Ph.D. DISSERTATION

A Study on Reconfigurable Intelligent Surfaces-Assisted Communications

지능형 반사평면 기반 무선 통신 기법 연구

BY

Jiao Wu

AUGUST 2023

DEPARTMENT OF ELECTRICAL AND
COMPUTER ENGINEERING
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Ph.D. DISSERTATION

A Study on Reconfigurable Intelligent Surfaces-Assisted Communications

지능형 반사평면 기반 무선 통신 기법 연구

BY

Jiao Wu

AUGUST 2023

DEPARTMENT OF ELECTRICAL AND
COMPUTER ENGINEERING
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

A Study on Reconfigurable Intelligent Surfaces-Assisted Communications

지능형 반사평면 기반 무선 통신 기법 연구

지도교수 심 병 효

이 논문을 공학박사 학위논문으로 제출함

2023년 8월

서울대학교 대학원

전기 정보 공학부

우 지 아 오

우지아오의 공학박사 학위 논문을 인준함

2023년 8월

위 원 장:	김 성 철
부위원장:	심 병 효
위 원:	이 경 한
위 원:	최 준 원
위 원:	이 병 주

Abstract

With the global deployment of the fifth-generation (5G) systems, research on the sixth generation (6G) systems for 2030 and beyond has been fully launched. 6G aims to integrate non-communication technologies like sensing, computing, and artificial intelligence (AI), while also building upon the capabilities and scenarios established by 5G applications. The driving forces behind the development of 6G are the increasing mobile traffic demands and the emergence of innovative applications such as holographic telepresence, extended reality (XR), digital twin, and autonomous systems. These applications impose stringent requirements on the key performance indicators (KPIs) of 6G, demanding approximately $10 \sim 100$ times higher rates, reliability, lower latency, improved mobility, and energy efficiency compared to 5G. As a result, 6G is expected to achieve a peak data rate of 1 terabit-per-second (Tbps) and significantly reduced latency, reaching sub-millisecond levels. To meet these demanding requirements, novel technologies are required, as the existing mechanisms and conventional approaches cannot suffice.

In the first part of this dissertation, we focus on the channel estimation framework for near-field reconfigurable intelligent surface (RIS)-assisted terahertz (THz) systems. RIS-assisted THz communications have garnered considerable attention as they offer the potential to support extremely high data rates in 6G wireless networks. By adjusting the wireless propagation environment of THz systems through phase shifts of reflecting elements, RIS can dramatically enhance overall throughput. However, accurately acquiring channel information is crucial to realizing the full potential of RIS-assisted THz systems. Conventional channel estimation techniques based on planar wavefront assumptions fail to deliver satisfactory performance in near-field RIS-assisted THz systems due to the spherical wavefront of the THz electromagnetic signal. In light of this challenge, we propose an efficient channel estimation technique for near-field

RIS-assisted wideband THz systems called the Polar-Domain Frequency-Dependent RIS-Assisted Channel Estimation (PF-RCE) scheme. Key idea of the proposed PF-RCE scheme is to estimate the sparse multipath components (i.e., angles, distances, and path gains) of the near-field THz channel by exploiting the polar-domain sparsity and common support properties.

In the second part of the dissertation, we investigate an energy-efficient power control and beamforming scheme for RIS-assisted Internet of Things (IoT) networks. RIS, composed of numerous low-cost reflecting elements arranged in a planar metasurface, has garnered significant attention for its ability to enhance both spectrum and energy efficiencies by reconfiguring the wireless propagation environment. In this work, we propose an optimization technique for RIS phase shifts and base station (BS) beamforming that minimizes the uplink transmit power of the RIS-aided IoT network. Key idea of the proposed scheme, referred to as Riemannian conjugate gradient-based joint optimization (RCG-JO), is to jointly optimize the RIS phase shifts and the BS beamforming vectors using the Riemannian conjugate gradient technique. By leveraging the product Riemannian manifold structure of the sets of unit-modulus phase shifts and unit-norm beamforming vectors, we convert the nonconvex uplink power minimization problem into an unconstrained problem and then find out the optimal solution over the product Riemannian manifold.

In the third part of the dissertation, we address a crucial challenge faced in wideband THz communication: the considerable array gain loss caused by the beam split effect. This phenomenon occurs when path components split into different spatial directions at various subcarrier frequencies, posing a significant hurdle for the practical application of conventional phase shift control and beamforming techniques in wideband THz systems. To overcome this issue, we propose a RIS-assisted wideband beamforming (RWB) technique, specifically designed to maximize the achievable data rate of RIS-assisted wideband THz systems. Key idea of RWB is to optimize the analog beamforming vector and the RIS phase shift vector alternately, while carefully designing the parameters of the beamforming network. This novel approach empowers us to fully unleash the potential of the wideband THz system, thereby maximizing its achievable data rate. By effectively managing the beam split effect, RWB paves the way for enhanced performance in wideband THz communication, playing a pivotal role in realizing the ambitious goals set for 6G wireless networks.

keywords: 6G wireless communications, reconfigurable intelligent surface, terahertz communications, channel estimation, wideband beamforming

student number: 2019-36083

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

With the success of the fifth generation new radio (5G NR), we are now witnessing the emergence of the sixth generation (6G) communication and its applications such as holographic telepresence, extended reality (XR), digital twin, and autonomous systems. The fundamental communication mechanism underlying these applications diverges significantly from the conventions of traditional communication systems. Rather than merely transmitting and receiving data, the focus now extends to delivering specialized services while addressing crucial aspects such as latency, energy efficiency, reliability, flexibility, and connection density. Key Performance Indicators (KPIs) for 6G present a significant leap, typically ranging from 10 to 100 times higher than previous generations in terms of data rate, reliability, latency, mobility, and energy efficiency. For example, the peak data rate is expected to reach 1 terabit-per-second (Tbps) and the latency is expected to be reduced to sub-millisecond levels. Given the unprecedented demands posed by these exacting requirements, the existing mechanisms and conventional approaches fall short in offering adequate support. This necessitates the development of an entirely new transmission approach to cater to 6G's unique challenges. Before we proceed further, we provide the fundamental of the key technologies and techniques

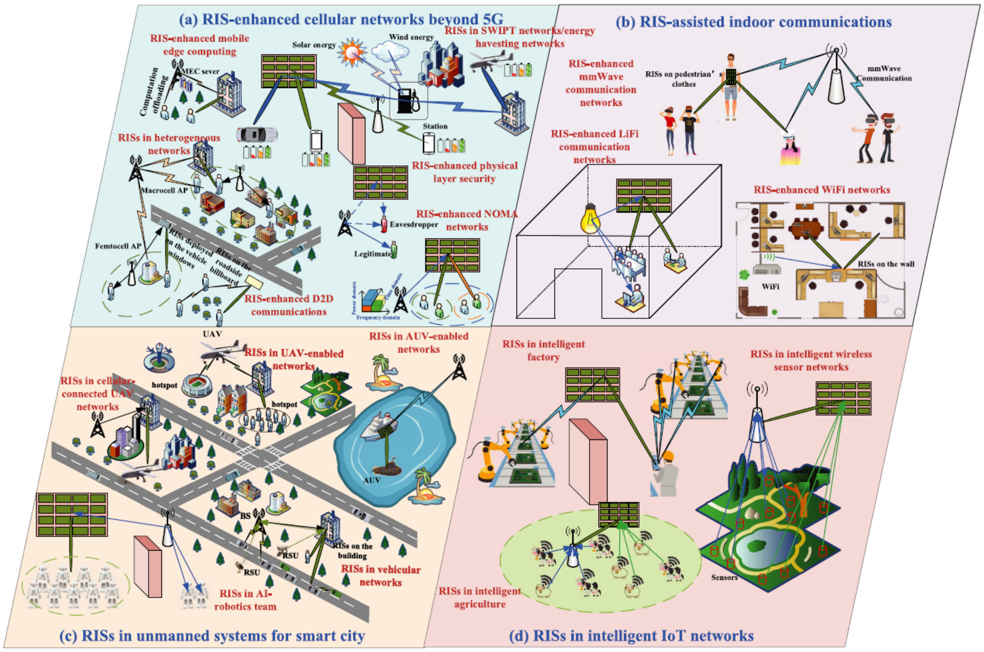

Figure 1.1: Employment of RIS in wireless communication networks.

that shape 6G wireless communications.

1.1.1 Basics of Reconfigurable Intelligent Surfaces

Reconfigurable intelligent surface (RIS), a planar metasurface consisting of a large array of reflecting elements, has received considerable interest for its potential to enhance the capacity and coverage of wireless networks. With the ability to dynamically manipulate the wireless communication environment, RIS has become a focal point of research in wireless communications, aiming to alleviate diverse challenges encountered within 6G wireless networks. The advantages of RIS are listed as follows.

- **Easy to deploy:** RISs are essentially passive devices crafted from electromagnetic (EM) materials. As illustrated in Fig. 1.1, the deployment of RIS extends to various structures, encompassing building facades, indoor walls, aerial platforms, roadside billboards, highway poles, vehicle windows, and even pedestrians' attire, owing to their cost-effectiveness.

- **Spectral efficiency enhancement:** RIS exhibits the capability to actively reshape the wireless propagation environment by mitigating power loss over extended distances. This is achieved through the passive reflection of incoming radio signals, enabling the creation of virtual line-of-sight (LoS) connections between base stations (BSs) and mobile users. Significant enhancement in throughput arises when the direct LoS link between BSs and users is obstructed by obstacles, such as tall buildings. Leveraging the intelligent deployment and design of RISs, it becomes feasible to construct a software-defined wireless environment. This, in turn, opens doors to potential improvements in the received signal-to-interference-plus-noise ratio (SINR).

- **Environment-friendly:** Unlike traditional relaying systems like amplify-and-forward (AF) and decode-and-forward (DF), which typically involve power amplification, RIS operates differently. It molds incoming signals by manipulating the phase shifts of individual reflecting elements. This approach renders the deployment of RISs not only more energy-efficient but also environmentally conscious when compared to conventional AF and DF systems.

- **Compatibility:** RIS facilitates full-duplex (FD) and full-band transmission by virtue of its nature as an electromagnetic wave reflector. Moreover, networks enhanced by RIS are seamlessly compatible with the standards and hardware of pre-existing wireless networks.

Due to the aforementioned attractive characteristics, RIS has been recognized as an effective solution for mitigating a wide range of challenges in 6G communications.

1.1.2 Basics of Terahertz Communications

Terahertz (THz) band ($0.1 \sim 10\,\mathrm{THz}$) communication is envisioned as one of the key enabling technologies to satisfy the exponential growth of data traffic volume, while

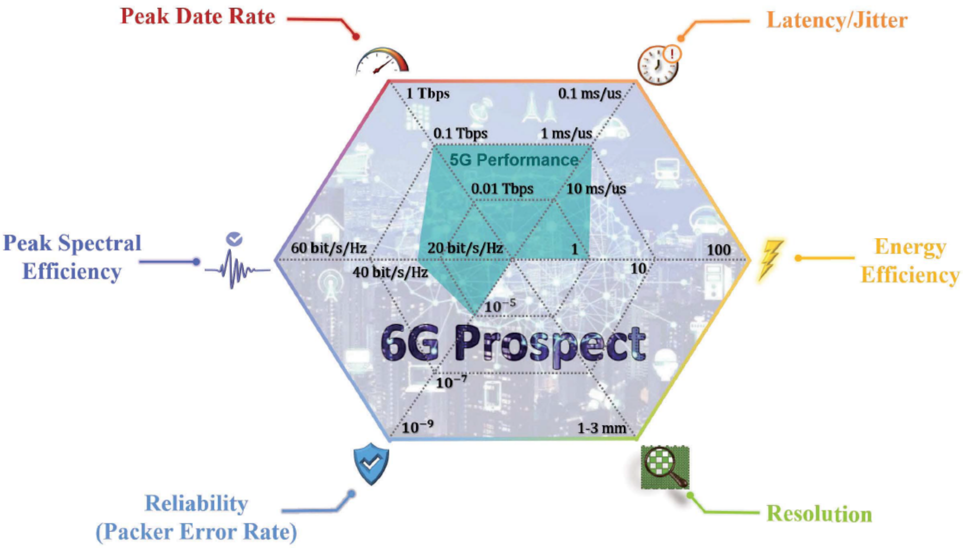

Figure 1.2: Extreme requirements of 6G.

meeting escalating demands in 6G and beyond wireless systems. As illustrated in Fig. 1.2, the anticipated landscape of 6G wireless systems encompasses peak data rates reaching 1 Tbps, accompanied by an envisaged peak spectral efficiency of 60 bps/Hz. Moreover, these systems target end-to-end reliability with a packet error rate of 10^{-9} and a latency as short as 0.1 ms. In addition, energy efficiency is expected to improve by over 100 times compared to 5G. The forthcoming Internet of Things (IoT) will provide 1 to 3 mm sensing resolutions, supporting billions of devices at an unprecedented scale.

Millimeter-wave (mmWave) communications (30–300 GHz) under 100 GHz have been officially adopted in recent 5G cellular systems. While the trend for higher carrier frequencies is apparent, achieving Tbps data rates and meeting the stringent Quality of Service (QoS) remain formidable for mmWave systems. Within the mm-wave systems under 100 GHz, constrained by a total consecutive bandwidth of less than 10 GHz, achieving a spectrum efficiency of 100 bps/Hz is an exceptionally demanding task, even with advanced physical layer techniques. The THz band presents itself as a pivotal wireless technology poised to address the future requirements of 6G wireless systems.

This is due to its four distinct strengths:

- Broad bandwidth: Offering contiguous bandwidth ranging from tens to hundreds of GHz.
- Ultra-fast symbol duration: Symbol durations measured in picoseconds.
- Antenna integration: Capable of integrating thousands of sub-millimeter-long antennas.
- Spectrum compatibility: Ease of coexistence with other regulated and standardized spectrums.

Traditionally, the THz band has been one of the least explored portions of the EM spectrum, largely due to the absence of efficient and practical THz transceivers and antennas. Nevertheless, it is acknowledged that THz communications confront a significant challenge arising from the pronounced signal directivity and severe signal attenuation. This often necessitates the establishment of LoS links to maintain communication quality.

1.1.3 Basics of Compressed Sensing Technique

Compressed sensing (CS) is a new paradigm to acquire, process, and recover sparse signals. This new approach is a very competitive alternative to conventional information processing operations including sampling, sensing, compression, estimation, and detection. Traditional way to acquire and reconstruct analog signals from sampled signals is based on Nyquist-Shannon's sampling theorem which states that the sampling rate should be at least twice the bandwidth of an analog signal to restore it from the discrete samples accurately. While these fundamental principles work well, they might be a bottleneck of resource overhead and also complexity in a situation where signals are sparse, meaning that the signals can be represented using a relatively small number of nonzero coefficients. At the heart of the CS lies the fact that a sparse signal vector can be recovered from the underdetermined linear system in a computationally efficient way. In other words, a small number of linear measurements of the signal contain enough information for its reconstruction. Main wisdom behind the CS is that essential knowledge in the large dimensional signals is just handful, and thus measurements with the size being proportional to the sparsity level of the input signal are enough to reconstruct the original signal. In the last decade, CS techniques have spread rapidly in many disciples such as medical imaging, machine learning, radar detection, seismology, computer science, statistics, and many others. Also, various wireless communication applications exploiting the sparsity of a target signal have been proposed in recent years. Notable examples, among many others, include channel estimation, interference cancellation, direction estimation, spectrum sensing, and symbol detection. To understand the principle of CS, we introduce a system given

$$\mathbf{y} = \mathbf{H}\mathbf{s}, \tag{1.1}$$

where $\mathbf{y} \in \mathbb{R}^{m \times 1}$ is the measurement vector, $\mathbf{H} \in \mathbb{R}^{m \times n}$ is the system matrix (a.k.a., the sensing matrix), and $\mathbf{s} \in \mathbb{R}^{n \times 1}$ is the desired signal vector. In the case of an overdetermined system ($m \geq n$) and the system matrix is a full rank matrix, one can recover $\mathbf{s}$ using a simple algorithm. However, when the system matrix is underdetermined, finding a solution is challenging and not straightforward. When the desired vector $\mathbf{s}$ is a non-sparse signal, one can apply a solution minimizing the l_2-norm of $\mathbf{s}$. That is,

$$\mathbf{s}^* = \arg\min ||\mathbf{s}||_2 \, s.t., \mathbf{y} = \mathbf{H}\mathbf{s}, \tag{1.2}$$

and one can obtain the estimated desired signal $\mathbf{s}^*$ as

$$\mathbf{s}^* = \mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{y}. \tag{1.3}$$

When the desired signal is a sparse vector, that is l_0-norm of the desired signal $\mathbf{s}$ $||\mathbf{s}||_0 = k$ where $k < n$, one can apply to find the l_0 from the measurement vector.

$$\mathbf{s}^* = \arg\min ||\mathbf{s}||_0 \, s.t., \mathbf{y} = \mathbf{H}\mathbf{s}. \tag{1.4}$$

Since this solution counts the number of nonzero elements in $\mathbf{s}$, one needs to search all possible combinations which is not practical for a large n and k. Alternative approach is to minimize l_1-norm as follow:

$$\mathbf{s}^* = \arg\min \|\mathbf{s}\|_1 \; s.t., \mathbf{y} = \mathbf{H}\mathbf{s}. \tag{1.5}$$

Using the l_1-norm minimization, one can apply convex optimization for finding the solution. While the linear programming to solve l_1-norm minimization problem is effective, it requires substantial computational complexity and is not feasible in practical scenarios.

To overcome this problem, a greedy algorithm has been proposed over the years. By this, one hopes to find the local optimal in each iteration expecting to find the global optimal in the end. Most popular algorithm is the orthogonal matching pursuit (OMP). In each iteration, a column maximally correlated with the observation vector is chosen. Once the solver selects the correct columns, now the system goes to the overdetermined system. There are three key observations from the CS recovery problem.

- When the sparsity k is smaller than the size of the desired signal vector, one can find the support of $\mathbf{s}$ more accurately.
- When the size of measurement vector m is given, one can recover the desired signal vector more accurately as k increases.
- When the fact that $k < n$ is given to the solver, one can find more accurate $\mathbf{s}$ than the solver without using the fact.

1.1.4 Basics of Riemannian Manifold Optimization Technique

Roughly speaking, a smooth manifold is a generalization of the Euclidean space on which a notion of differentiability exists. A smooth manifold together with an inner product, often called a Riemannian metric, forms a smooth Riemannian manifold. Since the smooth Riemannian manifold is a differentiable structure equipped with an inner product, we can use various ingredients such as Riemannian gradient, Hessian matrix, exponential map, and parallel translation, for solving optimization problems with quadratic cost function. Therefore, optimization techniques in the Euclidean vector space (e.g., steepest descent, Newton method, conjugate gradient method) can be readily extended to solve a problem in the smooth Riemannian manifold.

An essential concept in Riemannian manifold optimization is the tangent space. At each point on the manifold, there exists a tangent space, which is a vector space that approximates the manifold near that point. The tangent space captures the local geometry of the manifold and allows us to define operations like addition and scalar multiplication of tangent vectors.

To optimize a function on a Riemannian manifold, iterative methods such as gradient descent are commonly employed. However, in this context, the notion of gradient differs from that in Euclidean spaces. Instead, we use the Riemannian gradient, which measures the rate of change of the objective function with respect to the tangent vector at each point on the manifold. The Riemannian gradient takes into account the Riemannian metric and the curvature of the manifold, providing a direction for optimization.

Various optimization algorithms have been developed for Riemannian manifold optimization, each with its strengths and considerations. These algorithms include Riemannian gradient descent, conjugate gradient methods, trust-region methods, and quasi-Newton methods. They leverage the Riemannian geometry of the manifold and incorporate appropriate adjustments to handle its curvature, ensuring effective optimization.

Riemannian manifold optimization finds applications in diverse fields, including machine learning, computer vision, robotics, and physics. It enables efficient optimization on curved manifolds, allowing for the development of algorithms that can handle non-Euclidean data structures and complex geometries. By incorporating the intrinsic geometry of the manifold, Riemannian manifold optimization provides a powerful framework for solving optimization problems in these domains.

1.2 Contribution and Organization

In this dissertation, we introduce a RIS-assisted wireless communication system for 6G.

In Chapter 2, we introduce a novel channel estimation framework for near-field RIS-assisted THz systems. To support extremely high data rates in 6G wireless networks, RIS-assisted THz communications have gained much attention in recent years. By manipulating the phase shifts of reflecting elements, RIS can proactively adjust the wireless propagation environment of THz systems, thereby enhancing the overall throughput significantly. To realize the full potential of RIS-assisted THz systems, an acquisition of accurate channel information is of great importance. However, since the wavefront of the THz electromagnetic signal is spherical, the conventional channel estimation techniques using the planar wavefront assumption suffer from severe performance degradation in the near-field RIS-assisted THz systems. An aim of this work is to propose an effective channel estimation technique for near-field RIS-assisted wideband THz systems. Key idea of the proposed *polar-domain frequency-dependent RIS-assisted channel estimation (PF-RCE)* scheme is to estimate the sparse multipath components (i.e., angles, distances, and path gains) of the near-field THz channel by exploiting the polar-domain sparsity and common support properties.

In Chapter 3, we present an energy-efficient power control and beamforming scheme for RIS-assisted IoT networks. RIS, a planar metasurface consisting of a large number of low-cost reflecting elements, has received much attention due to the ability to improve both the spectrum and energy efficiencies by reconfiguring the wireless propagation environment. In this work, we propose an RIS phase shift and BS beamforming optimization technique that minimizes the uplink transmit power of the RIS-aided IoT network. Key idea of the proposed scheme, referred to as *Riemannian conjugate gradient-based joint optimization (RCG-JO)*, is to jointly optimize the RIS phase shifts and the BS beamforming vectors using the Riemannian conjugate gradient technique. By exploiting the product Riemannian manifold structure of the sets of unit-modulus phase shifts and unit-norm beamforming vectors, we convert the nonconvex uplink power minimization problem into the unconstrained problem and then find out the optimal solution over the product Riemannian manifold.

In Chapter 4, we propose an innovative approach to address the challenges of frequency-dependent beamforming for RIS-assisted wideband THz systems. A significant hurdle in wideband THz communication pertains to the severe array gain loss caused by the beam split effect that path components split into distinct spatial directions across different subcarrier frequencies. Therefore, the conventional phase shift control and beamforming techniques cannot be directly applied to wideband THz systems. In this work, we introduce a new paradigm called *RIS-assisted wideband beamforming (RWB)* technique that maximizes the achievable data rate of the RIS-assisted wideband THz systems. Key idea of RWB involves the alternating optimization of the analog beamforming vector and the RIS phase shift vector. By properly designing the parameters encompassing time delays, analog phase shifts, and RIS phase shifts of the beamforming network, we aim to the maximize the achievable data rate of the wideband THz system. To counteract the array gain loss caused by the beam split effect, we leverage a small yet strategic ensemble of true-time delay (TDD)-based phase shifters and analog phase shifters. This ensemble allows to simultaneously generate frequency-dependent beams aligning with the physical directions at different subcarrier frequencies. Using the frequency-dependent beamforming vector, we then exploit the Riemannian conjugate gradient (RCG) method to find out the phase shifts that maximize the achievable data rate of RIS-assisted wideband THz systems.

Chapter 5 summarizes the contribution of the dissertation and discusses the future research directions based on studies of this dissertation.

Chapter 2

Parametric Sparse Channel Estimation for RIS-Assisted Terahertz Systems

In this chapter, we study the channel estimation framework tailored for near-field RIS-assisted THz systems. To facilitate the demanding data rates expected in 6G wireless networks, there has been a significant surge of interest in RIS-assisted THz communications in recent years. Through the strategic manipulation of phase shifts in reflective elements, the RIS can proactively adjust the wireless propagation environment of THz systems, leading to a substantial enhancement in overall throughput. To realize the full potential of RIS-assisted THz systems, an acquisition of accurate channel information is of great importance. However, since the wavefront of the THz electromagnetic signal is spherical, the conventional channel estimation techniques relying on the assumption of planar wavefront suffers severe performance degradation in near-field RIS-assisted THz systems. An aim of this work is to introduce an efficient channel estimation technique designed for near-field RIS-assisted wideband THz systems. Key idea of the proposed polar-domain frequency-dependent RIS-assisted channel estimation (PF-RCE) scheme is to estimate the sparse multipath components (i.e., angles, distances, and path gains) of the near-field THz channel by exploiting the polar-domain sparsity and common support properties.

2.1 Introduction

To support the exponential growth of data traffic in 6G networks, terahertz (THz) communications have attracted considerable interest from both industry and academia [2]. By leveraging the broad spectrum available in the THz band ($0.1 \sim 10\,\mathrm{THz}$), THz communications can enable truly immersive mobile services, such as digital twin, holographic telepresence, and metaverse experiences. However, THz communications face a significant challenge posed by the strong directivity and severe signal attenuation of transmit signals, which often necessitate a line-of-sight (LoS) link to maintain the communication quality. Recently, reconfigurable intelligent surfaces (RIS) that proactively modify the wireless channel through intelligent signal reflection have emerged as a potential solution to provide an alternative LoS link [3]. To fully exploit the potential of RIS-assisted THz communications, the phases of RIS reflecting elements should be properly configured to reflect the surrounding environment [4]. To do so, an acquisition of the RIS-assisted channel information at the base station (BS) is of great importance [5].

Over the years, various channel acquisition techniques for RIS-assisted high frequency systems have been proposed [6–9]. In [6], a channel estimation technique using the tensor completion method has been proposed for RIS-assisted systems. In [7], a two-stage channel estimation technique for RIS-assisted multi-user systems has been proposed. Potential problem of these approaches is the huge pilot overhead caused by the full-dimensional RIS-assisted channel estimation. To reduce the dimension of the channel to be estimated, compressed sensing (CS)-based techniques that exploit the limited scattering property of THz/mmWave channels have been proposed [8–10]. Since the system models of these CS techniques are typically based on the far-field channel model where electromagnetic (EM) radiation is modeled as planar waves, the CS-based channel estimation techniques may not perform well in near-field THz systems where EM radiation is modeled as spherical waves. In fact, in the RIS-assisted THz systems where the extremely large number of reflecting elements is used, the array aperture is comparable to the propagation distance so that the THz channel can be categorized as a near-field channel[1].

Unfortunately, the estimation of near-field RIS-assisted THz channel is not easy since the angular-domain sparsity is not available due to the spherical wavefront. In the near-field region, the incident angle at each antenna is different, meaning that the near-field THz channel is no longer sparse in the angular-domain. Furthermore, since the phase delay between two adjacent antennas depends on both the angle and the distance, the near-field THz channel is modeled as a complex function of angle and distance. Due to this so-called *near-field effect*, the conventional channel estimation schemes relying on the angular-domain sparsity are ineffective for near-field RIS-assisted THz systems [11]. Yet another important issue in the wideband THz systems is that the difference between carrier and subcarrier frequencies is large so that the array response vector (a set of phase delays in antenna elements) of each subcarrier is different [12]. This phenomenon, so-called *frequency-wideband effect*, makes it difficult to estimate multiple subcarrier channels simultaneously. Therefore, to come up with a proper channel acquisition mechanism alleviating the near-field and frequency-wideband effects is crucial for the success of the near-field RIS-assisted wideband THz systems.

An aim of this chapter is to propose an efficient channel estimation technique for the near-field RIS-assisted wideband THz systems. Key idea of the proposed scheme, referred to as the *polar-domain frequency-dependent RIS-assisted channel estimation (PF-RCE)*, is to estimate the sparse multipath components (i.e., angles, distances, and path gains) of the near-field RIS-assisted THz channel by exploiting the polar-domain sparsity and common support properties. Since the near-field channel is a function of the angles and distances of a few dominant paths, the near-field RIS-assisted

[1]For example, in the 512-antenna THz MIMO systems operating at 0.1 THz, the Rayleigh distance (i.e., a boundary between the far-field region and near-field region) is around 400 m, which covers a large part of the THz cell area.

THz channel vector can be readily modeled as a sparse vector in the polar-domain, a coordinate system represented by the angle and distance. Also, since the signals of different subcarriers propagate through the same physical path in wideband systems, the non-zero element positions (i.e., support) of the sparse channel vectors are the same for all subcarriers. By leveraging this property, we can formulate a joint sparse recovery problem for the acquisition of the multipath components and solve it using the block-sparse recovery algorithm.

We note that there have been some works investigating the near-field effect in wireless communication systems [13–18]. In [13] and [14], theoretical limitations for RIS-assisted wideband localization and near-field localization have been investigated. In [15], a joint dictionary learning and sparse recovery algorithm for near-field channel estimation has been proposed. In [16] and [17], low-complexity near-field localization techniques that exploit RIS as a lens have been proposed. Also, in [18], a channel estimation technique for hybrid RIS-empowered multiple-input multiple-output (MIMO) systems has been proposed. A downside of these approaches is that the wideband effect which causes the channel to have a frequency-dependent sparse structure is not accounted for so the performance degradation would be severe in the near-field RIS-assisted wideband THz systems. Our work is distinct from previous works in the following aspects:

- While previous works have focused on the channel estimation of traditional cellular systems without RIS, our work deals with the channel estimation of RIS-assisted communication systems. To the best of our knowledge, this is the first work investigating the channel estimation of near-field wideband RIS-assisted THz systems while considering both near-field and wideband effects. This makes the proposed PF-RCE scheme particularly important for emerging RIS-assisted THz communication systems.

- We propose a novel polar bin design to enhance the sparse recovery performance of PF-RCE. Note that the performance of the sparse recovery algorithm depends heavily on the column correlation of the sensing matrix. The previous works use a simple polar bin generation strategy based on uniform quantization. In our work, we deliberately design the polar bin to minimize the column correlation of the sensing matrix.

- We propose an RIS phase shift control scheme to support the proposed channel estimation framework. While the conventional phase shift control schemes focus on the design of phase shifts maximizing the throughput, little work has been done on the design of phase shifts improving the channel estimation accuracy. By configuring the RIS phase shifts to achieve desirable properties of the sensing matrix (e.g., column orthogonality), we can improve the channel estimation accuracy without using additional pilot resources.

- We provide the empirical simulation results from which we demonstrate that PF-RCE outperforms the conventional channel estimation schemes by a large margin in terms of the normalized mean square error (NMSE). For example, when compared to the conventional far-field and narrowband CS-based schemes, PF-RCE achieves more than $4.2\,$dB and $7.8\,$dB NMSE gains, respectively. We also demonstrate that PF-RCE achieves the near-optimal sparse recovery performance (close to the oracle bound). Furthermore, we show that by employing the proposed polar bin design and the RIS phase shift control scheme, the NMSE gain of PF-RCE can be increased by $1.2\,$dB.

Notations: Lower and upper case symbols are used to denote vectors and matrices, respectively. The superscripts $(\cdot)^{*}$, $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^{\dagger}$ denote the conjugate, transpose, Hermitian transpose, and pseudo-inverse, respectively. $\|\mathbf{x}\|_2$ and $\|\mathbf{X}\|_{\mathrm{F}}$ are used as the Euclidean norm of a vector $\mathbf{x}$ and the Frobenius norm of a matrix $\mathbf{X}$, respectively. $\exp(x)$ denotes the exponential function of x, $\mathrm{vec}(\mathbf{X})$ denotes the vectorization of $\mathbf{X}$, and $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix whose diagonal elements are $\mathbf{x}$. $\mathbf{X} \odot \mathbf{Y}$ and $\mathbf{X} \otimes \mathbf{Y}$ denote the Hadamard and Kronecker products of $\mathbf{X}$ and $\mathbf{Y}$, respectively. In

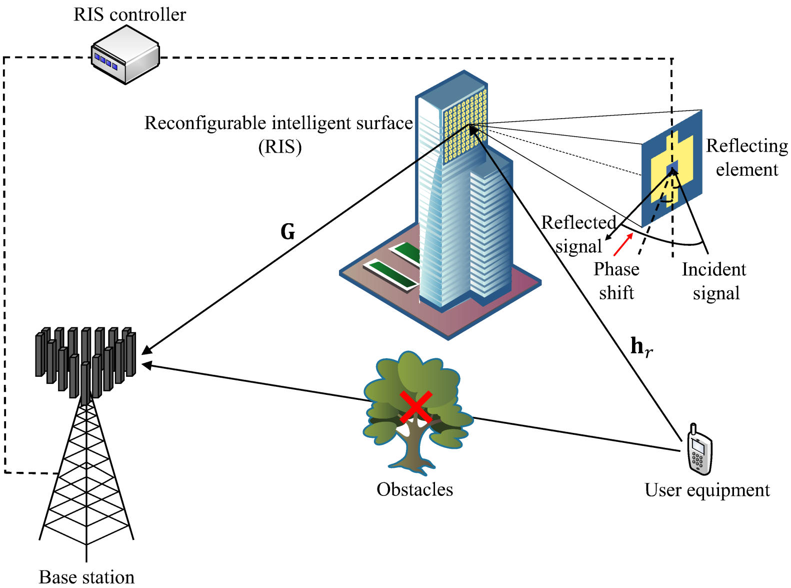

Figure 2.1: Near-field RIS-assisted THz system model.

addition, $\mathbf{X} * \mathbf{Y}$ and $\mathbf{X} \bullet \mathbf{Y}$ denote the column-wise and row-wise Khatri-Rao products of $\mathbf{X}$ and $\mathbf{Y}$, respectively.

