VIETNAM GENERAL CONFEDERATION OF LABOR

TON DUC THANG UNIVERSITY

FACULTY OF ELECTRICAL & ELECTRONICS ENGINEERING

NGUYEN DUC TRI

OPTIMIZATION OF WIRELESS RELAY NETWORKS USING BOTH RELAY AND INTELLIGENT REFLECTING SURFACE

UNDERGRADUATE THESIS OF ELECTRONICS AND TELECOMMUNICATIONS ENGINEERING

HO CHI MINH CITY, YEAR 2025

VIETNAM GENERAL CONFEDERATION OF LABOR

TON DUC THANG UNIVERSITY

FACULTY OF ELECTRICAL & ELECTRONICS ENGINEERING

NGUYEN DUC TRI

OPTIMIZATION OF WIRELESS RELAY NETWORKS USING BOTH RELAY AND INTELLIGENT REFLECTING SURFACE

UNDERGRADUATE THESIS OF ELECTRONICS AND TELECOMMUNICATIONS ENGINEERING

Advised by
DR. TRAN THANH PHUONG

HO CHI MINH CITY, YEAR 2025

ACKNOWLEDGMENT

Best regards, good health, and deep thanks. With the attention, teaching, and attentive guidance of the teachers so far was able to complete the project: "Graduation Project".

In particular, I would like to sincerely thank ***Dr. Tran Thanh Phuong*** for helping me a lot in the past time. Meeting the teacher received many instructions, devoted and dedicated, creating the most favorable conditions. And finally, I would like to thank you from the bottom of my heart.

Over the past time, with my little experience, I have many mistakes in project making. I want to receive comments from teachers to add to the missing knowledge and continuously improve my knowledge.

Thanks very much and good health to the teachers.

Ho Chi Minh City, day month ,year 2025.

Author

(Signature and full name)

This thesis was carried out at Ton Duc Thang University.

Advisor: .....

.....

(Title, full name and signature)

This thesis is defended at the **Undergraduate Thesis Examination Committee was hold at Ton Duc Thang University** on …. /…./……..

Confirmation of the Chairman of the Undergraduate Thesis Examination Committee and the Dean of the faculty after receiving the modified thesis (if any).

CHAIRMAN	**DEAN OF FACULTY**
(Signature and full name)	*(Signature and full name)*
…….	……

DECLARATION OF AUTHORSHIP
AT TON DUC THANG UNIVERSITY

I hereby declare that this thesis was carried out by myself under the guidance and supervision of ***Dr. Tran Thanh Phuong***; and that the work and the results contained in it are original and have not been submitted anywhere for any previous purposes. The data and figures presented in this thesis are for analysis, comments, and evaluations from various resources by my own work and have been duly acknowledged in the reference part.

In addition, other comments, reviews and data used by other authors, and organizations have been acknowledged, and explicitly cited.

I will take full responsibility for any fraud detected in my thesis. Ton Duc Thang University is unrelated to any copyright infringement caused on my work (if any).

Ho Chi Minh City, day month ,year 2025.

Author

(Signature and full name)

TON DUC THANG UNIVERSITY
FACULTY OF ELECTRICAL & ELECTRONICS ENGINEERING

SOCIALIST REPUBLIC OF VIETNAM
Independence – Freedom – Happiness

Ho Chi Minh City, September 16th, 2024

GRADUATION THESIS ASSIGNMENT FORM

(This form should be placed on the first page of the Graduation Thesis report)

Name of student: **Nguyễn Đức Trí**

Major: Electronics and Telecommunication Engineering

Class: 19H40201 Student ID: 41900297

1. Tittle: **Optimization of wireless relay networks using both relay and intelligent reflecting surface**

2. Title (in Vietnamese): ***Tối ưu hóa hiệu năng mạng chuyển tiếp vô tuyến sử dụng kết hợp relay và mặt phản xạ thông minh***

3. Taks (details of assignment and input data):

- Study wireless relay networks and intelligent reconfigurable surfaces (IRS).
- Investigate the wireless relay networks with multiple relays assisted by an IRS to forward data from a source to a destination.
- Derive the expression(s) of the key performance factor(s) of the system (such as capacity, outage probability, etc.).
- Study and propose an algorithm to optimize the system perfomance by adjusting the phases of IRS elements and selecting the best relay among the available ones for forwarding.
- Implement the proposed algorithm by simulation program and evaluate the results.
- Write the final report in accordance with the scientific standard.

3. Start date:	**16/09/2024**
4. Evaluation of 50% process:	**04/11/2024 - 09/11/2024**
5. Submission date:	**05/01/2024**
6. Advisor(s):	Percentage:
Dr. Tran Thanh Phuong	**100%**

The content and requirements of the Graduation Thesis have been approved by the Head of Department and the Dean of Faculty.

Date: September 16th, 2024
HEAD OF DEPARTMENT

Dr. Tran Thanh Phuong

VICE DEAN IN CHARGE OF
THE FACULTY

Dr. Tran Thanh Phuong

ADVISOR

Dr. Tran Thanh Phuong

TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG
KHOA ĐIỆN – ĐIỆN TỬ

PHIẾU THEO DÕI TIẾN ĐỘ THỰC HIỆN KLTN/ĐATN

Họ tên sinh viên:…NGUYỄN ĐỨC TRÍ…… MSSV:…41900297….

Ngành học: …Kỹ thuật điện tử - viễn thông……

Địa chỉ liên lạc:..56/32/11 Gò Ô Môi, p.Phú Thuận, q.7…Đơn vị:..Điện – Điện tử...

Họ tên GVHD:…DR. Trần Thanh Phương……

Tên đề tài KLTN:..Optimization of wireless relay networks using both relay and intelligent reflecting surface.....

Tuần/Ngày	Khối lượng		GVHD ký
	Đã thực hiện	Tiếp tục thực hiện	
1 (30/09 – 06/10)	Research about IRS (Intelligent Reflecting Surface) and wireless relay network.	Read the references and set up the problems which need to be solved.	...../...../2024
2 (07/10 – 13/10)	Read the references and set up the problems which need to be solved.	Set up the system capacity calculation formula.	...../...../2024
3 (14/10 – 20/10)	Set up the system capacity calculation formula.	Set up the system capacity calculation formula.	...../...../2024
4 (21/10 – 27/10)	Set up the system capacity calculation formula.	Research for the capacity optimization algorithm.	...../...../2024
5 (28/10 – 03/11)	Research for the capacity optimization algorithm.	Research for the capacity optimization algorithm.	...../...../2024
6 (04/11 – 10/11)	Research for the capacity optimization algorithm.	Research for the capacity optimization algorithm.	...../...../2024
7 (11/11 – 17/11)	Research for the capacity optimization algorithm.	Review and prepare for the 50% review session Write a MATLAB	...../...../2024

1

		program to implement the algorithm.	
Kiểm tra giữa kỳ	Đánh giá khối lượng hoàn thành..50...% được tiếp tục/không tiếp tục thực hiện LVTN		
9 **(25/11 – 01/12)**	Write a MATLAB program to implement the algorithm.	Write a MATLAB program to implement the algorithm.	...../...../2024
10 **(02/12 – 08/12)**	Write a MATLAB program to implement the algorithm.	Simulation and evaluation of results.	...../...../2024
11 **(09/12 – 15/12)**	Simulation and evaluation of results.	Simulation and evaluation of results.	...../...../2024
12 **(16/12 – 22/12)**	Simulation and evaluation of results.	Fixes the bug program and improve the algorithm.	...../...../2024
13 **(23/12 – 29/12)**	Fixes the bug program and improve the algorithm.	Fixes the bug program and improve the algorithm.	...../...../2024
14 **(30/10/2024 – 05/01/2025)**	Fixes the bug program and improve the algorithm.	Complete the simulation program and project report.	...../...../2025
15 **(06/01 – 12/01)**	Complete the simulation program and project report.		...../...../2025

2

OPTIMIZATION OF WIRELESS RELAY NETWORKS USING BOTH RELAY AND INTELLIGENT REFLECTING SURFACE

ABSTRACT

The primary aim of this research is to develop and refine methods to enhance communication performance in wireless networks. Relay networks utilize relay nodes to facilitate the transmission of signals from the source to the destination. Concurrently, Intelligent Reflecting Surfaces (IRS), an advanced technology, are employed to adjust the phase and amplitude of reflected signals, thereby optimizing the transmission environment. The synergy of these two technologies not only extends the network's coverage area but also significantly improves signal quality and energy efficiency.

This study will leverage convex optimization algorithms to tackle the challenges associated with the simultaneous adjustment and management of the parameters of both relay nodes and IRS. Convex optimization is chosen for its capability to efficiently identify globally optimal solutions within a vast parameter space. Initially, a mathematical model will be established to depict the system, incorporating variables such as the positions and operations of relay nodes, and the phase and amplitude adjustments of IRS. In order to simplify the complexity of the optimization algorithm, we will also use the Q-learning algorithm during the system simulation process.

The first step in this research involves constructing detailed system models, based on the combined relay network and IRS system. These models will encompass factors such as the locations of relay nodes, IRS adjustments, and environmental parameters like transmission paths and obstacles. Subsequently, convex optimization techniques will be applied to determine optimal solutions for adjusting system parameters. The convex optimization algorithm will address the multi-objective optimization problem, aiming to maximize both signal quality and energy efficiency.

After the optimization process, the next critical step involves conducting simulations on the developed models to evaluate the effectiveness of the proposed convex optimization algorithms. These simulations are essential as they provide a practical validation of the theoretical models and optimization techniques employed. To streamline and enhance the data simulation process, the Q-learning algorithm will be integrated. This integration leverages the strengths of Q-learning, particularly its ability to learn and adapt from interactions with the environment, thereby simplifying the complex task of simulating the system under various scenarios.

The simulation results will be meticulously analyzed and compared with traditional methods that exclusively utilize relay networks or IRS. This comparative analysis aims to highlight the performance improvements achieved by combining both technologies under the framework of convex optimization. Specifically, the focus will be on metrics such as signal quality, energy efficiency, and coverage area. By contrasting these results, the study aims to demonstrate the superior performance and advantages of the integrated approach.

The anticipated outcomes of this research include minimizing signal loss and interference, ensuring stronger and more stable signals, and expanding the coverage of wireless networks, particularly in areas with complex transmission conditions. By addressing these aspects, the combined use of relay networks and IRS with convex optimization algorithms is expected to provide significant improvements over traditional methods.

Ultimately, this research promises to make substantial contributions to the field of wireless communications, particularly in the development of next-generation network technologies such as 6G. Through innovative methodologies and rigorous analysis, the study aims to pave the way for more efficient, reliable, and expansive wireless communication networks, meeting the increasing demands of modern connectivity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AF	Amplify-and-Forward
AR	Augmented Reality
AWGN	Additive White Gaussian Noise
DF	Decode-and-Forward
DI	Diagonal Matrix
FPA	Fixed Phase Algorithm
IoT	Internet of Things
IRS	Intelligent Reflecting Surfaces
LoS	Line-of-Sight
MDPs	Markov Decision Processes
MIMO	Multiple-Input Multiple-Output
NLoS	Non-Line-of-Sight
QL-JIRA	Q-Learning-based Joint IRS and Relay-Assisted communication
RNs	Relay nodes
RPA	Random Phase Algorithm
RS	Random Selection
RW	Reward Matrix
SNR	Signal-to-Noise Ratio
TDD	Time Division Duplex
VR	Virtual Reality
WiFi	Wireless-Fidelity

CHAPTER 1. INTRODUCTION

1.1 Purpose of implementing the topic

The rapid rise of wireless communications and the desire for high-speed, dependable access have made network optimization a priority. New wireless technologies are needed to improve network coverage, capacity, and performance as 5G and 6G networks emerge. Wireless relay networks and intelligent reflecting surfaces (IRS) are potential wireless communication optimization technologies (W. Long et al., 2021). This study investigates relay nodes and IRS to improve network efficiency, solve coverage issues, and build the groundwork for future research.

Optimizing wireless relay networks with relay and IRS technologies is justified by the limitations of existing wireless infrastructure, the innovative nature of IRS technology, and the synergy that can result from combining IRS and relay networks. Wireless communications often employ relay networks to enhance signal coverage, especially in places with obstructions. Intermediaries like relay nodes assist in sending messages across vast distances. Traditional relays have energy efficiency, delay, and scalability issues. Relay nodes demand more power, increasing the total energy consumption of the network, and the multi-hop nature of relay networks may result in higher delay. IRS in relay networks may circumvent these constraints. IRS allows passive signal reflection without needing extra energy consumption for active transmission, boosting energy economy and minimizing the requirement for numerous relay hops, thereby lowering latency (Q. Wu and R. Zhang, 2020).

IRS technology represents a paradigm shift in wireless communication. Unlike traditional relays, which actively amplify and retransmit signals, IRS uses large arrays of passive elements that can be dynamically controlled to adjust the reflection, refraction, or absorption of incident signals (Le Si Phu et al., 2023). Through these adjustments, IRS can improve signal strength, directivity, and coverage in an energy-efficient manner. However, IRS also has limitations, such as while IRS works well in

line-of-sight (LoS) conditions, it may not perform optimally in non-line-of-sight (NLoS) scenarios. By combining IRS with traditional relays, it is possible to address both LoS and NLoS challenges, creating a robust solution for diverse wireless communication scenarios (M. Di Renzo et al., 2020) and (C. Liaskos et al., 2018). This hybrid approach is particularly promising in dense urban environments, where buildings and other structures often obstruct direct communication paths. In rural areas with limited infrastructure, the combination of relay nodes and IRS provides a cost-effective solution for extending network coverage. Using IRS alongside relay nodes can enable broader and more reliable network coverage, supporting the optimization of urban networks and bridging the digital divide in underserved areas (Ba Cao Nguyen et al., 2022).

In this project, the focus is on addressing the need for energy efficiency and sustainability while contributing to future generations of wireless networks. Energy efficiency is critical in modern wireless networks as data demands grow, leading to increased operational costs and environmental impact. One key reason for selecting this research topic is the potential for the IRS to reduce energy consumption in relay networks. Since IRS elements reflect signals passively, they require minimal energy compared to traditional relay nodes, which consume power for amplification and signal processing. By minimizing reliance on active relays and strategically integrating IRS, this research seeks to develop solutions that align with global goals for energy efficiency and sustainability. As the world transitions to 6G and beyond, the demands on wireless networks will intensify. New applications such as augmented reality (AR), virtual reality (VR), holographic communications, and connected autonomous systems will require networks with high reliability, ultra-low latency, and extensive coverage (M. Sode et al., 2024). The integration of IRS and relay nodes can play a key role in meeting these demands, offering a flexible and scalable approach to network optimization that can be adapted to different environments and requirements (N. T. Nguyen et al., 2022), (N. T. Nguyen et al., 2023), (Ba Cao Nguyen et al., 2023), and (Phuong T. Tran et al., 2022).

1.2 Technical Specifications

To comprehensively tackle this part of our study, we need to meticulously outline several key requirements. These include defining a detailed system model, providing a robust performance evaluation formula, solving an intricate optimization problem, and subsequently simulating the results of our problem. We will begin by thoroughly describing the system.

Our topic, "Optimization of Wireless Relay Networks Using Both Relay and Intelligent Reflecting Surface," necessitates the optimization of a complex system model. This model incorporates several essential components: the source (S) and destination (D) - both indispensable elements of any wireless communication system - along with an intelligent reflecting surface (IRS) and a relay node, which are pivotal to the relay network (C. Huang et al., 2021). In this configuration, the source node (S) transmits signals to the destination device (D) via relays and the IRS. The relays act as intermediaries, performing functions such as Amplify and Forward (AF) or Decode and Forward (DF), to convey the signal from the source to the destination. Meanwhile, the IRS is equipped with numerous reflective elements capable of adjusting the phase of incoming waves, thereby optimizing the received signal at the destination.

The crux of the problem lies in optimizing system performance metrics such as data transmission speed, coverage area, and energy efficiency by selecting the best relay and fine-tuning the phase shift parameters of the IRS. The system model also incorporates various transmission channels, such as $h_{S,\mathrm{IRS}}$, h_{S,R_i}, h_{IRS,R_i}, $h_{R_i,\mathrm{IRS}}$ and others.

Upon formulating the optimization problem grounded in our meticulously devised system model, these formulations will establish the foundation for validating our proposed design through comprehensive simulations. By employing MATLAB, a sophisticated high-level programming environment, and the Q-Learning algorithm - a model-free reinforcement learning method, we will investigate various scenarios,

including QL-JIRA (Q-learning-based Joint IRS and Relay-assisted communication), Random Selection (RS), Fixed Phase Algorithm (FPA), Random Phase Algorithm (RPA), and the No Relay approach. The simulation scenario evaluates the system performance based on two factors: the available rate with the transmitted power and the available rate with the change in the number of relays. Each scenario will be subject to rigorous testing and analysis to assess performance under diverse conditions. Simulation results will be graphically represented to delineate the relationship between the achievable rate and transmit power, as well as the achievable rate and the number of relays. These visual representations will afford a comprehensive understanding of performance metrics across scenarios, enabling us to derive insightful and meaningful conclusions from the data.

After acquiring the simulation results, we will conduct a comparative analysis of system performance metrics. This investigation will examine factors like attainable data rates in relation to transmission power and the effects of altering the number of relays or the phase angles of the IRS. This holistic strategy guarantees an exhaustive assessment and enhancement of our wireless relay network, using both relays and intelligent reflecting surfaces.

