Aletri, B. A. Yosuf, A. Adnan Qidan, T. El-Gorashi, and J. M. H. Elmirghani, "AI-Driven Resource Allocation in Optical Wireless Communication Systems," in *2023 23rd International Conference on Transparent Optical Networks (ICTON)*, pp. 1–5, July 2023.

[23] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal Multiple Access for 5G and Beyond," *Proceedings of the IEEE*, vol. 105, pp. 2347–2381, Dec 2017.

[24] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A Survey of Non-Orthogonal Multiple Access for 5G," *IEEE Communications Surveys Tutorials*, vol. 20, pp. 2294–2323, thirdquarter 2018.

[25] Z. Chen, X. Ma, C. Han, and Q. Wen, "Towards intelligent reflecting surface empowered 6G terahertz communications: A survey," *China Communications*, vol. 18, pp. 93–119, May 2021.

[26] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent Reflecting Surface-Aided Wireless Communications: A Tutorial," *IEEE Transactions on Communications*, vol. 69, pp. 3313–3351, May 2021.

[27] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network: Joint Active and Passive Beamforming Design," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec 2018.

[28] D. K. Pin Tan, J. He, Y. Li, A. Bayesteh, Y. Chen, P. Zhu, and W. Tong, "Integrated Sensing and Communication in 6G: Motivations, Use Cases, Requirements, Challenges and Future Directions," in *2021 1st IEEE International Online Symposium on Joint Communications Sensing (JCS)*, pp. 1–6, Feb 2021.

[29] K. Kim, J. Kim, and J. Joung, "A Survey on System Configurations of Integrated Sensing and Communication (ISAC) Systems," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 1176–1178, Oct 2022.

[30] Y. Liu, S. Zhang, X. Mu, Z. Ding, R. Schober, N. Al-Dhahir, E. Hossain, and X. Shen, "Evolution of NOMA Toward Next Generation Multiple Access (NGMA) for 6G," *IEEE Journal on Selected Areas in Communications*, vol. 40, pp. 1037–1071, April 2022.

[31] Z. Liu and L.-L. Yang, "Sparse or Dense: A Comparative Study of Code-Domain NOMA Systems," *IEEE Transactions on Wireless Communications*, vol. 20, pp. 4768–4780, Aug 2021.

[32] S. H. Amin, A. H. Mehana, S. S. Soliman, and Y. A. Fahmy, "User Capacity in Downlink MISO-NOMA Systems," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, Dec 2018.

[33] M. B. Dadi and B. C. Rehaimi, "Performance of downlink MIMO-NOMA system in 5G Networks," in *2023 IEEE International Workshop on Mechatronic Systems Supervision* $(IW_MSS), pp.\ 1 - -4, Nov2023.$

[34] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy-Efficient Resource Allocation for Downlink Non-Orthogonal Multiple Access Network," *IEEE Transactions on Communications*, vol. 64, pp. 3722–3732, Sep. 2016.

[35] R. Zhang, X. Kang, and Y.-C. Liang, "Minimum Throughput Maximization for Peer-Assisted NOMA-Plus-TDMA Symbiotic Radio Networks," *IEEE Wireless Communications Letters*, vol. 10, pp. 1847–1851, Sep. 2021.

[36] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A New Wireless Communication Paradigm through Software-Controlled Metasurfaces," *IEEE Communications Magazine*, vol. 56, pp. 162–169, Sep. 2018.

[37] J. Y. Dai, Q. Cheng, and T. J. Cui, "IRS Hardware Architectures," *Intelligent Reconfigurable Surfaces (IRS) for Prospective 6G Wireless Networks*, pp. 83–98, 2022.

[38] E. Björnson, Özdogan, and E. G. Larsson, "Intelligent Reflecting Surface Versus Decode-and-Forward: How Large Surfaces are Needed to Beat Relaying?," *IEEE Wireless Communications Letters*, vol. 9, pp. 244–248, Feb 2020.