2.2 RIS-Assisted Wideband THz System Model

In this section, we present the system model and the uplink channel estimation protocol for the near-field RIS-assisted wideband THz system. We also discuss the RIS-assisted THz channel model accounting for both the near-field and frequency-wideband effects.

2.2.1 RIS-Assisted Wideband THz System Model

We consider an RIS-assisted THz system where a N-antenna BS serves a single-antenna UE (see Fig. 2.1). We assume that the RIS consists of M reflecting elements

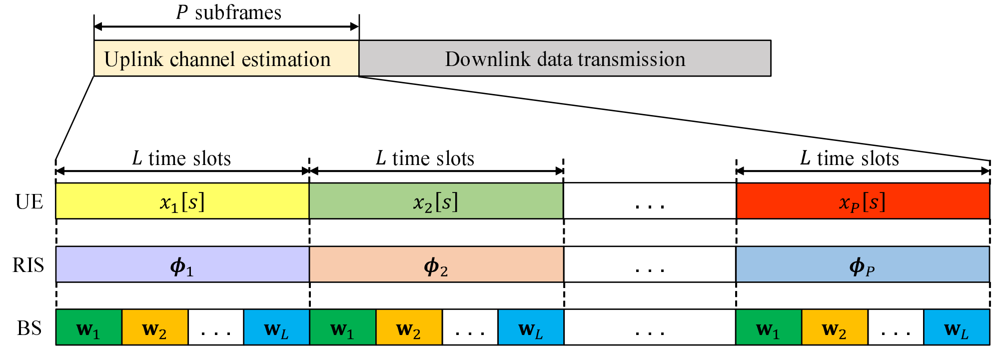

Figure 2.2: RIS-assisted wideband THz channel estimation protocol.

arranged in a uniform linear array (ULA)[2] and an RIS controller operated by the BS is connected to the RIS through a dedicated control link. Also, we consider the OFDM system with the carrier frequency f_c and S subcarriers. During the uplink channel estimation period, the UE transmits the uplink pilot signals and then the BS estimates the uplink channel information using the received pilot signals. To be specific, the uplink channel estimation period consists of P successive subframes, each of which is divided into L time slots (see Fig. 2.2). By exploiting the channel reciprocity of time division duplexing (TDD) systems, the BS uses the acquired uplink channel information in the phase shift control and the downlink data transmission.

In the RIS-assisted THz systems, other than the direct channel, the reflected channels (i.e., the UE-RIS channel $\mathbf{h}_r[s] \in \mathbb{C}^{M\times 1}$ and the RIS-BS channel $\mathbf{G}[s] \in \mathbb{C}^{N\times M}$) need to be considered. Here, we consider a practical scenario where the direct links are severely blocked due to obstacles (e.g., walls and corners), and thus the BS communicates with the UE only via the RIS-assisted links[3]. The effective uplink channel

[2]Note that the proposed channel estimation framework can be readily extended to the uniform planar array (UPA)-type RIS-assisted systems, in which the azimuth angles as well as the elevation angles at the RIS are used for the channel model. Since the polar-domain sparsity and the common support properties are valid in UPA-type RIS-assisted THz systems, we can estimate both azimuth and elevation angles using the proposed scheme.

[3]The proposed scheme can be readily extended to a scenario considering the direct BS-UE commu-

$\mathbf{h}[s] \in \mathbb{C}^{N\times 1}$ from the UE to the BS at the s-th subcarrier is

$$\mathbf{h}[s] = \mathbf{G}[s]\text{diag}(\boldsymbol{\phi})\mathbf{h}_r[s] \tag{2.1}$$
$$= \mathbf{G}[s]\text{diag}(\mathbf{h}_r[s])\boldsymbol{\phi} \tag{2.2}$$
$$= \mathbf{H}[s]\boldsymbol{\phi}, \tag{2.3}$$

where $\mathbf{H}[s] = \mathbf{G}[s]\text{diag}(\mathbf{h}_r[s]) \in \mathbb{C}^{N\times M}$ is the RIS-assisted channel matrix and $\boldsymbol{\phi} = [e^{j\varphi_1}, \cdots, e^{j\varphi_M}]^{\text{T}}$ is the phase shift vector with φ_m being the phase shift of the m-th RIS reflecting element.

Under this setup, the received pilot signal $\mathbf{r}_p[s] \in \mathbb{C}^{N\times 1}$ of UE associated with s-th subcarrier at p-th subframe is expressed as

$$\mathbf{r}_p[s] = \mathbf{H}[s]\boldsymbol{\phi}_p x_p[s] + \mathbf{n}_p[s], \quad p = 1, \cdots, P, \tag{2.4}$$

where $\boldsymbol{\phi}_p \in \mathbb{C}^{M\times 1}$ is the RIS phase shift vector, $x_p[s]$ is the pilot symbol, and $\mathbf{n}_p[s]$ is the Gaussian noise vector of s-th subcarrier at p-th subframe. During L successive time slots, the BS sequentially employs the combining vectors $\{\mathbf{w}_l\}_{l=1}^{L}$ to obtain the processed signal $y_{l,p}[s]$:

$$y_{l,p}[s] = \mathbf{w}_l^{\text{H}}\mathbf{H}[s]\boldsymbol{\phi}_p x_p[s] + n_{l,p}[s],\ l = 1, \cdots, L,\ p = 1, \cdots, P. \tag{2.5}$$

By combining the processed signals into a $L \times P$ pilot signal matrix $\mathbf{Y}[s]$, we obtain

$$\mathbf{Y}[s] = \mathbf{W}^{\text{H}}\mathbf{H}[s]\boldsymbol{\Phi}\mathbf{X}[s] + \mathbf{N}[s], \tag{2.6}$$

where $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_L] \in \mathbb{C}^{N\times L}$ is the combining matrix, $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_P] \in \mathbb{C}^{M\times P}$ is the RIS phase shift matrix (see Section IV.B for the detailed RIS phase shift design), and $\mathbf{X}[s] = \text{diag}(x_1[s], \cdots, x_P[s])$ is the pilot symbol matrix.

nication link. By switching off all the reflecting elements, the direct channel can be easily estimated via conventional channel estimation techniques.

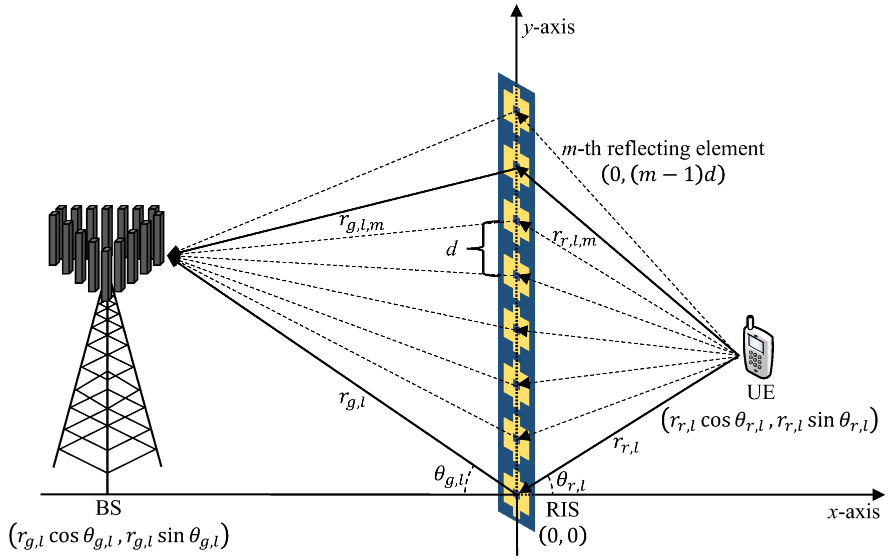

Figure 2.3: Illustration of the near-field RIS-assisted THz channel.

2.2.2 Near-Field Effect of RIS-Assisted THz Systems

The Rayleigh distance is widely used to quantify the boundary between the far-field and near-field regions. The Rayleigh distance D is proportional to the number of antennas M and the wavelength λ_c (i.e., $D = \frac{1}{2}M^2\lambda_c$) [1]. In the mmWave systems, D is only a few meters (e.g., $D \approx 5\,\text{m}$ when $M = 32$ and $f_c = 28\,\text{GHz}$) so that most of the signal transmissions take place in the far-field region. However, due to the extremely large number of RIS reflecting elements (e.g., $M = 128 \sim 1024$), D can be up to a few hundred meters (e.g., $D \approx 400\,\text{m}$ when $M = 512$ and $f_c = 0.1\,\text{THz}$) in the RIS-assisted THz systems, meaning that most of the coverage area can be classified as a near-field region.

In the near-field THz channel model, due to the spherical wavefront of the EM waves, the phase delay between two adjacent antennas is affected by both the angle and distance. To be specific, let r_m and r be the distances between the transmitter and the m-th receiving antenna and reference antenna, respectively, then the phase delay of the m-th receiving antenna is given by $e^{-j2\pi f \Delta r_m / c}$ where $\Delta r_m = r_m - r$ and f is the signal frequency. In the far-field channel, Δr_m^{far} is a function of the incident angle θ at the reference antenna:

$$\Delta r_m^{\text{far}} = -d(m-1)\sin\theta, \tag{2.7}$$

where d is the antenna spacing. Then, the far-field array response vector can be expressed as $\mathbf{a}^{\text{far}}(\theta) = [e^{-j2\pi f \Delta r_1^{\text{far}}/c}, \cdots, e^{-j2\pi f \Delta r_M^{\text{far}}/c}]^{\text{T}}$.

In the near-field channel, however, Δr_m depends on both angle and distance due to the spherical wavefront. When the coordinate of the reference antenna is set to $(0,0)$, then the coordinates of the transmitter and the m-th receiving antenna are $(r\cos\theta, r\sin\theta)$ and $(0, d(m-1))$, respectively, so that Δr_m^{near} consisting of far-field and near-field terms is given by (see Fig. 2.3)

$$\Delta r_m^{\text{near}} = \sqrt{(r\cos\theta)^2 + (r\sin\theta - (m-1)d)^2} - r \tag{2.8}$$

$$\overset{(a)}{\approx} \underbrace{-d(m-1)\sin\theta}_{\text{far-field term}} + \underbrace{\frac{d^2(m-1)^2}{2}\frac{\cos^2\theta}{r}}_{\text{near-field term}}, \tag{2.9}$$

where (a) follows from $\sqrt{1+x} \approx 1 + \frac{1}{2}x - \frac{1}{8}x^2$ [19]. Since Δr_m^{near} is a function of the incident angle θ and the distance r, the near-field array response vector can be expressed as $\mathbf{a}^{\text{near}}(\theta, r) = [e^{-j2\pi f \Delta r_1^{\text{near}}/c}, \cdots, e^{-j2\pi f \Delta r_M^{\text{near}}/c}]^{\text{T}}$.

2.2.3 RIS-Assisted Wideband THz Channel Model

In this work, we use the near-field geometric channel models for the RIS-BS channel $\mathbf{G}[s]$ and the UE-RIS channel $\mathbf{h}_r[s]$ [20].

First, the near-field RIS-BS channel $\mathbf{G}[s] \in \mathbb{C}^{N\times M}$ at the s-th subcarrier is expressed as

$$\mathbf{G}[s] = \sum_{l=1}^{P_g} \alpha_{l,s}\mathbf{a}_{\text{B},s}(\psi_{g,l})\mathbf{a}_{\text{R},s}^{\text{H}}(\theta_{g,l}, r_{g,l})e^{-j2\pi f_s r_{g,l}/c}, \tag{2.10}$$

where P_g is the number of propagation paths, $\alpha_{l,s}$ is the path gain, f_s is the baseband frequency of the s-th subcarrier, $\psi_{g,l}$ is the angle of arrival (AoA), $\theta_{g,l}$ is the angle of departure (AoD), and $r_{g,l}$ is the distance between the reference reflecting element and the scatterer or BS of the l-th path. Also, $\mathbf{a}_{\mathrm{B},s}(\psi_{g,l})$ and $\mathbf{a}_{\mathrm{R},s}(\theta_{g,l}, r_{g,l})$ are the far-field and near-field array response vectors at the BS and RIS, respectively, given by [20]

$$\mathbf{a}_{\mathrm{B},s}(\psi_{g,l}) = [e^{j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})d\sin\psi_{g,l}}, \cdots, e^{j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})d(N-1)\sin\psi_{g,l}}]^{\mathrm{T}}, \tag{2.11}$$

$$\mathbf{a}_{\mathrm{R},s}(\theta_{g,l}, r_{g,l}) = [e^{-j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})\Delta r_{g,l,1}}, \cdots, e^{-j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})\Delta r_{g,l,M}}]^{\mathrm{T}}, \tag{2.12}$$

where f_c is the carrier frequency, λ_c is the wavelength, and $\Delta r_{g,l,m}$ is the difference of the distance between the m-th reflecting element and the reference reflecting element given by

$$\Delta r_{g,l,m} \approx -d(m-1)\sin\theta_{g,l} + \frac{d^2(m-1)^2}{2}\frac{\cos^2\theta_{g,l}}{r_{g,l}}. \tag{2.13}$$

By defining $\boldsymbol{\psi}_g = [\psi_{g,1}, \cdots, \psi_{g,P_g}]^{\mathrm{T}}$, $\boldsymbol{\theta}_g = [\theta_{g,1}, \cdots, \theta_{g,P_g}]^{\mathrm{T}}$, and $\mathbf{r}_g = [r_{g,1}, \cdots, r_{g,P_g}]^{\mathrm{T}}$, we obtain the matrix-vector form of $\mathbf{G}[s]$:

$$\mathbf{G}[s] = \mathbf{A}_{\mathrm{B},s}(\boldsymbol{\psi}_g)\boldsymbol{\Lambda}_s\mathbf{A}_{\mathrm{R},s}^{\mathrm{H}}(\boldsymbol{\theta}_g, \mathbf{r}_g), \tag{2.14}$$

where $\mathbf{A}_{\mathrm{B},s}(\boldsymbol{\psi}_g) = [\mathbf{a}_{\mathrm{B},s}(\psi_{g,1}), \cdots, \mathbf{a}_{\mathrm{B},s}(\psi_{g,P_g})]$ is the far-field array response matrix at the BS, $\mathbf{A}_{\mathrm{R},s}(\boldsymbol{\theta}_g, \mathbf{r}_g) = [\mathbf{a}_{\mathrm{R},s}(\theta_{g,1}, r_{g,1}), \cdots, \mathbf{a}_{\mathrm{R},s}(\theta_{g,P_g}, r_{g,P_g})]$ is the near-field array response matrix at the RIS, and $\boldsymbol{\Lambda}_s = \mathrm{diag}(\alpha_{1,s}e^{-j2\pi f_s r_{g,1}/c}, \cdots, \alpha_{P_g,s}e^{-j2\pi f_s r_{g,P_g}/c})$ is the path gain matrix.

Second, the UE-RIS channel $\mathbf{h}_r[s]$ at the s-th subcarrier is expressed as

$$\mathbf{h}_r[s] = \sum_{l=1}^{P_r}\beta_{l,s}\mathbf{a}_{\mathrm{R},s}(\theta_{r,l}, r_{r,l})e^{-j2\pi f_s r_{r,l}/c}, \tag{2.15}$$

where P_r is the number of propagation paths, $\beta_{l,s}$ is the path gain, $\theta_{r,l}$ is the AoA, and $r_{r,l}$ is the distance between the reference reflecting element and the scatterer or UE of the l-th path. By defining $\boldsymbol{\theta}_r = [\theta_{r,1} \cdots \theta_{r,P_r}]^{\mathrm{T}}$ and $\mathbf{r}_r = [r_{r,1} \cdots r_{r,P_r}]^{\mathrm{T}}$, we obtain the matrix-vector form of $\mathbf{h}_r[s]$:

$$\mathbf{h}_r[s] = \mathbf{A}_{\mathrm{R},s}(\boldsymbol{\theta}_r, \mathbf{r}_r)\boldsymbol{\beta}_s, \tag{2.16}$$

where $\mathbf{A}_{\mathrm{R},s}(\boldsymbol{\theta}_r, \mathbf{r}_r) = [\mathbf{a}_{\mathrm{R},s}(\theta_{r,1}, r_{r,1}), \cdots, \mathbf{a}_{\mathrm{R},s}(\theta_{r,P_r}, r_{r,P_r})]$ is the near-field array response matrix at the RIS and $\boldsymbol{\beta}_s = [\beta_{1,s}e^{-j2\pi f_s r_{r,1}/c}, \cdots, \beta_{P_r,s}e^{-j2\pi f_s r_{r,P_r}/c}]^{\mathrm{T}}$ is the path gain vector.

Using (2.14) and (2.16), the RIS-assisted channel $\mathbf{H}[s]$ at the s-th subcarrier can be expressed as a function of multipath components (i.e., AoAs, AoDs, distances, and path gains) [21].

Lemma 1. *The RIS-assisted channel matrix* $\mathbf{H}[s] = \mathbf{G}[s]\text{diag}(\mathbf{h}_r[s])$ *can be expressed as [21]*

$$\mathbf{H}[s] = \mathbf{A}_{\text{B},s}(\boldsymbol{\psi}_g)(\mathbf{\Lambda}_s \otimes \boldsymbol{\beta}_s^{\text{T}})(\mathbf{A}_{\text{R},s}^*(\boldsymbol{\theta}_r, \mathbf{r}_r) \bullet \mathbf{A}_{\text{R},s}(\boldsymbol{\theta}_g, \mathbf{r}_g))^{\text{H}}. \tag{2.17}$$

Also, the vectorized RIS-assisted channel $\text{vec}(\mathbf{H}[s])$ *can be expressed as*

$$\text{vec}(\mathbf{H}[s]) = \big((\mathbf{A}_{\text{R},s}(\boldsymbol{\theta}_r, \mathbf{r}_r) \bullet \mathbf{A}_{\text{R},s}^*(\boldsymbol{\theta}_g, \mathbf{r}_g)) \otimes \mathbf{A}_{\text{B},s}(\boldsymbol{\psi}_g)\big)\text{vec}(\mathbf{\Lambda}_s \otimes \boldsymbol{\beta}_s^{\text{T}}). \tag{2.18}$$

Finally, by vectorizing $\mathbf{Y}[s]$ into $\mathbf{y}[s] = \text{vec}(\mathbf{Y}[s]) \in \mathbb{C}^{LP\times 1}$, we obtain the linear system:

$$\begin{aligned}
\mathbf{y}[s] &= ((\mathbf{\Phi}\mathbf{X}[s])^{\text{T}} \otimes \mathbf{W}^{\text{H}})\text{vec}(\mathbf{H}[s]) + \mathbf{n}[s] && (2.19)\\
&= ((\mathbf{\Phi}\mathbf{X}[s])^{\text{T}} \otimes \mathbf{W}^{\text{H}})\big((\mathbf{A}_{\text{R},s}(\boldsymbol{\theta}_r, \mathbf{r}_r) \bullet \mathbf{A}_{\text{R},s}^*(\boldsymbol{\theta}_g, \mathbf{r}_g)) \otimes \mathbf{A}_{\text{B},s}(\boldsymbol{\psi}_g)\big) \\
&\quad\ \text{vec}(\mathbf{\Lambda}_s \otimes \boldsymbol{\beta}_s^{\text{T}}) + \mathbf{n}[s] && (2.20)\\
&= \mathbf{\Psi}[s]\mathbf{g}[s] + \mathbf{n}[s], && (2.21)
\end{aligned}$$

where $\mathbf{\Psi}[s] = ((\mathbf{\Phi}\mathbf{X}[s])^{\text{T}} \otimes \mathbf{W}^{\text{H}})\big((\mathbf{A}_{\text{R},s}(\boldsymbol{\theta}_r, \mathbf{r}_r) \bullet \mathbf{A}_{\text{R},s}^*(\boldsymbol{\theta}_g, \mathbf{r}_g)) \otimes \mathbf{A}_{\text{B},s}(\boldsymbol{\psi}_g)\big) \in \mathbb{C}^{LP\times P_g^2 P_r}$ is the system matrix and $\mathbf{g}[s] = \text{vec}(\mathbf{\Lambda}_s \otimes \boldsymbol{\beta}_s^{\text{T}}) \in \mathbb{C}^{P_g^2 P_r \times 1}$ is the combined path gain vector.

2.2.4 Frequency-Wideband Effect of RIS-Assisted THz Systems

In the linear system (2.21), the system matrix $\mathbf{\Psi}[s]$ is expressed as a function of the array response matrices $\mathbf{A}_{\text{B},s}(\boldsymbol{\psi}_g)$, $\mathbf{A}_{\text{R},s}(\boldsymbol{\theta}_g, \mathbf{r}_g)$, and $\mathbf{A}_{\text{R},s}(\boldsymbol{\theta}_r, \mathbf{r}_r)$. Note that these array response matrices are functions of the ratio $\gamma_s = \frac{f_s}{f_c}$ between the carrier frequency f_c and the subcarrier frequency f_s (see (2.11) and (2.12)). In the mmWave band, f_s is relatively smaller than f_c so that $\gamma_s \approx 0$ in most cases. This means that the array response vectors of all subcarriers are almost identical. However, in the THz band, γ is larger or smaller than 0 due to the extremely large bandwidth so that the array response vector is expressed as a function of the subcarrier frequency. One can deduce from this discussion that $\boldsymbol{\Psi}[s]$ corresponding to one subcarrier is different from others [22]. Due to this so-called *frequency-wideband effect*, in the wideband THz systems, it is very difficult to estimate the multipath components of multiple subcarriers simultaneously.

2.3 Sparse Channel Estimation for Near-Field RIS-Assisted THz Systems

The primary goal of the proposed PF-RCE is to estimate the sparse multipath components, i.e., angles ($\boldsymbol{\psi}_g$, $\boldsymbol{\theta}_g$, $\boldsymbol{\theta}_r$), distances ($\mathbf{r}_g$, $\mathbf{r}_r$), and path gains ($\{\boldsymbol{\Lambda}_s, \boldsymbol{\beta}_s\}$), of the near-field RIS-assisted wideband THz channel. To this end, we map the angles and distances to the positions of non-zero elements of the sparse path gain vector. In doing so, we can convert the multipath components estimation problem to the problem to find out the support of the sparse path gain vector. Also, since the signals of different subcarriers propagate through the same physical path in the wideband systems, the sparse path gain vectors of all subcarriers have the common support. Thus, by exploiting the measurements of all subcarriers, we formulate a joint sparse recovery problem to find out the common support. Using the block-sparse recovery algorithm, we can effectively acquire the common support and the corresponding multipath components and then recover the near-field RIS-assisted THz channel from the acquired multipath components [23].

Overall process of PF-RCE is as follows:

- **Polar-domain sparse mapping**: We map the angles ($\boldsymbol{\psi}_g$, $\boldsymbol{\theta}_g$, $\boldsymbol{\theta}_r$) and distances ($\mathbf{r}_g$, $\mathbf{r}_r$) to the positions of non-zero elements (i.e., support $\boldsymbol{\Omega}$) of the sparse path gain vector.

- **Block-sparse representation**: We formulate the block-sparse linear system as

$\mathbf{y} = \bar{\mathbf{\Psi}}\bar{\mathbf{g}}+\mathbf{z}$ where $\bar{\mathbf{\Psi}}$ is the polar-domain sensing matrix and $\bar{\mathbf{g}}$ is the block-sparse path gain vector.

- **Block-sparse recovery and channel reconstruction**: Using the block-sparse recovery algorithm, we find out $\mathbf{\Omega}$ and $\bar{\mathbf{g}}$. After that, we acquire the angles ($\boldsymbol{\psi}_g$, $\boldsymbol{\theta}_g$, $\boldsymbol{\theta}_r$), distances ($\mathbf{r}_g$, $\mathbf{r}_r$), and path gains ($\{\mathbf{\Lambda}_s, \boldsymbol{\beta}_s\}$) and then reconstruct the RIS-assisted THz channel $\{\mathbf{H}[s]\}$.

2.3.1 Polar-Domain Sparse Mapping

In this step, from the set of quantized angle and distance pairs so-called polar bin $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})$, we generate the polar-domain dictionary matrices $\bar{\mathbf{A}}_{\mathrm{B},s}$ and $\bar{\mathbf{A}}_{\mathrm{R},s}$ and the corresponding sparse path gain vector $\bar{\mathbf{g}}[s]$ for each subcarrier. Using the polar-domain dictionary matrices and the sparse path gain vectors, we reformulate the linear system in (2.21) into a sparse linear system.

Specifically, the polar-domain BS and RIS dictionary matrices $\bar{\mathbf{A}}_{\mathrm{B},s} \in \mathbb{C}^{N\times Q_\theta}$ and $\bar{\mathbf{A}}_{\mathrm{R},s} \in \mathbb{C}^{M\times(Q_\theta Q_r)}$ generated from $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})$ (see Section IV.A for detailed polar bin design) are given by

$$\bar{\mathbf{A}}_{\mathrm{B},s} = [\mathbf{a}_{\mathrm{B},s}(\bar{\theta}_1), \cdots, \mathbf{a}_{\mathrm{B},s}(\bar{\theta}_{Q_\theta})], \tag{2.22}$$

$$\bar{\mathbf{A}}_{\mathrm{R},s} = [\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_1, \bar{r}_1), \cdots, \mathbf{a}_{\mathrm{R},s}(\bar{\theta}_{Q_\theta Q_r}, \bar{r}_{Q_\theta Q_r})], \tag{2.23}$$

where Q_θ and Q_r are the quantization levels of angle and distance, respectively. Using $\bar{\mathbf{A}}_{\mathrm{B},s}$ and $\bar{\mathbf{A}}_{\mathrm{R},s}$, the sparse representation of the near-field RIS-assisted channel matrix $\mathbf{H}[s]$ is given by

$$\mathbf{H}[s] = \bar{\mathbf{A}}_{\mathrm{B},s}(\bar{\mathbf{\Lambda}}_s \otimes \bar{\boldsymbol{\beta}}_s^{\mathrm{T}})(\bar{\mathbf{A}}_{\mathrm{R},s}^{*} \bullet \bar{\mathbf{A}}_{\mathrm{R},s})^{\mathrm{H}}, \tag{2.24}$$

where $\bar{\mathbf{\Lambda}}_s \in \mathbb{C}^{Q_\theta\times(Q_\theta Q_r)^2}$ and $\bar{\boldsymbol{\beta}}_s \in \mathbb{C}^{(Q_\theta Q_r)^2\times 1}$ are the sparse RIS-BS path gain matrix and UE-RIS path gain vector, respectively, such that $\|\bar{\mathbf{\Lambda}}_s\|_0 = P_g$ and $\|\bar{\boldsymbol{\beta}}_s\|_0 = P_r$. Then the vectorized RIS-assisted channel $\mathrm{vec}(\mathbf{H}[s])$ can be expressed as

$$\mathrm{vec}(\mathbf{H}[s]) = \big((\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^{*}) \otimes \bar{\mathbf{A}}_{\mathrm{B},s}\big)\mathrm{vec}(\bar{\mathbf{\Lambda}}_s \otimes \bar{\boldsymbol{\beta}}_s^{\mathrm{T}}). \tag{2.25}$$

Note that $\mathrm{vec}(\mathbf{H}[s])$ includes the term $\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^{*}$, a row-wise Khatri-Rao product of the RIS dictionary matrix $\bar{\mathbf{A}}_{\mathrm{R},s}$ and its conjugate $\bar{\mathbf{A}}_{\mathrm{R},s}^{*}$. Due to the property of the row-wise Khatri-Rao product, $\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^{*}$ contains a large number of duplicated columns[4].

Lemma 2. $\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^{*} \in \mathbb{C}^{M\times(Q_\theta Q_r)^2}$ *in* (2.25) *contains only* $(2Q_\theta-1)(2Q_r-1)$ *distinct columns where* Q_θ *and* Q_r *are the quantization levels of angle and distance* ($Q_\theta Q_r = Q$)*, respectively. When removing the duplicated columns,* $\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^{*}$ *is reduced to* $\mathbf{D}[s] = [\mathbf{d}_{1,1}, \cdots, \mathbf{d}_{2Q_\theta-1,2Q_r-1}] \in \mathbb{C}^{M\times(2Q_\theta-1)(2Q_r-1)}$ *where the* m*-th element of* $\mathbf{d}_{k,l}$ *is given by*

$$[\mathbf{d}_{k,l}]_m = \exp\bigg(j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})\big(\frac{2d(m-1)}{Q}(k-Q_\theta+(l-Q_r)Q_\theta) - \frac{d^2(m-1)^2}{2r_{\min}Q}((k-Q_\theta)Q_r+l-Q_r)\big)\bigg), \tag{2.26}$$

for $m=1,\cdots,M$*,* $k=1,\cdots,2Q_\theta-1$*, and* $l=1,\cdots,2Q_r-1$.

Proof. See Appendix A. □

By removing the duplicated column vectors of $\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^{*}$ and merging the corresponding elements of $\mathrm{vec}(\bar{\mathbf{\Lambda}}_s \otimes \bar{\boldsymbol{\beta}}_s^{\mathrm{T}})$, the vectorized RIS-assisted channel $\mathrm{vec}(\mathbf{H}[s])$ can be re-expressed as[5]

$$\mathrm{vec}(\mathbf{H}[s]) = \big((\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^{*}) \otimes \bar{\mathbf{A}}_{\mathrm{B},s}\big)\mathrm{vec}(\bar{\mathbf{\Lambda}}_s \otimes \bar{\boldsymbol{\beta}}_s^{\mathrm{T}}) \tag{2.27}$$

$$= (\mathbf{D}[s] \otimes \bar{\mathbf{A}}_{\mathrm{B},s})\bar{\mathbf{g}}[s] \tag{2.28}$$

$$= \bar{\mathbf{A}}[s]\bar{\mathbf{g}}[s], \tag{2.29}$$

where $\bar{\mathbf{A}}[s] = \mathbf{D}[s] \otimes \bar{\mathbf{A}}_{\mathrm{B},s} \in \mathbb{C}^{MN\times(2Q_\theta-1)(2Q_r-1)Q_\theta}$ is the total dictionary matrix.

[4]For example, the row-wise Khatri-Rao product of the 2-point DFT matrix $\mathbf{F}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and its conjugate $\mathbf{F}_2^{*}$ is $\mathbf{F}_2 \bullet \mathbf{F}_2^{*} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$ which consists of 2 duplicated column vectors.

[5]Let $\mathbf{y} = \mathbf{A}\mathbf{x}$ be a linear system where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_1]$ and $\mathbf{x} = [x_1, x_2, x_3, x_4]^{\mathrm{T}}$. By merging the duplicated columns of $\mathbf{A}$ and the corresponding elements of $\mathbf{x}$, we obtain $\mathbf{y} = \bar{\mathbf{A}}\bar{\mathbf{x}}$ where $\bar{\mathbf{A}} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$ and $\bar{\mathbf{x}} = [x_1 + x_4, x_2, x_3]^{\mathrm{T}}$.