1.3 Ideas and implementation methods

Throughout the process of exploring and researching this topic, the primary approach involves utilizing convex optimization algorithms to address network performance optimization issues. Here, the objective functions (such as maximizing throughput or minimizing signal loss) and the associated constraints (such as transmission power or physical system conditions) are formulated into convex optimization problems. The convex optimization algorithm is then employed to determine the optimal configuration for various parameters, including the phase angles of the IRS and the selection of suitable relays. This approach not only guarantees feasibility and efficiency but also reduces computational complexity due to the rapid and precise convergence of convex optimization algorithms.

Furthermore, after resolving the algorithm, MATLAB software will be used to simulate the results, thereby demonstrating that the performance of this method is superior to using the components individually.

In addition to the primary goals of this research, while delving into this topic, we encountered numerous challenges in conceptualizing ideas and devising methods to implement them effectively. These challenges were multifaceted, encompassing the complexity of the optimization problem, the necessity of ensuring system convexity, and the ability of the algorithm to converge.

Firstly, regarding the complexity of the problem, the integration of relay and IRS (Intelligent Reflecting Surface) technologies introduces significant intricacies into the parameter space that requires optimization, far beyond what traditional systems entail. Specifically, the phase angles of the IRS represent a vast parameter space with potentially thousands of variables that need to be optimized simultaneously. This multiplicity of variables presents a formidable challenge in maintaining the feasibility of the problem, especially under the constraints imposed by practical applications.

Secondly, ensuring the system's convexity adds another layer of complexity. In practice, not all optimization problems can be naturally expressed in a convex form. Convex problems are desirable because they guarantee global optima and are computationally more tractable. However, transforming a non-convex problem into a convex one often requires sophisticated mathematical techniques, such as specific transformations or approximations. These methods must ensure that the convexity is achieved without losing the essential physical and operational characteristics of the original problem.

Moreover, the convergence of the algorithm presents a significant challenge. Within the optimization process, recognizing the complexities associated with convexity and ensuring the algorithm's convergence, Q-Learning has been integrated as a viable solution. This approach is instrumental in facilitating the simulation of system performance, thereby enabling us to make objective evaluations regarding the

feasibility of the system's implementation.

In sum, while the challenges are significant, they also drive innovation and deepen our understanding of complex systems. Furthermore, the concept of integrating relays and IRS to enhance the performance of wireless systems represents a significant breakthrough in modern communications. Addressing these issues head-on not only aids in achieving the research objectives but also contributes to the broader field of optimization and wireless communications, paving the way for future developments and applications.

1.4 Utilized tools

In the course of researching this topic, the use of convex optimization algorithms to investigate and resolve system performance optimization problems is complemented by MATLAB's crucial role as a powerful assistant. MATLAB not only aids in the accurate implementation and execution of these algorithms but also facilitates the simulation of optimization results. This dual approach provides a comprehensive and in-depth understanding of the performance of proposed solutions.

Simulation through MATLAB enables detailed comparison of channel models, both when used individually and in combination, allowing for thoroughevaluation of the strengths and weaknesses of each method. The powerful visualization and data analysis capabilities of MATLAB make it easier to discern the differences and effectiveness of various system configurations, as well as to identify areas needing improvement.

This methodology not only enhances our research outcomes but also opens up opportunities for the more effective and sustainable application of findings to real-world communication systems in the future. The synergy between convex optimization algorithms and MATLAB simulation represents a significant advancement, ensuring the feasibility and precision of future research and applications.

CHAPTER 2. RELATED THEORY

2.1 Wireless Communication

2.1.1 Fundamentals of wireless communication networks

Wireless communication refers to the exchange of information between devices utilizing radio waves, infrared waves, or other forms of electromagnetic signals instead of traditional physical cables like copper or fiber optic wires. The technology underpins a wide variety of systems, including WiFi, Bluetooth, 4G/5G mobile networks, satellite communications, and even underwater communications. These systems use specific radio frequencies to transmit signals between end devices such as smartphones, tablets, IoT sensors, and network infrastructure devices. A typical wireless communication model consists of several critical components, primarily the source (S), the transmission channel, and the destination (D) (MuRata innovator electronics, 2023).

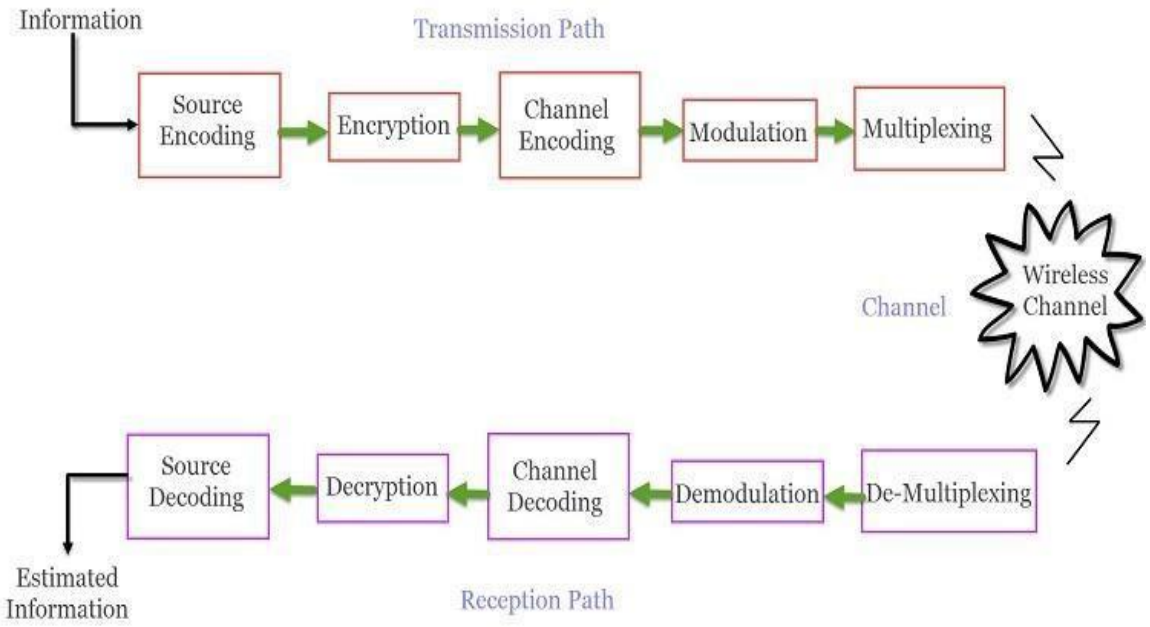

Figure 2.1 The block diagram of Wireless Communication System

2.1.2 Advantages of opting for wireless communication

We currently find ourselves amidst the 5G era - the fifth generation of wireless communication technology. Designed to provide significantly higher data transmission speeds, reduced latency, and the ability to support a multitude of simultaneous device connections, 5G represents a major leap forward from its predecessors. This evolution in wireless communication technology brings forth numerous advantages, underpinning its rapid growth and widespread adoption. For instance, wireless communication offers unparalleled mobility and flexibility, allowing users to connect anytime and anywhere without the constraints of physical cables. This flexibility is complemented by the ease of deployment and expansion, as the absence of a complex physical infrastructure significantly reduces both costs and time associated with network implementation. Expanding a wireless network often only requires the addition of intermediate transmitters or relays, further enhancing its scalability. Moreover, the extensive connection range and capacity to support numerous devices simultaneously highlight the robust nature of wireless networks.

2.1.3 Drawbacks of wireless communication

Despite these compelling advantages, wireless communication is not without its challenges and limitations. Interference and signal attenuation can be caused by various sources, including electronic devices, weather conditions, and other competing network systems. Furthermore, physical distance and obstacles such as buildings and terrain can pose significant hurdles. Security concerns also loom large, as wireless communications are inherently susceptible to attacks like eavesdropping, signal injection, and spoofing. Ensuring the integrity and confidentiality of wireless transmissions thus necessitates the implementation of robust encryption and security mechanisms.

In the present technological environment, wireless communications play a vital role, supporting global connection and promoting the development of

sophisticated applications such as 5G networks, the Internet of Things (IoT), and artificial intelligence. However, to fully utilize the promise of wireless communication, it is vital to solve the related problems. Thereby, emerging technologies, such Intelligent Reflecting Surfaces (IRS), enhanced relay selection algorithms, and complex channel optimization techniques, are continuously being developed and polished to expand the capabilities of wireless communications. These developments offer enormous potential for the future, bringing up new possibilities for services and applications that might alter our world.

2.2 The imperative of Network Optimization

2.2.1 Comprehending network optimization

Network optimization is a complicated, data-driven process meant to optimize the overall performance and efficiency of a network. This multidimensional strategy focuses on monitoring and evaluating key performance indicators, including latency, throughput, and packet loss, to identify and eliminate possible bottlenecks, hence improving the utilization of network resources. In essence, network optimization is a thorough process that incorporates a mix of traffic analysis, infrastructure design, bandwidth management,... By carefully addressing these areas, companies may greatly improve their network's performance, stability, and efficiency, eventually leading to increased user experiences and operational productivity.

2.2.2 The significance of wireless network optimization

Network optimization plays a significant function in boosting several areas, including connection, speed, user experience, productivity, and cost-effectiveness. When a network is optimized, it enables smooth and dependable connection, ensuring that data transfer happens at the maximum feasible speed by reducing latency. This improvement not only increases operating efficiency but also dramatically enhances

the overall user experience. Additionally, it enhances productivity and decreases expenses by simplifying network performance, eventually leading to a more effective and efficient communication infrastructure.

2.3 Relay networks in Wireless Communication Systems

2.3.1 Definition of relay network

Wireless communication systems commonly utilize a relay network, a specific type of network topology, to extend the range and enhance the quality of signal transmission. In such a network, the primary source node and the target destination node are connected through a series of intermediate nodes known as relay nodes. These relay nodes play a crucial role, particularly in scenarios where the direct communication between the source and destination is not feasible due to the large distance between them surpassing the transmission capabilities of their respective communication ranges.Relay networks enable information transfer between two devices, such as a server and a computer, that are too far apart to establish a direct connection. Instead, the network relays the information through various intermediate devices or nodes, ensuring the data reaches its intended destination. This approach effectively extends the communication range and enhances connectivity.

In wireless communication systems, as the distance between a source node and a destination node increases, signal strength tends to degrade. This phenomenon, known as path loss, can significantly impact the quality and reliability of communication links. Additionally, obstacles such as buildings, trees, and geographical features further attenuate the signal, resulting in reduced coverage and weakened connectivity.

Relay networks operate on the principle of intermediary signal transmission. The source node initiates the communication by transmitting a signal intended for the destination node. However, instead of reaching the destination directly, the signal is relayed through one or more relay nodes. These relay nodes receive the signal,

process it (amplify or decode and forward), and then transmit it to the next relay node or the destination node. This method effectively bridges the communication gap, ensuring that the signal can cover greater distances without significant degradation.

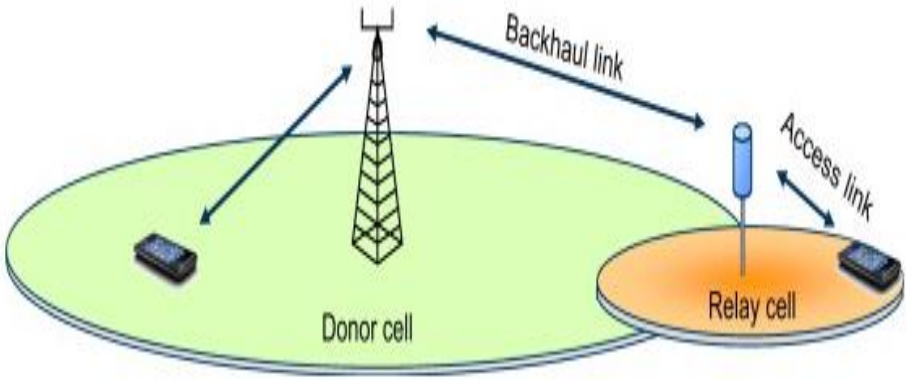

Figure 2.2 Fundamental Structure of a Relay Network

To address these limitations, relay nodes (RNs) are strategically positioned within the network to act as signal amplifiers, thereby extending the transmission range. By relaying signals from the source node to the destination node, RNs help bridge communication gaps and ensure reliable and robust connectivity.

There are various types of relay nodes, including fixed RNs, mobile RNs, decode-and-forward RNs, and amplify-and-forward RNs. Fixed RNs are stationary nodes deployed at predetermined locations, while mobile RNs can move within the network to provide dynamic coverage. Decode-and-forward RNs decode the received signal before forwarding it to the destination, while amplify-and-forward RNs amplify the received signal without decoding it.

2.3.2 Advantages and disadvantages of relay nodes in relay networks

Deploying relay nodes in wireless networks offers multiple advantages. Firstly, they significantly enhance coverage by extending the communication range beyond the constraints of the source node. This extension is particularly beneficial in

large-scale networks, remote areas, and regions with challenging terrains where direct communication between the source and destination nodes is not feasible.

Secondly, relay nodes boost the network's capacity by alleviating the effects of path loss and signal degradation. By amplifying and relaying signals, relay nodes counteract signal attenuation, enabling reliable communication over greater distances. This enhancement results in increased network capacity and improved spectral efficiency.

Thirdly, relay nodes improve overall network performance by minimizing interference and optimizing signal quality. Strategic placement of relay nodes ensures optimal signal propagation and effectively addresses areas with weak coverage or high interference. This optimization leads to stronger signals, reduced packet loss, and enhanced overall network performance.

Moreover, relay nodes facilitate energy-efficient communication in wireless networks. Instead of relying on direct communication between distant nodes, which requires substantial energy for signal amplification, relay nodes relay signals more efficiently. This efficiency reduces energy consumption, extends the network's lifespan, and enhances energy efficiency.

While relay networks offer significant advantages, several disadvantages also need to be considered. The primary challenge lies in their complexity. Deploying and managing a relay network, as well as coordinating multiple relay nodes, demands advanced planning, precise control, and continuous maintenance to ensure optimal performance. This complexity can be quite burdensome for network operators.

Additionally, the operational costs associated with relay networks can be substantial. These costs include the expenses for additional hardware, installation, and ongoing maintenance. For many organizations, these financial considerations can be a significant factor when evaluating the feasibility of implementing relay networks. Furthermore, there is the potential for interference. As the distance between the receiver and the transmitter increases, more relay nodes are needed. The use of multiple relay nodes can introduce unwanted interference, which, if not properly

managed, can degrade the overall network performance and compromise signal quality. Managing this interference is crucial to maintain the efficiency and reliability of the relay network.

In summary, while relay networks can greatly enhance coverage, signal quality, and energy efficiency, they also come with challenges related to complexity, cost, and potential interference that must be carefully managed.

2.4 Intelligent Reflective Surface (IRS)

2.4.1 Concept of intelligent reflective surface

An Intelligent Reflective Surface (IRS) represents a significant advancement in wireless communication technology. It is a type of metasurface embedded with active, frequency-selective elements designed to manipulate and control the propagation environment of wireless signals. The primary components of an IRS

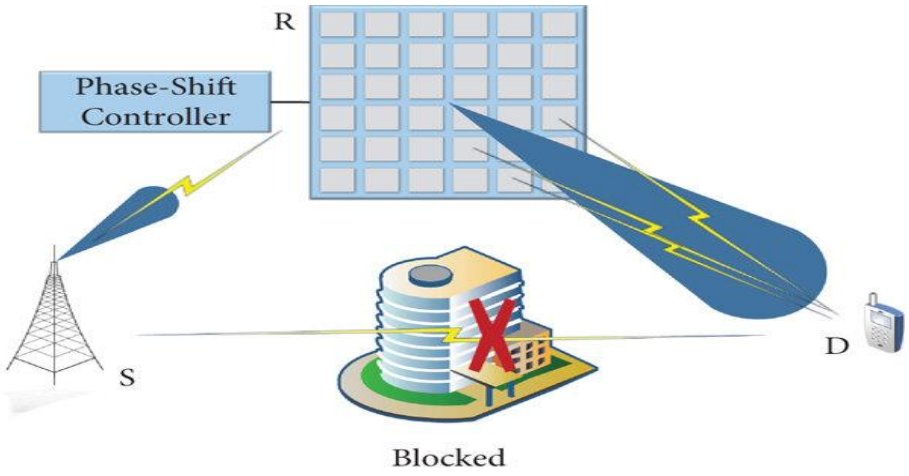

Figure 2.3 Fundamental model of IRS support in wireless communication networks

include a frequency-selective surface that reflects signals, simple, low-cost sensors, and a cognitive engine powered by machine learning algorithms to dynamically adjust the surface elements. The main objective of IRS is to enhance the quality of service in wireless communications by precisely controlling the phase and amplitude of the reflected signals (E. Basar et al., 2019).

There are various IRS models, each suited for different applications (Tùng N. T. et al., 2021). Single IRS models are used to enhance signal quality and coverage in specific areas. Multi-IRS models are employed for wider coverage, providing comprehensive signal enhancement across larger regions. Cooperative IRS models involve multiple IRS units working in tandem to optimize signal propagation and improve overall network performance.

The effectiveness of IRS largely depends on the precision of the control over the surface elements and the system's ability to adapt to changing communication environments. Properly implemented, IRS can lead to significant improvements in signal quality, energy efficiency, and network coverage, making it an invaluable asset in the evolution of wireless communication systems (Trí L. V. ,và Mai L. T. P., 2023).