[39] H. Ur Rehman, F. Bellili, A. Mezghani, and E. Hossain, "Joint Active and Passive Beamforming Design for IRS-Assisted Multi-User MIMO Systems: A VAMP-Based Approach," *IEEE Transactions on Communications*, vol. 69, pp. 6734–6749, Oct 2021.

[40] V. B. Shukla, V. Bhatia, and K. Choi, "Cascaded Channel Estimator for IRS-Aided mmWave Hybrid MIMO System," *IEEE Wireless Communications Letters*, vol. 13, pp. 622–626, March 2024.

[41] 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)*, pp. 1–5, June 2013.

[42] S. Tomida and K. Higuchi, "Non-Orthogonal Access with SIC in Cellular Downlink for User Fairness Enhancement," in *2011 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS)*, pp. 1–6, Dec 2011.

[43] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, pp. 2–14, January 1972.

[44] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, and H. V. Poor, "Application of Non-Orthogonal Multiple Access in LTE and 5G Networks," *IEEE Communications Magazine*, vol. 55, pp. 185–191, February 2017.

[45] E. Gelal, J. Ning, K. Pelechrinis, T.-S. Kim, I. Broustis, S. V. Krishnamurthy, and B. D. Rao, "Topology Control for Effective Interference Cancellation in Multiuser

MIMO Networks," *IEEE/ACM Transactions on Networking*, vol. 21, pp. 455–468, April 2013.

[46] 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, pp. 2181–2195, Oct 2017.

[47] C. Shannon, "The Zero Error Capacity of a Noisy Channel," *IRE Transactions on Information Theory*, vol. 2, pp. 8–19, Sep. 1956.

[48] Z. Chen, Z. Ding, X. Dai, and R. Zhang, "An Optimization Perspective of the Superiority of NOMA Compared to Conventional OMA," *IEEE Transactions on Signal Processing*, vol. 65, pp. 5191–5202, Oct 2017.

[49] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, and P. A. Hoeher, "Multiple-Antenna Techniques for Wireless Communications - A Comprehensive Literature Survey," *IEEE Communications Surveys Tutorials*, vol. 11, pp. 87–105, Second 2009.

[50] B. Clerckx, Y. Mao, R. Schober, E. A. Jorswieck, D. J. Love, J. Yuan, L. Hanzo, G. Y. Li, E. G. Larsson, and G. Caire, "Is NOMA Efficient in Multi-Antenna Networks? A Critical Look at Next Generation Multiple Access Techniques," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1310–1343, 2021.

[51] Y. Liu, H. Xing, C. Pan, A. Nallanathan, M. Elkashlan, and L. Hanzo, "Multiple-Antenna-Assisted Non-Orthogonal Multiple Access," *IEEE Wireless Communications*, vol. 25, pp. 17–23, April 2018.

[52] B. Kim, W. Chung, S. Lim, S. Suh, J. Kwun, S. Choi, and D. Hong, "Uplink NOMA with Multi-Antenna," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1–5, May 2015.

[53] S. Menaa and A. Khelil, "On the Performance of SISO, SIMO and MISO-NOMA Systems under Perfect and Imperfect SIC," in *2022 19th International Multi-Conference on Systems, Signals Devices (SSD)*, pp. 458–463, May 2022.

[54] A. Kucar, "Modeling of FDMA, TDMA, CDMA and Handover Mobile Radio Channels," in *IEEE 43rd Vehicular Technology Conference*, pp. 863–866, May 1993.

[55] M. Aldababsa, M. Toka, S. Gökçeli, G. K. Kurt, and O. Kucur, "A Tutorial on Non-Orthogonal Multiple Access for 5G and Beyond," *Wireless communications and mobile computing*, vol. 2018, no. 1, p. 9713450, 2018.

[56] M. E. Morocho-Cayamcela, H. Lee, and W. Lim, "Machine Learning for 5G/B5G Mobile and Wireless Communications: Potential, Limitations, and Future Directions," *IEEE Access*, vol. 7, pp. 137184–137206, 2019.

[57] Q. Wu and R. Zhang, "Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network," *IEEE Communications Magazine*, vol. 58, pp. 106–112, January 2020.