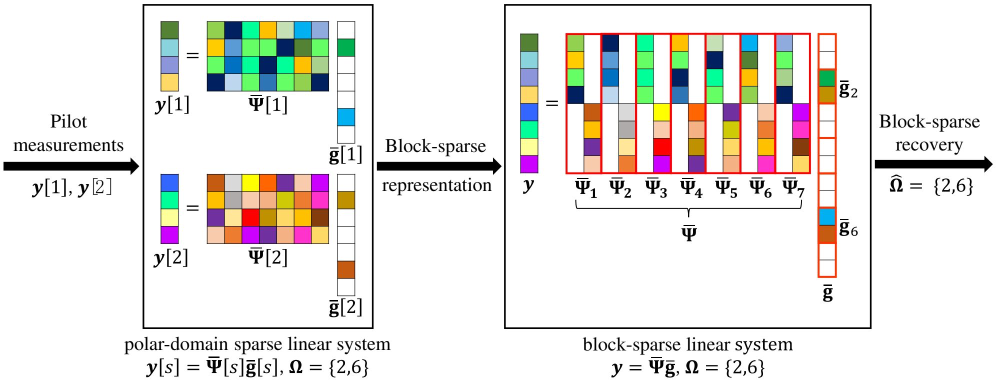

Figure 2.4: Overall description of the proposed PF-RCE algorithm.

By plugging (2.29) into (2.19), we obtain the sparse linear system:

$$\mathbf{y}[s] = ((\mathbf{\Phi}\mathbf{X}[s])^{\mathrm{T}} \otimes \mathbf{W}^{\mathrm{H}})\mathrm{vec}(\mathbf{H}[s]) + \mathbf{n}[s] \tag{2.30}$$

$$= ((\mathbf{\Phi}\mathbf{X}[s])^{\mathrm{T}} \otimes \mathbf{W}^{\mathrm{H}})(\mathbf{D}[s] \otimes \bar{\mathbf{A}}_{\mathrm{B},s})\bar{\mathbf{g}}[s] + \mathbf{n}[s] \tag{2.31}$$

$$= \bar{\mathbf{\Psi}}[s]\bar{\mathbf{g}}[s] + \mathbf{n}[s], \tag{2.32}$$

where $\bar{\mathbf{\Psi}}[s] \in \mathbb{C}^{LP \times Q_{\mathrm{tot}}}$ is the polar-domain sensing matrix given by

$$\bar{\mathbf{\Psi}}[s] = ((\mathbf{\Phi}\mathbf{X}[s])^{\mathrm{T}} \otimes \mathbf{W}^{\mathrm{H}})(\mathbf{D}[s] \otimes \bar{\mathbf{A}}_{\mathrm{B},s}), \tag{2.33}$$

and $\bar{\mathbf{g}}[s] \in \mathbb{C}^{Q_{\mathrm{tot}} \times 1}$ is the combined sparse path gain vector such that $\|\bar{\mathbf{g}}[s]\|_0 = P_g P_r$, and $Q_{\mathrm{tot}} = (2Q_\theta - 1)(2Q_r - 1)Q_\theta$.

2.3.2 Block-Sparse Representation

Since the angles and distances are the same for all subcarriers, the sparse path gain vectors $\{\bar{\mathbf{g}}[s]\}_{s=1}^{S}$ share the common support $\mathbf{\Omega}$. Based on this observation, we combine the i-th elements of $\{\bar{\mathbf{g}}[s]\}_{s=1}^{S}$ into a vector $\bar{\mathbf{g}}_i$. Note that the elements of $\bar{\mathbf{g}}_i$ are either all nonzero ($i \in \mathbf{\Omega}$) or all zero ($i \notin \mathbf{\Omega}$). Thus, by concatenating $\bar{\mathbf{g}}_1, \cdots, \bar{\mathbf{g}}_{Q_{\mathrm{tot}}}$, we can obtain the block-sparse path gain vector $\bar{\mathbf{g}}$ where the non-zero elements appear in a few blocks of the vector (see Fig. 2.4). Using this, we can formulate the block-sparse linear system in the form of $\mathbf{y} = \bar{\mathbf{\Psi}}\bar{\mathbf{g}} + \mathbf{z}$ ($\bar{\mathbf{\Psi}}$ is the block sensing matrix) and then find out $\mathbf{\Omega}$ and $\bar{\mathbf{g}}$ using the block-sparse recovery algorithm. Once $\mathbf{\Omega}$ is recovered, we can extract the angle and distance information from the corresponding column vectors of the dictionary matrix. Using the extracted angles, distances, and path gains, we can reconstruct the vectorized RIS-assisted channel vec$(\mathbf{H}[s])$ from which we acquire the near-field RIS-assisted THz channel $\mathbf{H}[s]$.

To be specific, we first combine the elements of $\{\bar{\mathbf{g}}[s]\}_{s=1}^{S}$ at the i-th position into $\bar{\mathbf{g}}_i \in \mathbb{C}^{S\times 1}$:

$$\bar{\mathbf{g}}_i = [\bar{g}_i[1], \cdots, \bar{g}_i[S]]^{\mathrm{T}}, \quad i = 1, \cdots, Q_{\text{tot}}, \tag{2.34}$$

where $\bar{g}_i[s]$ is the i-th element of $\bar{\mathbf{g}}[s]$. Since $\{\bar{\mathbf{g}}[s]\}_{s=1}^{S}$ have the common support $\mathbf{\Omega}$, the elements of $\bar{\mathbf{g}}_i$ are either all nonzero ($i \in \mathbf{\Omega}$) or all zero ($i \notin \mathbf{\Omega}$). This means that $\bar{\mathbf{g}} = [\bar{\mathbf{g}}_1^{\mathrm{T}}, \cdots, \bar{\mathbf{g}}_{Q_{\text{tot}}}^{\mathrm{T}}]^{\mathrm{T}}$ is a block-sparse vector where the nonzero elements appear in a few $S \times 1$ blocks. Similarly, we combine the i-th column vectors of $\{\bar{\mathbf{\Psi}}[s]\}_{s=1}^{S}$ into a block-diagonal matrix $\bar{\mathbf{\Psi}}_i \in \mathbb{C}^{LPS\times S}$ as

$$\bar{\mathbf{\Psi}}_i = \text{diag}(\bar{\boldsymbol{\psi}}_i[1], \cdots, \bar{\boldsymbol{\psi}}_i[S]), \quad i = 1, \cdots, Q_{\text{tot}}, \tag{2.35}$$

where $\bar{\boldsymbol{\psi}}_i[s] \in \mathbb{C}^{LPS\times 1}$ is the i-th column vector of $\bar{\mathbf{\Psi}}[s]$. Using the combined matrix is $\bar{\mathbf{\Psi}} = [\bar{\mathbf{\Psi}}_1, \cdots, \bar{\mathbf{\Psi}}_{Q_{\text{tot}}}]$ and $\bar{\mathbf{g}}$, we obtain the block-sparse linear system:

$$\mathbf{y} = [\mathbf{y}[1]^{\mathrm{T}}, \cdots, \mathbf{y}[S]^{\mathrm{T}}]^{\mathrm{T}} = \bar{\mathbf{\Psi}}\bar{\mathbf{g}} + \mathbf{n}, \tag{2.36}$$

where $\bar{\mathbf{\Psi}} = [\bar{\mathbf{\Psi}}_1, \cdots, \bar{\mathbf{\Psi}}_{Q_{\text{tot}}}] \in \mathbb{C}^{LPS\times Q_{\text{tot}}S}$ is the block sensing matrix and $\bar{\mathbf{g}} = [\bar{\mathbf{g}}_1^{\mathrm{T}}, \cdots, \bar{\mathbf{g}}_{Q_{\text{tot}}}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{Q_{\text{tot}}S\times 1}$ the block-sparse path gain vector[6]. The corresponding

[6]For example, when the common support is $\mathbf{\Omega} = \{2, 6\}$, $\bar{\mathbf{g}}[1] = [0, 1, 0, 0, 0, 2, 0]^{\mathrm{T}}$, and $\bar{\mathbf{g}}[2] = [0, 3, 0, 0, 0, 4, 0]^{\mathrm{T}}$, then $\bar{\mathbf{g}}_2 = [1, 3]^{\mathrm{T}}$, $\bar{\mathbf{g}}_6 = [2, 4]^{\mathrm{T}}$, and $\{\bar{\mathbf{g}}\}_i = [0, 0]^{\mathrm{T}}$ for $i \notin \mathbf{\Omega}$. In this case, $\bar{\mathbf{g}} = [0, 0, 1, 3, 0, 0, 0, 0, 0, 0, 2, 4, 0, 0]^{\mathrm{T}}$ is a block 2-sparse vector where the nonzero elements appear in 2 of 2×1 blocks.

block-sparse recovery problem to recover the block-sparse path gain vector $\bar{\mathbf{g}}$ is

$$\mathcal{P}_0 : \min_{\bar{\mathbf{g}}=[\bar{\mathbf{g}}_1^{\mathrm{T}},\cdots,\bar{\mathbf{g}}_{Q_{\text{tot}}}^{\mathrm{T}}]^{\mathrm{T}}} \|\mathbf{y} - \bar{\mathbf{\Psi}}\bar{\mathbf{g}}\|_2 \tag{2.37a}$$

$$\text{s.t.} \quad \sum_{i=1}^{Q_{\text{tot}}} \mathcal{I}(\|\bar{\mathbf{g}}_i\|_2) = P_g P_r, \tag{2.37b}$$

where $\mathcal{I}(x)$ is the indicator function such that $\mathcal{I}(x) = 1$ if $x \neq 0$ and $\mathcal{I}(x) = 0$ otherwise. Note that $\sum_{i=1}^{Q_{\text{tot}}} \mathcal{I}(\|\bar{\mathbf{g}}_i\|_2)$ represents the number of non-zero blocks in $\bar{\mathbf{g}}$.

2.3.3 Block-Sparse Recovery and Channel Reconstruction

In solving $\mathcal{P}_0$, one can use the block-sparse recovery algorithm such as block orthogonal least squares (BOLS) [24]. In the BOLS algorithm, an index of the submatrix of the sensing matrix is chosen at a time using a greedy strategy and then the residual is updated. To be specific, in the t-th iteration, an index $\hat{\omega}_t$ corresponding to the submatrix $\bar{\mathbf{\Psi}}_{\hat{\omega}_t}$ of the block sensing matrix $\bar{\mathbf{\Psi}}$ that leads to the most significant reduction in the residual power is chosen as

$$\hat{\omega}_t = \arg \min_{i=1,\cdots,Q_{\text{tot}}} \|\mathbf{P}^{\perp}_{\hat{\mathbf{\Omega}}_{t-1}\cup\{i\}} \mathbf{r}_{t-1}\|_2^2, \quad t = 1, \cdots, P_g P_r, \tag{2.38}$$

where $\hat{\mathbf{\Omega}}_t = \{\hat{\omega}_1, \cdots, \hat{\omega}_t\}$ and $\mathbf{r}_{t-1} = \mathbf{P}^{\perp}_{\hat{\mathbf{\Omega}}_{t-1}} \mathbf{y}$ is the residual. Also, $\mathbf{P}^{\perp}_{\hat{\mathbf{\Omega}}_t} = \mathbf{I}_{LPS} - \mathbf{P}_{\hat{\mathbf{\Omega}}_t}$ is the orthogonal complement of $\mathbf{P}_{\hat{\mathbf{\Omega}}_t}$ where $\mathbf{P}_{\hat{\mathbf{\Omega}}_t} = \bar{\mathbf{\Psi}}_{\hat{\mathbf{\Omega}}_t} \bar{\mathbf{\Psi}}^{\dagger}_{\hat{\mathbf{\Omega}}_t}$ is the orthogonal projection onto span$(\bar{\mathbf{\Psi}}_{\hat{\mathbf{\Omega}}_t})$ and $\bar{\mathbf{\Psi}}_{\hat{\mathbf{\Omega}}_t} = [\bar{\mathbf{\Psi}}_{\hat{\omega}_1}, \cdots, \bar{\mathbf{\Psi}}_{\hat{\omega}_t}]$. It is worth mentioning that since each $\bar{\mathbf{\Psi}}_{\hat{\omega}_t} = \text{diag}(\bar{\boldsymbol{\psi}}_{\hat{\omega}_t}[1], \cdots, \bar{\boldsymbol{\psi}}_{\hat{\omega}_t}[S])$ is a block-diagonal matrix, the orthogonal projection matrix $\mathbf{P}_{\hat{\mathbf{\Omega}}_t} \in \mathbb{C}^{LPS \times LPS}$ is also a block-diagonal matrix whose diagonal elements are the orthogonal project matrices $\{\mathbf{P}_{\hat{\mathbf{\Omega}}_t}[s]\}_{s=1}^{S}$ for S subcarriers:

$$\mathbf{P}_{\hat{\mathbf{\Omega}}_t} = \text{diag}(\mathbf{P}_{\hat{\mathbf{\Omega}}_t}[1], \cdots, \mathbf{P}_{\hat{\mathbf{\Omega}}_t}[S]), \tag{2.39}$$

where $\mathbf{P}_{\hat{\mathbf{\Omega}}_t}[s] = (\bar{\mathbf{\Psi}}[s])_{\hat{\mathbf{\Omega}}_t} (\bar{\mathbf{\Psi}}[s])^{\dagger}_{\hat{\mathbf{\Omega}}_t} \in \mathbb{C}^{LP \times LP}$ and $(\bar{\mathbf{\Psi}}[s])_{\hat{\mathbf{\Omega}}_t} = \left[\bar{\boldsymbol{\psi}}_{\hat{\omega}_1}[s], \cdots, \bar{\boldsymbol{\psi}}_{\hat{\omega}_t}[s]\right]$. Using (2.39), we can significantly reduce the computational complexity of PF-RCE since we only need to compute the low-dimensional projection matrices $\{\mathbf{P}_{\hat{\mathbf{\Omega}}_t}[s]\}_{s=1}^{S} \subseteq$

Table 2.1: Comparison of computational complexity.

	Computational Complexity
Proposed PF-RCE	$\mathcal{O}(SL^2P^2P_g^2P_r^2(2Q_\theta-1)(2Q_r-1))$
Conventional NB-CE	$\mathcal{O}(SL^2P^2P_g^2P_r^2(Q_\theta Q_r)^2)$
Conventional FF-CE	$\mathcal{O}(SL^2P^2P_g^2P_r^2Q_\theta^2)$

$\mathbb{C}^{LP\times LP}$ instead of the high-dimensional projection matrix $\mathbf{P}_{\hat{\mathbf{\Omega}}_t}\in\mathbb{C}^{LPS\times LPS}$. Also, using (2.39), one can re-express (2.38) as

$$\hat{\omega}_t=\arg\min_{i=1,\cdots,Q_{\text{tot}}}\sum_{s=1}^{S}\|(\mathbf{P}_{\hat{\mathbf{\Omega}}_{t-1}\cup\{i\}}[s])^{\perp}\mathbf{r}_{t-1}[s]\|_2^2, \tag{2.40}$$

where $t=1,\cdots,P_gP_r$ is the iteration index. The iteration is repeated until P_gP_r indices are selected.

Once we obtain the support $\hat{\mathbf{\Omega}}=\{\hat{\omega}_1,\cdots,\hat{\omega}_{P_gP_r}\}$, we can acquire the estimate of $\{\bar{\mathbf{g}}[s]\}$ as

$$(\bar{\mathbf{g}}^*[s])_{\hat{\mathbf{\Omega}}}=(\bar{\mathbf{\Psi}}[s])^{\dagger}_{\hat{\mathbf{\Omega}}}\mathbf{y}[s]. \tag{2.41}$$

Since each quantized angle and distance pair of the polar bin corresponds to the column vector of the dictionary matrix $\mathbf{D}[s]\otimes\bar{\mathbf{A}}_{\text{B},s}$, we can extract the angle and distance information from $\big(\mathbf{D}[s]\otimes\bar{\mathbf{A}}_{\text{B},s}\big)_{\mathbf{\Omega}}$. Using the extracted angles, distances, and path gains, we can reconstruct the vectorized near-field RIS-assisted THz channel (see (2.29)):

$$\text{vec}(\mathbf{H}^*[s])=\big(\mathbf{D}[s]\otimes\bar{\mathbf{A}}_{\text{B},s}\big)_{\hat{\mathbf{\Omega}}}(\bar{\mathbf{g}}^*[s])_{\hat{\mathbf{\Omega}}}. \tag{2.42}$$

Finally, one can acquire the near-field RIS-assisted THz channel $\mathbf{H}^*[s]$ from $\text{vec}(\mathbf{H}^*[s])$.

2.3.4 Computational Complexity Analysis

In this subsection, we analyze the computational complexity of the proposed PF-RCE algorithm. Specifically, the block-sparse recovery process in PF-RCE consists of four major steps: 1) calculating the orthogonal complements $\{(\mathbf{P}_{\hat{\mathbf{\Omega}}_{t-1}\cup\{i\}}[s])^{\perp}\}_{s=1}^{S}$ of the projection matrices, 2) calculating the residual power $\sum_{s=1}^{S}\|(\mathbf{P}_{\hat{\mathbf{\Omega}}_{t-1}\cup\{i\}}[s])^{\perp}\mathbf{r}_{t-1}[s]\|_2^2$ and finding out the index $\hat{\omega}_t$ minimizing the residual power, 3) updating the residual $\mathbf{r}_t$, and 4) estimating the block-sparse path gain vectors $\{(\bar{\mathbf{g}}^*[s])_{\hat{\mathbf{\Omega}}}\}_{s=1}^{S}$. The overall complexity $\mathcal{C}^{\text{PF-RCE}}$ of PF-RCE is expressed as

$$\mathcal{C}^{\text{PF-RCE}} = P_g P_r(\mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3) + \mathcal{C}_4, \tag{2.43}$$

where $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4$ are the computational complexities of the aforementioned steps. Note that $P_g P_r$ are multiplied at $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ since the block-sparse recovery process consists of $P_g P_r$ iterations. In the following lemma, we provide the overall complexity $\mathcal{C}^{\text{PF-RCE}}$ of PF-RCE.

Lemma 3. *The total computational complexity $\mathcal{C}^{\text{PF-RCE}}$ is given by*

$$\mathcal{C}^{\text{PF-RCE}} = \mathcal{O}(SL^2P^2P_g^2P_r^2(2Q_\theta - 1)(2Q_r - 1)), \tag{2.44}$$

where Q_θ and Q_r are the quantization levels of angle and distance, respectively.

Proof. See Appendix B. □

For comparison, we also discuss the complexities of the conventional narrow-band channel estimation (NB-CE) scheme that estimates the polar-domain channel parameters (i.e., angles, distances, and path gains) of each subcarrier channel separately [1] and the far-field channel estimation (FF-CE) scheme that estimates the angular-domain channel parameters (i.e., angles and path gains) of all subcarrier channels simultaneously [20]. It is clear from Table I that the computational complexity of PF-RCE is lower than that of NB-CE. This is because by exploiting (2.39), the residual power calculation of PF-RCE can be decomposed into S sub-calculations, resulting in a significant reduction of computational complexity. Also, while NB-CE uses the high-dimensional sensing matrix containing a large number of duplicated columns, PF-RCE uses the low-dimensional sensing matrix where the duplicated columns are removed (see Lemma 2). Also, we see that the computational complexity of PF-RCE is higher than that of

FF-CE, since PF-RCE estimates the whole multipath components (angles, distances, and path gains) whereas FF-CE estimates only the angles and path gains.

2.4 Practical Issues in Near-Field RIS-Assisted THz Channel Estimation

In this section, we discuss two practical issues related to RIS-assisted THz channel estimation. We first describe the polar bin design and then discuss the RIS phase shift control.

2.4.1 Polar Bin Design for RIS-assisted THz Channel Estimation

Note that the sparse recovery performance of PF-RCE relies heavily on the column correlation of the polar-domain sensing matrix $\bar{\mathbf{\Psi}}[s]$. Since $\bar{\mathbf{\Psi}}[s]$ is generated from the total dictionary matrix $\bar{\mathbf{A}}[s]$, the column correlation of $\bar{\mathbf{\Psi}}[s]$ is determined by that of $\bar{\mathbf{A}}[s]$. As shown in Lemma 2, $\bar{\mathbf{A}}[s]$ is a function of the polar bin, a set of quantized angle and distance pairs $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})$. So, by deliberately designing $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})$ such that the column correlation of $\bar{\mathbf{A}}[s]$ is minimized, we can reduce the column correlation of $\bar{\mathbf{\Psi}}[s]$, thereby improving the sparse recovery performance.

Recall that $\bar{\mathbf{A}}[s]$ can be expressed as $\bar{\mathbf{A}}[s] = \mathbf{D}[s] \otimes \bar{\mathbf{A}}_{\mathrm{B},s}$ (see (2.29)) where $\mathbf{D}[s]$ is the polar-domain RIS dictionary matrix (see Lemma 2) and $\bar{\mathbf{A}}_{\mathrm{B},s}$ is the angular-domain BS dictionary matrix (see (2.22)). Thus, the column correlation of $\bar{\mathbf{A}}[s]$ is the multiplication of the column correlations of $\mathbf{D}[s]$ and $\bar{\mathbf{A}}_{\mathrm{B},s}$ (i.e., $\mu(\bar{\mathbf{A}}[s]) = \mu(\mathbf{D}[s])\mu(\bar{\mathbf{A}}_{\mathrm{B},s})$):

$$\mu(\bar{\mathbf{A}}[s]) = \max_{(i,j)\neq(k,l)} |\bar{\mathbf{a}}_{i,j}^{\mathrm{H}} \bar{\mathbf{a}}_{k,l}|^2 \tag{2.45}$$

$$= \max_{(i,j)\neq(k,l)} |(\mathbf{d}_i \otimes \bar{\mathbf{a}}_{\mathrm{B},j})^{\mathrm{H}} (\mathbf{d}_k \otimes \bar{\mathbf{a}}_{\mathrm{B},l})|^2 \tag{2.46}$$

$$= \max_{i\neq k} |\mathbf{d}_i^{\mathrm{H}} \mathbf{d}_k|^2 \max_{j\neq l} |\bar{\mathbf{a}}_{\mathrm{B},j}^{\mathrm{H}} \bar{\mathbf{a}}_{\mathrm{B},l}|^2 \tag{2.47}$$

$$= \mu(\mathbf{D}[s])\mu(\bar{\mathbf{A}}_{\mathrm{B},s}), \tag{2.48}$$

where $\bar{\mathbf{a}}_{i,j}$, $\mathbf{d}_i$, and $\bar{\mathbf{a}}_{\mathrm{B},j}$ are the column vectors of $\bar{\mathbf{A}}[s]$, $\mathbf{D}[s]$, and $\bar{\mathbf{A}}_{\mathrm{B},s}$, respectively. To minimize $\mu(\bar{\mathbf{A}}[s])$, we need to minimize $\mu(\mathbf{D}[s])$ and $\mu(\bar{\mathbf{A}}_{\mathrm{B},s})$. Since $\bar{\mathbf{A}}_{\mathrm{B},s}$ is generated from the angular bin $\bar{\boldsymbol{\theta}}$ and the optimal angular bin minimizing $\mu(\bar{\mathbf{A}}_{\mathrm{B},s})$ can be easily obtained by uniformly discretizing $\sin\bar{\boldsymbol{\theta}}$ in $[-1,1)$, we only need to find out the optimal polar bin minimizing the column correlation of $\mathbf{D}[s]$.

To be specific, $\mu(\mathbf{D}[s])$ can be expressed as a function of polar bin $(\bar{\boldsymbol{\theta}},\bar{\mathbf{r}})$ as

$$\mu(\mathbf{D}[s]) = \max_{i\neq j} f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j), \tag{2.49}$$

where $f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j) = |\frac{1}{M}\mathbf{a}^{\mathrm{H}}_{\mathrm{R},s}(\bar{\theta}_i,\bar{r}_i)\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_j,\bar{r}_j)|^2$ and $\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_i,\bar{r}_i)$ is the near-field array response vector in (2.12). Then the optimization problem $\mathcal{P}_{\text{bin}}$ to find out the polar bin $(\bar{\boldsymbol{\theta}},\bar{\mathbf{r}})$ minimizing $\mu(\mathbf{D}[s])$ is formulated as

$$\mathcal{P}_{\text{bin}}: \min_{(\bar{\boldsymbol{\theta}},\bar{\mathbf{r}})} \max_{i\neq j} f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j), \tag{2.50a}$$

$$\text{s.t. } 0 \le \bar{\theta}_i < 2\pi, r_{\min} \le \bar{r}_i \le r_{\max}, \quad i = 1,\cdots,Q, \tag{2.50b}$$

where $[r_{\min}, r_{\max}]$ is the range of communication distance and $Q = Q_\theta Q_r$ is the number of quantized angle and distance pairs in $(\bar{\boldsymbol{\theta}},\bar{\mathbf{r}})$. Since $\mu(\mathbf{D}[s])$ is a nonlinear function of $(\bar{\boldsymbol{\theta}},\bar{\mathbf{r}})$, it is not easy to find out the optimal solution of $\mathcal{P}_{\text{bin}}$. As a remedy, we exploit the observation that the m-th element of $\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_i,\bar{r}_i)$ consists of the far-field term $d(m-1)\sin\bar{\theta}_i$ and the near-field term $\frac{d^2(m-1)^2}{2}\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}$ (see (2.9)). Based on this observation, we express $f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j)$ as a function of $|\sin\bar{\theta}_i - \sin\bar{\theta}_j|$ and $\left|\frac{\cos^2\bar{\theta}_i}{\bar{r}_i} - \frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\right|$ and then convert the column correlation minimization problem to the problem to maximize the sum of these two terms.

Proposition 1. *The normalized correlation* $f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j) = |\frac{1}{M}\mathbf{a}^{\mathrm{H}}_{\mathrm{R},s}(\bar{\theta}_i,\bar{r}_i)\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_j,\bar{r}_j)|^2$ *between the near-field array response vectors* $\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_i,\bar{r}_i)$ *and* $\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_j,\bar{r}_j)$ *can be approximated as a function of* $\alpha = \sin\bar{\theta}_i - \sin\bar{\theta}_j$ *and* $\beta = \frac{\cos^2\bar{\theta}_i}{\bar{r}_i} - \frac{\cos^2\bar{\theta}_j}{\bar{r}_j}$ *:*

$$\begin{aligned} f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j) \approx\ & \frac{1}{M^2 d\beta}\Big(C\Big(M\sqrt{d\beta}-\frac{\alpha}{\sqrt{d\beta}}\Big)+C\Big(\frac{\alpha}{\sqrt{d\beta}}\Big)\Big)^2 \\ & +\frac{1}{M^2 d\beta}\Big(S\Big(M\sqrt{d\beta}-\frac{\alpha}{\sqrt{d\beta}}\Big)+S\Big(\frac{\alpha}{\sqrt{d\beta}}\Big)\Big)^2, \end{aligned} \tag{2.51}$$

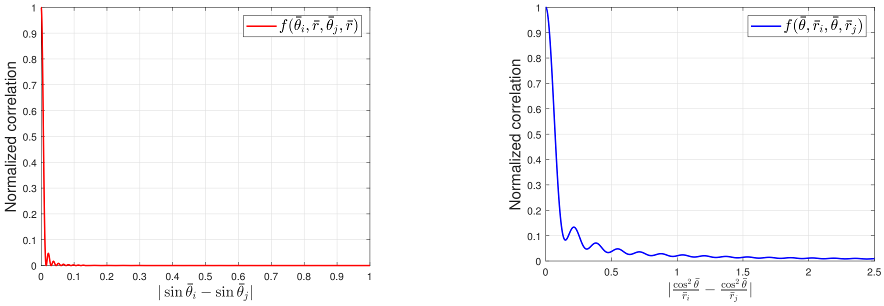

Figure 2.5: Correlation between the near-field array response vectors

where $C(x) = \int_0^x \cos(\frac{1}{2}\pi t^2)dt$ *and* $S(x) = \int_0^x \sin(\frac{1}{2}\pi t^2)dt$ *are the Fresnel integrals.*

Proof. See Appendix C. $\square$

As shown in Fig. 2.5, $f(\bar{\theta}_i, \bar{r}_i, \bar{\theta}_j, \bar{r}_j)$ can be approximated to the decreasing function of $|\sin\bar{\theta}_i - \sin\bar{\theta}_j|$ and $\left|\frac{\cos^2\bar{\theta}_i}{\bar{r}_i} - \frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\right|$. Thus, the problem to minimize $\max_{i\neq j} f(\bar{\theta}_i, \bar{r}_i, \bar{\theta}_j, \bar{r}_j)$ can be converted to the problem to maximize the sum of $\min_{i\neq j} |\sin\bar{\theta}_i - \sin\bar{\theta}_j|$ and $\min_{i\neq j} \left|\frac{\cos^2\bar{\theta}_i}{\bar{r}_i} - \frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\right|$:

$$\mathcal{P}'_{\text{bin}} : \max_{(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})} \left(\min_{i\neq j} |\sin\bar{\theta}_i - \sin\bar{\theta}_j|^2 + c \min_{i\neq j} \left|\frac{\cos^2\bar{\theta}_i}{\bar{r}_i} - \frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\right|^2 \right) \tag{2.52a}$$

$$\text{s.t. } 0 \leq \bar{\theta}_i < 2\pi, r_{\text{min}} \leq \bar{r}_i \leq r_{\text{max}}, \qquad i = 1, \cdots, Q, \tag{2.52b}$$

where $c > 0$ is the regularization parameter. Since $\bar{\boldsymbol{\theta}}$ and $\bar{\mathbf{r}}$ are coupled with each other in the objective function, it is very difficult to optimize them jointly. To find out a tractable solution of $\mathcal{P}'_{\text{bin}}$, we optimize $\bar{\boldsymbol{\theta}}$ and $\bar{\mathbf{r}}$ in an alternating fashion.

First, when $\bar{\boldsymbol{\theta}}$ is fixed, $\mathcal{P}'_{\text{bin}}$ is reduced to the distance quantization problem $\mathcal{P}_{\text{dis}}$:

$$\mathcal{P}_{\text{dis}} : \max_{\bar{\mathbf{r}}} \min_{i\neq j} \left|\frac{\cos^2\bar{\theta}_i}{\bar{r}_i} - \frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\right| \tag{2.53a}$$

$$\text{s.t. } r_{\text{min}} \leq \bar{r}_i \leq r_{\text{max}}, \quad i = 1, \cdots, Q. \tag{2.53b}$$

The optimality condition for $\mathcal{P}_{\text{dis}}$ is provided in the following lemma.

Lemma 4. *Given a bounded function* $f : \mathcal{X} \to [f_{\text{min}}, f_{\text{max}}]$ *and set of bounded non-negative weights* $\{c_i\}_{i=1}^{Q} \subseteq [c_{\text{min}}, c_{\text{max}}]$, *the optimal solution of* $\mathcal{P} : \max_{\{x_i\} \subseteq \mathcal{X}} \min_{i \neq j} |c_i f(x_i) - c_j f(x_j)|$ *is*

$$x_i^* = f^{-1}\left(\frac{1}{c_i}\left(c_{\text{min}} f_{\text{min}} + i\frac{c_{\text{max}} f_{\text{max}} - c_{\text{min}} f_{\text{min}}}{Q}\right)\right), \quad i = 1, \cdots, Q. \tag{2.54}$$

Proof. See Appendix D. □

Remark 1. *One can see that from Lemma 4 that the optimal solution* $\{x_i^*\}$ *of the problem* $\mathcal{P}$ *satisfies two properties: 1)* $\{c_i f(x_1^*)\}_{i=1}^{Q}$ *are all distinct values and 2)* $\{c_i f(x_i^*)\}_{i=1}^{Q}$ *are uniformly quantized in* $[c_{\text{min}} f_{\text{min}}, c_{\text{max}} f_{\text{max}}]$.