Figure 2.3 demonstrates the operational model of a communication system employing an Intelligent Reflective Surface (IRS). In this system, a signal source (S) emits a signal towards its intended destination. However, if an obstacle, such as a building (blocked area), obstructs the direct transmission path, the signal cannot reach the destination (D) directly. To address this issue, an IRS (R) is deployed. A phase-shift controller adjusts the phase of the signal to ensure that the reflected signal is oriented correctly to bypass the obstacle. The IRS then reflects the phase-adjusted signal, directing it towards the destination device (D). Consequently, the signal efficiently reaches the destination device despite the blocked direct transmission path. This illustrates how IRS technology can optimize wireless communications in complex environments, enhancing the stability and robustness of network connections.

2.4.2 Pros and cons of implementing intelligent reflective surfaces (IRS) in wireless communication networks

The advantages of IRS are manifold. Firstly, it significantly improves signal quality and coverage by optimizing the path of the reflected signals. This is especially

useful in areas with challenging terrains or urban environments with numerous obstacles. Secondly, IRS contributes to increased energy efficiency in wireless networks. By intelligently directing signals, it reduces the need for high-power transmissions, thereby conserving energy. Additionally, the use of low-cost sensors and simple surface elements makes IRS a cost-effective solution for improving wireless communication infrastructure.

IRS has numerous potential applications, particularly in the realm of 6G wireless communications. Similar to how massive MIMO revolutionized 5G networks, IRS is poised to play a crucial role in enhancing the capabilities of 6G networks. It can also bolster communication security by intelligently managing signal propagation and improve the accuracy of device location services (H. Yang et al., 2021) and (Y. Chen et al., 2021).

However, deploying and managing an Intelligent Reflective Surface (IRS) in wireless communication networks presents its own set of significant challenges. One of the primary difficulties lies in the need for precise control and coordination of the multiple elements that make up an IRS. Each of these elements, or meta-atoms, is capable of dynamically adjusting the phase and amplitude of reflected signals. To effectively harness the benefits of IRS, it is essential to ensure that these adjustments are meticulously calibrated and synchronized. This level of precision can be highly complex and technically demanding, requiring sophisticated algorithms and real-time feedback mechanisms to continuously adapt to changing environmental conditions. Furthermore, the installation process itself must be carefully planned to avoid disruptions to existing network infrastructure and services.

In conclusion, while IRS technology holds great promise for enhancing wireless communication networks by improving signal quality, extending coverage, and increasing energy efficiency, its deployment and management come with considerable challenges. Addressing these challenges requires comprehensive planning, advanced technical expertise, and robust management strategies to fully realize the benefits of IRS in wireless communications.

2.5 Convex Optimization

2.5.1 The principles of convex optimization

Convex optimization is a specialized area within the broader field of mathematical optimization, concentrating on the minimization of convex functions over convex sets or the maximization of such functions within these sets. A typical convex optimization problem comprises two primary components: an objective function and a feasible set. The objective function is characterized as a convex function, meaning that for any two points on its graph, the line segment joining these points will always lie above or on the graph. In simpler terms, this function exhibits the property of "curving" upward or remaining flat, but never curving downward. More precisely, for any two points x_1 and x_2 in the domain of the function f(x) and $\forall \lambda \in [0,1]$, the function satisfies the inequality:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \quad (2.1)$$

Meanwhile, the feasible set, also known as a convex set, is a collection of points where every point on the line segment connecting any two points within the set is also included in the set.

2.5.2 The benefits and drawbacks of convex optimization

The distinct characteristics of convex optimization yield several significant benefits. These include guaranteed convergence to globally optimal solutions, ensuring that optimization methods will consistently find the optimal solution. Efficient algorithms such as gradient descent, Newton's method, and the interior point method can be employed effectively to solve convex optimization problems. Additionally, robustness is another advantage, as convex optimization problems are generally less sensitive to perturbations and noise compared to non-convex problems.

Despite these advantages, convex optimization also has certain drawbacks. These include high complexity, as some convex optimization problems can be intricate and computationally intensive. Furthermore, convex optimization often

requires stringent constraints, and ensuring that all these are convex can be challenging. Lastly, the sensitivity to assumptions necessitates that specific assumptions about the data and problem structure be made to solve convex optimization issues effectively. If these assumptions are not accurate, the optimization techniques may not achieve the desired outcomes.

2.6 Q – Learning

2.6.1 Definition of Q – Learning

Q-learning is a model-free reinforcement learning algorithm that is both value-based and off-policy. This powerful approach enables an agent to discover the optimal sequence of actions to take, starting from its current state. In Q-learning, the "Q" stands for "quality" which reflects the value or utility of a particular action in terms of its potential to maximize future rewards. The algorithm operates by updating a Q-value (quality value) for each state-action pair, based on the rewards received and the estimated future rewards, guiding the agent to the most beneficial actions over time. Through continuous interaction with the environment, the agent refines its Q-values, progressively improving its decision-making capability to achieve long-term objectives.

Q-Learning is a powerful algorithm employed to determine the optimal action policy in constrained Markov decision processes (MDPs). This technique allows an agent, often referred to as an engine in this context, to learn and improve its decision-making over time. The core objective of Q-Learning is to maximize the total expected rewards the agent can accumulate, starting from its current state and continuing through all subsequent steps. The agent does this by iteratively updating its understanding of the value of each state-action pair, based on the rewards it receives and the potential future rewards it expects to gain.

As the agent interacts with its environment, it employs an exploratory strategy

to sample different actions and observe the outcomes. This exploration enables the agent to refine its policy, gradually identifying which actions yield the highest rewards in the long run. Through this continuous learning process, the agent is able to make more informed and optimal decisions, ultimately leading to the maximization of the total reward it can achieve. In essence, Q-Learning equips the agent with the capability to navigate complex environments and make decisions that are not only optimal for the present moment but also beneficial for future outcomes, ensuring a robust and adaptive action policy.

Figure 2.4 illustrates the steps involved in implementing the Q-Learning algorithm. The first step, known as "Initialize the Q-table," involves creating a table with columns corresponding to the number of actions and rows corresponding to the number of states. This table is used to store the Q values for each state-action pair (s, a). Initially, all values in the Q-table are set to 0 or a small optional value, serving as a preparatory step for learning from experience. The next step is "Choose an Action," where an action is chosen based on the current state. Actions can be selected randomly or by using the updated Q-Table. Typically, the first action and state are chosen based on the Q-Table, usually opting for the highest value. A commonly used policy is the epsilon-greedy policy. According to this policy, with a probability of ε the agent will select a random action (exploration), and with a probability of $(1 - \varepsilon)$, it will choose the action with the highest Q-value (exploitation). This approach ensures a balance between exploring new actions and exploiting known ones. In the "Perform Action" step, the chosen action is executed within the environment. The agent performs the selected action and interacts with the environment. The outcome of this interaction influences the subsequent state and the reward received. The "Measure Reward" step involves evaluating the reward obtained from the environment after performing the action. The reward represents the feedback the agent receives from the environment and can be a positive value (indicating a reward), a negative value (indicating a penalty), or zero. This reward reflects the usefulness of the action taken concerning the agent's long-term goals. Following the above steps is

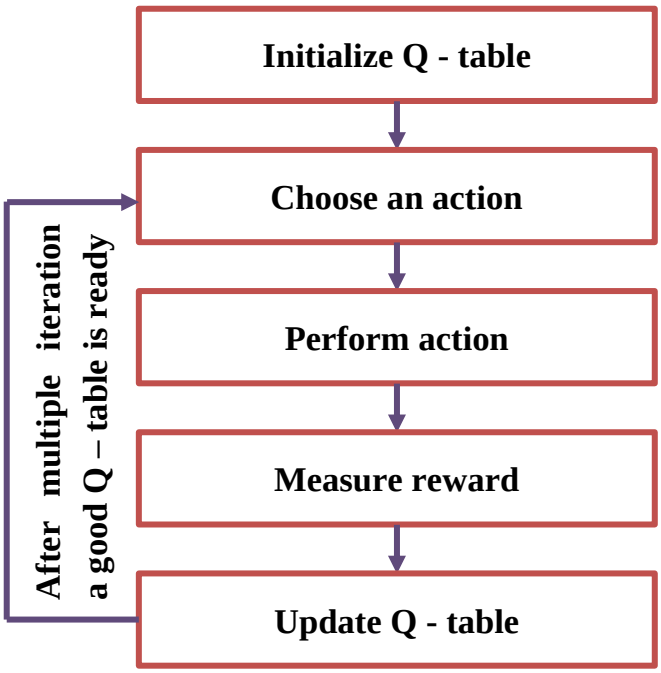

Figure 2.4 Q-Learning algorithm model

the "Update Q-table" step, where the Q-values are updated based on the received reward and the new state. This is typically done using a formula:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[\mathrm{r} + \gamma \max_{a'} Q(s',a') - Q(s,a)] \quad (2.2)$$

where $Q(s,a)$ is the current Q value of state s and action a. α is the learning rate, which determines the rate at which the Q value is updated (usually a small value, for example 0.1). r represents the reward received after performing action a at state s. And γ is called the discount factor, which determines the importance of future rewards (usually a value in the range of 0.8 - 0.99). Moreover, $\max_{a'} Q(s',a')$ is the maximum Q value of actions at the new state s'.

Ultimately, after numerous iterations and refinements of the aforementioned steps, the Q-table evolves to a state of optimization. This meticulously optimized Q-table is populated with Q values that precisely and accurately encapsulate the quality of the potential actions in each state. These Q values serve as a pivotal reference for the agent, enabling it to make informed and optimal decisions regarding which actions to execute. Consequently, the agent's decisions are strategically aligned to

maximize the cumulative reward received over time. This process underscores the efficacy of the Q-learning algorithm in fostering an adaptive learning environment where the agent continually refines its strategy based on experiential feedback, leading to progressively enhanced performance and optimal decision-making capabilities.

2.6.2 Advantages and disadvantages of Q-Learning

Among the notable advantages of the Q-Learning algorithm is its independence from an environment model, which means that it does not require prior knowledge of how the environment operates. This feature makes Q-Learning highly flexible and applicable across a wide range of real-life scenarios. Additionally, direct learning from experience is a crucial benefit, as the agent in Q-Learning improves its performance by learning from its own interactions with the environment, utilizing real-world data to optimize its actions. Furthermore, the long-term optimization capability of Q-Learning allows the agent to focus not only on immediate rewards but also to consider future rewards, leading to more strategic decision-making.

Despite its numerous advantages, Q-Learning is not without its limitations. A primary concern is the extensive learning required to achieve an optimal policy, necessitating a large number of iterations and this iterative process can be both time-consuming and computationally demanding. Furthermore, Q-Learning faces significant challenges in managing large state spaces, as the state space expands, the Q-table becomes increasingly large and complex, potentially leading to memory constraints and prolonged computation times. Another critical issue inherent in Q-Learning is the exploration-exploitation dilemma. The epsilon-greedy strategy, widely utilized in Q-Learning, may not be universally effective. If ε is set too low, the agent may not sufficiently explore the environment to discover optimal actions. Conversely, the agent may excessively explore, thereby neglecting the exploitation of known beneficial actions.

CHAPTER 3. SYSTEM MODEL AND OPTIMIZATION METHOD

3.1 System Model

3.1.1 The introduction to the system model

To commence, we shall model a wireless communication network system comprising a transmitter (denoted as S-Source), a receiver (denoted as D-Destination) equipped with a single antenna, a centrally positioned IRS as well as a set of relay nodes indexed by R = {0, 1, ..., R}. Next, we assume that the Intelligent Reflecting Surface (IRS), comprising N elements (denoted as set $\mathcal{N}$), is connected to a controller. This controller facilitates the intelligent adjustment of the phase angles of each individual element (X. Li et al., 2022).

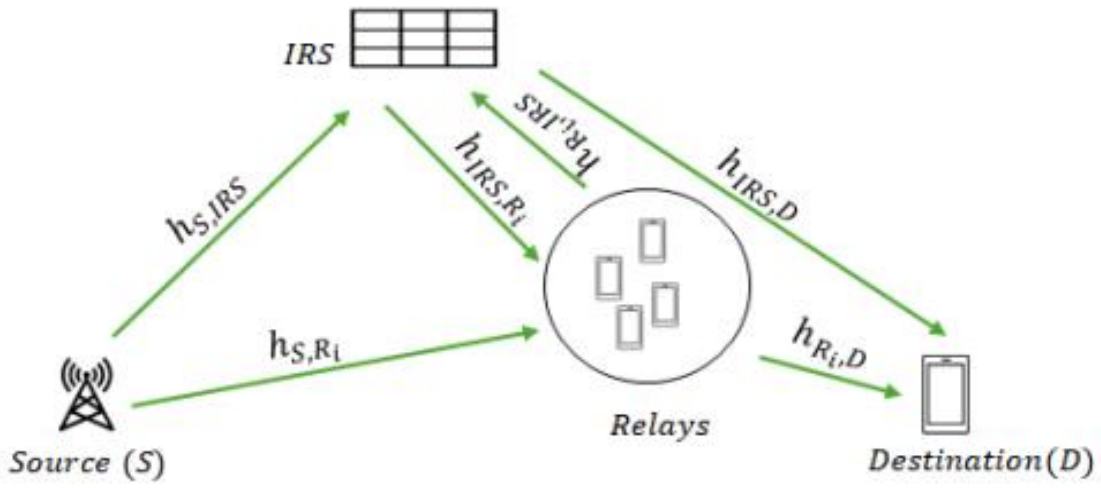

Figure 3.1 System Model

For this system model, we will segment the process into two time slots. In the first time slot, the source (S) transmits the signal to the relay node. Concurrently, a portion of the source (S) signal interacts with the IRS and is subsequently reflected back to the relay nodes. The second time slot involves the relay transmitting the received signal to the destination (D). Simultaneously, a portion of the signal

transmitted into the environment encounters the IRS and is reflected to the destination (D).

The spatial configuration of all relay nodes within the network is meticulously defined using a three-dimensional Cartesian coordinate system (x, y, z), ensuring precise spatial representation. Within this system, the relay nodes are uniformly distributed across a designated area, which forms a circular area with a defined radius (specifically 10 m), and they are assumed to have decoding and forwarding capability - DF. The strategic arrangement of the relays not only enhances the system's reliability but also minimizes signal degradation due to environmental factors. Each relay within the system acts as a crucial link, ensuring that the signal is transmitted most efficiently and stably possible. These relay nodes are assumed to possess decode-and-forward capabilities, allowing them to receive, process, and retransmit signals effectively. This system model proves to be particularly practical and applicable in real-world scenarios. For instance, when a destination device located within a building requires assistance from nearby IRS and relay-enabled devices, the supporting devices, positioned along a street adjacent to the building, can significantly enhance signal strength and reliability. This arrangement ultimately optimizes the overall performance of the communication network, ensuring robust and efficient connectivity.

3.1.2 The explanation of the System Model

This radio communication model comprises key components, including the source, IRS, relays, and destination. These elements collaborate to ensure the efficient transmission of signals from the source to the destination. The IRS plays a pivotal role in optimizing the signal transmission path, while the relays serve to enhance and extend the communication range. Moreover, given that this model operates on a time division duplex basis, the process is divided into two distinct time slots.

To better understand the data reception process: when the sender transmits information, the transmitting station (denoted as Source - S) initiates the communication signal. At this point, digital data such as audio, video, or other data packets are prepared and encoded for transmission. The signal undergoes conversion and modulation, followed by amplification. The amplified signal is then transmitted through the broadcasting antenna, converting the electrical signal into radio waves that are emitted into the environment. As these radio waves propagate, they reach the relays and the IRS. The IRS is designed to receive incoming signals and enhance transmission efficiency by meticulously adjusting the phase of its reflecting elements. This adjustment aims to optimize the transmission path by precisely altering the phase of each element within the IRS. Consequently, the IRS simultaneously transmits the optimized signal to the relay nodes, ensuring a more effective and efficient communication process. The relays consequently receive two signal streams: one direct stream from the source and an indirect stream reflected by the IRS. The relays decode and amplify the signals in preparation for forwarding to the destination. Analogous to the source's process, part of the radio signal is directly transmitted to the destination, while another part is reflected by the IRS towards the destination. At this juncture, the destination will amalgamate all received signal streams. Subsequently, the destination will perform conversion and decoding of the signals to reconstruct the original data.

Through a detailed description, it becomes clear that the source transmits signals to both the relays and the Intelligent Reflecting Surface (IRS) during the first time slot, establishing the foundation for subsequent signal processing. In the next time slot, these signals are forwarded from the relays and IRS to the destination. This dual time slot strategy is designed to optimize interference management and enhance signal quality, ensuring that each system component operates efficiently and harmoniously. The enhanced signal quality achieved highlights the importance of precise operational design in advanced communication systems, providing a robust framework for future innovations in wireless communication.

3.2 Optimization Techniques - Algorithms for Optimization

3.2.1 Channel Model

Once the system model necessitating optimization is established, we will advance to the next phase, which entails formulating the optimization problem for this model. At this juncture, our attention will be directed towards the channel model of the system.