[58] H. Yang, X. Cao, F. Yang, J. Gao, S. Xu, M. Li, X. Chen, Y. Zhao, Y. Zheng, and L. Sijia, "A programmable metasurface with dynamic polarization, scattering and focusing control," *Scientific Reports*, vol. 6, 10 2016.

[59] M. Di Renzo, F. Habibi Danufane, X. Xi, J. de Rosny, and S. Tretyakov, "Analytical Modeling of the Path-Loss for Reconfigurable Intelligent Surfaces – Anomalous Mirror or Scatterer ?," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, May 2020.

[60] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless Communications Through Reconfigurable Intelligent Surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[61] W. Ni, X. Liu, Y. Liu, H. Tian, and Y. Chen, "Resource Allocation for Multi-Cell IRS-Aided NOMA Networks," *IEEE Transactions on Wireless Communications*, vol. 20, pp. 4253–4268, July 2021.

[62] A. Ihsan, W. Chen, M. Asif, W. U. Khan, Q. Wu, and J. Li, "Energy-Efficient IRS-Aided NOMA Beamforming for 6G Wireless Communications," *IEEE Transactions on Green Communications and Networking*, vol. 6, pp. 1945–1956, Dec 2022.

[63] S. Kumar, P. Yadav, M. Kaur, *et al.*, *A survey on IRS NOMA integrated communication networks.* 2022.

[64] P. V. Reddy, S. Reddy, S. Reddy, R. D. Sawale, P. Narendar, C. Duggineni, and H. B. Valiveti, "Analytical Review on OMA vs. NOMA and Challenges Implementing NOMA," in *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 552–556, Oct 2021.

[65] Y. Liu, G. Pan, H. Zhang, and M. Song, "On the Capacity Comparison Between MIMO-NOMA and MIMO-OMA," *IEEE Access*, vol. 4, pp. 2123–2129, 2016.

[66] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 770–774, Nov 2013.

[67] D. Li, N. Zhao, Y. Chen, A. Nallanathan, Z. Ding, and M.-S. Alouini, "Joint Precoding Optimization for Secure Transmission in Downlink MISO-NOMA Networks," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, Sep. 2019.

[68] B. Su, Q. Ni, and B. He, "Robust Transmit Designs for Secrecy Rate Constrained MISO NOMA System," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–5, Sep. 2018.

[69] Y. Jeong, C. Lee, and Y. H. Kim, "Power Minimizing Beamforming and Power Allocation for MISO-NOMA Systems," *IEEE Transactions on Vehicular Technology*, vol. 68, pp. 6187–6191, June 2019.

[70] P. Xu and K. Cumanan, "Optimal Power Allocation Scheme for Non-Orthogonal Multiple Access With α -Fairness," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2357–2369, 2017.

[71] X. He, Z. Huang, H. Wang, and R. Song, "Sum Rate Analysis for Massive MIMO-NOMA Uplink System With Group-Level Successive Interference Cancellation," *IEEE Wireless Communications Letters*, vol. 12, pp. 1194–1198, July 2023.

[72] Y. Qi and M. Vaezi, "Secure Transmission in MIMO-NOMA Networks," *IEEE Communications Letters*, vol. 24, pp. 2696–2700, Dec 2020.

[73] J. Zhu, J. Wang, Y. Huang, K. Navaie, Z. Ding, and L. Yang, "On Optimal Beamforming Design for Downlink MISO NOMA Systems," *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 3008–3020, March 2020.

[74] F. Alavi, K. Cumanan, Z. Ding, and A. G. Burr, "Robust Beamforming Techniques for Non-Orthogonal Multiple Access Systems with Bounded Channel Uncertainties," *IEEE Communications Letters*, vol. 21, pp. 2033–2036, Sep. 2017.

[75] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. Karagiannidis, "A Minorization-Maximization Method for Optimizing Sum Rate in Non-Orthogonal Multiple Access Systems," *IEEE Transactions on Signal Processing*, vol. 64, 05 2015.