By plugging $f(x) = \frac{1}{x} \in \left[0, \frac{1}{r_{\text{min}}}\right]$ and $c_i = \cos^2 \bar{\theta}_i \in [0, 1]$ to (2.54), we obtain the optimal solution $\bar{\mathbf{r}}$ of $\mathcal{P}_{\text{dis}}$ as

$$\bar{r}_i = \frac{r_{\text{min}} Q}{i} \cos^2 \bar{\theta}_i, \quad i = 1, \cdots, Q. \tag{2.55}$$

Once $\bar{\mathbf{r}}$ is obtained, the second term of the objective function in $\mathcal{P}'_{\text{bin}}$ is fixed to $\min_{i \neq j} \left|\frac{\cos^2 \bar{\theta}_i}{\bar{r}_i} - \frac{\cos^2 \bar{\theta}_j}{\bar{r}_j}\right| = \frac{1}{r_{\text{min}} Q}$. Then $\mathcal{P}'_{\text{bin}}$ is reduced to the angle quantization problem $\mathcal{P}_{\text{ang}}$:

$$\mathcal{P}_{\text{ang}} : \max_{\bar{\boldsymbol{\theta}}} \min_{i \neq j} |\sin \bar{\theta}_i - \sin \bar{\theta}_j| \tag{2.56a}$$

$$\text{s.t. } 0 \leq \bar{\theta}_i < 2\pi, \quad i = 1, \cdots, Q. \tag{2.56b}$$

Similarly, by plugging $f(x) = \sin(x) \in [-1, 1]$ and $c_i = 1$ to (2.54), one can obtain the optimal solution $\bar{\boldsymbol{\theta}}$ of $\mathcal{P}_{\text{ang}}$ as

$$\bar{\theta}_{(i-1)Q_r + j} = \arcsin\Big(-1 + \frac{2}{Q}(i - 1 + (j-1)Q_\theta)\Big), \; i = 1, \cdots, Q_\theta, \; j = 1, \cdots, Q_r, \tag{2.57}$$

where Q_θ and Q_r are the quantization levels of angle and distance ($Q = Q_\theta Q_r$), respectively.

In Fig. 2.6, we plot $\{(\sin \bar{\theta}_i, \frac{\cos^2 \bar{\theta}_i}{\bar{r}_i}) \mid i = 1, \cdots, Q\}$ for the polar bin $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})$ generated by the proposed quantization method (see (2.55) and (2.57)) and the conventional

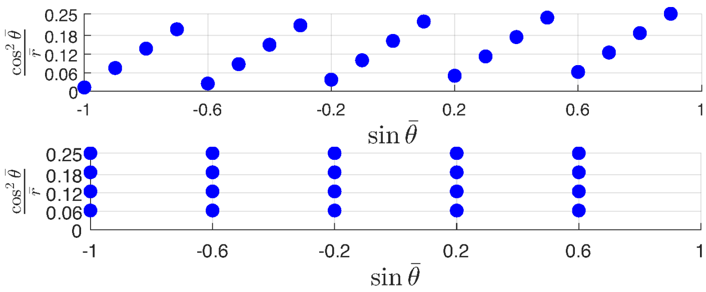

Figure 2.6: Comparison of polar bins generated by (a) the proposed method (see (2.55) and (2.57)) and (b) the conventional method in [1] when $Q_\theta = 5$, $Q_r = 4$, and $r_{\text{min}} = 4$. One can see that the minimum distance d_{min} between the points of the proposed method ($d_{\text{min}} = 0.1179$) is much larger than that of the conventional method ($d_{\text{min}} = 0.0625$).

method in [1]. Note that the conventional polar bin $(\{\bar{\theta}_i\}_{i=1}^{Q}, \{\bar{r}\}_{i=1}^{Q})$ are simply chosen from the uniform grid points of $[-1, 1] \times [0, \frac{1}{r_{\text{min}}}]$ ($[-1, 1]$ and $[0, \frac{1}{r_{\text{min}}}]$ are the ranges of $\sin\bar{\theta}_i$ and $\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}$), which may result in some elements sharing the same $\sin\bar{\theta}_i$ or $\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}$ value (see Fig. 2.6). In contrast, based on Lemma 4, the proposed polar bin design ensures that $\sin\bar{\theta}_i$ and $\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}$ values of each $(\bar{\theta}_i, \bar{r}_i)$ are not overlapped. This results in the inclined pattern observed in Fig. 2.6. One can also observe that the minimum distance between the points of the proposed scheme is much larger than that of the conventional method. Recall that the normalized correlation between the near-field array steering vectors $f(\bar{\theta}_i, \bar{r}_i, \bar{\theta}_j, \bar{r}_j)$ can be approximated to the decreasing function of the distance between the points. Hence, the column correlation $\mu(\mathbf{D}[s]) = \max_{i\neq j} f(\bar{\theta}_i, \bar{r}_i, \bar{\theta}_j, \bar{r}_j)$ of the proposed scheme is smaller than that of the conventional method, meaning that the sparse recovery performance can be improved via the proposed polar bin generation method.

2.4.2 Phase Shift Control for RIS-assisted THz Channel Estimation

In the proposed PF-RCE scheme, we convert the near-field RIS-assisted THz channel estimation problem to the block-sparse recovery problem and then solve it using the BOLS algorithm. From (2.33), one can see that the polar-domain sensing matrix of the block-sparse recovery problem is a function of the RIS phase shift matrix. Since the block-sparse recovery performance depends heavily on the correlations between the sub-matrices of the block sensing matrix (i.e., block-mutual coherence), the phase shift matrix should be properly designed such that the block-mutual coherence of the sensing matrix is minimized.

To be specific, the block-mutual coherence μ_b of $\bar{\mathbf{\Psi}} = [\bar{\mathbf{\Psi}}_1 \cdots \bar{\mathbf{\Psi}}_{Q_{\text{tot}}}]$ representing the maximum correlation among the sub-matrices $\{\bar{\mathbf{\Psi}}_i\}_{i=1}^{Q_{\text{tot}}}$ is defined as

$$\mu_b(\mathbf{\Phi}) = \max_{1 \leq i < j \leq Q_{\text{tot}}} \frac{\|\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})\|_2}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi})\|_2 \|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})\|_2}. \tag{2.58}$$

Then the block-mutual coherence minimization problem $\mathcal{P}_1$ to find out the optimal phase shift matrix $\mathbf{\Phi}$ minimizing $\mu_b(\mathbf{\Phi})$ is formulated as

$$\mathcal{P}_1 : \min_{\mathbf{\Phi} \in \mathcal{M}} \ \mu_b(\mathbf{\Phi}) \tag{2.59a}$$

$$\text{s.t.} \ \left|[\mathbf{\Phi}]_{m,p}\right| = 1, \quad m = 1, \cdots, M, \quad p = 1, \cdots, P. \tag{2.59b}$$

By defining the auxiliary matrix $\mathbf{\Sigma}_{i,j} = \frac{\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi})\|_2 \|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})\|_2}$, we obtain $\mu_b = \max_{1 \leq i < j \leq Q_{\text{tot}}} \|\mathbf{\Sigma}_{i,j}\|_2$ so that $\mathcal{P}_1$ is reformulated as

$$\mathcal{P}_2 : \min_{\mathbf{\Sigma}, \mathbf{\Phi}} \ \max_{1 \leq i < j \leq Q_{\text{tot}}} \|\mathbf{\Sigma}_{i,j}\|_2 \tag{2.60a}$$

$$\text{s.t.} \ \mathbf{\Sigma}_{i,j} = \frac{\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi})\|_2 \|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})\|_2}, \quad 1 \leq i < j \leq Q_{\text{tot}}, \tag{2.60b}$$

$$\left|[\mathbf{\Phi}]_{m,p}\right| = 1, \quad m = 1, \cdots, M, \quad p = 1, \cdots, P, \tag{2.60c}$$

where $\mathbf{\Sigma} = [\mathbf{\Sigma}_{1,2}, \cdots, \mathbf{\Sigma}_{Q_{\text{tot}}-1, Q_{\text{tot}}}]$ is the block-mutual coherence matrix. Due to the quadratic fractional structure of (2.60b) and the unit-modulus constraints (2.60c), $\mathcal{P}_2$ is modeled as a non-convex problem where the global optimal solution is very difficult to find. Also, since $\mathbf{\Sigma}$ and $\mathbf{\Phi}$ are coupled with each other in (2.60b), it is not easy to optimize them simultaneously.

To make the problem tractable, we employ the augmented Lagrangian relaxation technique that adds a quadratic penalty term to the objective function. To be specific, the modified objective function L, so-called the augmented Lagrangian, is defined as

$$
\begin{aligned}
&L(\mathbf{\Sigma}, \mathbf{\Phi}, \mathbf{\Lambda}) \\
&= \max_{1\le i<j\le Q_{\text{tot}}} \|\mathbf{\Sigma}_{i,j}\|_2 + \mathbb{1}_{\mathcal{M}}(\mathbf{\Phi}) \\
&\quad + 2 \sum_{1\le i<j\le Q_{\text{tot}}} \operatorname{Re}\left\{\operatorname{tr}\left(\mathbf{\Lambda}_{i,j}^{\mathrm{H}}\left(\mathbf{\Sigma}_{i,j} - \frac{\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi})\|_2 \|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})\|_2}\right)\right)\right\} \\
&\quad + \frac{\rho}{2} \sum_{1\le i<j\le Q_{\text{tot}}} \left\| \mathbf{\Sigma}_{i,j} - \frac{\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi})\|_2 \|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})\|_2} \right\|_{\mathrm{F}}^2,
\end{aligned}
\tag{2.61}
$$

where $\mathbb{1}_{\mathcal{M}}(\cdot)$ is the indicator function, $\mathcal{M} = \{\mathbf{\Phi} \in \mathbb{C}^{M\times P} : \left|[\mathbf{\Phi}]_{m,p}\right| = 1, \forall m, \forall p\}$ is the manifold of the phase shift matrix $\mathbf{\Phi}$ satisfying the unit-modulus constraints (2.60c), $\mathbf{\Lambda} = [\mathbf{\Lambda}_{1,2}, \cdots, \mathbf{\Lambda}_{Q_{\text{tot}}-1, Q_{\text{tot}}}]$ is the Lagrangian multiplier matrix, and $\rho > 0$ is the scaling factor. Using $L(\mathbf{\Sigma}, \mathbf{\Phi}, \mathbf{\Lambda})$, the dual problem $\mathcal{P}_3$ can be expressed as

$$
\mathcal{P}_3 : \max_{\mathbf{\Lambda}} \min_{\mathbf{\Sigma}, \mathbf{\Phi}} L(\mathbf{\Sigma}, \mathbf{\Phi}, \mathbf{\Lambda}). \tag{2.62}
$$

It is worth noting that $\mathcal{P}_3$ is an unconstrained problem, and thus it is much easier to handle than the primary problem $\mathcal{P}_2$. In fact, by exploiting the weak duality[7], the optimal value of $\mathcal{P}_3$ serves as a lower bound of the optimal value of $\mathcal{P}_2$. Unfortunately, it is still not easy to solve $\mathcal{P}_3$ since the augmented Lagrangian L is a joint function of $\mathbf{\Sigma}$, $\mathbf{\Phi}$, and $\mathbf{\Lambda}$. So, we solve the problem by alternately updating the block-mutual coherence matrix $\mathbf{\Sigma}$, the phase shift matrix $\mathbf{\Phi}$, and the Lagrangian multiplier matrix $\mathbf{\Lambda}$:

$$
\mathbf{\Sigma}^{(t+1)} = \arg\min_{\mathbf{\Sigma}} L(\mathbf{\Sigma}, \mathbf{\Phi}^{(t)}, \mathbf{\Lambda}^{(t)}), \tag{2.63}
$$

$$
\mathbf{\Phi}^{(t+1)} = \arg\min_{\mathbf{\Phi}\in\mathcal{M}} L(\mathbf{\Sigma}^{(t+1)}, \mathbf{\Phi}, \mathbf{\Lambda}^{(t)}), \tag{2.64}
$$

$$
\mathbf{\Lambda}^{(t+1)} = \mathbf{\Lambda}^{(t)} + \rho\left(\mathbf{\Sigma}_{i,j}^{(t+1)} - \frac{\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi}^{(t+1)})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi}^{(t+1)})}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi}^{(t+1)})\|_2 \|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi}^{(t+1)})\|_2}\right). \tag{2.65}
$$

First, the optimization problem $\mathcal{P}_{\mathbf{\Sigma}}$ corresponding to (2.63) is given by

$$
\mathcal{P}_{\mathbf{\Sigma}} : \min_{\mathbf{\Sigma}} \max_{1\le i<j\le Q_{\text{tot}}} \|\mathbf{\Sigma}_{i,j}\|_2 + \frac{\rho}{2} \sum_{1\le i<j\le Q_{\text{tot}}} \|\mathbf{\Sigma}_{i,j} - \mathbf{Z}_{i,j}^{(t)}\|_{\mathrm{F}}^2, \tag{2.66}
$$

[7]Note that the primary problem can be rewritten as $\mathcal{P}_2 = \min_{\mathbf{\Sigma},\mathbf{\Phi}} \max_{\mathbf{\Lambda}} L(\mathbf{\Sigma}, \mathbf{\Phi}, \mathbf{\Lambda})$. By using the max-min inequality such that $\max_x \min_y f(x, y) \le \min_y \max_x f(x, y)$ for any function $f(x, y)$, the weak duality is obtained.

where $\mathbf{Z}_{i,j}^{(t)} = \frac{\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi}^{(t)})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi}^{(t)})}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi}^{(t)})\|_2\|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi}^{(t)})\|_2} - \frac{\mathbf{\Lambda}_{i,j}^{(t)}}{\rho}$. Note that $\mathcal{P}_{\mathbf{\Sigma}}$ is a matrix ℓ_2-norm minimization problem which can be equivalently converted to the convex semidefinite program (SDP). By using the convex optimization tool (e.g., SDPT3 [25]), we can obtain the global optimal solution $\mathbf{\Sigma}^*$.

Second, the optimization problem $\mathcal{P}_{\mathbf{\Phi}}$ corresponding to (2.64) is given by

$$\mathcal{P}_{\mathbf{\Phi}} : \min_{\mathbf{\Phi}} \sum_{1\le i<j\le Q_{\text{tot}}} \left\| \frac{\bar{\mathbf{\Psi}}_i^{\mathrm{H}}(\mathbf{\Phi})\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})}{\|\bar{\mathbf{\Psi}}_i(\mathbf{\Phi})\|_2\|\bar{\mathbf{\Psi}}_j(\mathbf{\Phi})\|_2} - \mathbf{W}_{i,j}^{(t)} \right\|_{\mathrm{F}}^2 \tag{2.67a}$$

$$\text{s.t.} \quad \left|[\mathbf{\Phi}]_{m,p}\right| = 1, \quad m = 1, \cdots, M, \quad p = 1, \cdots, P, \tag{2.67b}$$

where $\mathbf{W}_{i,j}^{(t)} = \mathbf{\Sigma}_{i,j}^{(t+1)} + \frac{\mathbf{\Lambda}_{i,j}^{(t)}}{\rho}$. One major obstacle in solving $\mathcal{P}_{\mathbf{\Phi}}$ is the non-convex unit-modulus constraints (2.67b). To handle this issue, one can exploit the property that the set of unit-modulus phase shift matrices $\mathcal{M} = \{\mathbf{\Phi} \in \mathbb{C}^{M\times P} : \left|[\mathbf{\Phi}]_{m,p}\right| = 1, \forall m, \forall p\}$ has a smooth Riemannian manifold structure[8] [26]. Using this property, we can convert $\mathcal{P}_{\mathbf{\Phi}}$ to an unconstrained optimization problem on the Riemannian manifold [27]. Since the optimization over the Riemannian manifold is conceptually analogous to that in the Euclidean space, optimization tools of Euclidean space (e.g., conjugate gradient method) can be readily employed to solve the problem on the Riemannian manifold (e.g., Riemannian conjugate gradient (RCG) method)[9] [26, 28].

After updating $\mathbf{\Sigma}$ and $\mathbf{\Phi}$, as shown in (2.65), we update $\mathbf{\Lambda}$ using the dual ascent method [29]. The update procedures (2.63)-(2.65) are repeated until $\mathbf{\Phi}$ converges. Once the RIS phase shift matrix $\mathbf{\Phi}$ minimizing the block-mutual coherence μ_b is obtained, the BS employs $\mathbf{\Phi}$ for the configuration of RIS phase shifts to improve the block-sparse recovery performance of PF-RCE.

---

[8] A smooth Riemannian manifold is a generalization of the Euclidean space on which the notion of differentiability exists.

[9] When compared with the conventional conjugate gradient method, the RCG method requires two additional operations: 1) *projection operator* to find out Riemannian gradient on the tangent space from the Euclidean gradient and 2) *retraction operator* to make sure that the updated point lies on the manifold [26].

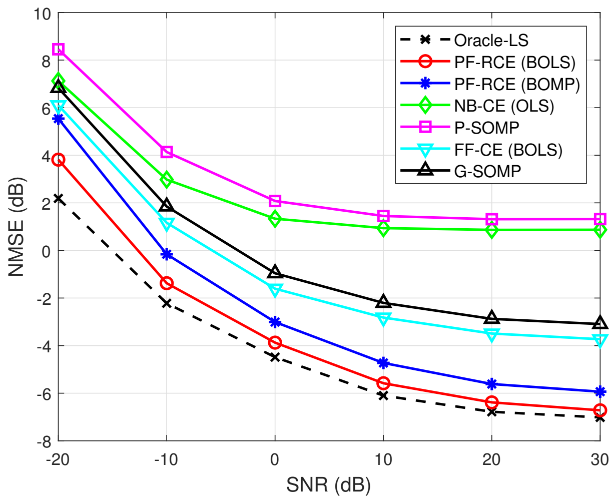

Figure 2.7: NMSE vs. SNR.

2.5 Simulation Results

2.5.1 Simulation Setup

In this section, we present numerical results to validate the effectiveness of the proposed PF-RCE technique. We consider the near-field RIS-assisted THz systems where a single-antenna UE transmits the uplink pilot signal to a $N = 16$-antenna BS with the aid of an RIS equipped with $M = 128$ passive reflecting elements. The RIS and UE are located randomly around the BS within the cell radius of $R = 20\,\text{m}$. For both BS-RIS and UE-RIS channels, we use the block-fading multipath channel model where each channel consists of $P_g = P_r = 2$ paths. The carrier frequency, bandwidth, and the number of subcarriers are set to $f_c = 0.1\,\text{THz}$, $B = 10\,\text{GHz}$, and $S = 128$, respectively. The uplink pilot transmission period includes $P = 12$ subframes, each of which consists of $L = 4$ time slots. Also, the quantization levels of angle and distance are set to $Q_\theta = Q_r = 128$. Throughout the simulations, we set the signal-to-noise-ratio (SNR) to 30 dB. At each point of the figures, the simulation results are averaged over $1,000$ independent channel realizations.

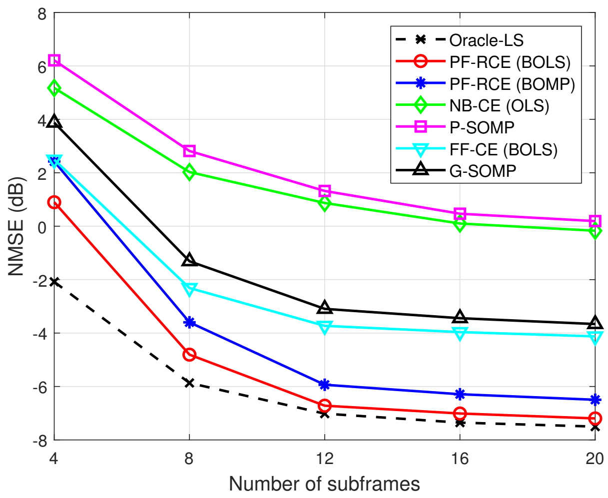

Figure 2.8: NMSE vs. number of subframes.

For comparison, we employ the following benchmark schemes: 1) the oracle-LS scheme where the BS-RIS and UE-RIS angle and distance information is perfectly known at the BS and only the BS-RIS-UE path gains are estimated via LS technique, 2) the narrowband channel estimation schemes, including the OLS-based narrowband channel estimation (NB-CE) algorithm and the polar-domain simultaneous OMP (P-SOMP) algorithm [1], and 3) the far-field channel estimation schemes, including the BOLS-based far-field channel estimation (FF-CE) algorithm and the generalized simultaneous OMP (G-SOMP) algorithm [20]. As a performance metric, we use the normalized mean square error (NMSE) defined as $\text{NMSE} = \sum_{s=1}^{S} \mathbb{E}\big[\|\mathbf{H}^*[s] - \mathbf{H}[s]\|_{\text{F}}^2 / \|\mathbf{H}[s]\|_{\text{F}}^2\big]$.

2.5.2 Simulation Results

In Fig. 2.7, we plot the NMSE as a function of SNR. Overall, we see that the proposed PF-RCE scheme outperforms the conventional channel estimation schemes by a large margin. For instance, when SNR $= 30$ dB, PF-RCE achieves around 3 dB and 3.6 dB NMSE gains over the conventional FF-CE scheme and G-SOMP scheme, respectively. Even when compared to the oracle-LS scheme, PF-RCE performs similarly

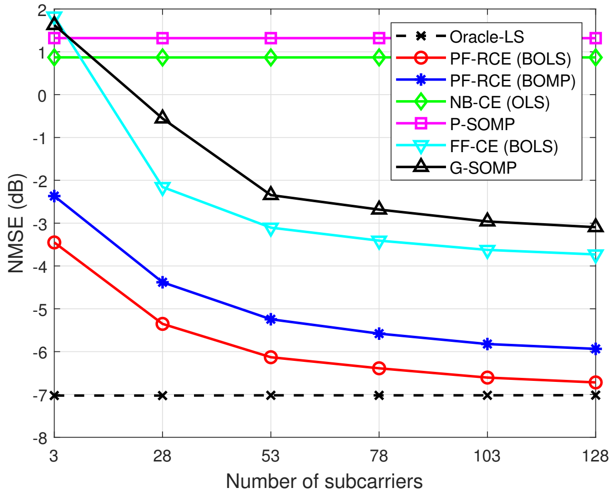

Figure 2.9: NMSE vs. number of subcarriers.

to the oracle-LS scheme. This is because PF-RCE estimates the whole multipath components (angles, distances, and path gains) by exploiting the polar-domain sparsity whereas the conventional far-field channel estimation schemes estimate only the angles and path gains.

In Fig. 2.8, we plot the NMSE as a function of the number of subframes P used for the uplink channel estimation. We find that the proposed PF-RCE achieves a significant pilot overhead reduction over the conventional channel estimation schemes. For instance, to achieve -2 dB NMSE, PF-RCE requires 6 subframes while the narrowband channel estimation schemes require more than 20 subframes. Even when compared to the far-field channel estimation schemes, PF-RCE achieves around 60% reduction on the pilot overhead. It is clear from the results that PF-RCE is effective in reducing the pilot overhead of the RIS-assisted THz systems especially when the ultra-massive antenna array is deployed at the RIS.

In Fig. 2.9, we plot the NMSE as a function of the number of subcarriers S. We observe that PF-RCE achieves a significant NMSE gain over the conventional schemes. We also see that while the NMSEs of the conventional narrowband channel estimation

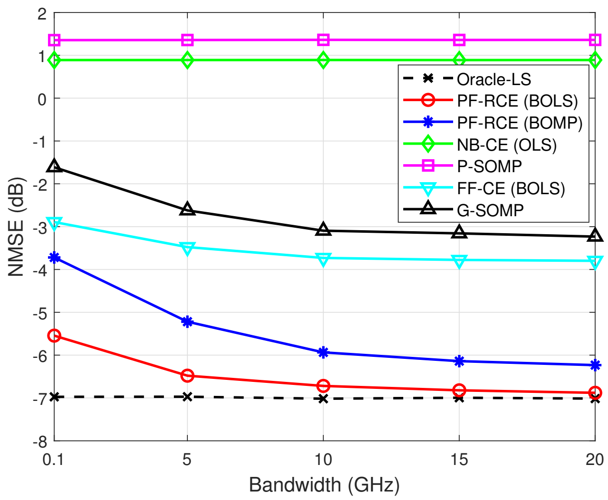

Figure 2.10: NMSE vs. system bandwidth.

schemes are similar in all regimes under test, the NMSE of PF-RCE decreases with the number of subcarriers. For example, when the number of subcarriers increases from $S = 3$ to $S = 128$, the NMSE of BOLS-based PF-RCE decreases from $-3.4\,$dB to around $-6.8\,$dB. The reason is that when the number of subcarriers increases, the performance gain of PF-RCE obtained from the joint sparse recovery of all subcarrier channels also increases but such is not the case for the narrowband channel estimation schemes since they estimate the multipath components of each subcarrier channel separately.

In Fig. 2.10, we plot the NMSE as a function of the system bandwidth B. Similar to Fig. 9, we see that the NMSE gain of PF-RCE over the narrowband channel estimation schemes increases with B. For example, when $B = 0.1\,$GHz, the NMSE gain of PF-RCE over NB-CE is around $6.4\,$dB but it increases up to $7.8\,$dB when $B = 20\,$GHz. In the narrowband systems, the difference between the carrier frequency and the subcarrier frequency is close to zero so that the sensing matrices of different subcarriers are the same. Thus, the NMSE gain of PF-RCE obtained from the joint sparse recovery of subcarrier channels is marginal. In the wideband THz systems, however, the sensing

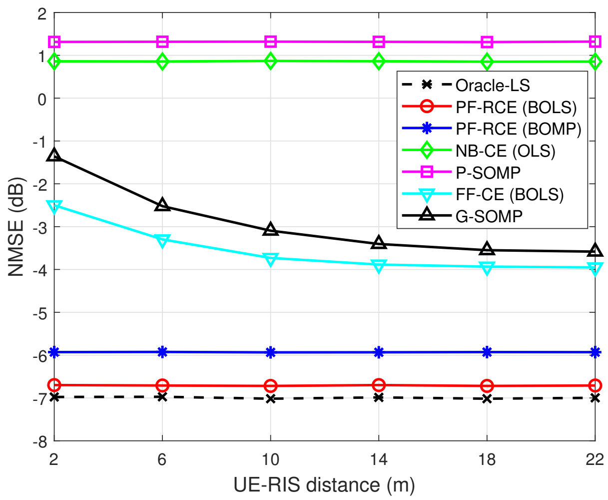

Figure 2.11: NMSE vs. UE-RIS distance.

matrices of different subcarriers are frequency-dependent, and thus PF-RCE can fully enjoy the benefit of joint sparse recovery.

In Fig. 3.1, we plot the NMSE as a function of the distance between the UE and RIS d_{UE}. We observe that the proposed PF-RCE scheme outperforms the conventional channel estimation schemes. In particular, we observe that the NMSE gain of PF-RCE over the far-field channel estimation schemes increases when the UE-RIS distance decreases. For instance, when the UE-RIS distance decreases from 22 m to 2 m, the NMSE gain of PF-RCE over FF-CE increases from 2.7 dB to 4.2 dB. When the UE is close to the RIS, the UE locates at the near-field region of the RIS so that the RIS-assisted THz channel is expressed as a function of the angles, UE-RIS distances, and path gains. Since the far-field channel estimation schemes estimate only the angles and path gains, they suffer from severe performance degradation. In contrast, using the polar-domain sparsity, PF-RCE finds out both the angle and distance from which the near-field RIS-assisted THz channel information can be acquired.

In Fig. 3.2, we plot the NMSE as a function of the number of RIS reflecting elements M. We observe that the performance gap between PF-RCE and the far-field channel

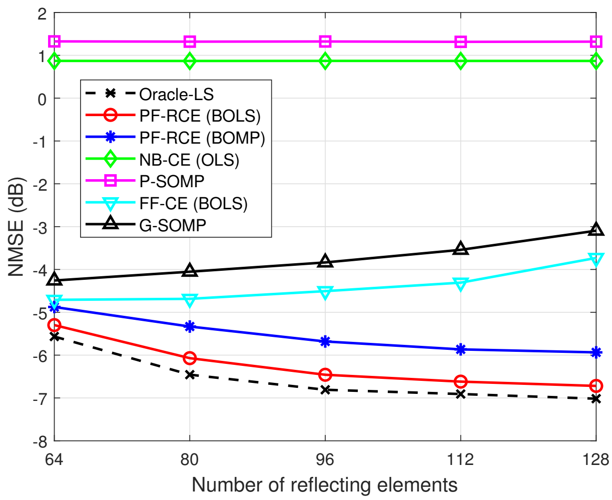

Figure 2.12: NMSE vs. number of reflecting elements.

estimation schemes increases with M. When $M = 64$, for instance, the NMSE gain of PF-RCE over FF-CE is 0.6 dB but it increases to 3.6 dB when $M = 128$. This is because when the number of RIS reflecting elements is large, the array aperture is comparable to the communication distance. In this case, the RIS-assisted THz channel is categorized as the near-field channel so that the performance degradation of the conventional schemes relying on the far-field channel model is considerable.

In Fig. 3.3, we compare the block-mutual coherence of polar-domain sensing matrix achieved by the proposed phase shift control scheme and the random phase shift scheme. We observe that by employing the proposed phase shift control scheme, the block-mutual coherence can be reduced by more than 33%. We also see that the proposed scheme converges within 10 iterations.

In Fig. 3.4, to demonstrate the effectiveness of the proposed polar bin design and RIS phase shift control, we compare the NMSEs of PF-RCE and the conventional schemes [1]. We observe that when SNR $= 30$ dB, the proposed polar bin generation scheme achieves around 0.5 dB NMSE gain over the conventional polar bin generation scheme. Moreover, by exploiting the proposed RIS phase shift control method, the

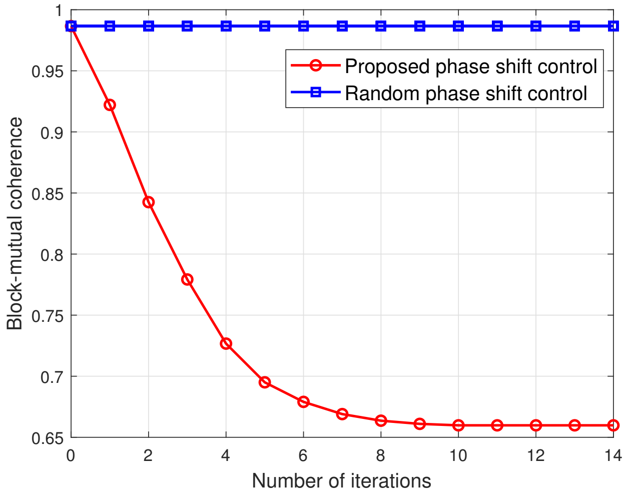

Figure 2.13: Block-mutual coherence vs. number of iterations.

NMSE gain of PF-RCE increases to 1.2 dB. This is because deliberately designed polar bin and RIS phase shifts can minimize the column correlation of the polar-domain sensing matrix in PF-RCE but the conventional schemes have no such mechanism to improve the sparse recovery performance.

In Fig. 3.5, we plot the NMSE as a function of the number of propagation paths. We observe that the proposed PF-RCE works well even when the number of paths is $P_g = P_r = 4$. Interestingly, we observe that the performance gap between the OLS-based and OMP-based schemes increases with the number of paths. This is because the OLS algorithm performs well even when the sparsity level is high but the OMP algorithm performs well only when the sparsity level is low.

2.6 Summary

In this chapter, we proposed an efficient channel estimation technique for near-field RIS-assisted wideband THz systems. The proposed PF-RCE scheme estimates the multipath components (angles, distances, and path gains) of the near-field RIS-

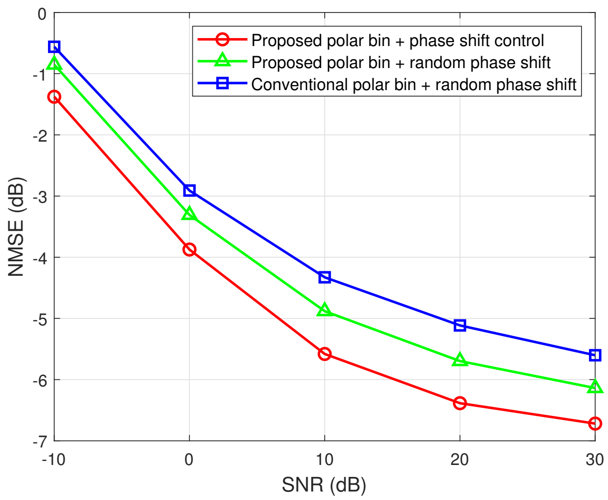

Figure 2.14: NMSE vs. SNR.

assisted THz channel by exploiting the polar-domain sparsity and common support properties. Since the number of multipath components is much smaller than that of the RIS reflecting elements, the pilot overhead can be reduced significantly. In PF-RCE, by exploiting the polar-domain sparsity, the multipath components estimation problem is converted into the sparse recovery problem in the polar-domain. Then using the common support property, the multipath components of all subcarriers are jointly estimated via the block-sparse recovery algorithm. We demonstrated from numerical evaluations that PF-RCE can accurately estimate the near-field RIS-assisted wideband THz channel with low pilot overhead. In our work, we focused on the development of the RIS-assisted THz channel estimation framework, but an extension to the beamforming design using the acquired channel information would also be an interesting research direction worth pursuing.