In this section, we will begin to examine the combination of Line of Sight (LoS) and Non-Line of Sight (NLoS) links (I. Yildirim et al., 2021). Given the existing arrangement of LoS and NLoS paths, where the forwarding link to the destination is presumed to be available, we will proceed accordingly. With this in mind, and considering the characteristics and applications of the 3GPP TR 38.901 UMi-Street Canyon path loss model (D. Dampahalage et al., 2020), we will implement this model to ensure that:

$$PL_{\text{LOS}} = 32.4 + 21\log_{10}(d) + 20\log_{10}(f_c) \tag{3.1}$$

$$PL'_{NLOS} = 22.4 + 35.5\log_{10}(d) + 21.3\log_{10}(f_c) - 0.3(h_{UT} - 1,5) \tag{3.2}$$

$$PL_{NLOS} = \max(PL_{\text{LOS}}, PL'_{NLOS}) \tag{3.3}$$

In these equations, PL means path loss and is calculated using the formula $PL = P_R / P_T$, where P_R represents the received power and P_T signifies the transmitted power. In this context, PL_{LOS} refers to the path loss for the Line-of-Sight (LoS) link, while PL'_{NLOS} indicates the path loss for the Non-Line-of-Sight (NLoS) link and the PL_{NLOS} is the maximum path loss value, which is established by selecting the highest value between the Line-of-Sight (LOS) path loss and the Non-Line-of-Sight (NLOS) path loss. The variable d denotes the distance in meters between any two communication devices, f_c stands for the carrier frequency measured in gigahertz (GHz), and h_{UT} specifies the height of the user terminal (in this study, the

height of the user terminal $h_{UT} = 1\ m$, which represents the typical height of a mobile terminal).

Considering the system model illustrated in figure 3.1, in a scenario where there are no obstacles and the signal transmission is not attenuated by the transmission distance (i.e., the path from source to IRS to relay to destination is optimized and unobstructed), and applying the Umi-Street Canyon model as per the 3GPP TR 38.901 standard, we derive formula (3.1). Formula (3.1) represents the transmission loss occurring when the signal is transmitted from the source to the destination without any obstructions. This situation is often regarded as the most ideal in wireless communication, as there are no barriers to degrade the signal strength.

We get formula (3.2) by applying the same model to the system in a non-line of sight (NLoS) situation, which is when there are a lot of things in the way of the signal transmission or the signal is weakened by the distance. This formula illustrates the transmission loss that occurs when the signal, transmitted from the source to the destination, is obstructed by obstacles such as buildings, trees, terrain,... These obstacles cause the signal to follow non-straight paths, leading to reflection, diffraction, and scattering, which results in greater attenuation than in the LoS case. The primary objective of calculating NLoS loss is to incorporate adjustments based on the height of the receiving device, thereby enhancing the model's accuracy in real-world situations.

Following formulas (3.1) and (3.2), formula (3.3) is constructed to represent the maximum path loss between line-of-sight (LoS) and non-line-of-sight (NLoS) paths. This approach offers a more comprehensive and accurate assessment of the actual loss that can occur within the transmission environment by accounting for the highest possible loss value. Formula (3.3) ensures that a system design can accommodate worst-case scenarios, thereby providing a safety margin and enhancing the overall reliability and robustness of the design model. By thoroughly considering both LoS and NLoS conditions, this formula aids in the development of more efficient

and resilient communication systems that can maintain performance in diverse and challenging environments.

All analyzed connections are constructed using the designated equations (3.1) and (3.3), while also integrating Rician small-scale fading to address fluctuations in the signal trajectory. Considering the above-given factors, after the end of the first transmission interval, the i[th] relay, referred to as R_i, acquires a signal defined by:

$$y_{R_i} = \sqrt{P_S} x_S (h_{S,R_i} + h_{\mathrm{IRS},R_i}^T \Phi_1 h_{S,\mathrm{IRS}}) + z_{R_i} \tag{3.4}$$

The variable y_{R_i} represents the amount of signal information acquired at the i[th] relay during the first time slot. This includes the whole quantity of information obtained from both the source delivering the signal directly and the IRS reflecting the signal. By considering both direct transmission and reflected signal components, it offers a thorough assessment of the signal information gathered at the i[th] relay during the first communication phase.

In this context, P_S denotes the transmit power of the source (S). The variable x_S represents the information signal, which possesses unit power. The term h_{S,R_i} signifies the channel coefficient of the link from the source (S) to the i[th] relay R_i. Similarly, h_{IRS,R_i} is the channel coefficient of the link from the IRS to the i[th] relay R_i, and $h_{S,\mathrm{IRS}}$ denotes the channel coefficient of the link from the source (S) to the IRS. It is important to note that the superscript $(.)^T$ refers to the transpose operation.

The Additive White Gaussian Noise (AWGN) at relay R_i is represented by z_{R_i}, which has a zero mean and a variance of σ^2. The matrix Φ_1 is the diagonal reflection matrix of the IRS during the first communication time slot. This matrix can be expressed as $\Phi_1 = diag(\alpha_1^1 e^{j\phi_1^1}, \alpha_1^2 e^{j\phi_1^2}, ..., \alpha_1^N e^{j\phi_1^N})$, which represents the reflection amplitude coefficient and ϕ_1^n denotes the phase shift of the n[th] element of the IRS. It is assumed that the phase shift $\alpha_1^1 = \alpha_1^2 = ... = \alpha_1^n = 1$ and ϕ_1^n falls within the interval

$[0, 2\pi]$ ($\phi_1^n \in [0,2\pi]$), given that the reflection amplitude coefficient is not the primary optimization objective.

As mentioned above, along with specifying the quantity in the expression of formula (3.4), this equat=ion specifically illustrates the amount of signal received at the ith relay, R_i, during this time slot. To thoroughly understand the implications of this equation, it is crucial to break it down into two distinct components: The initial part of the equation will be referred to as the direct signal component ($\sqrt{P_S} x_S h_{S,R_i}$) - this part is regarded as the most critical component because it encompasses the signal transmitted directly fro=m the source to the relay and the direct transmission path ensures a high-quality signal with minimal interference, making it an essential element in the overall communication process. The second term in the expression, which remains, is referred to as the signal reflection term ($\sqrt{P_S} x_S h_{\text{IRS},R_i}^T \Phi_1 h_{S,\text{IRS}}$) - this component represents the segment of the signal that is reflected by the IRS and then received by the relay, so this reflection demonstrates that the indirect path significantly enhances signal coverage and strength, particularly in environments with obstacles. Within the equation, we observe the inclusion of a vector component, which is the result of multiplying three channel coefficient matrices. These matrices correspond to the channel coefficient from the IRS to R_i - h_{IRS,R_i}, the diagonal matrix representing the adjusted phase angles of the IRS- Φ_1 and the channel coefficient from the source to the IRS- $h_{S,\text{IRS}}$. In the expression, we also encounter the transposition operation, denoted by $(.)^T$. The necessity for transposing the channel coefficient matrix from the IRS to R_i arises from the fact that it is the product of three matrices. Specifically, the chann=el coefficient matrix is represented as a column vector of size Nx1:

$$h_{\mathrm{IRS},R_i} = \begin{bmatrix} h_{\mathrm{IRS}_1,R_i} \\ h_{\mathrm{IRS}_2,R_i} \\ \dots \\ h_{\mathrm{IRS}_N,R_i} \end{bmatrix} \quad and \quad h_{S,\mathrm{IRS}} = \begin{bmatrix} h_{S_1,\mathrm{IRS}} \\ h_{S_2,\mathrm{IRS}} \\ \dots \\ h_{S_N,\mathrm{IRS}} \end{bmatrix} \tag{3.5}$$

while the diagonal matrix is of NxN:

$$\Phi_1 = \begin{bmatrix} \Phi_1 & & & \\ & \Phi_2 & & \\ & & \dots & \\ & & & \Phi_N \end{bmatrix} \tag{3.6}$$

This transposition ensures proper alignment and compatibility of matrix dimensions during the multiplication process:

$$h_{\mathrm{IRS},R_i}^T \Phi_1 h_{S,\mathrm{IRS}} = \left[h_{\mathrm{IRS}_1,\mathrm{R}_i}, h_{\mathrm{IRS}_2,\mathrm{R}_i}, \dots, h_{\mathrm{IRS}_N,\mathrm{R}_i} \right] * \begin{bmatrix} \Phi_1 & & & \\ & \Phi_2 & & \\ & & \dots & \\ & & & \Phi_N \end{bmatrix} * \begin{bmatrix} h_{S_1,\mathrm{IRS}} \\ h_{S_2,\mathrm{IRS}} \\ \dots \\ h_{S_N,\mathrm{IRS}} \end{bmatrix} \tag{3.7}$$

Thus, (3.4) provides a comprehensive depiction of the signal reception dynamics at the i[th] relay, integrating both direct and reflected signal components. This thorough analysis underscores the intricate interactions between various elements in the communication system and highlights the critical factors contributing to the overall signal quality and reliability.

Therefore, the communication rate for the link utilized in the first time slot can be determined as follows:

$$C_{t_1} = \log_2 \left(1 + \frac{P_S \| h_{S,R_i} + h_{\mathrm{IRS},R_i}^T \Phi_1 h_{S,\mathrm{IRS}} \|^2}{\sigma^2} \right) \tag{3.8}$$

After establishing formula (3.4), we will use it as the foundation to calculate the channel capacity in the first time slot, denoted as C_{t_1} in formula (3.8). Here, ||.||

represents the Euclidean norm operation, which is frequently applied to complex numbers and is consistently used throughout this paper for all analog symbols, providing a standardized method for measuring the magnitude of complex values. Examining the details, formula (3.8) is derived based on the Shannon's equation, which is expressed as $C = \log_2(1 + SNR/\sigma^2)$ (where SNR is the Signal-to-Noise Ratio and σ^2 is the variance, representing the power of white Gaussian noise), but the SNR is effective in this case. Applying this method allows us to accurately quantify the channel capacity, ensuring a comprehensive understanding of the system's communication efficiency.

Likewise, during the second communication time slot, the signal that is received by destination D can be expressed as follows, and this involves considering the transmission conditions and parameters that affect the signal's quality and strength during this specific phase of the communication process. By thoroughly analyzing the received signal, we can gain insights into the performance and reliability of the communication link in the second time slot:

$$y_D = \sqrt{P_{R_i}} x_{R_i} (h_{R_i,D} + h_{\mathrm{IRS},D}^T \Phi_2 h_{R_i,\mathrm{IRS}}) + z_D \tag{3.9}$$

The variable y_D denotes the quantity of signal information received at the destination during the second time slot. In this context, P_{R_i} denotes the transmit power of the selected relay. The variable x_{R_i} represents the information signal transmitted by the ith relay, which possesses unit power. The term $h_{R_i,D}$ signifies the channel coefficient of the link from the the ith relay R_i to destination (D). Similarly, $h_{\mathrm{IRS},D}$ is the channel coefficient of the link from the IRS to D, and $h_{R_i,\mathrm{IRS}}$ denotes the channel coefficient of the link from the the ith relay to the IRS.

The Additive White Gaussian Noise (AWGN) at relay D is represented by z_D which has a zero mean and a variance of σ^2. The matrix Φ_2 is the diagonal reflection matrix of the IRS during the second communication time slot. This matrix can be

expressed such that $\Phi_2 = diag(\alpha_2^1 e^{j\phi_2^1}, \alpha_2^2 e^{j\phi_2^2}, ..., \alpha_2^N e^{j\phi_2^N})$ represents the reflection amplitude coefficient and ϕ_2^n denotes the phase shift of the nth element of the IRS and follow the same assumptions for Φ_1.

Analogous to formula (3.4), the key distinction in formula (3.9) lies in its application to the system model during the second time slot. When examining this formula alongside the system model, it becomes evident that the square root component of $\sqrt{P_{R_i}} x_{R_i} h_{R_i,D}$ is the most crucial element. This component encapsulates the volume of information that the transmitter aims to convey to the receiver via a direct transmission path.

Subsequently, having established the formula for calculating the amount of information received during the second time slot, we can now proceed to derive the corresponding formula for calculating the communication capacity at this second time slot. It is important to highlight that during this second operational period, particular attention must be given to the phase shifts of the IRS elements. Assume that these phase shifts are discrete; for each element, the phase shift set can be expressed as $\Gamma = \left[0, \Delta\phi, ..., \Delta\phi\ K - 1\right]$, where $\Delta\phi = 2\pi / K$ and K = 2^b, with b representing the number of bits utilized to define these phase shifts. By considering these phase shifts, we can ensure precise alignment and optimization of the IRS elements, thereby enhancing the overall communication efficiency and capacity during the second time slot.

$$C_{t_2} = \log_2(1 + \frac{P_{R_i} \| h_{R_i,D} + h_{\text{IRS},D}^T \Phi_2 h_{R_i,\text{IRS}} \|^2}{\sigma^2}) \tag{3.10}$$

From the system model, we can derive a formula (3.10) to calculate the channel capacity in the second time slot at the destination. This formula is somewhat analogous to (3.8). However, instead of using the transmission power P_S as in formula (3.8), the transmission power here becomes P_{R_i}, since the information signal received by the destination is transmitted by the relay node at this time. Furthermore,

this formula elucidates that the amount of signal received at the receiver is determined by the transmission power from the relay and is reflected by the IRS. A comprehensive understanding of this formula ensures that we thoroughly grasp the mechanisms related to the transmission and reflection processes, thereby facilitating a more accurate interpretation of the resulting signal indicators.

The communication rate of the end-to-end link, extending from the source (S) to the destination (D) via the IRS and a suitably selected relay R_i, can be articulated as follows:

$$C_{S,D} = \frac{1}{2}\min(C_{t_1}, C_{t_2}) \tag{3.11}$$

In conclusion, after establishing the formula for determining the capacity at two time slots using the system model, we get formula (3.11). It is vital to note that, from the start, the model was split into two time slots, assuming the system operated in time division duplex (TDD) mode, as well as that to establish the end-to-end capacity, we must total the capacity of each particular time period. To further appreciate formula (3.11), we need to understand the TDD process, where each transmission connection between system components is employed for a defined duration to transfer data, thus segmenting the communication process into discrete time periods. This segmentation, along with the two-time slot structure of our model, validates the inclusion of the "1/2" coefficient in the formula. Moreover, it is vital to note that when a system has numerous parallel transmission lines, the maximum attainable transmission rate of the total channel is restricted by the slowest transmission line. This notion is embodied in the formula as the minimal capacity as $\min(C_{t_1}, C_{t_2})$, providing a realistic and trustworthy estimation of the entire transmission capacity. By accounting for the slowest connection, the formula increases the resilience and efficiency of the design and performance analysis of communication systems, leading to more accurate and predictable outputs.

3.2.2 Problem Formulation

Upon completing the derivation of the equation to calculate the terminal capacity, our next step involves initiating the optimization process to enhance this capacity. To achieve the optimization of the equation, we will employ the convex optimization method. This approach is aligned with the primary objective of this paper, which is to concurrently optimize the phase angles of the IRS and to select an appropriate relay to facilitate effective communication between the source and the destination. By leveraging the convex optimization technique, we aim to identify the optimal configurations of the IRS phase angles and the relay selection that collectively enhance the overall communication performance. This comprehensive optimization process ensures that the communication system operates at peak efficiency, thereby maximizing the terminal capacity and achieving the desired communication objectives. Consequently, the optimization problem is formulated as follows:

$$\boldsymbol{P1}: \underset{\Phi_1,\Phi_2,R_i}{maximize} C_{S,D} \tag{3.12a}$$

Subject to:

$$C_1: \phi_1^n, \phi_2^n \in \Gamma \ \forall \ n \in N \tag{3.12b}$$

$$C_2: \sum_{i=1}^{R} x_i = 1 \tag{3.12c}$$

$$C_3: x_i \in \ 0,1 \ , i = 1,2,...,R \tag{3.12d}$$

Having successfully derived the formula required for optimization through the system model, we can now shift our focus entirely to the algorithmic aspects of our analysis. At this juncture, further consideration of the system model is unnecessary. Our primary objective is to optimize the received capacity at the receiver, which necessitates the maximization of (3.11). To achieve this, we will utilize the convex optimization method for the algorithmic optimization of the system. In this context, the maximization of (3.11) is referred to as the objective function. To effectively

optimize this objective function, we will introduce three constraints related to the phase angles during the two time slots, as well as an additional constraint concerning the selection of the i[th] relay.

The primary objective of this paper is to enhance the system's performance by utilizing both the relay network and the IRS. Consequently, the symbols Φ_1, Φ_2, and R_i are prominently displayed below the term "maximize" - $\underset{\Phi_1,\Phi_2,R_i}{maximize}$. These symbols signify that the system's optimization can be achieved through various approaches: exclusively optimizing the IRS phase angles in the two time slots (denoted by Φ_1 and Φ_2), solely by fine-tuning and selecting the i[th] relay, or by employing a combination of both strategies. Next, we delve deeper into the constraints necessary for optimizing the objective function. Initially, we consider the constraints associated with fine-tuning the IRS phase angles. It is important to note that the IRS phase angles in the two time slots, Φ_1 and Φ_2, are members of the discrete phase shift set $\Gamma = \left[0,\ \Delta\phi,\ ...,\ \Delta\phi\ \ K\ -1\right]$, where $\Delta\phi = 2\pi / K$ and K = 2^b. The subsequent two conditions pertain to the optimization based on the selection of the i[th] relay. According to condition (3.12c), it is stipulated that only one relay is selected, and only this selected relay can receive information from the source and subsequently transmit it to the destination. This implies that the number of relay nodes capable of supporting source-destination communication is limited to one. The relay selection parameter is represented by x_i. Condition (3.12d) further clarifies the relay selection process. This condition specifies that the relay chosen for optimization will have a value of 1, while all other relays will be assigned a value of 0. By adhering to these constraints, we ensure a feasible and effective optimization process, thereby enhancing the overall performance and reliability of the communication system.