[76] M. Hadi and R. Ghazizadeh, "Resource Allocation in OMA-NOMA based Two-Tier Heterogeneous Networks," in *2020 8th Iranian Joint Congress on Fuzzy and intelligent Systems (CFIS)*, pp. 001–006, Sep. 2020.

[77] R. Makkar, D. Rawal, N. Sharma, and V. K. Chakka, "Performance Analysis of Hybrid NOMA-OMA Scheme for 5G NR System," in *2020 IEEE 17th India Council International Conference (INDICON)*, pp. 1–6, Dec 2020.

[78] A. B. Rozario and M. F. Hossain, "Hybrid TDMA-NOMA Based M2M Communications over Cellular Networks with Dynamic Clustering and 3D Channel Models," in *2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, pp. 1–6, Nov 2019.

[79] A. J. Muhammed, Z. Ma, P. D. Diamantoulakis, L. Li, and G. K. Karagiannidis, "Energy-Efficient Resource Allocation in Multicarrier NOMA Systems With Fairness," *IEEE Transactions on Communications*, vol. 67, pp. 8639–8654, Dec 2019.

[80] Z. Liu, C. Liang, Y. Yuan, K. Y. Chan, and X. Guan, "Resource Allocation Based on User Pairing and Subcarrier Matching for Downlink Non-Orthogonal Multiple Access Networks," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, pp. 679–689, March 2021.

[81] X. Wei, H. Al-Obiedollah, K. Cumanan, M. Zhang, J. Tang, W. Wang, and O. A. Dobre, "Resource Allocation Technique for Hybrid TDMA-NOMA System with

Opportunistic Time Assignment," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, June 2020.

[82] B. Kimy, S. Lim, H. Kim, S. Suh, J. Kwun, S. Choi, C. Lee, S. Lee, and D. Hong, "Non-orthogonal Multiple Access in a Downlink Multiuser Beamforming System," in *MILCOM 2013 - 2013 IEEE Military Communications Conference*, pp. 1278–1283, Nov 2013.

[83] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[84] S. Ali, E. Hossain, and D. I. Kim, "Non-Orthogonal Multiple Access (NOMA) for Downlink Multiuser MIMO Systems: User Clustering, Beamforming, and Power Allocation," *IEEE Access*, vol. 5, pp. 565–577, 2017.

[85] F. Alavi, K. Cumanan, M. Fozooni, Z. Ding, S. Lambotharan, and O. A. Dobre, "Robust Energy-Efficient Design for MISO Non-Orthogonal Multiple Access Systems," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7937–7949, 2019.

[86] D. Ni, L. Hao, Q. T. Tran, and X. Qian, "Transmit Power Minimization for Downlink Multi-Cell Multi-Carrier NOMA Networks," *IEEE Communications Letters*, vol. 22, pp. 2459–2462, Dec 2018.

[87] N. Glei and R. B. Chibani, "Energy-efficient resource allocation for noma systems," in *2019 16th International Multi-Conference on Systems, Signals Devices (SSD)*, pp. 648–651, March 2019.

[88] H. M. Al-Obiedollah, K. Cumanan, J. Thiyagalingam, A. G. Burr, Z. Ding, and O. A. Dobre, "Energy Efficient Beamforming Design for MISO Non-Orthogonal Multiple Access Systems," *IEEE Transactions on Communications*, vol. 67, pp. 4117–4131, June 2019.

[89] J. Luo, J. Tang, D. K. C. So, G. Chen, K. Cumanan, and J. A. Chambers, "A Deep Learning-Based Approach to Power Minimization in Multi-Carrier NOMA With SWIPT," *IEEE Access*, vol. 7, pp. 17450–17460, 2019.

[90] E. Björnson, Özdogan, and E. G. Larsson, "Reconfigurable Intelligent Surfaces: Three Myths and Two Critical Questions," *IEEE Communications Magazine*, vol. 58, pp. 90–96, December 2020.

[91] Z. Wang, L. Liu, S. Zhang, and S. Cui, "Massive MIMO Communication With Intelligent Reflecting Surface," *IEEE Transactions on Wireless Communications*, vol. 22, pp. 2566–2582, April 2023.