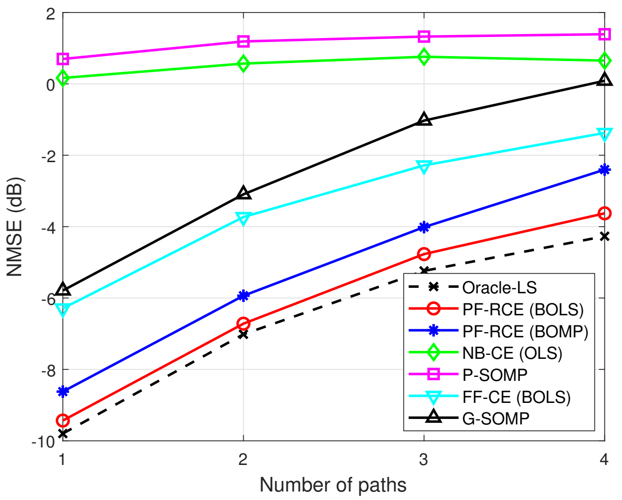

Figure 2.15: NMSE vs. number of paths.

2.7 Proofs

2.7.1 Proof of Lemma 2

From the definition of row-wise Khatri-Rao product, $\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^*$ can be re-expressed as

$$\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^* = [\mathrm{diag}(\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_1, \bar{r}_1))\bar{\mathbf{A}}_{\mathrm{R},s}^*, \cdots, \mathrm{diag}(\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_Q, \bar{r}_Q))\bar{\mathbf{A}}_{\mathrm{R},s}^*]. \tag{2.68}$$

One can see that the column vectors of $\bar{\mathbf{A}}_{\mathrm{R},s} \bullet \bar{\mathbf{A}}_{\mathrm{R},s}^*$ have the form of $\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_i, \bar{r}_i) \odot \mathbf{a}_{\mathrm{R},s}^*(\bar{\theta}_j, \bar{r}_j)$. Recall that $\mathbf{a}_{\mathrm{R},s}(\bar{\theta}, \bar{r}) = [e^{-j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})\Delta r_1(\bar{\theta},\bar{r})}, \cdots, e^{-j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})\Delta r_M(\bar{\theta},\bar{r})}]^{\mathrm{T}}$ where

$$\Delta r_m(\bar{\theta}, \bar{r}) \approx -d(m-1)\sin\bar{\theta} + \frac{1}{2}d^2(m-1)^2\frac{\cos^2\bar{\theta}}{\bar{r}}. \tag{2.69}$$

Thus, the natural logarithm of the m-th element $a_m(i,j)$ of $\mathbf{a}_{\mathrm{R},s}(\bar{\theta}_i,\bar{r}_i)\odot\mathbf{a}^*_{\mathrm{R},s}(\bar{\theta}_j,\bar{r}_j)$ is given by

$$\begin{aligned}\ln a_m(i,j) =& -j\frac{2\pi}{\lambda_c}\Big(1+\frac{f_s}{f_c}\Big)(\Delta r_m(\bar{\theta}_i,\bar{r}_i)-\Delta r_m(\bar{\theta}_j,\bar{r}_j)) \qquad (2.70)\\ =&j\frac{2\pi}{\lambda_c}\Big(1+\frac{f_s}{f_c}\Big)\Big(d(m-1)(\sin\bar{\theta}_i-\sin\bar{\theta}_j)\\ &-\frac{1}{2}d^2(m-1)^2\Big(\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}-\frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\Big)\Big). \qquad (2.71)\end{aligned}$$

By plugging (2.55) and (2.57) into (2.71), we obtain

$$\begin{aligned}\ln a_m(i,j) =&j\frac{2\pi}{\lambda_c}\Big(1+\frac{f_s}{f_c}\Big)\Big(\frac{2d(m-1)}{Q}\big(x_i-x_j+(y_i-y_j)Q_\theta\big)\\ &-\frac{d^2(m-1)^2}{2r_{\min}Q}\big((x_i-x_j)Q_r+y_i-y_j\big)\Big)\end{aligned}$$

where $i=(x_i-1)Q_r+y_i$ and $j=(x_j-1)Q_r+y_j$ for $1\le x_i,x_j\le Q_\theta$ and $1\le y_i,y_j\le Q_r$. Also, Q_θ and Q_r are the quantization levels of angle and distance ($Q=Q_\theta Q_r$), respectively. Based on $1-Q_\theta\le x_i-x_j\le Q_\theta-1$ and $1-Q_r\le y_i-y_j\le Q_r-1$, the number of distinct column vectors of $\bar{\mathbf{A}}_{\mathrm{R},s}\bullet\bar{\mathbf{A}}^*_{\mathrm{R},s}$ is $(2Q_\theta-1)(2Q_r-1)$ in total.

Let $\mathbf{D}[s]\in\mathbb{C}^{M\times(2Q_\theta-1)(2Q_r-1)}$ be the matrix composed by the distinct columns of $\bar{\mathbf{A}}_{\mathrm{R},s}\bullet\bar{\mathbf{A}}^*_{\mathrm{R},s}$. By denoting $k=x_i-x_j+Q_\theta$ and $l=y_i-y_j+Q_r$, the (k,l)-th column vector $\mathbf{d}_{k,l}$ of $\mathbf{D}[s]$ is

$$\begin{aligned}[\mathbf{d}_{k,l}]_m = \exp\Big(&j\frac{2\pi}{\lambda_c}(1+\frac{f_s}{f_c})(\frac{2d(m-1)}{Q}(k-Q_\theta+(l-Q_r)Q_\theta)\\ &-\frac{d^2(m-1)^2}{2r_{\min}Q}((k-Q_\theta)Q_r+l-Q_r))\Big), \qquad (2.72)\end{aligned}$$

for $m=1,\cdots,M$, $k=1,\cdots,2Q_\theta-1$, and $l=2Q_r-1$. Thus, we obtain the desired results.

2.7.2 Proof of Lemma 2

First, since the computational complexity of calculating the pseudo-inverse matrix of a $M\times N$ matrix is $\mathcal{O}(M^2N)$ ($M\ge N$) and $1\le t\le P_gP_r$, the computational complexity $\mathcal{C}_1$ for calculating $\mathbf{P}^{\perp}_{\hat{\mathbf{\Omega}}_{t-1}\cup\{i\}}$ for $i = 1, \cdots, Q_{\text{tot}}$ is

$$\mathcal{C}_1 = \mathcal{O}(SL^2P^2P_gP_rQ_{\text{tot}}). \tag{2.73}$$

Second, the computational complexity $\mathcal{C}_2$ for calculating the residual power and finding out the index $\hat{\omega}_t$ minimizing the residual power is

$$\mathcal{C}_2 = \mathcal{O}(SL^2P^2Q_{\text{tot}}). \tag{2.74}$$

Third, since the residual is updated as $\mathbf{r}_t = [((\mathbf{P}_{\hat{\mathbf{\Omega}}_t}[1])^{\perp}\mathbf{y}[1])^{\text{T}}, \cdots, ((\mathbf{P}_{\hat{\mathbf{\Omega}}_t}[S])^{\perp}\mathbf{y}[S])^{\text{T}}]^{\text{T}}$, the computational complexity $\mathcal{C}_3$ required for the residual update is

$$\mathcal{C}_3 = \mathcal{O}(SL^2P^2). \tag{2.75}$$

Lastly, recall that we obtain the block-sparse path gain vectors using the LS estimation as $(\bar{\mathbf{g}}^*[s])_{\hat{\mathbf{\Omega}}}) = (\bar{\mathbf{\Psi}}[s])^{\dagger}_{\hat{\mathbf{\Omega}}}\mathbf{y}[s]$ for $s = 1, \cdots, S$. Thus, the computational complexity $\mathcal{C}_4$ required for the block-sparse path gain vector estimation is

$$\mathcal{C}_4 = \mathcal{O}(SL^2P^2P_gP_r). \tag{2.76}$$

Using (2.73)-(2.76), we obtain the overall complexity $\mathcal{C}^{\text{PF-RCE}}$ of PF-RCE:

$$\mathcal{C}^{\text{PF-RCE}} = P_gP_r(\mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3) + \mathcal{C}_4 \tag{2.77}$$

$$= \mathcal{O}(SL^2P^2P_g^2P_r^2Q_{\text{tot}}). \tag{2.78}$$

2.7.3 Proof of Proposition 2

Using the definition $\mathbf{a}(\bar{\theta},\bar{r}) = \big[1,\cdots,e^{-j\frac{2\pi f}{c}(-d(M-1)\sin\bar{\theta}+\frac{1}{2}d^2(M-1)^2\frac{\cos^2\bar{\theta}}{\bar{r}})}\big]^{\mathrm{T}}$, the correlation function $f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j) = |\frac{1}{M}\mathbf{a}^{\mathrm{H}}(\bar{\theta}_i,\bar{r}_i)\mathbf{a}(\bar{\theta}_j,\bar{r}_j)|^2$ can be expressed as

$$
\begin{aligned}
&f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j)\\
&=\Big|\frac{1}{M}\sum_{m=1}^{M}\exp\Big(j\frac{2\pi f}{c}\big(-d(m-1)(\sin\bar{\theta}_i-\sin\bar{\theta}_i)\\
&\quad+\frac{1}{2}d^2(m-1)^2\big(\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}-\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}\big)\big)\Big)\Big|^2 \qquad (2.79)\\
&=\Big|\frac{1}{M}\sum_{m=0}^{M-1}\exp\Big(j\frac{1}{2}\pi\Big(\sqrt{\frac{2fd^2}{c}\big(\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}-\frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\big)}\Big)^2\\
&\quad\Big(m-\frac{\sin\bar{\theta}_i-\sin\bar{\theta}_j}{d\big(\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}-\frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\big)}\Big)^2\Big)\Big|^2 \qquad (2.80)\\
&=\Big|\frac{1}{M}\sum_{m=0}^{M-1}e^{j\frac{1}{2}\pi w^2(m-v)^2}\Big|^2, \qquad (2.81)
\end{aligned}
$$

where $w=\sqrt{\frac{2fd^2}{c}\big(\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}-\frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\big)}$ and $v=\frac{\sin\bar{\theta}_i-\sin\bar{\theta}_j}{d\big(\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}-\frac{\cos^2\bar{\theta}_j}{\bar{r}_j}\big)}$. By employing the approximation $\sum_{m=0}^{M-1}e^{j\frac{1}{2}\pi w^2(m-v)^2}\approx\int_0^M e^{j\frac{1}{2}\pi w^2(x-v)^2}dx$, (2.81) can be re-expressed as

$$
\begin{aligned}
&\Big|\frac{1}{M}\sum_{m=0}^{M-1}\exp\Big(j\frac{1}{2}\pi w^2(m-v)^2\Big)\Big|^2\\
&\approx\Big|\frac{1}{M}\int_0^M\exp\Big(j\frac{1}{2}\pi w^2(x-v)^2\Big)dx\Big|^2 \qquad (2.82)\\
&=\frac{1}{M^2w^2}\Big(\Big|\int_{-wv}^{w(M-v)}\cos\Big(\frac{1}{2}\pi t^2\Big)dt\Big|^2+\Big|\int_{-wv}^{w(M-v)}\sin\Big(\frac{1}{2}\pi t^2\Big)dt\Big|^2\Big), \qquad (2.83)
\end{aligned}
$$

Finally, using the Fresnel integrals $C(x)=\int_0^x\cos(\frac{1}{2}\pi t^2)dt$ and $S(x)=\int_0^x\sin(\frac{1}{2}\pi t^2)dt$ and $d=\frac{\lambda_c}{2}=\frac{c}{2f}$, $f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j)$ can be approximated as

$$
\begin{aligned}
f(\bar{\theta}_i,\bar{r}_i,\bar{\theta}_j,\bar{r}_j)\approx&\frac{1}{M^2d\beta}\Big(C\Big(M\sqrt{d\beta}-\frac{\alpha}{\sqrt{d\beta}}\Big)+C\Big(\frac{\alpha}{\sqrt{d\beta}}\Big)\Big)^2\\
&+\frac{1}{M^2d\beta}\Big(S\Big(M\sqrt{d\beta}-\frac{\alpha}{\sqrt{d\beta}}\Big)+S\Big(\frac{\alpha}{\sqrt{d\beta}}\Big)\Big)^2, \qquad (2.84)
\end{aligned}
$$

where $\alpha=\sin\bar{\theta}_i-\sin\bar{\theta}_j$ and $\beta=\frac{\cos^2\bar{\theta}_i}{\bar{r}_i}-\frac{\cos^2\bar{\theta}_j}{\bar{r}_j}$.

2.7.4 Proof of Lemma 3

For a given $\{x_i^*\}$ in (2.54), the objective function $\min_{i\neq j}|c_i f(x_i^*) - c_j f(x_j^*)|$ of $\mathcal{P}$ is given by

$$\min_{i\neq j} |c_i f(x_i^*) - c_j f(x_j^*)| = \frac{c_{\max} f_{\max} - c_{\min} f_{\min}}{Q-1}. \tag{2.85}$$

Now assume that there exists $\{x_i\}$ which provides a higher objective function value than $\{x_i^*\}$, i.e., $\min_{i\neq j}|c_i f(x_i) - c_j f(x_j)| > \frac{c_{\max} f_{\max} - c_{\min} f_{\min}}{Q-1}$. Also, without the loss of generality, assume that $x_1, \cdots, x_Q$ are ordered as $c_1 f(x_1) \leq \cdots \leq c_Q f(x_Q)$. Then we obtain

$$c_{i+1} f(x_{i+1}) - c_i f(x_i) > \frac{c_{\max} f_{\max} - c_{\min} f_{\min}}{Q-1}, \tag{2.86}$$

By combining (2.86) for $i = 1, \cdots, Q-1$, we obtain

$$c_Q f(x_Q) - c_1 f(x_1) > c_{\max} f_{\max} - c_{\min} f_{\min}. \tag{2.87}$$

This contradicts the fact that $c_i f(x_i) \in [c_{\min} f_{\min}, c_{\max} f_{\max}]$. Thus, $\{x_i^*\}$ provides the maximum objective function value of $\mathcal{P}$, meaning that $\{x_i^*\}$ is the optimal solution of $\mathcal{P}$.

Chapter 3

Energy-Efficient Power Control and Beamforming for Reconfigurable Intelligent Surface-Aided Uplink IoT Networks

In this chapter, we study an energy-efficient power control and beamforming scheme for RIS-assisted IoT networks. RIS, a planar metasurface consisting of a large number of low-cost reflecting elements, has received much attention due to the ability to improve both the spectrum and energy efficiencies by reconfiguring the wireless propagation environment. In this work, we propose an RIS phase shift and BS beamforming optimization technique that minimizes the uplink transmit power of the RIS-aided IoT network. Key idea of the proposed scheme, referred to as Riemannian conjugate gradient-based joint optimization (RCG-JO), is to jointly optimize the RIS phase shifts and the BS beamforming vectors using the Riemannian conjugate gradient technique. By exploiting the product Riemannian manifold structure of the sets of unit-modulus phase shifts and unit-norm beamforming vectors, we convert the nonconvex uplink power minimization problem into the unconstrained problem and then find out the optimal solution over the product Riemannian manifold.

3.1 Introduction

As a paradigm to embrace the connection of massive number of devices, such as sensors, wearables, health monitors, and smart appliances, internet of things (IoT) has received much attention recently [30]. Since most of IoT devices are battery-limited, saving the device power is crucial for the dissemination of IoT networks. Recently, reconfigurable intelligent surface (RIS) has been emerging as a potential solution to enhance the sustainability of the IoT network [31]. In a nutshell, RIS is a planar metasurface consisting of a large number of low-cost passive reflecting elements, each of which can independently adjust the phase shift on the incident signal [32, 33]. Due to the capability to modify the wireless channel by controlling the phase shift of each reflecting element, RIS offers various benefits; When the direct link between the IoT device and the base station (BS) is blocked by obstacles, RIS can provide a virtual line-of-sight (LoS) link between them, ensuring the reliable link connection without requiring heavy transmit modules [34]. Due to its simple structure, lightweight, and conformal geometry, RIS can be easily deployed in the desired location. Also, with a small change in wireless standard, RIS can be easily integrated into existing communication systems.

Over the years, various efforts have been made to improve the energy efficiency of IoT networks [35–40]. In [35], a joint beamforming technique to minimize the downlink power consumption of wireless networks has been proposed. In [36], a joint power control and beamforming scheme for the RIS-aided device-to-device (D2D) network has been proposed. In [37], the energy-efficient downlink power control and phase shift designs for the Terahertz and MISO systems have been presented. In [38], an RIS phase shift control technique to maximize the downlink energy efficiency of RIS-aided systems has been proposed. In [39], an alternating optimization-based energy efficiency maximization scheme for the uplink RIS-aided systems has been proposed. Also, achievable sum throughput of an RIS-aided wireless powered sensor network (WPSN) has been investigated in [40].

While most of existing works focused on the downlink power control of RIS-aided systems, not much work has been made for the power minimization on the uplink side. Since IoT devices are battery-limited, it is of importance to come up with an energy-efficient uplink power control mechanism based on the RIS technique.

A major problem of the RIS-aided uplink power control is that the constraints on the RIS phase shifts and the BS beamforming vectors are nonconvex. This is because RIS can only change the phase shift of an incident signal so that the passive beamforming coefficients of reflecting elements are subject to the unit-modulus constraints [31]. This, together with the fact that the active beamforming vectors of BS are subject to the unit-norm constraints [41], makes it very difficult to find out the proper RIS phase shifts and BS beamforming vectors.

An aim of this chapter is to propose an approach that minimizes the uplink transmit power of an RIS-aided IoT network. Key idea of the proposed scheme, referred to as *Riemannian conjugate gradient-based joint optimization (RCG-JO)*, is to jointly optimize the RIS phase shifts and the BS beamforming vectors using the Riemannian conjugate gradient (RCG) algorithm. Specifically, by exploiting the smooth product Riemannian manifold structure of the sets of unit-modulus phase shifts and unit-norm beamforming vectors, we convert the uplink power minimization problem into the unconstrained problem on the Riemannian manifold. Since the optimization over the Riemannian manifold is conceptually analogous to that in the Euclidean space, optimization tools of the Euclidean space, such as the gradient descent method, can be readily used to solve the optimization problem on the Riemannian manifold [42]. In our approach, we employ the Riemannian conjugate gradient (RCG) method to find out the RIS phase shifts and BS beamforming vectors minimizing the uplink transmit power of an RIS-aided IoT network.

In recent years, there have been some efforts exploiting the manifold optimization technique for the RIS phase shifts control [4, 43, 44]. In [4], a Riemannian manifold-based alternating optimization technique for the design of RIS phase shifts and BS beamforming vectors has been proposed. In this scheme, the RIS phase shifts and the BS beamforming vectors are optimized on the Riemannian manifolds in an alternative fashion. In [43, 44], the manifold optimization-based RIS phase shifts control schemes have been proposed to maximize the downlink sum rate.

Our work is distinct from the previous works in the following aspects:

- We propose a Riemannian conjugate gradient-based joint optimization (RCG-JO) algorithm to find out the RIS phase shifts and the BS beamforming vectors minimizing the uplink transmit power of an RIS-aided IoT network. RCG-JO jointly optimizes the RIS phase shifts and the BS beamforming vectors on the product Riemannian manifold of the sets of unit-modulus RIS phase shifts and unit-norm BS beamforming vectors.

- We propose a channel estimation technique for the uplink RIS-aided IoT networks. Specifically, the RIS receives the pilot signals at the active reflecting elements and then feeds them back to the BS. Using the pilot signals, we estimate the partial RIS-aided channel information from which we recover the full RIS-aided channel information via the low-rank matrix completion (LRMC) algorithm.

- We present the convergence analysis of RCG-JO and demonstrate from the numerical experiments that RCG-JO converges to a fixed point within a few number of iterations.

- We provide the empirical simulation results from which we demonstrate that RCG-JO outperforms the conventional power control schemes by a large margin in terms of the uplink transmit power and the computational complexity. For example, when compared to the conventional power control scheme without RIS, RCG-JO saves around 94% of the uplink transmit power. Even when compared to the semidefinite relaxation (SDR)-based power control scheme, RCG-JO saves more than 44% of the uplink transmit power and achieves around 98% reduction in the computational complexity.

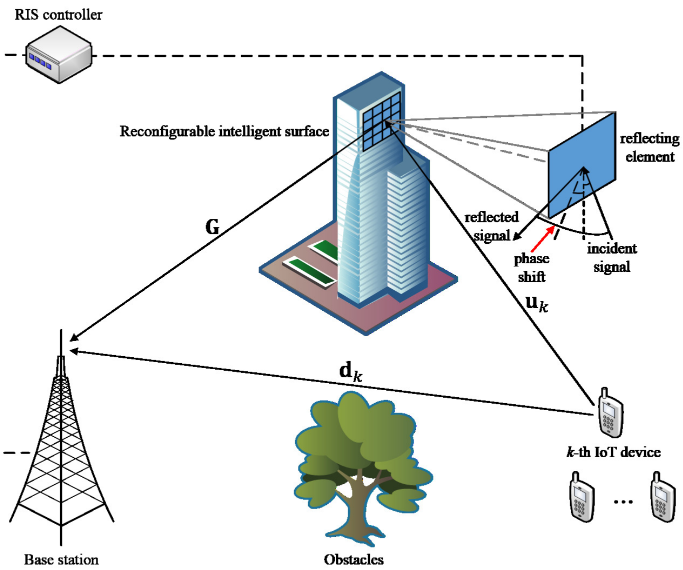

Figure 3.1: Illustration of the RIS-aided uplink IoT network.

Notations: Lower and upper case symbols are used to denote vectors and matrices, respectively. The superscripts $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$, and $(\cdot)^{\mathrm{H}}$ denote the conjugate, transpose, and conjugate transpose, respectively. The operation $\odot$ denotes the Hadamard product. $\|\mathbf{x}\|$ and $\|\mathbf{X}\|_{\mathrm{F}}$ denote the Euclidean norm of a vector $\mathbf{x}$ and the Frobenius norm of a matrix $\mathbf{X}$, respectively. $|\mathbf{x}|$ represents a vector of element-wise absolute values of $\mathbf{x}$. $\mathrm{tr}(\mathbf{X})$ is the trace of $\mathbf{X}$. $\mathrm{diag}(\mathbf{x})$ and $\mathrm{ddiag}(\mathbf{X})$ form diagonal matrices whose diagonal elements are $\mathbf{x}$ and diagonal elements of $\mathbf{X}$, respectively. In addition, $\mathrm{blkdiag}(\mathbf{X}_1, \mathbf{X}_2)$ denotes a block-diagonal matrix with $\mathbf{X}_1$ and $\mathbf{X}_2$ on the block-diagonal.

3.2 RIS-Aided Uplink IoT System Model

In this section, we present the system model, the channel estimation, and the data transmission protocols of RIS-aided uplink IoT networks. We then formulate the power minimization problem.

3.2.1 RIS-Aided Uplink IoT Network

We consider a single-input multi-output (SIMO) uplink IoT network where K devices with a single antenna transmit signals to the BS equipped with M antennas (see Fig. 3.1). In this network, an RIS consisting of a planar array of $N = N_x \times N_y$ reflecting elements is deployed to assist the uplink transmission. For example, each RIS reflecting element can adjust the phase of the incident signal independently using positive-intrinsic-negative (PIN) diodes [31]. The phase shifts of RIS reflecting elements are configured through the dedicated control link. In the setup, the effective uplink channel between the k-th IoT device and the BS is given by

$$\mathbf{h}_k = \mathbf{d}_k + \mathbf{G}\text{diag}(\boldsymbol{\theta})\mathbf{u}_k \tag{3.1}$$

$$= \mathbf{d}_k + \mathbf{G}\text{diag}(\mathbf{u}_k)\boldsymbol{\theta} \tag{3.2}$$

$$= \mathbf{d}_k + \mathbf{G}\mathbf{H}_k\boldsymbol{\theta}, \tag{3.3}$$

where $\mathbf{d}_k \in \mathbb{C}^{M\times 1}$ is the direct channel from the k-th IoT device to the BS, $\mathbf{u}_k \in \mathbb{C}^{N\times 1}$ is the channel from the k-th IoT device to the RIS, and $\mathbf{G} \in \mathbb{C}^{M\times N}$ is the channel from the RIS to BS (see Fig. 1). Also, $\mathbf{H}_k = \text{diag}(\mathbf{u}_k)$ and $\boldsymbol{\theta} = [\mu_1\theta_1, \cdots, \mu_N\theta_N]^{\mathrm{T}}$ is the phase shift vector of RIS. In addition, $\mu_n \in [0, 1]$ is the reflection amplitude coefficient[1] and $\theta_n = e^{j\phi_n}$ is the passive beamforming coefficient where $\phi_n \in [0, 2\pi)$ is the phase shift[2].

[1]In this work, we assume the ideal phase shift model where the reflection amplitude and phase shift are independent. In our system, the RIS controller coordinates the switching between two working modes, i.e., receiving mode ($\mu_n = 0$) for environment sensing (e.g., channel estimation) and reflecting mode ($\mu_n = 1$) for scattering the incident signals (e.g., data transmission).

[2]To characterize the performance limits of RIS, we assume the phase shifts vary continuously in $[0, 2\pi)$. Note that the proposed scheme can be readily extended to practical systems with finite level of phase shifts via discrete phase quantization [45].

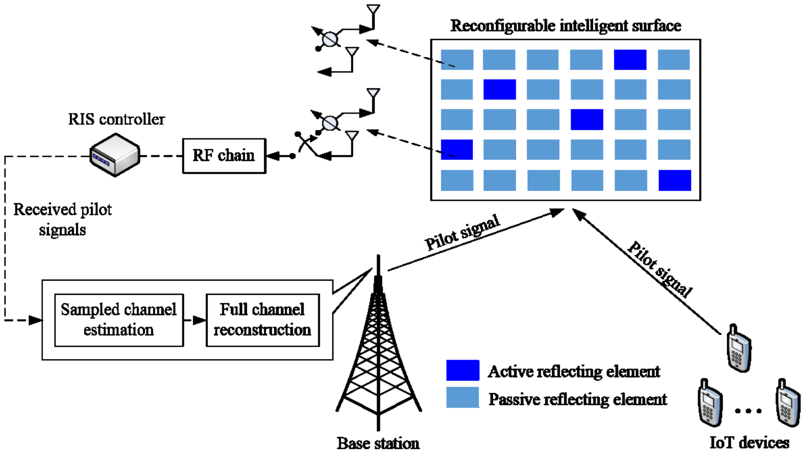

Figure 3.2: Illustration of uplink RIS-aided channel estimation consisting of two major procedures: 1) estimation of the partial channel information observed at the active reflecting elements and 2) reconstruction of the full RIS-aided channel via the LRMC algorithm.

3.2.2 Uplink Channel Estimation

In this subsection, we propose the uplink channel estimation process in time-division duplexing (TDD)-based RIS-aided IoT networks[3]. To enjoy the full potential of RIS, the BS needs to acquire not only the conventional direct channel $\mathbf{d}_k$ but also the channels reflected by the RIS (i.e., $\mathbf{G}$ and $\mathbf{u}_k$). When compared to the direct channel estimation, the estimation of RIS-aided channels is far more difficult since RIS contains a large number of reflecting elements. Recently, the RIS architectures exploiting active reflecting elements that can reflect and also receive the signal have been proposed [31, 34]. Using the pilot signals fed back from the active reflecting elements, the BS can directly estimate the RIS-aided channel components. However, due to the practical limitations such as the implementation cost of RF chains, hardware complexity,

[3]In the 5G systems and beyond, TDD will be a competitive duplexing option due to the improved spectrum efficiency, better adaptation quality to asymmetric uplink/downlink traffics, low transceiver cost, and better support of the massive MIMO [46].

and power consumption, only a few active reflecting elements can be employed. That is, the BS can acquire only partial information of the RIS-aided channels from the active reflecting elements.

To estimate the whole channel from the partial information, we exploit the property that the RIS-aided channels are dominated by LoS paths. Since there is only one propagation path, the channel matrix can be readily modeled as a low-rank matrix. When the channel matrix has a low-rank property, BS can reconstruct the full channel matrix from the partial observations using the LRMC algorithm [47]. Once the full RIS-aided channel information is obtained, BS can perform the uplink power allocation and the RIS phase shift control.

Sampled Channel Estimation

We assume that the RIS consists of $\bar{N}$ active reflecting elements where the index set[4] of active reflecting elements is Ω (see Fig. 3.2). The uplink channel estimation process consists of $(M+K)$ time slots where the m-th BS antenna transmits downlink pilot signal to the RIS at the m-th time slot ($m = 1, \cdots, M$) and then the k-th IoT device transmits uplink pilot signal to the RIS at the $(M+k)$-th time slot ($k = 1, \cdots, K$).

Let x_m^{BS} be the downlink pilot signal of the m-th BS antenna at the m-th time slot and x_k^{D} be the downlink pilot signal of the k-th IoT device at the $(M+k)$-th time slot. Then the received signal matrices $\mathbf{Y}_m^{\text{BS}}, \mathbf{Y}_k^{\text{D}} \in \mathbb{C}^{N_x \times N_y}$ of RIS from the m-th BS antenna and the k-th IoT device are

$$\mathbf{Y}_m^{\text{BS}} = (x_m^{\text{BS}})^* \mathbf{P}_\Omega(\mathbf{G}_m) + \mathbf{N}_m^{\text{BS}}, \quad m = 1, \cdots, M, \tag{3.4}$$

$$\mathbf{Y}_k^{\text{D}} = x_k^{\text{D}} \mathbf{P}_\Omega(\mathbf{U}_k) + \mathbf{N}_k^{\text{D}}, \qquad k = 1, \cdots, K, \tag{3.5}$$

where $\mathbf{G}_m \in \mathbb{C}^{N_x \times N_y}$ is the channel matrix from the m-th BS antenna to the RIS such that $\text{vec}(\mathbf{G}_m)$ is the m-th row vector of $\mathbf{G}$, $\mathbf{U}_k \in \mathbb{C}^{N_x \times N_y}$ is the channel matrix

[4] Note that $\Omega \subseteq \{1, \cdots, N_x\} \times \{1, \cdots, N_y\}$. For example, $\Omega = \{(1,2), (4,3), (5,6)\}$ when $N_x = N_y = 8$ and $\bar{N} = 3$.

from the k-th IoT device to the RIS such that $\text{vec}(\mathbf{U}_k) = \mathbf{u}_k$, and $\mathbf{N}_m^{\text{BS}}$ and $\mathbf{N}_k^{\text{D}}$ are the additive Gaussian noises. Also, $\mathbf{P}_\Omega(\mathbf{G}_m)$ and $\mathbf{P}_\Omega(\mathbf{U}_k)$ are the sampled matrices[5] of $\mathbf{G}_m$ and $\mathbf{U}_k$. Note that the sampling operator is used since only active reflecting elements can receive the pilot signals. Then the BS can easily acquire the sampled channel matrices $\mathbf{P}_\Omega(\mathbf{G}_m)$ and $\mathbf{P}_\Omega(\mathbf{U}_k)$ using the conventional least squares (LS) and minimum mean square error (MMSE) estimation techniques.