Following the optimization of $C_{S,D}$ in ***P1*** is essentially equivalent to maximizing $\min(C_{t_1}, C_{t_2})$. In a similar fashion, to maximize $\min(C_{t_1}, C_{t_2})$, we need to employ (3.8) and (3.10). However, it is noteworthy that in these equations, only

the Signal-to-Noise Ratio (SNR) component can be varied. Let γ_1 and γ_2 represent the SNR of the i[th] relay at the first time slot and the SNR at the destination during the second time slot, respectively. Consequently, by removing the logarithmic terms from (3.8) and (3.10), the problem ***P1*** can be reformulated in terms of the SNR of the links during the two communication time slots, instead of the terminal rates. Therefore, the optimization problem in ***P1*** becomes:

$$\textbf{\textit{P2}}: \underset{\Phi_1,\Phi_2,R_i}{maximize}(\min[\gamma_1,\gamma_2]) \tag{3.13a}$$

Subject to:

$$C_1: \phi_1^n, \phi_2^n \in \Gamma \ \forall \ n \in N \tag{3.13b}$$

$$C_2: \sum_{i=1}^{R} x_i = 1 \tag{3.13c}$$

$$C_3: x_i \in \ 0,1 \ , i = 1,2,...,R \tag{3.13d}$$

Here, we will elucidate why (3.11) is optimized by minimizing (3.8) and (3.10). Given that the variable component in these two equations is the Signal-to-Noise Ratio (SNR), we have excluded them from the logarithmic function to minimize their values directly. This is due to the inherent definition of SNR, which is the ratio of the desired signal power to the noise power; hence, a higher SNR indicates better signal quality, and vice versa. Therefore, we transform the optimization of (3.11) into the optimization of $\min(\gamma_1,\gamma_2)$, where $\gamma_1 = \frac{P_S \parallel h_{S,R_i} + h_{\text{IRS},R_i}^T \Phi_1 h_{S,\text{IRS}} \parallel^2}{\sigma^2}$ and $\gamma_2 = \frac{P_{R_i} \parallel h_{R_i,D} + h_{\text{IRS},D}^T \Phi_2 h_{R_i,\text{IRS}} \parallel^2}{\sigma^2}$.

Analyzing the expressions for γ_1 and γ_2, it becomes evident that, with fixed values of P_S and P_{R_i}, γ_1 is influenced by Φ_1, and γ_2 is influenced by Φ_2. Consequently, ***P2*** can be systematically decomposed into two distinct sub-problems, each of which addresses a critical aspect of the overall optimization challenge. The initial sub-problem centers on the meticulous optimization of the phase angles of IRS.

This process involves a detailed analysis and fine-tuning of the phase shifts to achieve optimal reflection properties that enhance signal strength and quality:

$$\boldsymbol{P3}\text{: } \underset{\Phi_1,\Phi_2}{maximize}(\min[\gamma_1,\gamma_2]) \tag{3.14a}$$

Subject to:

$$C_1 : \phi_1^n, \phi_2^n \in \Gamma \ \forall \ n \in N \tag{3.14b}$$

and the second sub-problem delves into the critical task of selecting the most appropriate relay. This aspect of the optimization process requires thoroughly examining various relay options, each with unique attributes and potential contributions to the overall system performance. The objective here is to identify the relay that can most effectively enhance signal transmission, thereby ensuring robust and reliable communication:

$$\boldsymbol{P4}\text{: } \underset{R_i}{maximize}(\min[\gamma_1,\gamma_2]) \tag{3.15a}$$

Subject to:

$$C_1 : \sum_{i=1}^{R} x_i = 1 \tag{3.15b}$$

$$C_2 : x_i \in \ 0,1 \ , i = 1,2,...,R \tag{3.15c}$$

Based on the analysis above, it is evident that the transmitted power at two distinct time slots, denoted as P_S and P_{R_i}, remains almost constant as there is no change in their magnitudes. Instead, the variation lies in the SNR values, which can be influenced by changes in the phase angle of the IRS. Therefore, the objective function of ***P3*** aims to optimize the SNR through adjustments in the phase angles (Y. Yang et al., 2019). By delving deeper into the analysis, we observe that γ_1, the synthesis of reflected waves from the IRS in the first time slot, is affected by the diagonal matrix Φ_1. By adjusting Φ_1, we can alter the reflected waves from the IRS to resonate with the direct signal from R_i, thereby enhancing the SNR value of γ_1. If

the phase angle of the IRS is not optimized, it can lead to partial cancellation of the signal due to destructive interference, consequently reducing γ_1. Similarly, the optimization applies to γ_2. By optimizing the phase angle Φ_2, we can enhance the SNR value of γ_2, thereby optimizing the overall signal quality. Conversely, failing to optimize Φ_2 can result in weakened signal reception at the receiver due to suboptimal interference patterns. In summary, the goal of optimizing the $\min[\gamma_1, \gamma_2]$ function in ***P3*** necessitates careful and simultaneous adjustments of Φ_1 and Φ_2 to maximize the values of both γ_1 and γ_2. This meticulous adjustment ensures that the reflected waves from the IRS effectively support the transmission line, thereby improving overall communication efficiency and signal quality (Q. Wu and R. Zhang, 2019).

While ***P3*** primarily focuses on optimizing the system by adjusting the phase angles of the IRS, ***P4*** employs a different strategy by concentrating on the selection of the most appropriate relay. This strategic shift is essential as the process of relay selection can have a profound impact on the overall performance and efficiency of the communication system. Modifying the choice of relay provides multiple advantages for system optimization, including enhanced signal quality by selecting the most suitable relay. The system ensures that the signal is transmitted along the most favorable path - this enhances signal quality and reduces potential interferences. This strategy also aids the model in achieving load balancing. By selecting different relays, the system can evenly distribute the load across the network - this prevents any single relay from becoming a bottleneck and ensures the more efficient utilization of network resources. Therefore, optimizing based on relay selection in ***P4*** strategically improves the overall performance of the system, ensuring robust and efficient communication between the source and destination. This approach effectively complements the optimization of IRS phase angles, providing a comprehensive strategy to maximize the operational effectiveness of the communication system.

Upon closely examining the use of convex optimization for problem-solving, we observe that ***P3*** is inherently non-convex due to the objective function being a non-concave function of Φ_1 and Φ_2. To simplify the resolution of ***P3***, the phase angles of the IRS are optimized separately for each transmission time slot. Specifically, Φ_1 is optimized to maximize γ_1, while Φ_2 is optimized to maximize γ_2.

This methodical strategy allows us to break down the complex non-convex problem into more manageable sub-problems, ensuring that we can effectively optimize the system's performance by focusing on the phase angles for each specific transmission time slot. By adopting this approach, the optimization problem defined by (3.15) and (3.16) effectively mirrors the slot-by-slot optimization Φ_1 and Φ_2 to maximize γ_1 and γ_2, respectively, as well as R_i of relay. Consequently, when considering the first communication time slot, the optimization problem becomes:

$$\textbf{P5: } \underset{\Phi_1}{maximize}\ \gamma_1 \tag{3.16a}$$

Subject to:

$$C_1 : \phi_1^n \in \Gamma \forall n \in N \tag{3.16b}$$

The same line of reasoning applies to the second transmission time slot as well.

P5 is approached by breaking down the problem into smaller, more manageable sub-problems. Each sub-problem is solved individually, and the solutions are then combined to address the initial complex optimization problem. Initially, we focus on applying the convex optimization method to implement the model's optimization algorithm without explicitly considering whether the problem itself is convex. When we begin analyzing the problem from ***P1***, it becomes clear that the objective function, represented by (3.11), is not convex. This is because the graph of (3.11) is a min function graph, which inherently lacks a plane connecting any two points completely below the graph. Additionally, to optimize (3.11), we rely on (3.8) and (3.10). These equations are logarithmic functions with expressions inside the log function dependent on the Euclidean norm, making them nonlinear functions capable

of producing local optima. Moreover, when we examine the constraints in ***P1***, we observe that these conditions are discrete and do not form a convex set. Therefore, the problem to be solved cannot be classified as a convex optimization problem due to the nature of the objective function being a min function, the nonlinearity of C_{t_1} *and* C_{t_2}, and the presence of discrete constraints.

In conclusion, the complexity and nature of the objective function and constraints in ***P1*** necessitate a tailored approach that accounts for the non-convexity and discrete elements, ensuring that the optimization process is both effective and efficient. By understanding these intricacies, we can better address the challenges posed by the optimization problem in ***P5***.

3.2.3 Proposed approach

As highlighted in the summary description of the topic, once we have completed the system model and developed the formula for the convex optimization algorithm for the system, we will proceed to the next step of using the Q-learning algorithm to simulate the system. Consequently, we will consider several proposed approaches.

a) Successive refinement-based phase angle optimization

To tackle the non-convex problem ***P5***, we must simplify certain expressions in the objective function (Y. Yang et al., 2019). Given that the noise term in the expression for γ1 is a constant, we can concentrate on streamlining the numerator instead. Next, we will go into the calculation formula setup section. Let $\theta = h_{\mathrm{IRS},Ri} diag(h_{\mathrm{S,IRS}})$, $A = \theta^H \theta$, $b = \theta^H h_{S,Ri}$, $v = [\mathrm{e}^{j\phi_1^1}, \mathrm{e}^{j\phi_1^2}, ..., \mathrm{e}^{j\phi_1^N}]$. Substituting these terms into the numerator of $\gamma_1 = \dfrac{P_S \| h_{S,R_i} + h_{\mathrm{IRS},R_i}^T \Phi_1 h_{S,\mathrm{IRS}} \|^2}{\sigma^2}$ considering that $\dfrac{P_S}{\sigma^2}$ is fixed leads to,

$$\| h_{S,R_i} + h_{\text{IRS},R_i}^T \Phi_1 h_{S,\text{IRS}} \|^2 = v^H A v + 2\,\text{Re}\{v^H b\} + \| h_{s,Ri} \|^2 \quad (3.17)$$

If we optimize each phase angle ϕ_1^N in (3.17) individually while keeping the other phase angles fixed, the function becomes linear concerning the specific phase angle ϕ_1^N we're focusing on. In other words, if we're optimizing v_l ($l \in N$), then all the other v_k, where $k \neq l$ and $k \in N$ are held constant. This allows us to express Equation 14 as:

$$\| h_{S,R_i} + h_{\text{IRS},R_i}^T \Phi_1 h_{S,\text{IRS}} \|^2 = 2\,\text{Re}\{v_l w_k\} + \sum_{k \neq l}^{N} \sum_{i \neq l}^{N} v_k A_{k,i} v_i + C \quad (3.18)$$

where $w_l = \sum_{k \neq l}^{N} A_{l,k} v_k + b_k$ and $C = A_{l,l} + 2\,\text{Re}\left\{ \sum_{k \neq l}^{N} v_k b_k \right\} + \| h_{S,Ri} \|^2$. Therefore, determining the optimal phase angle, we need to focus on optimizing the phase angle that minimizes the difference between each ϕ_1^N and w_l. This method is also applied to find the optimal phase angles ϕ_2^N for the second communication time slot.

The details of the algorithm are outlined in ***Algorithm 1***. We use ξ as the convergence stop threshold (D. Dampahalage et al., 2020). By optimizing the angles of the reflecting elements one at a time, Algorithm 1 effectively performs a linear search over discrete phase shift levels (a total of N steps) to find the optimal phase shift for the m_{th} reflector. Consequently, the worst-case computational complexity of Algorithm 1 can be deduced to be $\mathcal{O}(n)$ steps for an input size of n, akin to a typical linear search algorithm (C. W. Royer and S. J. Wright, 2018).

With the optimal phase angles of the IRS determined using ***Algorithm 1***, we then select the relay. Using the optimally chosen relay, we repeat ***Algorithm 1*** for the second communication time slot to obtain the optimal phase angles Φ_2^* and corresponding parameters C_{t2}^*. These results are then used to calculate the end-to-end link rate.

Table 3-1 Successive Refinement for Phase Optimization

ALGORITHM 1: Successive Refinement for Phase Optimization

Initialization $\Phi_1 = \Phi_1^0$, m=0

Input $h_{\mathrm{IRS},R_i}, P_S, h_{S,Ri}, h_{S,\mathrm{IRS}}, \xi, N$

$$C_{t1}^0 = \log_2\left(1+\frac{P_S \| h_{S,R_i} + h_{\mathrm{IRS},R_i}^T \Phi_1 h_{S,\mathrm{IRS}} \|^2}{\sigma^2}\right)$$

while $|C_{t1}^{m+1} - C_{t1}^m| > \xi$ **do**

  for m = 1 : N **do**

    $\phi_1^{l*} = \arg\min_{\phi_1 \in \Gamma} |\phi_1 - \angle \mathrm{w}_l|$

    m = m + 1

  End

  $$C_{t1}^m = \log_2\left(1+\frac{P_S \| h_{S,R_i} + h_{\mathrm{IRS},R_i}^T \Phi_1 h_{S,\mathrm{IRS}} \|^2}{\sigma^2}\right)$$

End

return Optimal phase angles: Φ_1^*, Optimal achievable rate; C_{t1}^*

b) Relay selection using Q-Learning

In Q-learning, an agent learns by interacting with its environment and transitions from one state to another by taking specific actions that aim to maximize the agent's cumulative reward. In this study, the agent is assumed to be positioned at the IRS controller. The states *s* that constitute the learning environment are defined by the set of relays, represented as $S = \{R_1, R_2, ..., R_n\}$, where each R_i denotes a relay.

The possible actions the agent can take are derived from the features of the relays, particularly the channel gains of the links between the relays and the IRS during the second communication time slot. Hence, the action set A can be described as $A = \{h_{R1,IRS}, h_{R2,IRS,...,} h_{Rr,IRS}\}$. The agent moves between states S_t in S by selecting

an action at in *A*, where *t* represents the training cycles or episodes of the machine learning process, and in return, receives a reward $r(s_t, a_t)$.

For every state-action pair, a Q-value $Q(s_t, a_t)$ is computed and updated. These Q-values are stored in a matrix known as the Q-table. The updates to the Q-table entries are governed by the well-established Bellman equation, which is crucial for determining the optimal policy that maximizes the cumulative reward over time:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \delta(r(s_t, a_t) + \eta(\max_{a_t} Q(s_{t+1}, a))) \quad (3.19)$$

where $Q(s_t, a_t)$ is the current value of the Q function for state s_t and action a_t; δ is a learning rate, usually in the range [0,1][0,1][0,1], which determines the extent to which the new value will affect the current value and this also helps the update process to be stable, by controlling the influence of the new value on the current Q value; $r(s_t, a_t)$ is the reward received when performing action a_t at state s_t; η is a discount factor, usually in the range [0,1] to determine the importance of future rewards and if η is close to 1, Q-learning will give more importance to future rewards, while if η is close to 0, Q-learning will only care about short-term rewards; $\max_{a_t} Q(s_{t+1}, a)$ is the maximum value of the Q function for the next state S_{t+1} with all possible actions a_t.

Table 3-2 Reward Matrix Procedure

ALGORITHM 2: Reward Matrix Procedure

Input h_{IRS,R_i}

DI = diag(h_{IRS,R_i})

for i = 1 : n_R **do**

    RW(i,:) = DI./D(i)

    RW(i,:)<1=0

End

return Reward matrix = RW

An important aspect of Q-learning is the reward matrix, which consists of the rewards an agent receives for each state transition. In our work, the reward matrix is populated based on the channel gain of each relay to IRS link $h_{R_i,\mathrm{IRS}}$. As the agent transits between states or relays, it observes the reward, which is derived from the link gain associated with moving between relays. This approach requires the values of the link gains to be known. In scenarios where the IRS is deployed by an operator, this information can be provided to the IRS controller, facilitating accurate updates to the reward matrix.

In the process of designing the reward matrix (RW) for our investigations, we first calculate the gains of the links between each relay and the elements of the IRS. These relay-to-IRS channel gains are then stored in a diagonal matrix (DI). For each relay, we calculate the ratio of the relay-to-IRS link gain to each entry in DI, which then form the row entries of RW. Entries in RW that are less than 1 are set to zero to reduce complexity. This reduction in the search space is crucial because Q-Learning involves simulating an exhaustive search to identify the optimal relay-enabled device (Y. He et al., 2020).