[92] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent Reflecting Surface-Assisted Multi-User MISO Communication: Channel Estimation and Beamforming Design," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 661–680, 2020.

[93] J. Tang, Z. Peng, Z. Zhou, D. K. So, X. Zhang, and K.-K. Wong, "Energy-Efficient Resource Allocation for IRS-aided MISO System with SWIPT," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, pp. 3217–3222, 2022.

[94] Z. Chu, P. Xiao, D. Mi, W. Hao, Y. Xiao, and L.-L. Yang, "Multi-IRS Assisted Multi-Cluster Wireless Powered IoT Networks," *IEEE Transactions on Wireless Communications*, vol. 22, pp. 4712–4728, July 2023.

[95] W.-B. Li and Y. Shin, "Deep Learning for Intelligent Reflecting Surfaces Aided MIMO Systems," in *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 902–905, Oct 2021.

[96] A. Taha, Y. Zhang, F. B. Mismar, and A. Alkhateeb, "Deep Reinforcement Learning for Intelligent Reflecting Surfaces: Towards Standalone Operation," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, May 2020.

[97] H. Albinsaid, K. Singh, A. Bansal, S. Biswas, C.-P. Li, and Z. J. Haas, "Multiple Antenna Selection and Successive Signal Detection for SM-Based IRS-Aided Communication," *IEEE Signal Processing Letters*, vol. 28, pp. 813–817, 2021.

[98] Z. Zhou, N. Ge, Z. Wang, and L. Hanzo, "Joint transmit precoding and reconfigurable intelligent surface phase adjustment: A decomposition-aided channel esti-

mation approach," *IEEE Transactions on Communications*, vol. 69, pp. 1228–1243, Feb 2021.

[99] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Processing Letters*, vol. 27, pp. 905–909, 2020.

[100] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A Framework of Robust Transmission Design for IRS-Aided MISO Communications With Imperfect Cascaded Channels," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5092–5106, 2020.

[101] Z. Li, W. Chen, Q. Wu, H. Cao, K. Wang, and J. Li, "Robust beamforming design and time allocation for irs-assisted wireless powered communication networks," *IEEE Transactions on Communications*, vol. 70, pp. 2838–2852, April 2022.

[102] D. Pereira-Ruisánchez, Fresnedo, D. Pérez-Adán, and L. Castedo, "DRL-Based Sequential Scheduling for IRS-Assisted MIMO Communications," *IEEE Transactions on Vehicular Technology*, vol. 73, pp. 8445–8459, June 2024.

[103] Z. Ding and H. Vincent Poor, "A Simple Design of IRS-NOMA Transmission," *IEEE Communications Letters*, vol. 24, pp. 1119–1123, May 2020.

[104] A. Rafieifar, H. Ahmadinejad, and A. Falahati, "IRS-aided NOMA in a Cell Free Massive MIMO System," in *2022 30th International Conference on Electrical Engineering (ICEE)*, pp. 874–879, May 2022.

[105] J. Zhang, W. Wang, J. Tang, N. Zhao, K.-K. Wong, and X. Wang, "Robust Secure Transmission for IRS-Aided NOMA Networks With Hybrid Beamforming," *IEEE Transactions on Wireless Communications*, vol. 23, pp. 3086–3101, April 2024.

[106] A. Waraiet, K. Cumanan, Z. Ding, and O. A. Dobre, "Robust Design for IRS-Assisted MISO-NOMA Systems: A DRL-Based Approach," *IEEE Wireless Communications Letters*, vol. 13, pp. 592–596, March 2024.

[107] H. Wang, C. Liu, Z. Shi, Y. Fu, and R. Song, "On power minimization for irs-aided downlink noma systems," *IEEE Wireless Communications Letters*, vol. 9, pp. 1808–1811, Nov 2020.

[108] B. Zheng, Q. Wu, and R. Zhang, "Intelligent reflecting surface-assisted multiple access with user pairing: Noma or oma?," *IEEE Communications Letters*, vol. 24, pp. 753–757, April 2020.