Full Channel Reconstruction

After the sampled channel estimation, the BS reconstructs the full channel matrices $\{\mathbf{G}_m\}_{m=1}^{M}$ and $\{\mathbf{U}_k\}_{k=1}^{K}$ from $\{\mathbf{P}_\Omega(\mathbf{G}_m)\}_{m=1}^{M}$ and $\{\mathbf{P}_\Omega(\mathbf{U}_k)\}_{k=1}^{K}$ via the LRMC algorithm [47].

To be specific, $\mathbf{G}_m$ can be reconstructed by solving the rank minimization problem:

$$\min_{\mathbf{X}\in\mathbb{C}^{N_x\times N_y}} \quad \text{rank}(\mathbf{X}) \tag{3.6a}$$

$$\text{s.t.} \quad \mathbf{P}_\Omega(\mathbf{X}) = \mathbf{P}_\Omega(\mathbf{G}_m). \tag{3.6b}$$

The solution $\mathbf{X}^*$ of (3.6) is the estimate of $\mathbf{G}_m$. Since the rank minimization problem is NP-hard, this problem is computationally intractable. To deal with the problem, we replace the non-convex objective function with its convex surrogate. The nuclear norm $\|\mathbf{X}\|_*$, the sum of the singular values of $\mathbf{X}$, has been widely used as a convex surrogate of rank$(\mathbf{X})$[6]:

$$\min_{\mathbf{X}\in\mathbb{C}^{N_x\times N_y}} \quad \|\mathbf{X}\|_* \tag{3.7a}$$

$$\text{s.t.} \quad \mathbf{P}_\Omega(\mathbf{X}) = \mathbf{P}_\Omega(\mathbf{G}_m). \tag{3.7b}$$

[5] $\mathbf{P}_\Omega$ is the sampling operator such that $[\mathbf{P}_\Omega(\mathbf{A})]_{x,y} = [\mathbf{A}]_{x,y}$ if $(x,y)\in\Omega$ and otherwise zero for a matrix $\mathbf{A}$.

[6] It has been shown that the nuclear norm is the convex envelope of rank function on the set $\{\mathbf{X} : \|\mathbf{X}\| \leq 1\}$ [47].

The nuclear norm minimization problem (3.7) can also be recast as a semidefinite programming (SDP) [47]:

$$\min_{\mathbf{Z}} \ \mathrm{tr}(\mathbf{Z}) \tag{3.8a}$$

$$\text{s.t.} \ \mathrm{tr}(\mathbf{A}_{x,y}^{\mathrm{H}}\mathbf{Z}) = [\mathbf{G}_m]_{x,y}, \quad (x,y) \in \Omega, \tag{3.8b}$$

$$\mathbf{Z} \succeq \mathbf{0}, \tag{3.8c}$$

where $\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Z}_2 \end{bmatrix} \in \mathbb{C}^{(N_x+N_y)\times(N_x+N_y)}$ for the Hermitian matrices $\mathbf{Z}_1 \in \mathbb{C}^{N_x \times N_x}$ and $\mathbf{Z}_2 \in \mathbb{C}^{N_y \times N_y}$, $\mathbf{A}_{x,y} = \mathbf{e}_x \mathbf{e}_{y+N_x}^{\mathrm{T}}$ is the linear sampling matrix, and $\mathbf{e}_x$ is the x-th column vector of $\mathbf{I}_{N_x+N_y}$. Since (3.8) is a convex problem, the solutions $\mathbf{Z}^*$ and $\mathbf{X}^*$ of (3.8) can be obtained using the convex optimization solvers (e.g., CVX [48]). Finally, $\mathbf{G}_m$ is obtained by $\mathbf{G}_m = \mathbf{X}^*$.

Similarly, $\mathbf{U}_k$ can be recovered from $\mathbf{P}_\Omega(\mathbf{U}_k)$ by solving the low-rank matrix completion problem. Finally, the whole RIS-aided channels $\mathbf{G}$ and $\mathbf{u}_k$ are obtained as $\mathbf{G} = [\mathrm{vec}(\mathbf{G}_1) \cdots \mathrm{vec}(\mathbf{G}_M)]^{\mathrm{T}}$ and $\mathbf{u}_k = \mathrm{vec}(\mathbf{U}_k)$.

3.2.3 Uplink Data Transmission

After the channel estimation, BS performs the uplink power allocation and the RIS phase shift control using the acquired channel information. Then BS sends the uplink transmit power indicator p_k and the RIS phase shift vector $\boldsymbol{\theta}$ to the k-th IoT device and RIS, respectively. Finally, each IoT device transmits the data to BS through the uplink channel.

Let $x_k = \sqrt{p_k} s_k$ be the data signal of the k-th IoT device where s_k and $p_k (\geq 0)$ are the normalized data symbol and the transmit power of the k-th IoT device, respectively.

Then, the received signal at BS from the k-th IoT device y_k is

$$y_k = \mathbf{w}_k^{\mathrm{H}}\Big(\mathbf{h}_k x_k + \sum_{j\neq k}^{K} \mathbf{h}_j x_j + \mathbf{n}_k\Big) \tag{3.9}$$

$$= \sqrt{p_k}\mathbf{w}_k^{\mathrm{H}}(\mathbf{d}_k + \mathbf{G}\mathbf{H}_k\boldsymbol{\theta})s_k + \sum_{j\neq k}^{K} \sqrt{p_j}\mathbf{w}_k^{\mathrm{H}}(\mathbf{d}_j + \mathbf{G}\mathbf{H}_j\boldsymbol{\theta})s_j + \mathbf{w}_k^{\mathrm{H}}\mathbf{n}_k, \tag{3.10}$$

where $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ is the normalized BS beamforming vector for the k-th IoT device, i.e., $\|\mathbf{w}_k\| = 1$, and $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma_k^2\mathbf{I})$ is the additive Gaussian noise. In this setting, the uplink achievable rate R_k of the k-th IoT device is given by

$$R_k = \log_2\Big(1 + \frac{p_k|\mathbf{w}_k^{\mathrm{H}}(\mathbf{d}_k + \mathbf{G}\mathbf{H}_k\boldsymbol{\theta})|^2}{\sum_{j\neq k}^{K} p_j|\mathbf{w}_k^{\mathrm{H}}(\mathbf{d}_j + \mathbf{G}\mathbf{H}_j\boldsymbol{\theta})|^2 + \sigma_k^2}\Big). \tag{3.11}$$

3.2.4 Uplink Power Minimization Problem Formulation

The uplink power minimization problem to optimize the RIS phase shift vector $\boldsymbol{\theta}$, the BS beamforming matrix $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_K]$, and the device power vector $\mathbf{p} = [p_1 \cdots p_K]$ is formulated as

$$\mathcal{P}_1 : \min_{\boldsymbol{\theta},\mathbf{W},\mathbf{p}} \sum_{k=1}^{K} p_k \tag{3.12a}$$

$$\text{s.t.} \quad \frac{p_k|\mathbf{w}_k^{\mathrm{H}}(\mathbf{d}_k + \mathbf{G}\mathbf{H}_k\boldsymbol{\theta})|^2}{\sum_{j\neq k}^{K} p_j|\mathbf{w}_k^{\mathrm{H}}(\mathbf{d}_j + \mathbf{G}\mathbf{H}_j\boldsymbol{\theta})|^2 + \sigma_k^2} \geq 2^{R_k^{\min}} - 1,$$

$$\forall k \in \mathcal{K}, \tag{3.12b}$$

$$|\theta_n| = 1, \qquad \forall n \in \mathcal{N}, \tag{3.12c}$$

$$\|\mathbf{w}_k\| = 1, \qquad \forall k \in \mathcal{K}, \tag{3.12d}$$

$$0 \leq p_k \leq p_k^{\max}, \quad \forall k \in \mathcal{K}, \tag{3.12e}$$

where $\mathcal{K}$ and $\mathcal{N}$ are the sets of IoT devices and RIS reflecting elements and $R_k^{\min}$ and $p_k^{\max}$ are the rate requirement and the maximum transmit power of the k-th IoT device, respectively. Note that (3.12c) is the unit-modulus constraint of the RIS phase shift and (3.12d) is the unit-norm constraint of the BS beamforming vector. Due to

Algorithm 1 Riemannian conjugate gradient-based joint optimization algorithm

Input: Number of iterations T, rate requirement $R_k^{\min}$, maximum transmit power $p_k^{\max}$

Output: Uplink transmit power $\mathbf{p}$, RIS phase shift vector $\boldsymbol{\theta}$, BS beamforming matrix $\mathbf{W}$

Initialize: $t = 1$, $\mathbf{p}_t = \mathbf{p}_{\text{ini}}$, $\boldsymbol{\theta}_t = \boldsymbol{\theta}_{\text{ini}}$, $\mathbf{W}_t = \mathbf{W}_{\text{ini}}$

while P_{total} does not converge **do**

    Optimize $\boldsymbol{\theta}_t$ and $\mathbf{W}_t$ simultaneously using Algorithm 2 when $\mathbf{p}_t$ is fixed

    Optimize $\mathbf{p}_t$ by solving an LP problem when $\boldsymbol{\theta}_t$ and $\mathbf{W}_t$ are fixed

    $t = t + 1$

end

the nonconvexity of the unit-modulus and unit-norm constraints, $\mathcal{P}_1$ is a non-convex problem. This, together with the quadratic fractional and coupled structure of the rate function in (3.12b), makes $\mathcal{P}_1$ very difficult to solve.

3.3 Riemannian Conjugate Gradient-based joint optimization Algorithm

The primal goal of the proposed RCG-JO technique is to find out the RIS phase shifts and the BS beamforming vectors minimizing the uplink transmit power of RIS-aided IoT networks. As mentioned, main obstacles in solving the uplink power minimization problem are the nonconvex unit-modulus constraint of the RIS phase shift and unit-norm constraint of the BS beamforming vector. To handle these issues, we exploit the smooth product Riemannian manifold structure of the sets of unit-modulus phase shifts and unit-norm beamforming vectors. Since the product of two manifolds is also a Riemannian manifold with smooth structure, we can readily convert the uplink power minimization problem $\mathcal{P}_1$ to an unconstrained problem on the product Riemannian manifold. After that, by using the differential geometry tools, we can design the gradient descent algorithm on the Riemannian manifold and use it to obtain the optimal RIS phase shifts and the BS beamforming vectors minimizing the uplink transmit power of the RIS-aided IoT network.

In a nutshell, the proposed RCG-JO algorithm consists of two major steps: 1) Lagrangian relaxation to move the complicated rate constraint to the objective function and 2) alternating optimization of $\mathbf{p}$ and $(\boldsymbol{\theta}, \mathbf{W})$ to minimize the modified objective function on the product Riemannian manifold. To be specific, in the alternating optimization step, we first fix the device power $\mathbf{p}$ and then jointly optimize the phase shift vector $\boldsymbol{\theta}$ and the BS beamforming matrix $\mathbf{W}$ using the RCG method. Once the optimal $\boldsymbol{\theta}$ and $\mathbf{W}$ are obtained, the optimization problem of $\mathbf{p}$ is formulated as a linear programming (LP) problem where the optimal solution can be easily obtained using the convex optimization technique. We repeat these procedures until the objective function $P_{\text{total}} = \sum_{k=1}^{K} p_k$ converges (see Algorithm 1).

3.3.1 Notions on Riemannian Manifolds

In this subsection, we briefly introduce properties and operators of the optimization on Riemannian manifold. Roughly speaking, a smooth manifold is a generalization of the Euclidean space on which the notion of differentiability exists [49]. The tangent space $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$ at a point $\mathbf{X}$ of a manifold $\mathcal{Y}$ is the set of the tangent vectors of all the curves at $\mathbf{X}$, where the curve γ of $\mathcal{Y}$ is a mapping from $\mathbb{C}$ to $\mathcal{Y}$. Put it formally, for a given point $\mathbf{X} \in \mathcal{Y}$, the tangent space of $\mathcal{Y}$ at $\mathbf{X}$ is defined as $\mathcal{T}_{\mathbf{X}}\mathcal{Y} = \{\gamma'(0) : \gamma \text{ is a curve in } \mathcal{Y}, \gamma(0) = \mathbf{X}\}$ (see Fig. 3.3(a)). A manifold $\mathcal{Y}$ together with a smoothly varying inner product $g = \langle \cdot, \cdot \rangle : \mathcal{T}_{\mathbf{X}}\mathcal{Y} \times \mathcal{T}_{\mathbf{X}}\mathcal{Y} \to \mathbb{C}$ on the tangent space $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$ forms a smooth Riemannian manifold, denoted as $(\mathcal{Y}, g)$ where g is termed as the Riemannian metric.

When the Riemannian manifold $(\mathcal{Y}, g)$ is the cartesian product of two Riemannian manifolds $(\mathcal{Y}_1, g_1)$ and $(\mathcal{Y}_2, g_2)$, then the Riemannian metric g is defined as $g = g_1 + g_2$. In the following lemma, we show that the tangent space on the product Riemannian manifold is the direct sum of the tangent spaces on each Riemannian manifold.

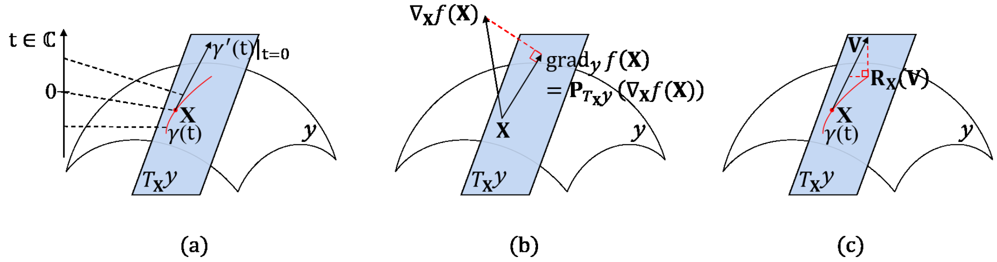

Figure 3.3: Illustration of (a) the tangent space $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$, (b) the Riemannian gradient $\mathrm{grad}_{\mathcal{Y}} f(\mathbf{X})$, and (c) the retraction operator $\mathbf{R}_{\mathbf{X}}(\mathbf{V})$ at the point $\mathbf{X}$ in the Riemannian manifold $\mathcal{Y}$. Note that Euclidean gradient $\nabla_{\mathbf{X}} f(\mathbf{X})$ is a direction for which the objective function is reduced in $\mathbb{C}^P$ whereas the Riemannian gradient $\mathrm{grad}_{\mathcal{Y}} f(\mathbf{X})$ is a direction for which the objective function is reduced in $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$.

Lemma 5. *For the point $\mathbf{X} = \mathbf{X}_1 \oplus \mathbf{X}_2$ where $\mathbf{X}_1 \in \mathcal{Y}_1$, $\mathbf{X}_2 \in \mathcal{Y}_2$, the tangent space at $\mathbf{X}$ of the product Riemannian manifold $\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2$ is given by*

$$\mathcal{T}_{\mathbf{X}}\mathcal{Y} = \mathcal{T}_{\mathbf{X}_1}\mathcal{Y}_1 \oplus \mathcal{T}_{\mathbf{X}_2}\mathcal{Y}_2. \tag{3.13}$$

Given a smooth objective function f on the Riemannian manifold $\mathcal{Y}$, minimization of f requires the notion of gradient at every $\mathbf{X} \in \mathcal{Y}$. To be specific, the Riemannian gradient of f at $\mathbf{X}$, denoted by $\mathrm{grad}_{\mathcal{Y}} f(\mathbf{X})$, is defined as a unique vector on $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$ that yields the steepest-descent of f. Put it formally, $\mathrm{grad}_{\mathcal{Y}} f(\mathbf{X})$ is obtained by projecting $\nabla_{\mathbf{X}} f(\mathbf{X})$ onto $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$ (see Fig. 3.3(b)).

Using Lemma 1, the Riemannian gradient $\mathrm{grad}_{\mathcal{Y}} f(\mathbf{X})$ is given by

$$\mathrm{grad}_{\mathcal{Y}} f(\mathbf{X}) = \mathrm{grad}_{\mathcal{Y}_1} f_1(\mathbf{X}_1) \oplus \mathrm{grad}_{\mathcal{Y}_2} f_2(\mathbf{X}_2), \tag{3.14}$$

where $\mathrm{grad}_{\mathcal{Y}_i} f_i(\mathbf{X}_i) \in \mathcal{T}_{\mathbf{X}_i}\mathcal{Y}_i$ is the Riemannian gradient of f_i at $\mathbf{X}_i \in \mathcal{Y}_i$ for $i = 1, 2$.

Definition 1. *An orthogonal projection onto the tangent space $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$ is a mapping $\mathbf{P}_{\mathcal{T}_{\mathbf{X}}\mathcal{Y}} : \mathbb{C}^{a \times b} \to \mathcal{T}_{\mathbf{X}}\mathcal{Y}$ such that for a given $\mathbf{U} \in \mathbb{C}^{a \times b}$, $\langle \mathbf{U} - \mathbf{P}_{\mathcal{T}_{\mathbf{X}}\mathcal{Y}}(\mathbf{U}), \mathbf{V} \rangle = 0$ for all $\mathbf{V} \in \mathcal{T}_{\mathbf{X}}\mathcal{Y}$. In particular, the orthogonal projection onto the tangent space of product* *manifold* $\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2$ *is*

$$\mathbf{P}_{\mathcal{T}_{\mathbf{X}}\mathcal{Y}}(\mathbf{U}) = \mathbf{P}_{\mathcal{T}_{\mathbf{X}_1}\mathcal{Y}_1}(\mathbf{U}_1) \oplus \mathbf{P}_{\mathcal{T}_{\mathbf{X}_2}\mathcal{Y}_2}(\mathbf{U}_2), \tag{3.15}$$

where $\mathbf{P}_{\mathcal{T}_{\mathbf{X}_i}\mathcal{Y}_i}(\mathbf{U}_i)$ *is the projection of* $\mathbf{U}_i$ *onto* $\mathcal{T}_{\mathbf{X}_i}\mathcal{Y}_i$ *for* $i = 1, 2$.

In order to express the concept of moving in the direction of a tangent space while staying on the manifold, we need an operation called *retraction*. As illustrated in Fig. 3.3(c), the retraction is a mapping from $\mathcal{T}_{\mathbf{X}}\mathcal{Y}$ to $\mathcal{Y}$ that preserves the gradient at $\mathbf{X}$ [49].

Definition 2. *The retraction* $\mathbf{R}_{\mathbf{X}}(\mathbf{V})$ *of a matrix* $\mathbf{V} \in \mathcal{T}_{\mathbf{X}}\mathcal{Y}$ *onto* $\mathcal{Y}$ *is defined as*

$$\mathbf{R}_{\mathbf{X}}(\mathbf{V}) = \arg\min_{\mathbf{Z}\in\mathcal{Y}} \|\mathbf{X} + \mathbf{V} - \mathbf{Z}\|_{\mathrm{F}}. \tag{3.16}$$

In particular, the retraction onto the product manifold $\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2$ *is*

$$\mathbf{R}_{\mathbf{X}}(\mathbf{V}) = \mathbf{R}_{\mathbf{X}_1}(\mathbf{V}_1) \oplus \mathbf{R}_{\mathbf{X}_2}(\mathbf{V}_2), \tag{3.17}$$

where $\mathbf{R}_{\mathbf{X}_i}(\mathbf{V}_i)$ *is the retraction of* $\mathbf{V}_i \in \mathcal{T}_{\mathbf{X}_i}\mathcal{Y}_i$ *onto* $\mathcal{Y}_i$ *for* $i = 1, 2$.

3.3.2 Joint RIS Phase Shift and BS Beamforming Optimization on Product Manifold

In the first step, we fix the device power and then jointly optimize the RIS phase shifts and BS beamforming vectors using the RCG method.

When the device power is given, $\mathcal{P}_1$ is reduced to

$$\begin{aligned}
\mathcal{P}_2 : \text{Find} \quad & (\boldsymbol{\theta}, \mathbf{W}) && (3.18\text{a})\\
\text{s.t.} \quad & \frac{p_k}{2^{R_k^{\min}} - 1} A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k) - \sum_{j\neq k}^{K} p_j\, A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k) \geq \sigma_k^2, && \\
& \forall k \in \mathcal{K}, && (3.18\text{b})\\
& |\theta_n| = 1, \qquad \forall n \in \mathcal{N}, && (3.18\text{c})\\
& \|\mathbf{w}_k\| = 1, \quad \forall k \in \mathcal{K}, && (3.18\text{d})
\end{aligned}$$

where $A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k) = |\mathbf{w}_k^{\mathrm{H}}(\mathbf{d}_j + \mathbf{G}\mathbf{H}_j\boldsymbol{\theta})|^2$ for $j, k \in \mathcal{K}$. Since the rate expression in (3.18b) is a joint quadratic function of $\boldsymbol{\theta}$ and $\mathbf{W}$, (3.18b) is a nonconvex constraint. To handle this issue, we use the Lagrangian relaxation to move the complicated rate constraints to the objective function. To be specific, the modified objective function is given by

$$L(\boldsymbol{\theta}, \mathbf{W}, \boldsymbol{\lambda}) = \sum_{k=1}^{K} \lambda_k \Big(- \frac{p_k}{2^{R_k^{\min}} - 1} A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sum_{j \neq k}^{K} p_j A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sigma_k^2 \Big), \quad (3.19)$$

where $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_K]^{\mathrm{T}}$ is the Lagrangian multiplier obtained by solving the corresponding dual problem. Using $L(\boldsymbol{\theta}, \mathbf{W}, \boldsymbol{\lambda})$, $\mathcal{P}_2$ is relaxed into

$$\begin{aligned}
\mathcal{P}_3 : \min_{\boldsymbol{\theta}, \mathbf{W}, \boldsymbol{\lambda}} \quad & \sum_{k=1}^{K} \lambda_k \Big(- \frac{p_k}{2^{R_k^{\min}} - 1} A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sum_{j \neq k}^{K} p_j A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sigma_k^2 \Big) && (3.20a) \\
\text{s.t.} \quad & |\theta_n| = 1, \qquad \forall n \in \mathcal{N}, && (3.20b) \\
& \|\mathbf{w}_k\| = 1, \qquad \forall k \in \mathcal{K}, && (3.20c) \\
& \lambda_k \geq 0, \qquad \forall k \in \mathcal{K}. && (3.20d)
\end{aligned}$$

The relaxed problem $\mathcal{P}_3$ looks simpler than $\mathcal{P}_2$, but it is still nonconvex and difficult to solve since the objective function of $\mathcal{P}_3$ is a joint quadratic function of $\boldsymbol{\theta}$ and $\mathbf{W}$. Additionally, we need to deal with the unit-modulus constraints (3.20b) and the unit-norm constraints (3.20c). To handle the problem, we jointly optimize $\boldsymbol{\theta}$ and $\mathbf{W}$ on the product manifold of the unit-modulus phase shifts and the unit-norm beamforming vectors using the RCG method. Once we obtain $\boldsymbol{\theta}$ and $\mathbf{W}$, we update the Lagrangian multiplier $\boldsymbol{\lambda}$. We repeat these procedures until $\boldsymbol{\theta}$ and $\mathbf{W}$ converge.

Joint RIS Phase Shift and BS Beamforming Optimization

For a given $\boldsymbol{\lambda}$, $\mathcal{P}_3$ is reduced to the unconstrained problem on the product manifold:

$$\mathcal{P}_{(\boldsymbol{\theta}, \mathbf{W})} : \min_{(\boldsymbol{\theta}, \mathbf{W}) \in \mathcal{M}_{\boldsymbol{\theta}} \times \mathcal{M}_{\mathbf{W}}} L(\boldsymbol{\theta}, \mathbf{W}), \quad (3.21)$$

where $\mathcal{M}_{\boldsymbol{\theta}}$ and $\mathcal{M}_{\mathbf{W}}$ are the complex circle manifold and complex oblique manifold given by

$$\mathcal{M}_{\boldsymbol{\theta}} = \{\boldsymbol{\theta} \in \mathbb{C}^{N\times 1} : |\theta_n| = 1,\ \forall n \in \mathcal{N}\}, \tag{3.22}$$

$$\mathcal{M}_{\mathbf{W}} = \{\mathbf{W} \in \mathbb{C}^{M\times K} : \text{ddiag}(\mathbf{W}^{\text{H}}\mathbf{W}) = \mathbf{I}_K\}, \tag{3.23}$$

with the inner products defined as $g_{\boldsymbol{\theta}}(\mathbf{z}_1, \mathbf{z}_2) = \langle \mathbf{z}_1, \mathbf{z}_2 \rangle = \text{Re}\{\mathbf{z}_1^{\text{H}}\mathbf{z}_2\}$ and $g_{\mathbf{W}}(\mathbf{Z}_1, \mathbf{Z}_2) = \langle \mathbf{Z}_1, \mathbf{Z}_2 \rangle = \text{Re}\{\text{tr}(\mathbf{Z}_1^{\text{H}}\mathbf{Z}_2)\}$, respectively.

By combining $\boldsymbol{\theta}$ and $\mathbf{W}$ into $\boldsymbol{\Sigma} = \text{blkdiag}(\boldsymbol{\theta}, \mathbf{W})$, $\mathcal{P}_{(\boldsymbol{\theta},\mathbf{W})}$ is re-expressed as

$$\mathcal{P}_{\boldsymbol{\Sigma}} : \min_{\boldsymbol{\Sigma}\in\mathcal{M}} \ L(\boldsymbol{\Sigma}), \tag{3.24}$$

where $\mathcal{M} = \mathcal{M}_{\boldsymbol{\theta}} \times \mathcal{M}_{\mathbf{W}}$ is the product manifold with the inner product defined as $g_{\boldsymbol{\Sigma}} = g_{\boldsymbol{\theta}} + g_{\mathbf{W}}$. In the following lemma, we provide the tangent space of the product manifold $\mathcal{M}$.

Lemma 6. *The tangent space $\mathcal{T}_{\boldsymbol{\Sigma}}\mathcal{M}$ of the product manifold $\mathcal{M}$ at the point $\boldsymbol{\Sigma}$ is given by*

$$\mathcal{T}_{\boldsymbol{\Sigma}}\mathcal{M} = \mathcal{T}_{\boldsymbol{\theta}}\mathcal{M}_{\boldsymbol{\theta}} \oplus \mathcal{T}_{\mathbf{W}}\mathcal{M}_{\mathbf{W}}, \tag{3.25}$$

where $\mathcal{T}_{\boldsymbol{\theta}}\mathcal{M}_{\boldsymbol{\theta}} = \{\mathbf{z} \in \mathbb{C}^{N\times 1} : \text{Re}\{\mathbf{z}^* \odot \boldsymbol{\theta}\} = \mathbf{0}_N\}$ is the tangent space of $\mathcal{M}_{\boldsymbol{\theta}}$ at $\boldsymbol{\theta}$ and $\mathcal{T}_{\mathbf{W}}\mathcal{M}_{\mathbf{W}} = \{\mathbf{Z} \in \mathbb{C}^{M\times K} : \text{ddiag}(\text{Re}\{\mathbf{W}^{\text{H}}\mathbf{Z}\}) = \mathbf{0}_K\}$ is the tangent space of $\mathcal{M}_{\mathbf{W}}$ at $\mathbf{W}$.

In order to minimize the objective function $L(\boldsymbol{\Sigma})$ in $\mathcal{P}_{\boldsymbol{\Sigma}}$, we need a Riemannian gradient which is obtained by projecting the Euclidean gradient of $L(\boldsymbol{\Sigma})$ at $\boldsymbol{\Sigma}$ onto $\mathcal{T}_{\boldsymbol{\Sigma}}\mathcal{M}$.

Lemma 7. *The orthogonal projection $\mathbf{P}_{\mathcal{T}_{\boldsymbol{\Sigma}}\mathcal{M}}(\bar{\mathbf{U}})$ of $\bar{\mathbf{U}} = \text{blkdiag}(\mathbf{u}, \mathbf{U})$ onto $\mathcal{T}_{\boldsymbol{\Sigma}}\mathcal{M}$ is*

$$\mathbf{P}_{\mathcal{T}_{\boldsymbol{\Sigma}}\mathcal{M}}(\bar{\mathbf{U}}) = \mathbf{P}_{\mathcal{T}_{\boldsymbol{\theta}}\mathcal{M}_{\boldsymbol{\theta}}}(\mathbf{u}) \oplus \mathbf{P}_{\mathcal{T}_{\mathbf{W}}\mathcal{M}_{\mathbf{W}}}(\mathbf{U}), \tag{3.26}$$

where $\mathbf{P}_{\mathcal{T}_{\boldsymbol{\theta}}\mathcal{M}_{\boldsymbol{\theta}}}(\mathbf{u}) = \mathbf{u} - \text{Re}\{\boldsymbol{\theta}^* \odot \mathbf{u}\} \odot \boldsymbol{\theta}$ is the orthogonal projection of $\mathbf{d}$ onto $\mathcal{T}_{\boldsymbol{\theta}}\mathcal{M}_{\boldsymbol{\theta}}$ and $\mathbf{P}_{\mathcal{T}_{\mathbf{W}}\mathcal{M}_{\mathbf{W}}}(\mathbf{U}) = \mathbf{U} - \mathbf{W}\,\text{ddiag}(\text{Re}\{\mathbf{W}^{\text{H}}\mathbf{U}\})$ is the orthogonal projection of $\mathbf{U}$ onto $\mathcal{T}_{\mathbf{W}}\mathcal{M}_{\mathbf{W}}$ [49].

To make sure that the point $\mathbf{\Sigma}$ is updated in the direction of the tangent space $\mathcal{T}_{\mathbf{\Sigma}}\mathcal{M}$ while staying on $\mathcal{M}$, a *retraction* operation, a mapping from $\mathcal{T}_{\mathbf{\Sigma}}\mathcal{M}$ to $\mathcal{M}$, is needed.

Lemma 8. *The retraction* $\mathbf{R}_{\mathbf{\Sigma}}(\bar{\mathbf{V}})$ *of* $\bar{\mathbf{V}} = \text{blkdiag}(\mathbf{v}, \mathbf{V}) \in \mathcal{T}_{\mathbf{\Sigma}}\mathcal{M}$ *is*

$$\mathbf{R}_{\mathbf{\Sigma}}(\bar{\mathbf{V}}) = \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{v}) \oplus \mathbf{R}_{\mathbf{W}}(\mathbf{V}), \tag{3.27}$$

where $\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{v}) = (\boldsymbol{\theta} + \mathbf{v}) \odot \frac{1}{|\boldsymbol{\theta}+\mathbf{v}|}$ *is the retraction of* $\mathbf{v} \in \mathcal{T}_{\boldsymbol{\theta}}\mathcal{M}_{\boldsymbol{\theta}}$ *and* $\mathbf{R}_{\mathbf{W}}(\mathbf{V}) = \frac{(\mathbf{W}+\mathbf{V})}{\|\text{ddiag}((\mathbf{W}+\mathbf{V})^{\text{H}}(\mathbf{W}+\mathbf{V}))\|_{\text{F}}}$ *is the retraction of* $\mathbf{V} \in \mathcal{T}_{\mathbf{W}}\mathcal{M}_{\mathbf{W}}$ *[49].*

To find out $\mathbf{\Sigma}$ minimizing $L(\mathbf{\Sigma})$ on the product manifold $\mathcal{M}$, we exploit the RCG method, an extension of the conjugate gradient (CG)[7] method to the Riemannian manifold. In this approach, the update equations of the conjugate direction $\mathbf{D}$ and the point $\mathbf{\Sigma}$ are given by

$$\mathbf{D}_i = -\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i) + \beta_i \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}}(\mathbf{D}_{i-1}), \tag{3.28}$$

$$\mathbf{\Sigma}_{i+1} = \mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i), \tag{3.29}$$

where $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)$ is the Riemannian gradient of $L(\mathbf{\Sigma}_i)$ at $\mathbf{\Sigma}_i$, β_i is the Fletcher-Reeves conjugate gradient parameter, and α_i is the step size [50].