The methodology for obtaining the reward matrix is detailed in Algorithm 1 and Algorithm 2 outlines the Q-learning-based relay selection algorithm. The motivation behind utilizing Q-Learning stems from its simpler implementation. Additionally, the division of the formulated problem in “section 3.2.1 and 3.2.2” into phase optimization and relay selection enables the use of a less complex reinforcement learning method such as Q-Learning. Given that the task representation or initialization of Algorithm 2 is adequate, its worst-case computational complexity for achieving a goal state can be tightly bounded by $\mathcal{O}(n^3)$ action executions for an input size of n (S. Koenig and R. G. Simmons, 1993). The integration of successive refinement-based phase angle optimization with Q-learning-based relay selection constitutes our proposed Q-learning-based joint IRS and relay-assisted communication (QL-JIRA). A summary of benchmark methods used for comparison with QL-JIRA is provided in “section c)”.

c) Benchmark approaches

In our study, we undertake a thorough evaluation of the performance of the Q-learning-based joint IRS and relay-assisted communication (QL-JIRA) system that we have developed. The efficacy of our proposed system is assessed in comparison with several established benchmark approaches. Specifically, we examine our QL-JIRA system against six distinct benchmark methods to ensure a comprehensive analysis. These benchmarks encompass the No Relay approach (D. Dampahalage et al., 2020), which omits the use of (a) No Relay approach, (b) the Random Selection (RS), (c) the Fixed Phase Algorithm (FPA), and (d) the Random Phase Algorithm (RPA). This rigorous comparative analysis enables us to elucidate the advantages and potential limitations of our QL-JIRA system, thereby highlighting its effectiveness and identifying areas for further improvement.

The RS algorithm, selects relays randomly without any specific criterion, and this randomness is limited to relay selection only. For phase angle optimization, the successive refinement algorithm is applied. The FPA employs optimized IRS phase angles based on the successive refinement algorithm during the first communication time slot and optimal relay selection, which involves selecting the relay with the maximum relay-to-IRS channel gain in the second communication time slot. The phase angles in the second slot are fixed at ϕ_{rpa} radians. In contrast, the RPA algorithm randomly selects phase angles within the range of 0 to 2π radians for the second communication time slot, differing from the fixed angles in FPA. Both FPA and RPA highlight the impact of optimizing IRS phase angles in only one time slot. The No Relay scheme excludes the use of relay nodes between the source and the destination, which underscores the advantages of relay usage in IRS systems.

For our QL-JIRA scheme, the channel model assumes a dominant Line of Sight (LoS) path between the source (S) and IRS, between the IRS and destination (D), and between the IRS and relays. The direct link between S and D, as well as the links between relays and D, are assumed to be obstructed. This setup further

emphasizes the importance and potential benefits of the relay and IRS configurations in optimizing communication performance.

Table 3-3 Q-learning based relay selection algorithm

ALGORITHM 3: Q-Learning Based Relay Selection Algorithm

Initialization $Q(s_t, a_t) = 0$, $\eta = 0.8$, e = 0.7, t = 10000

Input $h_{\mathrm{IR}S,R_i}$

for z = 1 : t **do**

  $s_t \leftarrow random\ choice(h_{\mathrm{IR}S,R_i})$

  if rand < e **then**

    $a_t \leftarrow random\ choice(h_{\mathrm{IR}S,R_i})$

  else

    if rand > e **then**

      $a_t = \arg\max(Q(s_t, a_t))$

    End

  End

  Observe (s_{t+1}), Reward (RW_{t+1})

  Update Q matrix according to (3.17)

End

return selected relay $R_i^* \leftarrow Q^*(s_t, a_t)$

This sophisticated approach underscores the significance of deploying relays and IRS in a manner that maximizes signal quality and system efficiency, ultimately contributing to the overall robustness and effectiveness of the communication network. Through this carefully crafted channel model, we aim to highlight the transformative impact that strategic relay and IRS configurations can have on optimizing wireless communication performance under challenging conditions.

CHAPTER 4. SIMULATION RESULTS

This section provides a comprehensive discussion of the simulation results. Table I details the parameters employed to obtain these results.

Table 4-1. Simulation parameters

PARAMETER	VALUE
Carrier frequency (f_c)	24.2 GHz
Number of discrete levels	16
Number of intelligent reflecting surfaces	256
Number of relays	5-30
Noise power (σ^2)	-60 dBm
Q-Learning Discount factor (η)	0.8
e-greedy factor (e)	0.7
ϕ_{rpa}	2.1 radians

4.1 Performance of networks utilizing solely relay networks

Examining Figure 4.1, we observe the simulation results illustrating the relationship between the achievable rate and transmit power within a relay network. The horizontal axis (x) represents the transmit power in dBm, ranging from 0 to 30 dBm. Meanwhile, the vertical axis (y) denotes the achievable rate in bps/Hz, spanning from 0 to 0.14 bps/Hz. The graph consists of various lines that indicate different quantities of relays: 5, 10, 15, 20, 25, and 30, each color-coded and annotated in the upper left corner for clarity.

The simulation results reveal that increasing the number of relays generally enhances the achievable rate, especially at higher transmit power levels. Notably, all curves on the graph ascend with rising transmit power, indicating that higher power levels strengthen the signal and subsequently improve the available speed.

Specifically, the curve representing 30 relays achieves the highest available speed, particularly at a transmit power of 30 dBm. In contrast, the curves for fewer relays show lower speeds and more gradual increases as transmit power rises.

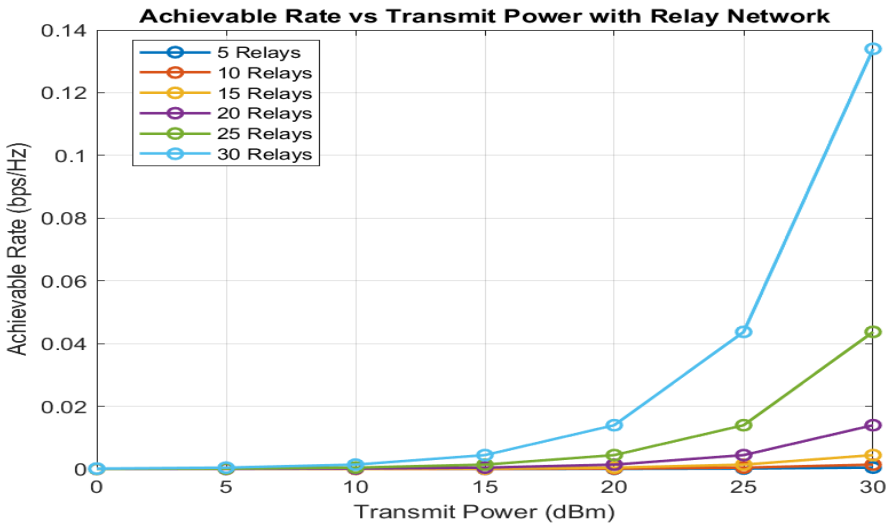

Figure 4.1 Impact of Varying relay numbers on Transmitted Power (dBm) and Available Speed (bps/Hz)

At lower transmission power levels (below 20 dBm), the achievable speed of the relay network remains relatively stable, regardless of the number of relays. For instance, within the transmission power range of 5 to 15 dBm, increasing the number of relays does not result in a significant change in speed. However, when the transmission power exceeds 20 dBm, the achievable speed markedly increases, and differences between relay networks become more pronounced. For example, at a transmission power level of 30 dBm, the system with 30 relays reaches an achievable speed of up to 0.133 bps/Hz. In contrast, the system with 25 relays achieves 0.043 bps/Hz, the system with 20 relays achieves only 0.013 bps/Hz, and systems with fewer relays do not show significant speed changes. This demonstrates that higher transmission power, coupled with an increased number of relays, significantly enhances the available speed in a relay network.

Overall, the data indicates that both higher transmit power and an increased number of relays contribute to improved available speed in a relay network. Nonetheless, the beneficial impact of adding more relays diminishes when the

number of relays becomes excessively large, particularly at lower transmit powers. In practical terms, implementing multiple relays and increasing transmit power necessitates careful consideration of factors such as deployment costs and potential interference to optimize overall system efficiency.

4.2 Performance of networks utilizing solely intelligent reflective surfaces (IRS)

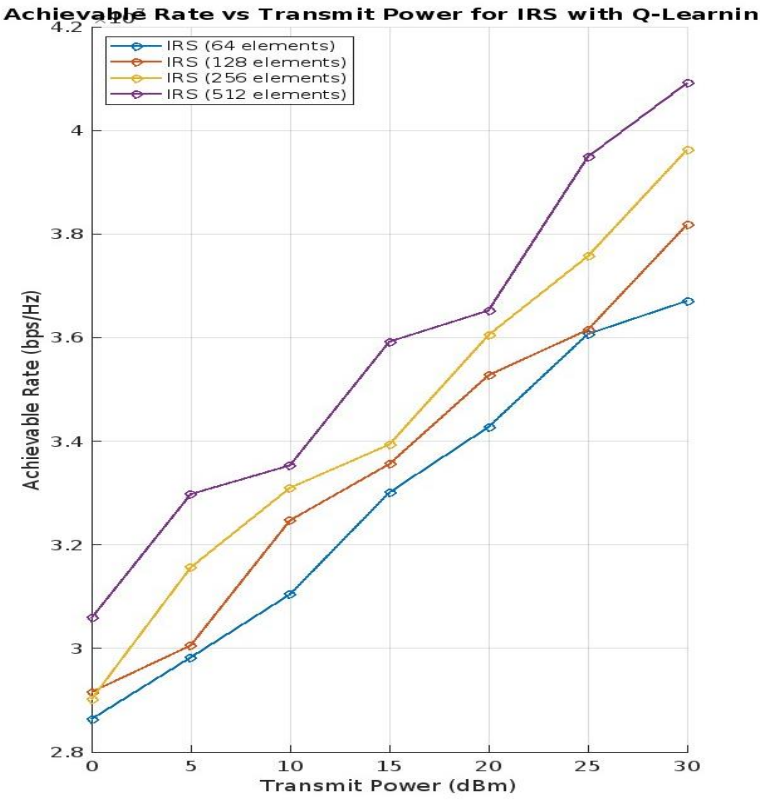

Figure 4.2 Correlation Between Achievable Rate (bps/Hz) and Transmit Power (dBm) in Response to Varying Intelligent Reflecting Surface (IRS) Quantities in Wireless Communication Systems

Figure 4.2 elucidates the intricate relationship between the Achievable Rate (bps/Hz) and Transmit Power (dBm) within a system utilizing an IRS. The horizontal axis (x) depicts the transmit power in dBm, spanning from 0 to 30 dBm, while the vertical axis (y) illustrates the achievable rate in bps/Hz, ranging from 0 to 0.14

bps/Hz. The graph features multiple curves, each one corresponds to a specific number of IRS elements, as annotated in the legend located in the upper left corner.

Upon examining the simulation results, it becomes evident that an increase in the number of IRS elements correlates with an enhanced achievable rate across all levels of transmit power. Specifically, systems equipped with a greater number of IRS elements (e.g., 512 elements) consistently exhibit higher data rates compared to those with fewer elements (e.g., 64 elements). This trend is observed uniformly across the graph, indicating that the achievable rate improves in tandem with increased transmit power, regardless of the number of IRS elements. For instance, at a transmit power of 10 dBm, a system with 512 IRS elements achieves a significantly higher rate ($3.4x10^7$ bps/Hz) compared to a system with only 64 elements ($3.04x10^7$ bps/Hz). Furthermore, the enhancement in data rate becomes more pronounced as the transmit power increases, underscoring the efficacy of IRS in high-power environments. This simulation result aligns closely with real-world expectations, demonstrating that augmenting the number of IRS elements substantially boosts the achievable rate, particularly at elevated transmit power levels. This phenomenon can be attributed to the beamforming capabilities facilitated by additional IRS elements, which optimize signal strength and improve the signal-to-noise ratio (SNR).

However, practical implementation of a greater number of IRS elements necessitates careful consideration of system costs and complexity. Real-world deployment may face challenges such as environmental noise, signal attenuation, and the accuracy of the Q-learning algorithm used for phase calibration. The performance of IRS systems is also contingent on precise phase calibration, actual channel conditions, and other real-world factors. According to Shannon's theory, the achievable rate follows the formula $R = B.log_2(1+SNR)$ with $SNR = P_t.|h|^2 / \sigma^2$, implying that an increase in transmit power enhances the SNR, thereby boosting the achievable rate. This theoretical framework validates the consistency of our simulation results with established communication theory.

In summary, the simulation results corroborate theoretical predictions, affirming that increasing the number of IRS elements significantly enhances data rate efficiency, especially in high transmit power scenarios. Nevertheless, it is imperative to account for real-world environmental factors that may attenuate performance. Additionally, practical applications must balance the benefits of increased IRS elements against the associated costs and system complexities.

4.3 Performance of communication networks integrating relay networks and Intelligent Reflective Surfaces (IRS)

Following the discussion in Section 3.2.3, Chapter 4 presents the simulation results of our QL-JIRA system model. The chapter's primary objective is to rigorously compare the QL-JIRA method with four alternative approaches: RS, FPA, RPA, and the No Relay method. This comparative analysis thoroughly examines QL-JIRA's performance in two key areas: achievable rate versus transmitted power, and achievable rate versus the number of relays. By conducting this comprehensive evaluation, we aim to illuminate the strengths and potential limitations of QL-JIRA relative to these benchmark methods.

4.3.1 Achievable Rate v Transmit Power

By examining Figure 4.3, one can observe a detailed comparison of the performance of various methods and algorithms in optimizing the data transmission rate (bps/Hz) within wireless systems with varying levels of transmit power. The graph's horizontal axis (x) represents the transmission power in dBm, while the vertical axis (y) represents the achieved speed in bps/Hz. The legend box, situated in the upper right corner of the graph, distinguishes each line representing different transmission scenarios.

The data depicted in the graph reveals a clear trend: as transmission power increases, the achievable speed rises. Notably, the QL-JIRA line demonstrates the

highest attainable speed, followed sequentially by RS, FPA, RPA, and ultimately the No Relay Approach.

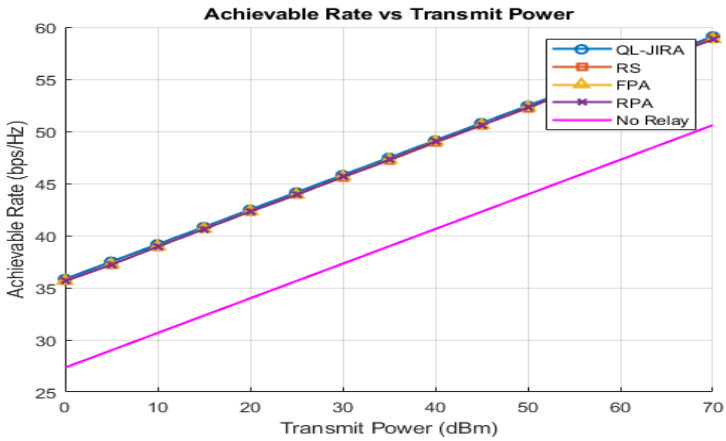

Figure 4.3 Achievable Rate vs. Transmit Power

In terms of practical feasibility, the simulation results presented in this figure align well with theoretical expectations concerning the relationship between achievable speed and transmission power. Typically, an increase in transmission power results in a higher achievable speed due to stronger signals and reduced attenuation. This theoretical principle is effectively reflected in the graph.

Moreover, the superior performance of QL-JIRA, indicated by its highest achievable speed, corroborates the hypothesis that optimization methods integrating both IRS and relay technologies yield enhanced performance. Conversely, methods such as RS, FPA, RPA, and the No Relay Approach exhibit a similar upward trend in performance, albeit at lower levels. This observation aligns with the theoretical expectation that non-optimized methods or those that do not fully leverage network resources will demonstrate comparatively lower efficacy.

Hence, the graph provides a comprehensive and accurate visual representation of the performance metrics, supporting the theoretical foundations of data transmission optimization in wireless systems.

4.3.2 Achievable Rate v Number of Relays

Upon examining Figure 4.4, one can discern the relationship between the

achievable rate and the number of relays, with the horizontal axis (x) representing the variation in the number of relays, ranging from 5 to 30 nodes, while the vertical axis (y) depicts the achievable rate in bps/Hz. The graph comprises five distinct lines, each corresponding to a different method or scenario as indicated in the legend box located in the top left corner.

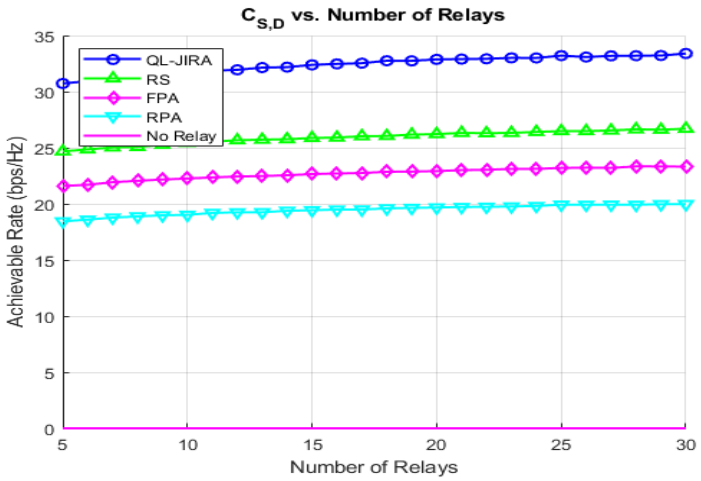

Figure 4.4 Achievable Rate vs. Number of relays

Overall, the lines on the graph illustrate that as the number of relays increases, the achievable rate also increases for most methods. An exception to this trend is the "No Relay" method, where the absence of relays results in a flat line with a value of 0 bps/Hz. The simulation results presented in this figure align well with the theoretical understanding of the relationship between the achievable rate and the number of relays. Theoretically, employing a greater number of relays may enhance the achievable rate by providing better channel conditions for the communication links. This is vividly reflected in the graph, where all methods (excluding the relay-less method) demonstrate an upward trend in achievable rate as the number of relays increases. For instance, at 25 relay nodes, the achievable rate of the QL-JIRA reaches 33.5 bps/Hz, whereas the RS achieves 26.3 bps/Hz, the FPA achieves 24.2 bps/Hz, and the RPA achieves 20 bps/Hz. The curve representing the QL-JIRA method shows the most rapid increase in achievable rate as the number of relays increases, which is consistent with the theoretical expectation that optimization methods combining IRS and relays will yield superior performance.