[109] F. Fang, Y. Xu, Q.-V. Pham, and Z. Ding, "Energy-Efficient Design of IRS-NOMA Networks," *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 14088–14092, Nov 2020.

[110] T.-H. Nguyen, H. Park, K. Seol, S. So, and L. Park, "Applications of Deep Learning and Deep Reinforcement Learning in 6G Networks," in *2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 427–432, July 2023.

[111] A. Waraiet and K. Cumanan, "Outage-Constrained Robust Resource Allocation Framework for IRS-Empowered NOMA Systems: A DRL-Based Joint Design," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 2748–2764, 2024.

[112] M. Nazir, A. Sabah, S. Sarwar, A. Yaseen, and A. Jurcut, "Power and Resource Allocation in Wireless Communication Network," *Wireless Personal Communications*, vol. 119, no. 4, pp. 3529–3552, 2021.

[113] S. Stanczak, M. Wiczanowski, and H. Boche, *Fundamentals of Resource Allocation in Wireless Networks: Theory and Algorithms*, vol. 3. Springer Science & Business Media, 2009.

[114] J. Cui, Z. Ding, and P. Fan, "Power Minimization Strategies in Downlink MIMO-NOMA Systems," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2017.

[115] V. G. Douros and G. C. Polyzos, "Review of some fundamental approaches for power control in wireless networks," *Computer Communications*, vol. 34, no. 13, pp. 1580–1592, 2011.

[116] M. Ebrahimi, M. A. Maddah-ali, and A. K. Khandani, "Power Allocation and Asymptotic Achievable Sum-Rates in Single-Hop Wireless Networks," in *2006 40th Annual Conference on Information Sciences and Systems*, pp. 498–503, March 2006.

[117] P. C. Weeraddana, M. Codreanu, M. Latva-aho, A. Ephremides, and C. Fischione, "Weighted Sum-Rate Maximization in Wireless Networks: A Review," *Foundations and Trends® in Networking*, vol. 6, no. 1–2, pp. 1–163, 2012.

[118] C. W. Tan, M. Chiang, and R. Srikant, "Fast Algorithms and Performance Bounds for Sum Rate Maximization in Wireless Networks," *IEEE/ACM Transactions on Networking*, vol. 21, pp. 706–719, June 2013.

[119] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[120] Y. Polyanskiy, *Channel coding: Non-asymptotic fundamental limits*. Princeton University, 2010.

[121] 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, pp. 1183–1210, May 2021.

[122] A. Muhammad, M. Elhattab, M. A. Arfaoui, and C. Assi, "Optimizing Information Freshness in RIS-Assisted Non-Orthogonal Multiple Access-Based IoT Networks," *IEEE Networking Letters*, vol. 5, pp. 71–75, June 2023.

[123] Z.-Q. Luo and W. Yu, "An Introduction to Convex Optimization for Communications and Signal Processing," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1426–1438, Aug 2006.

[124] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.

[125] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, Version 2.1," 2014.

[126] J. Lofberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, pp. 284–289, Sep. 2004.

[127] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Current Solu-

tions, and Open Challenges," *IEEE Communications Surveys Tutorials*, vol. 22, pp. 1251–1275, Secondquarter 2020.

[128] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Machine Learning for Wireless Networks with Artificial Intelligence: A Tutorial on Neural Networks," *arXiv preprint arXiv:1710.02913*, vol. 9, 2017.

[129] B. Mahesh, "Machine Learning Algorithms-A Review," *International Journal of Science and Research (IJSR).[Internet]*, vol. 9, no. 1, pp. 381–386, 2020.

[130] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, "Ten Challenges in Advancing Machine Learning Technologies toward 6G," *IEEE Wireless Communications*, vol. 27, pp. 96–103, June 2020.

[131] J. Wang, R. Li, J. Wang, Y.-q. Ge, Q.-f. Zhang, and W.-x. Shi, "Artificial Intelligence and Wireless Communications," *Frontiers of Information Technology & Electronic Engineering*, vol. 21, no. 10, pp. 1413–1425, 2020.