We note that the RCG method is distinct from the conventional CG method in three respects: 1) the projection of the previous conjugate direction $\mathbf{D}_{i-1}$ onto the tangent space $\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}$ is needed before performing a linear combination of $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma})$ and $\mathbf{D}_{i-1}$ since they lie on two different spaces $\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}$ and $\mathcal{T}_{\mathbf{\Sigma}_{i-1}}\mathcal{M}$, 2) the Riemannian gradient $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma})$ is used instead of the Euclidean gradient $\nabla_{\mathbf{\Sigma}} L(\mathbf{\Sigma})$ since we need to find out the search direction on the tangent space of $\mathcal{M}$, and 3) the retraction is required to ensure that the updated point $\mathbf{\Sigma}_{i+1}$ lies on $\mathcal{M}$. Specifically, the Riemannian gradient $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}) = \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}}\mathcal{M}}(\nabla_{\mathbf{\Sigma}} L(\mathbf{\Sigma}))$ is obtained by projecting the Euclidean

[7]The update equation of the conventional CG method is $\mathbf{\Sigma}_{i+1} = \mathbf{\Sigma}_i + \alpha_i \mathbf{D}_i$ where α_i is the step size and $\mathbf{D}_i$ is the conjugate direction. In addition, the conjugate direction is updated as $\mathbf{D}_i = -\nabla_{\mathbf{\Sigma}} L(\mathbf{\Sigma}) + \beta_i \mathbf{D}_{i-1}$ where $\nabla_{\mathbf{\Sigma}} L(\mathbf{\Sigma})$ is the Euclidean gradient of $L(\mathbf{\Sigma})$ at $\mathbf{\Sigma}$ and β_i is conjugate update parameter.

gradient onto the tangent space. In the following lemma, we provide the closed-form expression of $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma})$ of $L(\mathbf{\Sigma})$ on $\mathcal{M}$.

Lemma 9. *The Riemannian gradient* $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma})$ *of* $L(\mathbf{\Sigma})$ *on* $\mathcal{M}$ *is*

$$\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}) = \text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\boldsymbol{\theta}) \oplus \text{grad}_{\mathcal{M}_{\mathbf{W}}} L(\mathbf{W}). \tag{3.30}$$

Specifically, the Riemannian gradient $\text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\boldsymbol{\theta})$ *of* $L(\boldsymbol{\theta})$ *on* $\mathcal{M}_{\boldsymbol{\theta}}$ *is*

$$\begin{aligned} \text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\boldsymbol{\theta}) &= \mathbf{P}_{\mathcal{T}_{\boldsymbol{\theta}}\mathcal{M}_{\boldsymbol{\theta}}}(\nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta})) \\ &= \nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}) - \text{Re}\{\boldsymbol{\theta}^* \odot \nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta})\} \odot \boldsymbol{\theta}, \end{aligned} \tag{3.31}$$

where $\nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta})$ *is the Euclidean gradient of* $L(\boldsymbol{\theta})$ *given by (see* (3.19)*)*

$$\nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}) = \sum_{k=1}^{K} \lambda_k \Big(- \frac{p_k}{2^{R_k^{\min}} - 1} \frac{\partial A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k)}{\partial \boldsymbol{\theta}} + \sum_{j \neq k}^{K} p_j \frac{\partial A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k)}{\partial \boldsymbol{\theta}} \Big), \tag{3.32}$$

and

$$\frac{\partial A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k)}{\partial \boldsymbol{\theta}} = \mathbf{H}_j^{\mathrm{H}} \mathbf{G}^{\mathrm{H}} \mathbf{w}_k \mathbf{w}_k^{\mathrm{H}} (\mathbf{d}_j + \mathbf{G}\mathbf{H}_j \boldsymbol{\theta}). \tag{3.33}$$

Also, the Riemannian gradient $\text{grad}_{\mathcal{M}_{\mathbf{W}}} L(\mathbf{W})$ *of* $L(\mathbf{W})$ *on* $\mathcal{M}_{\mathbf{W}}$ *is*

$$\text{grad}_{\mathcal{M}_{\mathbf{W}}} L(\mathbf{W}) = \mathbf{P}_{\mathcal{T}_{\mathbf{W}}\mathcal{M}_{\mathbf{W}}}(\nabla_{\mathbf{W}} L(\mathbf{W})) \tag{3.34}$$

$$= \nabla_{\mathbf{W}} L(\mathbf{W}) - \mathbf{W}\text{ddiag}(\text{Re}\{\mathbf{W}^H \nabla_{\mathbf{W}} L(\mathbf{W})\}), \tag{3.35}$$

where $\nabla_{\mathbf{W}} L(\mathbf{W}) = \left[\frac{\partial L(\mathbf{W})}{\partial \mathbf{w}_1}, \cdots, \frac{\partial L(\mathbf{W})}{\partial \mathbf{w}_K}\right]$ *is the Euclidean gradient of* $L(\mathbf{W})$ *given by (see* (3.19)*)*

$$\frac{\partial L(\mathbf{W})}{\partial \mathbf{w}_k} = \lambda_k \Big(\frac{-p_k}{2^{R_k^{\min}} - 1} \frac{\partial A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k)}{\partial \mathbf{w}_k} + \sum_{j \neq k}^{K} p_j \frac{\partial A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k)}{\partial \mathbf{w}_k} \Big), \tag{3.36}$$

and

$$\frac{\partial A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k)}{\partial \mathbf{w}_k} = (\mathbf{d}_j + \mathbf{G}\mathbf{H}_j \boldsymbol{\theta})(\mathbf{d}_j + \mathbf{G}\mathbf{H}_j \boldsymbol{\theta})^{\mathrm{H}} \mathbf{w}_k. \tag{3.37}$$

Once $\mathbf{\Sigma}$ is determined, we can obtain the phase shift vector $\boldsymbol{\theta}$ and beamforming matrix $\mathbf{W}$ by decomposing $\mathbf{\Sigma}$. The optimization steps of $\boldsymbol{\theta}$ and $\mathbf{W}$ are summarized in Algorithm 2.

Lagrangian Multiplier Update

After updating the RIS phase shift vector $\boldsymbol{\theta}$ and the BS beamforming matrix $\mathbf{W}$, we update the Lagrangian multiplier $\boldsymbol{\lambda}$. To be specific, $\boldsymbol{\lambda}$ is updated in the direction of maximizing the dual function $D(\boldsymbol{\lambda}) = \min_{\boldsymbol{\theta},\mathbf{W}} L(\boldsymbol{\theta}, \mathbf{W}, \boldsymbol{\lambda})$ as

$$\boldsymbol{\lambda} = \arg\max_{\boldsymbol{\lambda}\succeq\mathbf{0}} D(\boldsymbol{\lambda}). \tag{3.38}$$

Since $D(\boldsymbol{\lambda})$ is the optimal value of the optimization problem $\mathcal{P}_3$, we cannot take derivative with respect to $\boldsymbol{\lambda}$, meaning that we cannot directly use the conventional gradient ascent method in finding out the optimal value of $\boldsymbol{\lambda}$.

To address this problem, we use the subgradient method, a generalized concept of gradient method for the non-smooth convex functions [51]. In particular, the subgradient $\mathbf{g} = [g_1, \cdots, g_K]^{\mathrm{T}}$ of the dual function $D(\boldsymbol{\lambda})$ is given by

$$g_k = -\frac{p_k}{2^{R_k^{\min}} - 1} A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sum_{j\neq k}^{K} p_j A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sigma_k^2. \tag{3.39}$$

Using g_k, we obtain the update equation of $\boldsymbol{\lambda}$:

$$\boldsymbol{\lambda}_{i+1} = \max(\boldsymbol{\lambda}_i + \eta_i \mathbf{g}_i,\ \mathbf{0}), \tag{3.40}$$

where η_i is the step size [51].

3.3.3 Uplink Transmit Power Minimization

Once the RIS phase shift vector $\boldsymbol{\theta}$ and the BS beamforming matrix $\mathbf{W}$ are obtained, we next find out the device power vector $\mathbf{p}$ minimizing the total uplink transmit power. When $\boldsymbol{\theta}$ and $\mathbf{W}$ are given, the original uplink transmit power minimization problem $\mathcal{P}_1$ is reduced to

$$\mathcal{P}_4 : \min_{\mathbf{p}} \sum_{k=1}^{K} p_k \tag{3.41a}$$

$$\text{s.t. } \frac{p_k}{2^{R_k^{\min}} - 1} A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k) - \sum_{j\neq k}^{K} p_j A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k) \geq \sigma_k^2, \quad \forall k \in \mathcal{K}, \tag{3.41b}$$

$$0 \leq p_k \leq p_k^{\max}, \quad \forall k \in \mathcal{K}. \tag{3.41c}$$

Algorithm 2 Optimization of $(\boldsymbol{\theta}, \mathbf{W})$ on product manifold $\mathcal{M}$

Input: Tolerance ϵ, number of iterations T, $c_1 = 0.0001$, $c_2 = 0.1$

Output: $(\boldsymbol{\theta}, \mathbf{W})$

Initialize: $i = 1$, $\boldsymbol{\Sigma}_1 = \text{blkdiag}(\boldsymbol{\theta}_1, \mathbf{W}_1) \in \mathcal{M}$, $\mathbf{D}_1 = -\text{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_1)$, $\beta_1 = 0$

while $i \leq T$ **do**

  $\text{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_i) = \mathbf{P}_{\mathcal{T}_{\boldsymbol{\Sigma}_i}\mathcal{M}}(\nabla_{\boldsymbol{\Sigma}} L(\boldsymbol{\Sigma}_{\mathbf{i}}))$

  $\beta_i = \frac{\|\text{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_i)\|^2}{\|\text{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_{i-1})\|^2}$

  $\mathbf{D}_i = -\text{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_i) + \beta_i \mathbf{P}_{\mathcal{T}_{\boldsymbol{\Sigma}_i}\mathcal{M}}(\mathbf{D}_{i-1})$

  Find a step size $\alpha_i > 0$ such that

  $L(\mathbf{R}_{\boldsymbol{\Sigma}_i}(\alpha_i \mathbf{D}_i)) \leq L(\boldsymbol{\Sigma}_i) + c_1 \alpha_i \langle \text{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_i), \mathbf{D}_i \rangle$

  $|\langle \text{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}_i}(\alpha_i \mathbf{D}_i)), \mathbf{P}_{\mathcal{T}_{\mathbf{R}_{\boldsymbol{\Sigma}_i}(\alpha_i \mathbf{D}_i)}\mathcal{M}}(\mathbf{D}_i) \rangle| \leq -c_2 \langle \text{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_i), \mathbf{D}_i \rangle$

  $\boldsymbol{\Sigma}_{i+1} = \mathbf{R}_{\boldsymbol{\Sigma}_i}(\alpha_i \mathbf{D}_i)$

  if $\|\boldsymbol{\Sigma}_{i+1} - \boldsymbol{\Sigma}_i\|^2 \leq \epsilon$ **then**

    $\boldsymbol{\theta}_{i+1} = \boldsymbol{\Sigma}_{i+1}(1 : N, 1)$

    $\mathbf{W}_{i+1} = \boldsymbol{\Sigma}_{i+1}(N + 1 : N + M, 2 : K + 1)$

    Exit from the while loop

  end

  $i = i + 1$

end

Since $\mathcal{P}_4$ is an LP optimization problem, the optimal solution can be easily obtained using the convex optimization tools (e.g., CVX [48]).

3.4 Convergence and Computational Complexity Analysis of RCG-JO

In this section, we analyze the convergence and the computational complexity of RCG-JO.

3.4.1 Convergence Analysis of RCG-JO Algorithm

Recall that the proposed RCG-JO technique consists of two major iterations: 1) inner iteration for jointly optimizing $\boldsymbol{\theta}$ and $\mathbf{W}$ on the product Riemannian manifold (Algorithm 2) and 2) outer iteration for alternately updating $(\boldsymbol{\theta}, \mathbf{W})$ and $\mathbf{p}$ (Algorithm 1). In case of Algorithm 1, due to the alternating optimization operations, it is very difficult to prove its convergence analytically so that we demonstrate its convergence from the numerical results. In case of Algorithm 2, we show that it converges to a local minimizer in this subsection.

We first explain the strong Wolfe conditions used to determine the step size α_i.

Definition 3. *A step size α_i is said to satisfy the strong Wolfe conditions, restricted to the conjugate direction $\mathbf{D}_i$, if the following two inequalities hold [49]:*

$$L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)) \leq L(\mathbf{\Sigma}_i) + c_1 \alpha_i \langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle, \tag{3.42}$$

$$|\langle \text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)), \mathbf{P}_{\mathcal{T}_{\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)}\mathcal{M}}(\mathbf{D}_i) \rangle| \leq -c_2 \langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle. \tag{3.43}$$

The first condition, known as Armijo's rule, ensures that the step size decreases the cost function $L(\mathbf{\Sigma})$ sufficiently. The second condition, known as curvature condition, ensures that the Riemannian gradient converges to zero. Note that since $\text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)) \in \mathcal{T}_{\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)}\mathcal{M}$, the second Wolfe condition (3.43) can be converted to

$$|\langle \text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)), \mathbf{D}_i \rangle| \leq -c_2 \langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle. \tag{3.44}$$

It has been shown that the step size satisfying the strong Wolfe conditions always exists if $\text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}}(\mathbf{D}))$ is Lipschitz continuous along $\mathbf{D}$ [52, Proposition 1]. In the following proposition, we show the Lipschitz continuity of the Riemannian gradient.

Proposition 2. *The objective function $L(\mathbf{\Sigma})$ is bounded below, meaning that there exists a constant $L^* \in \mathbb{R}$ such that for all $\mathbf{\Sigma} \in \mathcal{M}$, $L^* \leq L(\mathbf{\Sigma})$. Also, the Riemannian gradient grad$_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}}(\mathbf{D}))$ is Lipschitz continuous along $\mathbf{D}$. That is, for every $\mathbf{\Sigma} \in \mathcal{M}$, there exists a $K > 0$ such that for all $\mathbf{D} \in \mathcal{T}_{\mathbf{\Sigma}}\mathcal{M}$,*

$$\|\text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}}(\mathbf{D})) - \text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}}(\mathbf{0}))\| \leq K \|\mathbf{D}\|. \tag{3.45}$$

Proof. See Appendix A. $\square$

Since the Riemannian gradient does not converge if its derivative is unbounded, the Lipschitz continuity of the Riemannian gradient is crucial for the convergence of RCG method. Using the strong Wolfe conditions and the Lipschitz continuity of Riemannian gradient, we can show that the angle between the Riemannian gradient and the conjugate direction is bounded.

Theorem 1. *(Zoutendijk condition) Let $\mathbf{\Sigma}_i \in \mathcal{M}$ be the point and $\mathbf{D}_i \in \mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}$ be the conjugate direction of i-th RCG iteration. Then*

$$\sum_{i=1}^{\infty} \frac{\langle \mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle^2}{\|\mathbf{D}_i\|^2} < \infty. \tag{3.46}$$

Proof. See Appendix B. $\square$

Next, in the following proposition, we prove that the inner product of Riemannian gradient $\mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma})$ and the conjugate direction $\mathbf{D}_i$ is bounded.

Proposition 3. *Let $\mathbf{\Sigma}_i \in \mathcal{M}$ be the point and $\mathbf{D}_i \in \mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}$ be the conjugate direction of i-th RCG iteration with the conjugate gradient parameter $\beta_i = \frac{\|\mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2}{\|\mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1})\|^2}$. If the step size α_i satisfies the strong Wolfe conditions* (3.42) *and* (3.43) *with $c_2 < 1/2$, then for every $i \in \mathbb{N}$,*

$$-\frac{1}{1-c_2} \leq \frac{\langle \mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle}{\|\mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2} \leq \frac{2c_2 - 1}{1-c_2}. \tag{3.47}$$

Proof. See Appendix C. $\square$

Finally, by combining Theorem 1 and Proposition 2, we show the convergence of RCG method.

Theorem 2. *(Convergence of RCG method) Let $\mathbf{\Sigma}_i \in \mathcal{M}$ be the point and $\mathbf{D}_i \in \mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}$ be the conjugate direction of i-th RCG iteration with the conjugate gradient parameter $\beta_i = \frac{\|\mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2}{\|\mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1})\|^2}$. If the step size α_i satisfies the strong Wolfe conditions* (3.42) *and* (3.43) *with $c_2 < 1/2$, then*

$$\liminf_{i \to \infty} \|\mathrm{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\| = 0. \tag{3.48}$$

Proof. See Appendix D. □

Remark 2. *Theorem 2 shows that the RCG iteration of proposed scheme converges to a local minimizer $\mathbf{\Sigma}^*$ of objective function $L(\mathbf{\Sigma})$ on the product manifold $\mathcal{M}$.*

3.4.2 Computational Complexity Analysis of RCG-JO Algorithm

In our analysis, we measure the complexity in terms of the number of floating point operations (flops). We first provide the complexity analysis of the joint optimization of $\boldsymbol{\theta}$ and $\mathbf{W}$ in Algorithm 2.

Lemma 10. *The total computational complexity $\mathcal{C}_{(\boldsymbol{\theta},\mathbf{W})}$ of* Algorithm 2 *is given by*

$$\mathcal{C}_{(\boldsymbol{\theta},\mathbf{W})} = \mathcal{O}(K^2N^2M + K^2N^3 + K^2M^2). \tag{3.49}$$

Proof. See Appendix E. □

After updating the RIS phase shift vector $\boldsymbol{\theta}$ and the BS beamforming matrix $\mathbf{W}$, we update the Lagrangian multiplier $\boldsymbol{\lambda}$ using the subgradient method. The complexity of updating $\boldsymbol{\lambda}$ is $\mathcal{C}_{\boldsymbol{\lambda}} = \mathcal{O}(K^2N^2M)$.

Once $\boldsymbol{\theta}$, $\mathbf{W}$, and $\boldsymbol{\lambda}$ are updated, the optimization of the uplink power vector $\mathbf{p}$ is achieved by solving an LP problem in (3.41). Note that the numbers of flops to compute $A_{j,k}(\boldsymbol{\theta},\mathbf{W})$ and to solve the LP problem are N^2M and K^3, respectively. Thus, the overall complexity to optimize $\mathbf{p}$ is $\mathcal{C}_{\mathbf{p}} = \mathcal{O}(K^2N^2M + K^3)$.

In conclusion, the computational complexity $\mathcal{C}_{\text{RCG-JO}}$ of the proposed RCG-JO scheme is

$$\begin{aligned}\mathcal{C}_{\text{RCG-JO}} &= \mathcal{C}_{(\boldsymbol{\theta},\mathbf{W})} + \mathcal{C}_{\boldsymbol{\lambda}} + \mathcal{C}_{\mathbf{p}} \\ &= \mathcal{O}(K^2N^2M + K^3 + K^2N^3 + K^2M^2). \end{aligned} \tag{3.50}$$

For comparison, we also discuss the complexity of the conventional SDR-based scheme, which consists of three major steps: 1) optimization of $\boldsymbol{\theta}$ using SDR for fixed $\mathbf{w}$ and $\mathbf{p}$, 2) optimization of $\mathbf{w}$ using SDR for fixed $\boldsymbol{\theta}$ and $\mathbf{p}$, and 3) optimization of $\mathbf{p}$

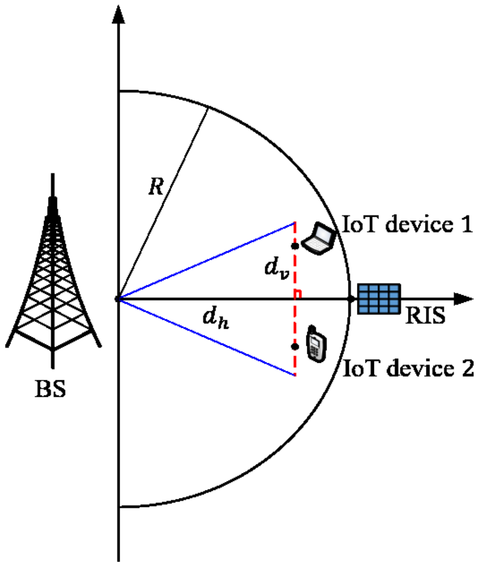

Figure 3.4: Locations of BS, RIS, and IoT devices for $K = 2$.

by solving an LP problem with the obtained $\boldsymbol{\theta}$ and $\mathbf{W}$. It has been shown that the worst-case complexity of the SDR method for optimizing $\boldsymbol{\theta} \in \mathbb{C}^{N\times 1}$ is $\mathcal{O}(N^6)$ [35]. Similarly, the worst-case complexity of optimizing $\mathbf{W} \in \mathbb{C}^{M\times K}$ using SDR is $\mathcal{O}(K^6M^6)$. In addition, the optimization of $\mathbf{p}$ is achieved by solving the LP problem, so that the complexity is $\mathcal{O}(K^2N^2M + K^3)$. In summary, the overall complexity $\mathcal{C}_{\text{SDR}}$ of the SDR-based scheme is

$$\mathcal{C}_{\text{SDR}} = \mathcal{O}(K^2N^2M + N^6 + K^6M^6). \tag{3.51}$$

3.5 Simulation Results

3.5.1 Simulation Setup

In this section, we present the numerical results to evaluate the performance of proposed RCG-JO algorithm. Our simulation setup is based on the uplink IoT network where $K = 2$ single-antenna devices transmit signals to the BS equipped with $M = 4$ receiving antennas (see Fig. 3.4). This uplink transmission is assisted by the

RIS equipped with $N = 64$ reflecting elements, which is randomly located at the circumference of the circle centered at the BS with the radius $R = 65\,\text{m}$. Also, K devices are uniformly distributed on a line that is perpendicular to the line connecting the BS and the RIS. The maximal vertical distance between the IoT devices and the line is $d_v = 3\,\text{m}$. We set the horizontal distance between the BS and IoT devices to $d_h = 57\,\text{m}$. Throughout the simulations, we set the rate requirement $R_k^{\min}$ and noise power σ_k^2 to $0.3\,\text{bps/Hz}$ and $-60\,\text{dBm}$, respectively.

We use the Rician fading channel models for $\mathbf{d}_k$, $\mathbf{u}_k$ and $\mathbf{G}$ with the Rician factors $\kappa_{\mathbf{d}} = 0$, $\kappa_k = 10$, and $\kappa_{\mathbf{G}} = \infty$, respectively. First, the channel matrix from the RIS to BS $\mathbf{G}$ is given by

$$\mathbf{G} = \sqrt{\beta_{\text{G}}}\Big(\sqrt{\frac{\kappa_{\text{G}}}{\kappa_{\text{G}}+1}}\mathbf{G}^{\text{LoS}} + \sqrt{\frac{1}{\kappa_{\text{G}}+1}}\mathbf{G}^{\text{NLoS}}\Big), \tag{3.52}$$

where $\mathbf{G}^{\text{LoS}} = \mathbf{a}_{\text{BS}}(\vartheta_{\mathbf{G}})\,\mathbf{a}_{\text{RIS}}^{\text{H}}(\psi_{\mathbf{G}}, \varphi_{\mathbf{G}})$ is the LoS component with $\vartheta_{\mathbf{G}}$, $\psi_{\mathbf{G}}$, and $\varphi_{\mathbf{G}}$ being the angle of arrival (AoA) at the BS, the azimuth and elevation of the angles of departure (AoDs) at the RIS, respectively, $\mathbf{G}^{\text{NLoS}}$ is the NLoS component generated from complex Gaussian distribution, β_{G} is the path loss between RIS and BS, and $\kappa_{\text{G}}(\geq 0)$ is the Rician factor. Here, $\mathbf{a}_{\text{BS}}(\vartheta_{\mathbf{G}}) = [1, \cdots, e^{-j\pi(M-1)\sin\vartheta_{\mathbf{G}}}]^{\text{T}} \in \mathbb{C}^{M\times 1}$ is the BS array steering vector and $\mathbf{a}_{\text{RIS}}(\psi_{\mathbf{G}}, \varphi_{\mathbf{G}}) = \mathbf{a}_{\text{RIS},x}(\psi_{\mathbf{G}}, \varphi_{\mathbf{G}}) \otimes \mathbf{a}_{\text{RIS},y}(\psi_{\mathbf{G}}, \varphi_{\mathbf{G}}) \in \mathbb{C}^{N\times 1}$ is the RIS array steering vector where $\mathbf{a}_{\text{RIS},x}(\psi_{\mathbf{G}}, \varphi_{\mathbf{G}}) = [1, \cdots, e^{-j\pi(N_x-1)\cos\psi_{\mathbf{G}}\sin\varphi_{\mathbf{G}}}]^{\text{T}}$ and $\mathbf{a}_{\text{RIS},y}(\psi_{\mathbf{G}}, \varphi_{\mathbf{G}}) = [1, \cdots, e^{-j\pi(N_y-1)\sin\psi_{\mathbf{G}}\sin\varphi_{\mathbf{G}}}]^{\text{T}}$. Second, the channel vector from the k-th IoT device to the RIS $\mathbf{u}_k$ is expressed as

$$\mathbf{u}_k = \sqrt{\beta_k}\Big(\sqrt{\frac{\kappa_k}{\kappa_k+1}}\mathbf{u}_k^{\text{LoS}} + \sqrt{\frac{1}{\kappa_k+1}}\mathbf{u}_k^{\text{NLoS}}\Big), \tag{3.53}$$

where $\mathbf{u}_k^{\text{LoS}} = \mathbf{a}_{\text{RIS}}(\psi_k, \varphi_k)$ is the LoS component with ψ_k and φ_k being the azimuth and elevation of the AoAs at the RIS, respectively, $\mathbf{u}_k^{\text{NLoS}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ is the NLoS component, β_k is the path loss, and $\kappa_k(\geq 0)$ is the Rician factor. The channel vector from the k-th IoT device to BS $\mathbf{d}_k$ is modeled similarly with $\mathbf{u}_k$. In general, RIS is deployed at the position where the LoS links with the BS and IoT devices are guaranteed, so $\mathbf{G}$ and $\mathbf{u}_k$ are dominated by the LoS paths. The path loss is $\beta = C_0(d/D_0)^{-\alpha}$, where

Table 3.1: System parameters

Parameters	**Values**
Number of devices (K)	2
Number of BS antennas (M)	4
Number of RIS reflecting elements (N)	64
BS-RIS distance (R)	65 m
BS-devices horizontal distance (d_h)	57 m
BS-devices vertical distance (d_v)	3 m
Carrier frequency (f)	2.5 GHz
Noise power (σ_k^2)	-60 dBm
Rate requirement of devices ($R_k^{\min}$)	0.3 bps/Hz
Maximum transmission power ($p_k^{\max}$)	1 W

d is the distance, α is the path loss exponent, and $C_0 = -30\,\mathrm{dB}$ is the path loss at the reference distance $D_0 = 1\,\mathrm{m}$ [35]. For the channels $\mathbf{d}_k$, $\mathbf{u}_k$ and $\mathbf{G}$, we set α to be 3.8, 2.8, and 2, respectively. In each point of the plots, the simulation results are averaged over $1,000$ independent channel realizations. The simulation parameters are summarized in Table II.

For comparison, we test the following benchmark techniques: 1) SDR-based scheme where $\boldsymbol{\theta}$ and $\mathbf{W}$ are optimized alternately using SDR[8] [35], 2) difference-of-convex (DC)-based scheme where $\boldsymbol{\theta}$ and $\mathbf{W}$ are optimized alternately using the DC programming [53], 3) deep reinforcement learning (DRL)-based scheme where $\boldsymbol{\theta}$ and $\mathbf{W}$ are jointly optimized by leveraging the DRL technique [54], 4) random phase shifts where $\boldsymbol{\theta}$ is randomly generated and the maximum ratio transmission (MRT) is used for $\mathbf{W}$, and 5) conventional system without RIS where $\mathbf{W}$ is generated randomly.

[8] Note that in the SDR-based scheme, after solving the relaxed SDP problem, the Gaussian randomization technique is employed to find out the feasible rank-one solution.

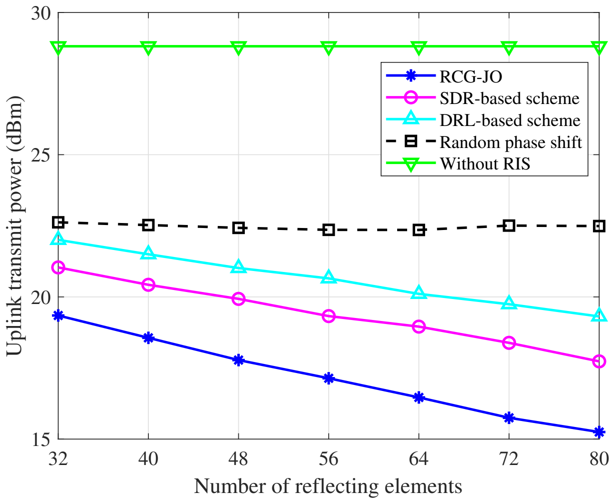

Figure 3.5: Uplink transmit power vs. number of reflecting elements N.

3.5.2 Simulation Results

In Fig. 3.5, we plot the uplink transmit power as a function of the number of RIS reflecting elements N. From the simulation results, we observe that RCG-JO outperforms the conventional schemes using SDR and random phase shifts. For example, when the number of RIS reflecting elements is $N = 64$, the proposed scheme achieves 44% and 74% reduction in power over the SDR-based scheme and the conventional scheme using random phase shifts, respectively. Also, we see that the performance gap between RCG-JO and the SDR-based scheme increases gradually with N. Furthermore, we see that the power saving gain of the conventional scheme using random phase shifts does not change with N. This is because without the optimization of RIS phase shifts and BS beamforming vectors, the RIS reflected signal power is comparable to the signal power transmitted from the direct link, so that the gain obtained from the joint active and passive beamforming is marginal. Also, we compare the uplink transmit power of RCG-JO with the DRL-based scheme. We observe that RCG-JO achieves 57% of power reduction over the DRL-based scheme. Note, for the DRL-based approach, it is not easy to find out the optimal decision (i.e., RIS phase shifts and BS beamforming

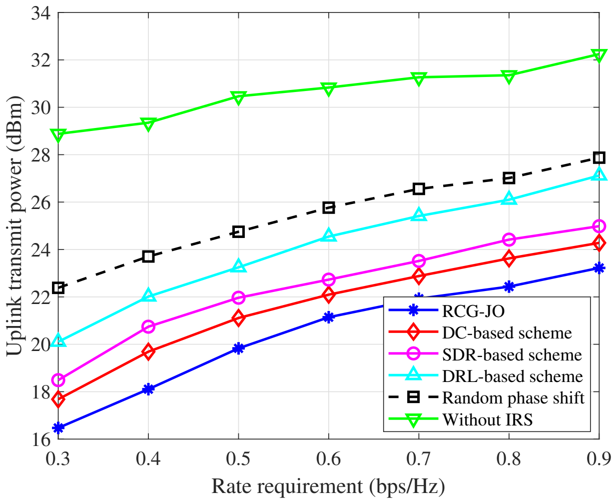

Figure 3.6: Uplink transmit power vs. rate requirement of devices $R_k^{\min}$.

vectors) minimizing the uplink transmit power and at the same time satisfying the rate requirements of the IoT devices. This is because the goal of DRL is to learn the decision policy maximizing the cumulative reward so that the minimization of uplink transmit power and the rate requirements might not be satisfied simultaneously.