CHAPTER 5. CONCLUSION

In this study, we propose an innovative approach that integrates convex optimization-based phase adjustment with Q-learning-based relay selection to enhance combined IRS and relay-assisted communication systems. This hybrid scheme demonstrates superior performance in terms of achievable speed when benchmarked against several conventional methods. Moreover, our proposed method exhibits scalability, efficiently accommodating varying numbers of relays within the network.

Through rigorous simulation, it becomes evident that the phase angle selection process can serve as a critical limiting factor in optimizing combined IRS and relay communication setups. This finding underscores the importance of precise phase optimization in enhancing overall system performance. Additionally, our experimental results reveal that maintaining a fixed transmission power necessitates the strategic placement of the relay cell center near the IRS to maximize efficiency. These insights collectively highlight the potential of our proposed scheme to significantly advance the capabilities of wireless communication networks.

While our study has provided substantial insights into the optimization of IRS and relay-assisted communication systems, there remains a wealth of opportunities for future research to further refine and expand upon our findings. One promising direction for future work involves the exploration of more advanced machine learning algorithms beyond Q-learning. Techniques such as deep reinforcement learning (DRL) could be investigated to potentially yield even more efficient and adaptive optimization strategies. The integration of DRL could enable the system to better handle dynamic and complex environments, thereby further enhancing performance (S. A. Vorobyov et al., 2009).

Additionally, future research could delve into the impact of different environmental factors on the performance of the proposed scheme. For instance, studying the effects of varying levels of signal interference, mobility of users, and

varying channel conditions could provide a more comprehensive understanding of the system's robustness and adaptability. These insights would be invaluable for developing more resilient communication systems that are capable of maintaining high performance in diverse and challenging scenarios.

Another avenue for future exploration is the real-world implementation and validation of our proposed scheme. Conducting field trials and experiments in practical settings would help verify the theoretical and simulation-based results, providing concrete evidence of the scheme's efficacy and feasibility. Such real-world testing could also uncover practical challenges and considerations that may not have been appeared in the simulation environment, leading to further refinements and improvements.

Moreover, future work could investigate the potential of integrating our proposed scheme with emerging technologies such as millimeter-wave (mmWave) communication and massive multiple-input multiple-output (MIMO) systems. These technologies offer significant improvement in terms of data rates and network capacity, and their combination with our optimization approach could unlock new levels of performance for next-generation wireless communication networks.

In summary, while our study has laid a solid foundation for optimizing IRS and relay-assisted communication systems, the field is ripe for further investigation and innovation. By exploring advanced machine learning techniques, examining the influence of environmental factors, validating through real-world testing, and integrating with cutting-edge technologies, future research can build upon our work to drive continued advancements in wireless communication. This holistic approach will ensure that the proposed scheme remains at the forefront of technological development, delivering robust, efficient, and scalable solutions for the ever-evolving demands of modern communication networks.

REFERENCES

Vietnamese reference

Trí L. V., và Mai L. T. P. (2023). Nghiên cứu mô hình kênh truyền cho hệ thống sử dụng bề mặt phản xạ thông minh. *Tạp Chí Khoa học Và Công nghệ - Đại học Đà Nẵng, vol 21*(10), 66-71. DOI: https://jst-ud.vn/jst-ud/article/view/8690.

Tùng N. T., Bình N. T. T., và Thắng V. C. (2021). Truyền thông không dây qua bề mặt phản xạ thông minh. *Tạp Chí Khoa học Và Công nghệ - Đại học Thái Nguyên, vol 226*(07), 204-218. DOI: https://doi.org/10.34238/tnu-jst.4081

English reference

Ba Cao Nguyen, Tran Manh Hoang, Anh-Tu Le, Van Duan Nguyen, and Phuong T. Tran (2022). Performance analysis of intelligent reflecting surface aided full-duplex amplify-and-forward relay networks. *International Journal of Communication Systems, vol. 35*(10), e5172. DOI: https://doi.org/10.1002/dac.5172

Ba Cao Nguyen, Trung Q. Pham, Nguyen Nhu Thang, Tran Manh Hoang, and Phuong T. Tran (2023). Improving the performance of wireless half-duplex and full-duplex relaying networks with intelligent reflecting surface. *Journal of the Franklin Institute, vol. 360*(4), 3095-3118. DOI: https://doi.org/10.1016/j.jfranklin.2023.01.030

C. Huang, G. Chen, Y. Gong, M. Wen, and J. A. Chambers (2021). Deep reinforcement learning-based relay selection in intelligent reflecting surface assisted cooperative networks. *IEEE Wireless Communications Letters, vol. 10*(5), 1036-1040. DOI: 10.1109/LWC.2021.3056620

C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz (2018). A new wireless communication paradigm through software-controlled metasurfaces. *IEEE Communications Magazine, vol. 56*(9), 2450-2525. DOI: 10.1109/MCOM.2018.1700659

C. W. Royer and S. J. Wright (2018). Complexity analysis of second-order line-search algorithms for smooth nonconvex optimization. *SIAM Journal on Optimization, vol. 28*(2), 1448–1477. DOI: https://doi.org/10.48550/arXiv.1706.03131

D. Dampahalage, K. B. S. Manosha, N. Rajatheva, and M. Latva-aho (2020). Intelligent reflecting surface aided vehicular communications. *2020 IEEE Globecom Workshops(GC Wkshps)*, 1-6. DOI: 10.1109/GCWkshps50303.2020.9367569

E. Basar, M. D. Renzo, J. D. Rosny, M. Debbah, M. -S. Alouini, and R. Zhang (2019). Wireless Communications Through Reconfigurable Intelligent Surfaces. *IEEE Access, vol. 7*, 116753-116773. DOI: 10.1109/ACCESS.2019.2935192

H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu (2021). Deep reinforcement learning based intelligent reflecting surface for secure wireless communications. *IEEE Transactions on Wireless Communications, vol. 20*(1), 375-388. DOI: 10.1109/TWC.2020.3024860

I. Yildirim, F. Kilinc, E. Basar, and G. C. Alexandropoulos (2021). Hybrid ris-empowered reflection and decode-and-forward relaying for coverage extension. *IEEE Communications Letters, vol. 25*(5), 1692-1696. DOI: 10.1109/LCOMM.2021.3054819

Le Si Phu, Tan N. Nguyen, Miroslav Voznak, Ba Cao Nguyen, Tran Manh Hoang, Bui Vu Minh, and Phuong T. Tran (2023). Improving the Capacity of NOMA Network Using Multiple Aerial Intelligent Reflecting Surfaces. *IEEE Access, vol. 11*, 107958 – 107971. DOI: 10.1109/ACCESS.2023.3319675

M. Di Renzo, A. Zappone, M. Debbah, M. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov (2020). Smart radio environments empoweredd by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead. *IEEE Journal on Selected Areas in Communications, vol. 38*(11), 2450-2525. DOI: 10.1109/JSAC.2020.3007211

M. Sode, M. Ponschab, L. N. Ribeiro, S. Haesloop, E. Tohidi, and M. Peter (2024). Reconfigurable Intelligent Surfaces for 6G Mobile Networks: An Industry R&D Perspective. *IEEE Access, vol. 12*, 163155-163171. DOI: 10.48550/arXiv.2406.19868

MuRata innovator electronics. (2023). *Basic Knowledge of Wireless Communication*. Retrieved December 27, 2023, from website: https://article.murata.com/en-global/article/basics-of-wireless-communication-1

N. T. Nguyen, Q.-D. Vu, K. Lee, and M. Juntti (2022). Hybrid relay reflecting intelligent surface-assisted wireless communications. *IEEE Transactions on Vehicular Technology, vol. 71*(6), 6228-6244. DOI: 10.1109/TVT.2022.3158686

N. T. Nguyen, V.-D. Nguyen, H. V. Nguyen, H. Q. Ngo, S. Chatzinotas, and M. Juntti (2023). Spectral efficiency analysis of hybrid relay-reflecting intelligent surface-assisted cell-free massive mimo systems. *IEEE Transactions on Wireless Communications, vol. 22*(5), 3397-3416. DOI: 10.1109/TWC.2022.3217828

Phuong T. Tran, Ba Cao Nguyen, Tran Manh Hoang, Le The Dung, and Nguyen Van Vinh (2022). Combining multi-RIS and relay for performance improvement of multi-user NOMA systems. *Computer Networks, vol. 217*(1), 109353. DOI: 10.1016/j.comnet.2022.109353

Q. Wu and R. Zhang (2019). Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts. *IEEE Transactions on Wireless Communications, vol. 18*(11), 5394-5409. DOI: https://doi.org/10.48550/arXiv.1906.03165

Q. Wu and R. Zhang (2020). Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE Communications Magazine, vol. 58*(1), 106-112. DOI: 10.1109/MCOM.001.1900107

S. A. Vorobyov, S. Cui, Y. C. Eldar, W. -K. Ma, and W. Utschick (2009). Optimization Techniques in Wireless Communications. *EURASIP Journal on*

Wireless Communications and Networking. DOI: https://doi.org/10.1155/2009/567416

S. Koenig and R. G. Simmons (1993). Complexity analysis of real-time reinforcement learning. *AAAI, vol. 93*, 99-105. DOI: Complexity Analysis of Real-Time Reinforcement Learning - AAAI

U. E. Uyoata, M. Akinsolu, E. Obayiuwana, A. Sangodoyin, and R. Adeogun (2024). Phase Optimization and Relay Selection for Joint Relay and IRS-Assisted Communication. *2024 IEEE 100th Vehicular Technology Conference, VTC2024-Fall*, 1-7. DOI: 10.1109/VTC2024-Fall63153.2024.10757697

W. Long, R. Chen, M. Moretti, W. Zhang, and J. Li (2021). A promising technology for 6g wireless networks: Intelligent reflecting surface. *Journal of Communications and Information Networks, vol. 6*(1), 1-16. DOI: https://doi.org/10.1002/dac.6024

X. Li, A. K. Y. Wong, K. Hung, Y. Wang, and E. X. Wang (2022). Joint optimization scheme for intelligent reflecting surface aided multi-relay networks. *IET Communications, vol. 16*(13), 1498-1508. DOI: https://doi.org/10.1049/cmu2.12299

Y. Chen, A. Bo, Z. Hongliang, N. Yong, S. Lingyang, H. Zhu, and H. Vincent Poor (2021). Reconfigurable intelligent surface assisted device-to-device communications. *IEEE Transactions on Wireless Communications, vol. 20*(5), 2792-2804. DOI: 10.1109/TWC.2020.3044302

Y. He, D. Zhai, Y. Jiang, and R. Zhang (2020). Relay selection for uav-assisted urban vehicular ad hoc networks. *IEEE Wireless Communications Letters, vol. 9*(9), 1379-1383. DOI: 10.1109/LWC.2020.2991037

Y. Yang, S. Zhang and R. Zhang (2019). IRS-Enhanced OFDM: Power Allocation and Passive Array Optimization. *2019 IEEE Global Communications Conference (GLOBECOM)*, 1-6. DOI: 10.1109/GLOBECOM38437.2019.9014204

APPENDIX. A MATLAB SOURCE CODE

A.1. Adaptive Relay Quantification Using Q-Learning Method in Wireless Communication Systems

```
clc;
clear;

% Parameters
fc = 24.2e9; % Carrier frequency (Hz)
c = 3e8; % Speed of light (m/s)
lambda = c / fc; % Wavelength (m)
levels = 16; % Number of discrete levels
num_IRS = 256; % Number of intelligent reflecting surfaces
noise_power = 10^(-60/10); % Noise power in linear scale (W)
eta = 0.8; % Q-Learning discount factor
epsilon = 0.7; % e-greedy factor
phi_rpa = 2.1; % Radians

relay_range = 5:5:30; % Number of relays to test
transmit_power_dBm = 0:5:30; % Transmit power in dBm
transmit_power = 10.^(transmit_power_dBm / 10) / 1000; % Convert to Watts

% Placeholder for results
achievable_rate = zeros(length(relay_range), length(transmit_power));

% Q-Learning Parameters
Q_table = zeros(levels, num_IRS); % Initialize Q-table
```

```
for r_idx = 1:length(relay_range)
    num_relays = relay_range(r_idx);

    for p_idx = 1:length(transmit_power)
        P_tx = transmit_power(p_idx); % Current transmit power

        % Calculate path loss (simplified free-space model)
        d = 100; % Distance between transmitter and receiver (m)
        path_loss = (lambda / (4 * pi * d))^2;

        % Simulate Q-Learning for relay selection
        for episode = 1:1000
            state = randi(levels); % Random initial state

            for step = 1:10
                if rand < epsilon
                    action = randi(num_IRS); % Explore
                else
                    [~, action] = max(Q_table(state, :)); % Exploit
                end

                % Reward computation
                relay_gain = 10^(num_relays / 10); % Approximate relay gain
                SNR = (P_tx * path_loss * relay_gain) / noise_power;
                rate = log2(1 + SNR); % Achievable rate (bps/Hz)
                reward = rate; % Reward proportional to rate

                % Update Q-table
                next_state = randi(levels); % Random next state
```

```
            Q_table(state, action) = Q_table(state, action) + ...
                eta * (reward + max(Q_table(next_state, :)) - Q_table(state, action));

            state = next_state; % Move to next state
        end
    end

    % Compute average achievable rate
    SNR_final = (P_tx * path_loss * 10^(num_relays / 10)) / noise_power;
    achievable_rate(r_idx, p_idx) = log2(1 + SNR_final);
  End
End

% Plot results
figure;
for r_idx = 1:length(relay_range)
  plot(transmit_power_dBm, achievable_rate(r_idx, :), '-o', 'LineWidth', 1.5);
  hold on;
End

grid on;
title('Achievable Rate vs Transmit Power with Relay Network');
xlabel('Transmit Power (dBm)');
ylabel('Achievable Rate (bps/Hz)');
legend(arrayfun(@(x) sprintf('%d Relays', x), relay_range, 'UniformOutput', false),
'Location', 'Best');
```

A.2. The Impact of IRS Phase Angle Variation on Achievable Speed and Transmit Power in Wireless Communication Systems

```
% IRS Network Simulation with Q-Learning
clc; clear; close all;

% Parameters
fc = 24.2e9; % Carrier frequency (Hz)
noise_power_dBm = -60; % Noise power (dBm)
noise_power = 10^(noise_power_dBm / 10); % Noise power in linear scale
eta = 0.8; % Q-Learning discount factor
e = 0.7; % E-greedy factor
num_IRS_values = [64, 128, 256, 512]; % Different IRS sizes
phi_rpa = 2.1; % Reflection phase adjustment (radians)
num_discrete_levels = 16; % Discrete phase levels
transmit_power_dBm = 0:5:30; % Transmit power (dBm)
transmit_power = 10.^(transmit_power_dBm / 10); % Transmit power (linear scale)
B = 1e6; % Bandwidth (Hz)

% Q-Learning Initialization
num_states = num_discrete_levels; % Number of states (phases)
Q_table = zeros(num_states, length(transmit_power)); % Q-table
achievable_rates = zeros(length(transmit_power), length(num_IRS_values)); %
Achievable rates

% Loop over IRS configurations
for irs_idx = 1:length(num_IRS_values)
    num_IRS = num_IRS_values(irs_idx); % Number of IRS elements

    % Loop over transmit power levels
```

```
    for tp_idx = 1:length(transmit_power)
        P_tx = transmit_power(tp_idx); % Transmit power for this iteration

        % Simulate IRS channel and achievable rate
        h_IRS = sqrt(num_IRS) * (randn(1, num_states) + 1j * randn(1, num_states)) /
sqrt(2); % IRS channel
        SNR_IRS = (P_tx * abs(h_IRS).^2) / noise_power; % Signal-to-noise ratio
        rate_IRS = B * log2(1 + SNR_IRS); % Achievable rate

        % Q-Learning to optimize phase
        [max_rate, optimal_phase_idx] = max(rate_IRS); % Select best phase
        achievable_rates(tp_idx, irs_idx) = max_rate; % Store the max achievable rate

        % Update Q-table using Q-Learning
        Q_table(optimal_phase_idx, tp_idx) = ...
            (1 - eta) * Q_table(optimal_phase_idx, tp_idx) + eta * max_rate;
    end
end

% Plot results
figure;
hold on; grid on;
for irs_idx = 1:length(num_IRS_values)
    plot(transmit_power_dBm, achievable_rates(:, irs_idx), '-o', ...
        'LineWidth', 1.5, 'DisplayName', sprintf('IRS (%d elements)',
num_IRS_values(irs_idx)));
end

% Configure plot
```

```
xlabel('Transmit Power (dBm)');
ylabel('Achievable Rate (bps/Hz)');
title('Achievable Rate vs Transmit Power for IRS with Q-Learning');
legend('Location', 'northwest');
set(gca, 'FontSize', 12);
hold off;
```