[132] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An introduction," *Robotica*, vol. 17, no. 2, pp. 229–235, 1999.

[133] A. G. Barto, "Reinforcement Learning," in *Neural systems for control*, pp. 7–30, Elsevier, 1997.

[134] M. Morales, *Grokking Deep Reinforcement Learning.* Manning Publications, 2020.

[135] J. Vermorel and M. Mohri, "Multi-armed Bandit Algorithms and Empirical Evaluation," in *Machine Learning: ECML 2005* (J. Gama, R. Camacho, P. B. Brazdil, A. M. Jorge, and L. Torgo, eds.), (Berlin, Heidelberg), pp. 437–448, Springer Berlin Heidelberg, 2005.

[136] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[137] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing Atari with Deep Reinforcement Learning," *arXiv preprint arXiv:1312.5602*, 2013.

[138] N. Heess, D. Silver, and Y. W. Teh, "Actor-Critic Reinforcement Learning with Energy-Based Policies," in *Proceedings of the Tenth European Workshop on Reinforcement Learning* (M. P. Deisenroth, C. Szepesvári, and J. Peters, eds.), vol. 24 of *Proceedings of Machine Learning Research*, (Edinburgh, Scotland), pp. 45–58, PMLR, 30 Jun–01 Jul 2013.

[139] Rachna and J. Malhotra, "Multi antenna techniques for the enhancement of mobile wireless systems: Challenges and opportunities," in *2014 International Conference on Advances in Engineering Technology Research (ICAETR - 2014)*, pp. 1–5, Aug 2014.

[140] B. Sklar, "Rayleigh fading channels in mobile digital communication systems .I. Characterization," *IEEE Communications Magazine*, vol. 35, no. 7, pp. 90–100, 1997.

[141] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage Constrained Robust Transmit Optimization for Multiuser MISO Downlinks: Tractable Approximations by Conic Optimization," *IEEE Transactions on Signal Processing*, vol. 62, pp. 5690–5705, Nov 2014.

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

[143] Z.-q. Luo, W.-k. Ma, A. M.-c. So, Y. Ye, and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.

[144] L. Xia, Z. Yang, J. Cui, Y. Wu, Z. Dong, and Z. Ding, "Transmit Power Minimization for IRS-Assisted Cooperative NOMA Networks With SWIPT," in *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Oct 2021.

[145] M. A. Abd-Elmagid, H. S. Dhillon, and N. Pappas, "AoI-Optimal Joint Sampling and Updating for Wireless Powered Communication Systems," *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 14110–14115, Nov 2020.

[146] C. Hu and Y. Dong, "Age of Information of Two-Way Data Exchanging Systems with Power-Splitting," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 295–306, 2019.

[147] S. M. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Resource Allocation for Downlink NOMA Systems: Key Techniques and Open Issues," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 40–47, 2018.

[148] M. Al-Mekhlafi, M. A. Arfaoui, M. K. Elhattab, C. M. Assi, and A. Ghrayeb, "Joint Resource Allocation and Phase Shift Optimization for RIS-Aided eMBB/URLLC Traffic Multiplexing," *IEEE Transactions on Communications*, vol. 70, pp. 1304–1319, 2021.

[149] T. Yoshizawa, S. B. M. Baskaran, and A. Kunz, "Overview of 5G URLLC System and Security Aspects in 3GPP," in *2019 IEEE Conference on Standards for Communications and Networking (CSCN)*, pp. 1–5, 2019.

[150] W. R. Ghanem, V. Jamali, and R. Schober, "Joint Beamforming and Phase Shift Optimization for Multicell IRS-aided OFDMA-URLLC Systems," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–7, March 2021.

[151] W. R. Ghanem, V. Jamali, Y. Sun, and R. Schober, "Resource Allocation for Multi-User Downlink MISO OFDMA-URLLC Systems," *IEEE Transactions on Communications*, vol. 68, pp. 7184–7200, Nov 2020.