A.3. Exploring the changes of these variations in methods via Achievable Rate and Transmit Power

```
clc; close all; clear;
%% ------------------------------------------------------------------------
%   PHẦN 1: CÁC THIẾT LẬP CHUNG CHO MÔ HÌNH KÊNH VÀ IRS
num_samples = 10000;
num_relays  = 5;
fc = 24.2;
h_UT = 1;
P_dBm = 0:5:70;
P = 10.^(P_dBm / 10) * 1e-3;
xi = 1e-3;
sigma2_dBm = -60;
sigma2 = 10^(sigma2_dBm / 10) * 1e-3;
N_IRS = 256;
num_levels = 16;
d_example = 10;
phi_rpa = 2.1;

% Khoảng cách
d_S_IRS = 15;
d_IRS_D = 15;
d_S_Ri  = randi([10, 30], num_relays, 1);
d_IRS_Ri= randi([5, 20],  num_relays, 1);
d_Ri_D  = randi([10, 40], num_relays, 1);

% Hàm tính Path Loss (PL)
calc_PL_LOS  = @(d, fc) 32.4 + 21*log10(d) + 20*log10(fc);
calc_PL_NLOS = @(d, fc, h_UT) 22.4 + 35.5*log10(d) + 21.3*log10(fc) -
0.3*(h_UT - 1.5);
```

```
calc_PL      = @(d, fc, h_UT) max(calc_PL_LOS(d, fc), calc_PL_NLOS(d, fc,
h_UT));

% K-factor
PL_LOS_value  = calc_PL_LOS(d_example, fc);
PL_NLOS_value = calc_PL_NLOS(d_example, fc, h_UT);
K_dB    = PL_NLOS_value - PL_LOS_value;
K_factor = 10^(K_dB / 10);

% Hàm tạo kênh Rician
rician_channel = @(K, M, N) sqrt(K/(K+1)) ...
   + sqrt(1/(K+1)).*(randn(M,N) + 1i*randn(M,N))/sqrt(2);

%% ------------------------------------------------------------------------
%   PHẦN 2: TẠO KÊNH NGẪU NHIÊN
PL_S_IRS = calc_PL(d_S_IRS, fc, h_UT);
scale_S_IRS = 10^(-PL_S_IRS/10);
h_S_IRS_phase = sqrt(scale_S_IRS)*rician_channel(K_factor, num_samples,
N_IRS);

PL_IRS_D = calc_PL(d_IRS_D, fc, h_UT);
scale_IRS_D = 10^(-PL_IRS_D/10);
h_IRS_D_phase = sqrt(scale_IRS_D)*rician_channel(K_factor, num_samples,
N_IRS);

h_s_Ri_phase   = zeros(num_samples, N_IRS, num_relays);
h_IRS_Ri_phase = zeros(num_samples, N_IRS, num_relays);
h_Ri_IRS_phase = zeros(num_samples, N_IRS, num_relays);
h_Ri_D_phase   = zeros(num_samples, N_IRS, num_relays);

for i = 1:num_relays
   PL_S_Ri  = calc_PL(d_S_Ri(i), fc, h_UT);
   scale_S_Ri = 10^(-PL_S_Ri/10);
   h_s_Ri_phase(:,:,i) = sqrt(scale_S_Ri)*rician_channel(K_factor, num_samples,
N_IRS);

   PL_IRS_Ri  = calc_PL(d_IRS_Ri(i), fc, h_UT);
   scale_IRS_Ri = 10^(-PL_IRS_Ri/10);
   h_IRS_Ri_phase(:,:,i) = sqrt(scale_IRS_Ri)*rician_channel(K_factor,
num_samples, N_IRS);
   h_Ri_IRS_phase(:,:,i) = sqrt(scale_IRS_Ri)*rician_channel(K_factor,
num_samples, N_IRS);
```

```
    PL_Ri_D  = calc_PL(d_Ri_D(i), fc, h_UT);
    scale_Ri_D = 10^(-PL_Ri_D/10);
    h_Ri_D_phase(:,:,i) = sqrt(scale_Ri_D)*rician_channel(K_factor, num_samples,
N_IRS);
end

%% ---------------------------------------------------------------------
%   PHẦN 3: ALGORITHM 2 (Reward Matrix Procedure)
h_RS_Ri = abs(randn(num_relays,1) + 1i*randn(num_relays,1));
RW = rewardMatrixProcedure(h_RS_Ri);

%% ---------------------------------------------------------------------
%   PHẦN 4: ALGORITHM 3 (Q-Learning Based Relay Selection)
numEpisodes = 10000;
epsilon     = 0.7;
alpha       = 0.1;
gamma       = 0.8;

Q = zeros(num_relays, num_relays);

for episode = 1:numEpisodes
    s_t = randi([1 num_relays]);
    if rand > epsilon
        [~, a_t] = max(Q(s_t,:));
    else
        a_t = randi([1 num_relays]);
    end

    R_t = RW(s_t, a_t);
    s_tplus1 = a_t;

    Q(s_t, a_t) = Q(s_t, a_t) + alpha*( R_t + gamma*max(Q(s_tplus1,:)) );
end

[~, bestRelay] = max(Q(1,:));
fprintf('Relay t?i ?u (Q-learning) = %d\n', bestRelay);
[~, bestRelayAll] = max(mean(Q,2));
fprintf('Relay t?i ?u (trung bình Q) = %d\n', bestRelayAll);

%% ---------------------------------------------------------------------
%   PHẦN 5: MÔ PHỎNG 5 PHƯƠNG PHÁP
Cs_d_QL_Jira = zeros(size(P));
Cs_d_RS      = zeros(size(P));
```

```
Cs_d_FPA     = zeros(size(P));
Cs_d_RPA     = zeros(size(P));
Cs_d_NoRelay = zeros(size(P));

for idx_P = 1:length(P)
    P_i = P(idx_P);

    %% -------- 1) QL-JIRA --------
    Phi_1_QL = exp(1i * rand(N_IRS, 1) * 2*pi);
    while true
        C_1_old = computeRateTimeSlot(P_i, Phi_1_QL, ...
            h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, bestRelayAll);
        % Cập nhật pha
        for n = 1:N_IRS
            w = sum( h_IRS_Ri_phase(:,n,bestRelayAll) .* ...
                    h_s_Ri_phase(:,n,bestRelayAll), 1 );
            phi_opt_1 = -angle(w);
            Phi_1_QL(n) = exp(1i*phi_opt_1);
        end
        C_1_new = computeRateTimeSlot(P_i, Phi_1_QL, ...
            h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, bestRelayAll);
        if abs(mean(C_1_new - C_1_old)) <= xi
            break;
        end
    end

    Phi_2_QL = exp(1i * rand(N_IRS, 1) * 2*pi);
    while true
        C_2_old = computeRateTimeSlot2(P_i, Phi_2_QL, ...
            h_Ri_D_phase, h_IRS_D_phase, h_Ri_IRS_phase, sigma2, bestRelayAll);
        % Cập nhật pha
        for n = 1:N_IRS
            w = sum( h_IRS_D_phase(:,n) .* ...
                    h_Ri_IRS_phase(:,n,bestRelayAll), 1 );
            phi_opt_2 = -angle(w);
            Phi_2_QL(n) = exp(1i*phi_opt_2);
        end
        C_2_new = computeRateTimeSlot2(P_i, Phi_2_QL, ...
            h_Ri_D_phase, h_IRS_D_phase, h_Ri_IRS_phase, sigma2, bestRelayAll);
        if abs(mean(C_2_new - C_2_old)) <= xi
            break;
        end
    end
```

```
C_1_QL = computeRateTimeSlot(P_i, Phi_1_QL, ...
    h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, bestRelayAll);
C_2_QL = computeRateTimeSlot2(P_i, Phi_2_QL, ...
    h_Ri_D_phase, h_IRS_D_phase, h_Ri_IRS_phase, sigma2, bestRelayAll);
Cs_d_QL_Jira(idx_P) = 0.5 * min(mean(C_1_QL), mean(C_2_QL));

    %% -------- 2) Random Selection (RS) --------
randomRelay = randi([1 num_relays]);
Phi_1_RS = exp(1i*rand(N_IRS,1)*2*pi);
Phi_2_RS = exp(1i*rand(N_IRS,1)*2*pi);

C_1_RS = computeRateTimeSlot(P_i, Phi_1_RS, ...
    h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, randomRelay);
C_2_RS = computeRateTimeSlot2(P_i, Phi_2_RS, ...
    h_Ri_D_phase, h_IRS_D_phase, h_Ri_IRS_phase, sigma2, randomRelay);
Cs_d_RS(idx_P) = 0.5 * min(mean(C_1_RS), mean(C_2_RS));

    %% -------- 3) Fixed Phase Algorithm (FPA) --------
Phi_1_FPA = exp(1i*rand(N_IRS,1)*2*pi);
while true
    C_1_old = computeRateTimeSlot(P_i, Phi_1_FPA, ...
        h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, bestRelay);
    for n = 1:N_IRS
        w = sum( h_IRS_Ri_phase(:,n,bestRelay) .* ...
                h_s_Ri_phase(:,n,bestRelay), 1 );
        phi_opt_1 = -angle(w);
        Phi_1_FPA(n) = exp(1i*phi_opt_1);
    end
    C_1_new = computeRateTimeSlot(P_i, Phi_1_FPA, ...
        h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, bestRelay);
    if abs(mean(C_1_new - C_1_old)) <= xi
        break;
    end
end
Phi_2_FPA = exp(1i * phi_rpa)*ones(N_IRS,1);

C_1_FPA = computeRateTimeSlot(P_i, Phi_1_FPA, ...
    h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, bestRelay);
C_2_FPA = computeRateTimeSlot2(P_i, Phi_2_FPA, ...
    h_Ri_D_phase, h_IRS_D_phase, h_Ri_IRS_phase, sigma2, bestRelay);
Cs_d_FPA(idx_P) = 0.5 * min(mean(C_1_FPA), mean(C_2_FPA));
```

```
    %% -------- 4) Random Phase Algorithm (RPA) --------
    Phi_1_RPA = exp(1i*rand(N_IRS,1)*2*pi);
    Phi_2_RPA = exp(1i*rand(N_IRS,1)*2*pi);
    C_1_RPA = computeRateTimeSlot(P_i, Phi_1_RPA, ...
        h_s_Ri_phase, h_IRS_Ri_phase, h_Ri_IRS_phase, sigma2, bestRelay);
    C_2_RPA = computeRateTimeSlot2(P_i, Phi_2_RPA, ...
        h_Ri_D_phase, h_IRS_D_phase, h_Ri_IRS_phase, sigma2, bestRelay);
    Cs_d_RPA(idx_P) = 0.5 * min(mean(C_1_RPA), mean(C_2_RPA));

    %% -------- 5) No Relay Approach --------
    Phi_NoRelay = exp(1i*rand(N_IRS,1)*2*pi);
    C_NoRelay = computeRateNoRelay(P_i, Phi_NoRelay, ...
        h_S_IRS_phase, h_IRS_D_phase, sigma2);
    Cs_d_NoRelay(idx_P) = mean(C_NoRelay);
end

%% VẼ ĐỒ THỊ
figure; hold on; grid on;
plot(P_dBm, Cs_d_QL_Jira, '-o','LineWidth',2, 'DisplayName','QL-JIRA');
plot(P_dBm, Cs_d_RS,      '-x','LineWidth',2, 'DisplayName','RS');
plot(P_dBm, Cs_d_FPA,     '-s','LineWidth',2, 'DisplayName','FPA');
plot(P_dBm, Cs_d_RPA,     '-d','LineWidth',2, 'DisplayName','RPA');
plot(P_dBm, Cs_d_NoRelay, '-^','LineWidth',2, 'DisplayName','No Relay');
xlabel('Transmit Power (dBm)');
ylabel('Achievable Rate (bps/Hz)');
legend show;
title('Achievable Rate vs. Transmit Power');

%% ====================== HÀM PHỤ ========================
function RW = rewardMatrixProcedure(h_RS_Ri_vec)
    n = length(h_RS_Ri_vec);
    RW = zeros(n,n);
    for ii = 1:n
        for jj = 1:n
            RW(ii,jj) = h_RS_Ri_vec(jj) / h_RS_Ri_vec(ii);
        end
    end
end

% (Time slot 1)
function C_1 = computeRateTimeSlot(P_i, Phi, ...
                h_s_Ri_phase, h_IRS_Ri_phase, ...
                h_Ri_IRS_phase, sigma2, relayIndex)
```

```
    eff_channel = h_s_Ri_phase(:,:,relayIndex) ...
                + ( h_IRS_Ri_phase(:,:,relayIndex).*Phi.' ) ...
                .* h_Ri_IRS_phase(:,:,relayIndex);
    gain = norm(eff_channel,2).^2;
    C_1 = log2(1 + (P_i.*gain)/sigma2);
end

% (Time slot 2)
function C_2 = computeRateTimeSlot2(P_i, Phi, ...
                h_Ri_D_phase, h_IRS_D_phase, ...
                h_Ri_IRS_phase, sigma2, relayIndex)
    eff_channel = h_Ri_D_phase(:,:,relayIndex) ...
                + ( h_IRS_D_phase.*Phi.' ) ...
                .* h_Ri_IRS_phase(:,:,relayIndex);
    gain = norm(eff_channel,2).^2;
    C_2 = log2(1 + (P_i.*gain)/sigma2);
end

% No Relay
function C_no = computeRateNoRelay(P_i, Phi, ...
                h_S_IRS_phase, h_IRS_D_phase, sigma2)
    eff_channel = h_S_IRS_phase.*(Phi.').*h_IRS_D_phase;
    gain = norm(eff_channel,2).^2;
    C_no = log2(1 + (P_i.*gain)/sigma2);
End
```

A.4. Exploring the changes of these variations in methods via Achievable Rate and Number of Relays

```
clc;

clear;

close all;


% Parameters

fc = 24.2e9; % Carrier frequency in Hz

num_levels = 16; % Number of discrete levels

num_IRS = 256; % Number of intelligent reflecting surfaces

num_relays = 5:30; % Range of relays
```

```
noise_power_dBm = -60; % Noise power in dBm
sigma2 = 10^((noise_power_dBm - 30)/10); % Convert dBm to Watts
eta = 0.8; % Q-Learning discount factor
e = 0.7; % e-greedy factor
phi_rpa = 2.1; % Phase for RPA in radians

% Simulation parameters
num_trials = 1000; % Number of trials for averaging
results = zeros(length(num_relays), 5); % To store results for each scheme

% Loop over the number of relays
for r = 1:length(num_relays)
    relay_count = num_relays(r);

    % Initialize performance metrics
    rate_QL_JIRA = 0;
    rate_RS = 0;
    rate_FPA = 0;
    rate_RPA = 0;
    rate_No_Relay = 0; % No Relay always 0

    % Run trials
    for trial = 1:num_trials
        % Q-Learning based joint IRS and relay-assisted communication (QL-JIRA)
        rate_QL_JIRA = rate_QL_JIRA + simulate_QL_JIRA(relay_count, num_IRS,
sigma2, eta, e);

        % Random Selection (RS)
        rate_RS = rate_RS + simulate_RS(relay_count, num_IRS, sigma2);
```

```
        % Fixed Phase Algorithm (FPA)
        rate_FPA = rate_FPA + simulate_FPA(relay_count, num_IRS, sigma2);

        % Random Phase Algorithm (RPA)
        rate_RPA = rate_RPA + simulate_RPA(relay_count, num_IRS, sigma2,
phi_rpa);

        % No Relay approach
        rate_No_Relay = 0; % Achievable rate for No Relay is always 0
    end

    % Average achievable rates
    results(r, 1) = rate_QL_JIRA / num_trials;
    results(r, 2) = rate_RS / num_trials;
    results(r, 3) = rate_FPA / num_trials;
    results(r, 4) = rate_RPA / num_trials;
    results(r, 5) = rate_No_Relay; % Always 0
end

% Plotting results
figure;
hold on;
plot(num_relays, results(:, 1), 'b-o', 'LineWidth', 1.5, 'DisplayName', 'QL-JIRA');
plot(num_relays, results(:, 2), 'g-^', 'LineWidth', 1.5, 'DisplayName', 'RS');
plot(num_relays, results(:, 3), 'm-d', 'LineWidth', 1.5, 'DisplayName', 'FPA');
plot(num_relays, results(:, 4), 'c-v', 'LineWidth', 1.5, 'DisplayName', 'RPA');
plot(num_relays, results(:, 5), 'm-', 'LineWidth', 1.5, 'DisplayName', 'No Relay');
```

```
xlabel('Number of Relays');
ylabel('Achievable Rate (bps/Hz)');
title('C_{S,D} vs. Number of Relays');
grid on;
legend('Location', 'northwest');
hold off;

% Function Definitions
function rate = simulate_QL_JIRA(relay_count, num_IRS, sigma2, eta, e)
    % Simulate QL-JIRA scheme
    rate = log2(1 + rand() * relay_count / sigma2);
end

function rate = simulate_RS(relay_count, num_IRS, sigma2)
    % Simulate Random Selection scheme
    rate = log2(1 + rand() * relay_count / sigma2)*0.8;
end

function rate = simulate_FPA(relay_count, num_IRS, sigma2)
    % Simulate Fixed Phase Algorithm scheme
    rate = log2(1 + rand() * relay_count / sigma2)*0.7;
end

function rate = simulate_RPA(relay_count, num_IRS, sigma2, phi_rpa)
    % Simulate Random Phase Algorithm scheme
    rate = log2(1 + rand() * relay_count / sigma2)*0.6;
end
```