We next evaluate the uplink transmit power of the proposed RCG-JO algorithm and benchmark schemes as a function of the rate requirement of IoT devices $R_k^{\min}$. As shown in Fig. 3.6, we observe that the uplink transmit power increases when the rate requirement of IoT devices becomes more strict. Also, when compared to the case without RIS, the rate requirement of IoT devices can be satisfied with lower uplink transmit power in the proposed scheme. For instance, when the rate requirement of devices is $R_k^{\min} = 0.4$ bps/Hz, RCG-JO achieves around 92% of uplink power reduction over the conventional scheme without RIS. This is because through the joint active and passive beamforming at the BS and RIS, we can improve the signal power and reduce the interference so that the rate requirements of IoT devices can be satisfied with lower transmit power. Furthermore, we see that RCG-JO significantly reduces the uplink transmit power over the benchmark schemes using SDR, DRL, and random

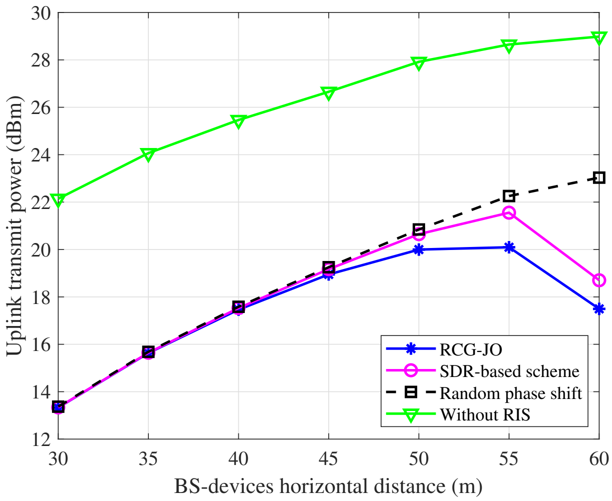

Figure 3.7: Uplink transmit power vs. BS-devices horizontal distance d_h.

phase shift. When $R_k^{\min} = 0.8\,$bps/Hz, for example, RCG-JO saves more than 65% of uplink transmit power over the conventional scheme using random phase shift. Even when compared to the DC-based scheme, RCG-JO achieves 24% of the uplink power saving gain. Note that in the DC-based scheme, due to the non-smoothness of matrix spectral norm, the subgradient method should be employed to minimize the objective function. Since the subgradient does not guarantee the steepest-descent of objective function, the performance degradation over the conventional gradient descent method is unavoidable. Whereas, in the proposed RCG-JO, we can directly minimize the objective function on the smooth Riemannian manifold using the Riemannian gradient descent method.

In Fig. 3.7, we plot the uplink transmit power as a function of the horizontal distance between the BS and IoT devices d_h. We observe that RCG-JO outperforms the benchmark schemes as d_h increases. For example, when the horizontal distance between the BS and devices is $d_h = 55\,$m, RCG-JO saves 29% and 40% of the uplink transmit power over the SDR-based scheme and the conventional scheme using random phase shift, respectively. In the system where the RIS is not employed, we see that

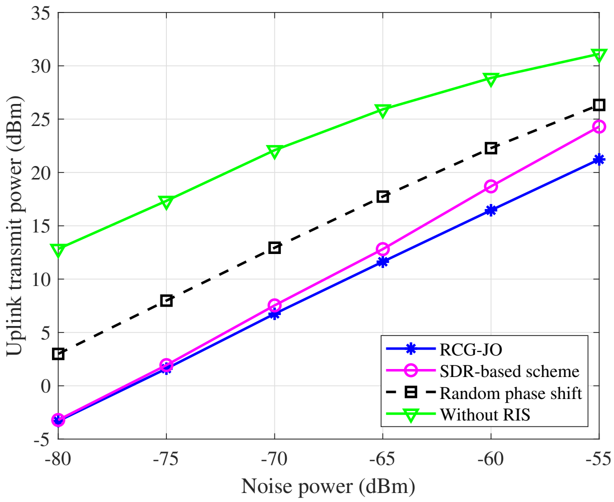

Figure 3.8: Uplink transmit power vs. noise power σ_k^2.

the uplink transmit power is considerable due to the signal attenuation, in particular when the devices are located far away from the BS. Whereas, in the RIS-aided IoT networks, we see that the uplink transmit power increases initially and then decreases as the horizontal distance increases. Basically, when the IoT device is far from the BS, large uplink transmit power is needed to satisfy the rate requirement of a device. In this case, the IoT device is getting close to the RIS so that the RIS-aided channel gain increases gradually. As a result, the rate requirement can be satisfied with relatively lower uplink transmit power.

In Fig. 3.8, we investigate the uplink transmit power of the proposed scheme and benchmark schemes as a function of the noise power σ_k^2. From the simulation results, we observe that the uplink transmit power increases with the noise power σ_k^2 and RCG-JO outperforms the benchmark schemes. For example, when the noise power $\sigma_k^2 = -55\,\text{dBm}$, the proposed scheme achieves 51% and 70% of uplink power reductions over the conventional schemes using SDR and the random phase shift, respectively. We also see that by using the RIS, the uplink transmit power of RCG-JO is significantly reduced (more than 80%) over the conventional scheme without RIS.

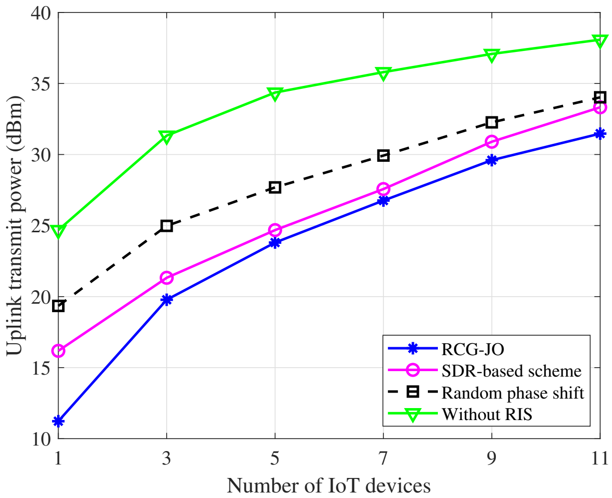

Figure 3.9: Uplink transmit power vs. number of IoT devices K.

In Fig. 3.9, we evaluate the uplink transmit power of RCG-JO and benchmark schemes as a function of the number of IoT devices K. We observe that RCG-JO outperforms the benchmark schemes in all tested scenarios. For example, when $K = 11$, RCG-JO saves 35% and 78% of the uplink transmit power over the SDR-based scheme and the conventional scheme without RIS, respectively. We also see that the power saving gain of RCG-JO over the SDR-based scheme increases with K. For instance, when we change K from 5 to 11, the power saving gain of RCG-JO over the SDR-based scheme increases from 18% to 35%. While RCG-JO solves the unconstrained optimization problem on the product manifold, the SDR-based scheme needs to find out the feasible rank-one solution after solving the SDP problem so that the performance degradation is inevitable. This, together with the result of computational complexity analysis in Section 3.4.2, demonstrates the effectiveness of RCG-JO.

To evaluate the effectiveness of proposed channel estimation technique, we investigate the uplink transmit power of RCG-JO using the perfect channel information, the estimated channel information, and the sampled channel information when $N = 100$. In Fig. 3.10, we observe that when the percentage of active reflecting elements is larger

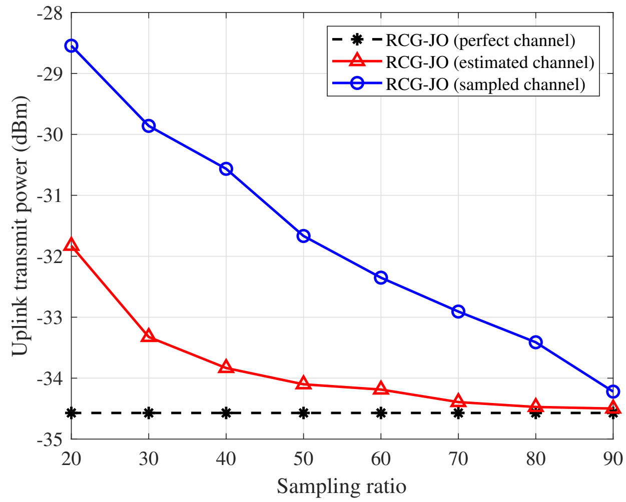

Figure 3.10: Uplink transmit power vs. percentage of active reflecting elements.

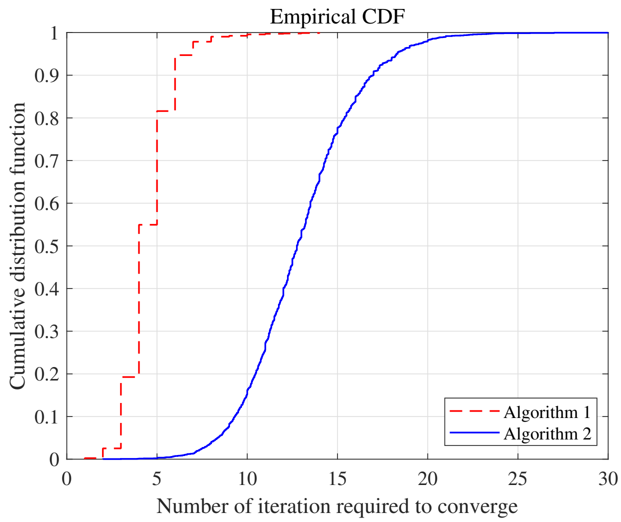

Figure 3.11: Cumulative distribution of the number of iterations required to converge.

than 20%, RCG-JO using the estimated channel information performs close to RCG-JO using the genie channel information. This is because the RIS-aided channel matrix can be readily modeled as a low-rank matrix and thus the LRMC algorithm can effectively reconstruct the full channel matrix (see Section 3.2.2).

In Fig. 3.11, we investigate the cumulative distributions of the number of iterations required for the convergence of outer iteration (Algorithm 1) and inner iteration (Algorithm 2). We test the number of iterations required for the convergence when using $10,000$ independent channel realizations. In all tested cases, we observe that Algorithm 1 converges within 15 iterations and Algorithm 2 converges within 30 iterations.

3.6 Summary

In this chapter, we proposed an RIS phase shift and BS beamforming optimization technique to minimize the uplink transmit power of an RIS-aided IoT network. Key idea of the proposed RCG-JO algorithm is to jointly optimize the RIS phase shifts and BS beamforming vectors using the Riemannian conjugate gradient method. By exploiting the product Riemannian manifold structure of the sets of unit-modulus RIS phase shift and unit-norm BS beamforming vector, we converted the uplink power minimization problem to an unconstrained problem on the Riemannian manifold. Then, we employed the Riemannian conjugate gradient method to find out the optimal RIS phase shifts and the BS beamforming vectors simultaneously. We demonstrated from the performance analysis and numerical evaluations that the proposed RCG-JO algorithm is effective in saving the uplink transmit power of RIS-aided IoT networks. In our work, we assumed the ideal phase shift model where the reflection amplitude and phase shift are independent, but an extension to the realistic scenarios with the phase-dependent reflection amplitude would also be an interesting research direction worth pursuing.

3.7 Proofs

3.7.1 Proof of Proposition 1

One can easily see that due to the unit-modulus constraint of $\boldsymbol{\theta}$ (3.12c) and the unit-norm constraint of $\mathbf{W}$ (3.12d), the objective function $L(\boldsymbol{\Sigma})$ is bounded.

Recall that the Riemannian gradient can be decomposed as $\text{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}}(\bar{\mathbf{D}}) = \text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d})) \oplus \text{grad}_{\mathcal{M}_{\mathbf{W}}} L(\mathbf{R}_{\mathbf{W}}(\mathbf{D}))$ where $\bar{\mathbf{D}} = \mathbf{d} \oplus \mathbf{D}$. Thus, we firstly find out the constants $K_{\boldsymbol{\theta}}$ and $K_{\mathbf{W}}$ satisfying the Lipschitz conditions for $\text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}))$ and $\text{grad}_{\mathcal{M}_{\mathbf{W}}} L(\mathbf{R}_{\mathbf{W}}(\mathbf{D}))$, respectively. We then obtain the constant $K = \max(K_{\boldsymbol{\theta}}, K_{\mathbf{W}})$ satisfying the Lipschitz condition for $\text{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}}(\bar{\mathbf{D}}))$.

When $\mathbf{W}$ is given, $L(\boldsymbol{\theta})$ is a quadratic function with respect to $\boldsymbol{\theta}$:

$$\begin{aligned} L(\boldsymbol{\theta}) &= \sum_{k=1}^{K} \lambda_k \Big(- \frac{p_k}{2^{R_k^{\min}} - 1} A_{k,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sum_{j \neq k}^{K} p_j A_{j,k}(\boldsymbol{\theta}, \mathbf{w}_k) + \sigma_k^2 \Big) \\ &= \boldsymbol{\theta}^{\mathrm{H}} \mathbf{B} \boldsymbol{\theta} + \mathbf{b}^{\mathrm{H}} \boldsymbol{\theta} + \boldsymbol{\theta}^{\mathrm{H}} \mathbf{b} + b, \end{aligned} \tag{3.54}$$

where $\mathbf{B} = \sum_{k=1}^{K} \lambda_k \big(- \frac{p_k}{2^{R_k^{\min}} - 1} \mathbf{H}_k^{\mathrm{H}} \mathbf{G}^{\mathrm{H}} \mathbf{w}_k \mathbf{w}_k^{\mathrm{H}} \mathbf{G} \mathbf{H}_k + \sum_{j \neq k}^{K} p_j \mathbf{H}_j^{\mathrm{H}} \mathbf{G}^{\mathrm{H}} \mathbf{w}_k \mathbf{w}_k^{\mathrm{H}} \mathbf{G} \mathbf{H}_j \big)$ and $\mathbf{b} = \sum_{k=1}^{K} \lambda_k \big(- \frac{p_k}{2^{R_k^{\min}} - 1} \mathbf{H}_k^{\mathrm{H}} \mathbf{G}^{\mathrm{H}} \mathbf{w}_k \mathbf{w}_k^{\mathrm{H}} \mathbf{d}_k + \sum_{j \neq k}^{K} p_j \mathbf{H}_j^{\mathrm{H}} \mathbf{G}^{\mathrm{H}} \mathbf{w}_k \mathbf{w}_k^{\mathrm{H}} \mathbf{d}_j \big)$. Then $\text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}))$ is expressed as

$$\begin{aligned} \text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d})) &= \mathbf{P}_{\mathcal{T}_{\boldsymbol{\theta}} \mathcal{M}_{\boldsymbol{\theta}}} (\nabla_{\boldsymbol{\theta}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}))) \\ &= \mathbf{P}_{\mathcal{T}_{\boldsymbol{\theta}} \mathcal{M}_{\boldsymbol{\theta}}} (\mathbf{B} \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}) + \mathbf{b}) \\ &= \mathbf{B} \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}) + \mathbf{b} - \mathrm{Re}\{\boldsymbol{\theta} \odot (\mathbf{B} \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}) + \mathbf{b})\} \odot \boldsymbol{\theta}. \end{aligned} \tag{3.55}$$

Using the triangle inequality, we have

$$\begin{aligned} &\|\text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d})) - \text{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{0}))\| \\ &\leq \|\mathbf{B}(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}) - \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{0}))\| + \|\,\mathrm{Re}\{\boldsymbol{\theta} \odot (\mathbf{B}(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}) - \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{0})))\} \odot \boldsymbol{\theta}\| \\ &\leq 2\|\mathbf{B}(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}) - \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{0}))\| \\ &\leq 2\|\mathbf{B}\|\|\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}) - \mathbf{R}_{\boldsymbol{\theta}}(\mathbf{0})\|. \end{aligned} \tag{3.56}$$

Since $\mathbf{B}$ is a quadratic function of $\mathbf{W}$ and the elements of $\mathbf{W}$ are bounded by the unit-norm constraints, $\|\mathbf{B}\|$ is bounded on $\mathbf{W} \in \mathcal{M}_{\mathbf{W}}$.

Now, what we need to show is that the retraction operator $\mathbf{R}_{\boldsymbol{\theta}}$ is Lipschitz continuous. To do so, we prove that each element of $\mathbf{R}_{\boldsymbol{\theta}}$ is Lipschitz continuous. Let θ_n and d_n be the n-th elements of $\boldsymbol{\theta}$ and $\mathbf{d}$, respectively. Since $|\theta_n| = 1$, we have

$$\begin{aligned} |R_{\theta_n}(d_n) - R_{\theta_n}(0)| &= \left| \frac{\theta_n + d_n}{|\theta_n + d_n|} - \theta_n \right| \\ &= |\theta_n| \left| \frac{1 + d_n\theta_n^{-1}}{|\theta_n||1 + d_n\theta_n^{-1}|} - 1 \right| \\ &= \left| \frac{1 + d_n\theta_n^{-1}}{|1 + d_n\theta_n^{-1}|} - 1 \right|. \end{aligned} \tag{3.57}$$

Without the loss of generality, we assume that $\theta_n = 1$. If $|d_n| \geq \frac{1}{2}$, then we have

$$\left| \frac{1 + d_n}{|1 + d_n|} - 1 \right| \leq \left| \frac{1 + d_n}{|1 + d_n|} \right| + 1 = 2 \leq 4|d_n|. \tag{3.58}$$

Also, if $|d_n| < \frac{1}{2}$, then we have

$$\begin{aligned} \left| \frac{1 + d_n}{|1 + d_n|} - 1 \right|^2 &= \frac{|1 + d_n| - \mathrm{Re}\{1 + d_n\}}{|1 + d_n|} \\ &= \frac{2\,\mathrm{Im}^2\{1 + d_n\}}{|1 + d_n|(|1 + d_n| + \mathrm{Re}\{1 + d_n\})} \\ &\leq \frac{2|d_n|^2}{|1 + d_n|^2} \\ &\leq 8|d_n|^2. \end{aligned} \tag{3.59}$$

Combining (3.56), (3.58), and (3.59), we see that $K_{\boldsymbol{\theta}} = 8 \sup_{\mathbf{W} \in \mathcal{M}_{\mathbf{W}}} \|\mathbf{B}\|$ satisfies the Lipschitz condition for $\mathrm{grad}_{\mathcal{M}_{\boldsymbol{\theta}}} L(\mathbf{R}_{\boldsymbol{\theta}}(\mathbf{d}))$. Similarly, we can obtain $K_{\mathbf{W}}$ using the fact that the column vectors of $\mathbf{W}$ have unit-norm. Finally, we obtain the constant $K = \max(K_{\boldsymbol{\theta}}, K_{\mathbf{W}})$ satisfying the Lipschitz condition for $\mathrm{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}}(\bar{\mathbf{D}}))$.

3.7.2 Proof of Theorem 1

Using the second Wolfe condition (3.44) and the Lipschitz continuity (3.45), we have

$$\begin{aligned} (c_2 - 1)\langle \mathrm{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_i), \mathbf{D}_i \rangle &\leq \langle \mathrm{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}_i}(\alpha_i \mathbf{D}_i)) - \mathrm{grad}_{\mathcal{M}} L(\boldsymbol{\Sigma}_i), \mathbf{D}_i \rangle \\ &= \langle \mathrm{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}_i}(\alpha_i \mathbf{D}_i)) - \mathrm{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}_i}(\mathbf{0})), \mathbf{D}_i \rangle \\ &\leq \|\mathrm{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}_i}(\alpha_i \mathbf{D}_i)) - \mathrm{grad}_{\mathcal{M}} L(\mathbf{R}_{\boldsymbol{\Sigma}_i}(\mathbf{0}))\| \|\mathbf{D}_i\| \\ &\leq \alpha_i K \|\mathbf{D}_i\|^2. \end{aligned} \tag{3.60}$$

Then from the first Wolfe condition (3.42) and (3.60), we have

$$\begin{aligned}
L(\mathbf{\Sigma}_{i+1}) &= L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)) \\
&\leq L(\mathbf{\Sigma}_i) + c_1 \alpha_i \langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)), \mathbf{D}_i \rangle \\
&\leq L(\mathbf{\Sigma}_i) - c_1 \frac{1-c_2}{K} \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle^2}{\|\mathbf{D}_i\|^2}. \qquad (3.61)
\end{aligned}$$

Finally, by combining (3.61) for $i = 1, \cdots, I$, we have

$$\begin{aligned}
\sum_{i=1}^{I} \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle^2}{\|\mathbf{D}_i\|^2} &\leq \frac{K(L(\mathbf{\Sigma}_1) - L(\mathbf{\Sigma}_{I+1}))}{c_1(1-c_2)} \\
&\overset{(a)}{\leq} \frac{K(L(\mathbf{\Sigma}_1) - L^*)}{c_1(1-c_2)}, \qquad (3.62)
\end{aligned}$$

where (a) is from the fact that $L(\mathbf{\Sigma})$ is bounded below by L^*. By taking the limit $I \to \infty$ of (3.62), we obtain the desired result (3.46).

3.7.3 Proof of Proposition 2

We use the mathematical induction to prove Proposition 2. When $i = 1$, $\mathbf{D}_1 = -\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_1)$ so that we can easily see that (3.47) holds. Now suppose that (3.47) holds for $i \geq 1$. By using the update equation of $\mathbf{D}_i$ in (3.28) and the Fletcher-Reeves conjugate gradient parameter $\beta_i = \frac{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1})\|^2}$, we have

$$\begin{aligned}
\frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1}), \mathbf{D}_{i+1} \rangle}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1})\|^2} &= \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1}), -\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1}) + \beta_{i+1} \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_{i+1}} \mathcal{M}}(\mathbf{D}_i) \rangle}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1})\|^2} \\
&= -1 + \beta_{i+1} \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1}), \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_{i+1}} \mathcal{M}}(\mathbf{D}_i) \rangle}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1})\|^2} \\
&= -1 + \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)), \mathbf{D}_i \rangle}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2}. \qquad (3.63)
\end{aligned}$$

Then, using the second Wolfe condition (3.44) and (3.63), we obtain

$$\begin{aligned}
-1 + c_2 \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2} &\leq \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1}), \mathbf{D}_{i+1} \rangle}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i+1})\|^2} \\
&\leq -1 - c_2 \frac{\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{D}_i \rangle}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2}. \qquad (3.64)
\end{aligned}$$

By employing the induction hypothesis (3.47) for i, we see that (3.47) holds for $i + 1$, which establishes the proposition 2.

Appendix D

Proof of Theorem 2

Using the update equation of $\mathbf{\Sigma}_i$ in (3.29), (3.47), and the second Wolfe condition (3.43), we obtain

$$\begin{aligned}|\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}}(\mathbf{D}_{i-1})\rangle| &= |\langle \text{grad}_{\mathcal{M}} L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_{i-1}\mathbf{D}_{i-1})), \mathbf{P}_{\mathcal{T}_{\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_{i-1}\mathbf{D}_{i-1})}\mathcal{M}}(\mathbf{D}_{i-1})\rangle| \\ &\le -c_2\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1}), \mathbf{D}_{i-1}\rangle \\ &\le \frac{c_2}{1-c_2}\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1})\|^2. \qquad (3.65)\end{aligned}$$

Then from the update equation of conjugate direction $\mathbf{D}_i$ in (3.28), we have

$$\begin{aligned}\|\mathbf{D}_i\|^2 &= \|-\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i) + \beta_i \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}}(\mathbf{D}_{i-1})\|^2 \\ &= \|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2 - 2\beta_i\langle \text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i), \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}}(\mathbf{D}_{i-1})\rangle + \beta_i^2\|\mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}}(\mathbf{D}_{i-1})\|^2 \\ &\le \|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2 + \frac{2c_2}{1-c_2}\beta_i\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1})\|^2 + \beta_i^2\|\mathbf{D}_{i-1}\|^2 \\ &= \|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2 + \frac{2c_2}{1-c_2}\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2 + \frac{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^4}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1})\|^4}\|\mathbf{D}_{i-1}\|^2 \\ &= \frac{1+c_2}{1-c_2}\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2 + \frac{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^4}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1})\|^4}\|\mathbf{D}_{i-1}\|^2. \qquad (3.66)\end{aligned}$$

By sequentially applying (3.66) until $i = 1$, we obtain

$$\|\mathbf{D}_i\|^2 \le \frac{1+c_2}{1-c_2}\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^4 \sum_{j=1}^{i} \frac{1}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_j)\|^2}. \qquad (3.67)$$

If we assume $\liminf_{i\to\infty} \|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\| \neq 0$, meaning that $\epsilon(>0)$ such that $\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\| \ge \epsilon$ for all $i \in \mathcal{N}$ exists, then we have

$$\|\mathbf{D}_i\|^2 \le \frac{1+c_2}{1-c_2}\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^4 \frac{i}{\epsilon^2}. \qquad (3.68)$$

This implies that

$$\sum_{i=1}^{\infty} \frac{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^4}{\|\mathbf{D}_i\|^2} \ge \frac{1-c_2}{1+c_2}\epsilon^2 \sum_{i=1}^{\infty}\frac{1}{i} = \infty. \qquad (3.69)$$

However, Theorem 1 and Proposition 2 imply that $\sum_{i=1}^{\infty} \frac{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^4}{\|\mathbf{D}_i\|^2} < \infty$, which is a contradiction to (3.69). Therefore, we otain the desired result $\liminf_{i\to\infty} \|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\| = 0$.

3.7.4 Proof of Lemma 6

In the first step of Algorithm 2, we compute $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma})$, which is given by

$$\begin{aligned}
&\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i) \\
&= (\nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}_i) - \text{Re}\{\boldsymbol{\theta}_i^* \odot \nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}_i)\} \odot \boldsymbol{\theta}_i) \\
&\quad \oplus (\nabla_{\mathbf{W}} L(\mathbf{W}_i) - \mathbf{W}_i \, \text{ddiag}(\text{Re}\{\mathbf{W}_i^{\text{H}} \nabla_{\mathbf{W}} L(\mathbf{W}_i)\})). \qquad (3.70)
\end{aligned}$$

Note that the numbers of flops required for computing the Euclidean gradients $\nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}_i)$ and $\nabla_{\mathbf{W}} L(\mathbf{W}_i)$ are $K^2N^2M + K^2N^3$ and $K^2N^2M + K^2M^2$, respectively. Thus, the complexity $\mathcal{C}_1$ of computing $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)$ is

$$\mathcal{C}_1 = \mathcal{O}(K^2N^2M + K^2N^3 + K^2M^2). \qquad (3.71)$$

Using the Riemannian gradient $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)$, we then compute the RCG coefficient $\beta_i = \frac{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)\|^2}{\|\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_{i-1})\|^2}$. The complexity $\mathcal{C}_2$ of computing β_i is

$$\mathcal{C}_2 = \mathcal{O}(K^2N + K^2M). \qquad (3.72)$$

Then the complexity $\mathcal{C}_3$ for updating the Riemannian conjugate direction $\mathbf{D}_i = -\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i) + \beta_i \mathbf{P}_{\mathcal{T}_{\mathbf{\Sigma}_i}\mathcal{M}}(\mathbf{D}_{i-1})$ is

$$\mathcal{C}_3 = \mathcal{O}(K^2M + KN). \qquad (3.73)$$

Next, we find out the step size α_i via the line search which consists of the computation of the following elements: 1) $\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)$, 2) $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)$, 3) $L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i))$, and 4) $L(\mathbf{\Sigma}_i)$. To compute the retraction $\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i)$, the required number of flops is $K^2M + N$. To compute $\text{grad}_{\mathcal{M}} L(\mathbf{\Sigma}_i)$, the required number of flops is $K^2N^2M + K^2N^3 + K^2M^2$. Also, the numbers of flops required for computing $L(\mathbf{R}_{\mathbf{\Sigma}_i}(\alpha_i \mathbf{D}_i))$

and $L(\mathbf{\Sigma}_i)$ are both K^2N^2M. Thus, the complexity $\mathcal{C}_4$ of the line search for updating α_i is

$$\mathcal{C}_4 = \mathcal{O}(K^2N^2M + K^2N^3 + K^2M^2). \tag{3.74}$$

Finally, the complexity $\mathcal{C}_5$ of updating $\mathbf{\Sigma}_i$ is

$$\mathcal{C}_5 = \mathcal{O}(K^2M + N). \tag{3.75}$$

In summary, the complexity $\mathcal{C}_{(\boldsymbol{\theta},\mathbf{W})}$ of Algorithm 2 is

$$\begin{aligned}\mathcal{C}_{(\boldsymbol{\theta},\mathbf{W})} &= \mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3 + \mathcal{C}_4 + \mathcal{C}_5 \\ &= \mathcal{O}(K^2N^2M + K^2N^3 + K^2M^2). \end{aligned} \tag{3.76}$$

Chapter 4

Towards Reconfigurable Intelligent Surfaces-Assisted Wideband Beamforming for 6G

In this chapter, we introduce a novel frequency-dependent beamforming scheme for RIS-assisted wideband THz systems. One major challenge of the wideband THz communication is the severe array gain loss caused by the beam split effect that the path components split into different spatial directions at different subcarrier frequencies. Therefore, the conventional phase shift control and beamforming techniques cannot be directly applied to wideband THz systems. In this work, we propose a *RIS-assisted wideband beamforming (RWB)* technique maximizing the achievable data rate of the RIS-assisted wideband THz systems. Key idea of RWB is to alternately optimize the analog beamforming vector and the RIS phase shift vector by properly designing the parameters of the beamforming network such that the average data rate of the wideband THz system is maximized. We demonstrate from the numerical evaluations that RWB achieves a significant data rate gain over the conventional schemes.

4.1 Introduction

To support the exponential growth of data traffic in 6G networks, THz massive multiple-input multiple-output (MIMO) communications have received much attention

[55, 56]. By exploiting the large available spectrum in THz frequency band ($0.1 \sim 10\,\text{THz}$) and the high array gain achieved by a large number of antennas, THz massive MIMO communications can support immersive mobile services such as digital twin, holographic telepresence, and metaverse realized by XR devices [2]. One notable feature of the THz band signal is that due to the strong directivity and severe attenuation, the communication link quality relies heavily on the existence of a LoS link [57]. To deal with this problem, RIS, a planar metasurface consisting of a large number of low-cost reflecting elements, has been widely used [58]. By smartly controlling the phase shifts of reflecting elements, RIS can provide a virtual LoS link and thus compensate for the severe path loss of THz communications. To realize the full potential of RIS-assisted THz massive MIMO communications, the proper design of the beamforming and RIS phase shifts is of great importance.

Recently, various beamforming techniques for RIS-assisted THz systems have been proposed [59–61]. In [59], a joint RIS phase shifts control and beamforming scheme for multi-hop RIS-assisted THz systems has been proposed. In [60], an angular domain beamforming technique for holographic RIS-assisted THz systems has been proposed. Also, a deep learning-based hybrid beamforming technique to maximize the sum rate of RIS-assisted THz systems has been proposed in [61]. Note that these beamforming schemes generate directional beams aligned with the spatial directions of the path components, which can realize the full array gain of narrowband systems but will lead to severe array gain loss for wideband systems. In fact, in RIS-assisted wideband THz systems, the path components split into totally separated spatial directions at different subcarrier frequencies, and thus the phase-controlled beams generated by the traditional frequency-independent phase shifters can only realize high array gain around the central frequency while suffering from the severe array gain loss at most subcarrier frequencies. This phenomenon, so-called *beam split effect*, will result in a serious data rate loss and counteract the data rate gain benefiting from the bandwidth increase [20]. Therefore, to come up with a beamforming method mitigating the beam split effect is crucial for the success of RIS-assisted wideband THz systems.

An aim of this chapter is to put forth an efficient beamforming technique maximizing the achievable data rate of the RIS-assisted wideband THz system. The proposed scheme, henceforth referred to as *RIS-assisted wideband beamforming (RWB)*, alternately optimizes the analog beamforming vector and the RIS phase shift vector by properly designing the parameters (time delays, analog phase shifts, and RIS phase shifts) of the beamforming network such that the achievable data rate of the wideband THz system is maximized. To compensate for the array gain loss caused by the beam split effect, we exploit a small number of true-time delay (TDD)-based phase shifters and analog phase shifters to simultaneously generate frequency-dependent beams aligning with the spatial directions at different subcarriers. For the frequency-invariant phase shift vector, we exploit the Riemannian conjugate gradient (RCG) method to find out the phase shifts that maximize the average data rate of the RIS-assisted wideband THz systems.

We demonstrate from the numerical evaluations that RWB achieves a significant achievable data rate gain over the conventional schemes. For example, when compared with the RCG-based frequency-independent beamforming (FIB) scheme, RWB achieves more than 290% data rate gain. Even when compared with the RCG-based delay-phase precoding (DPP) scheme, RWB achieves around 12% data rate gain.

4.2 RIS-Assisted Wideband THz Systems

4.2.1 RIS-Assisted THz System Model

We consider an RIS-assisted THz OFDM system where a N-antenna BS serves a single-antenna UE. An RIS consisting of M reflecting elements is deployed to assist the downlink transmission. We assume the analog beamforming architecture at the BS where an RF chain is connected with N analog phase shifters. The number of OFDM subcarriers is S, the carrier frequency is f_c, and the bandwidth is B. Also, we assume that the direct BS-UE link is blocked by obstacles and thus the effective channel $\mathbf{h}_i \in \mathbb{C}^N$ between the BS and the UE is

$$\mathbf{h}_i = \mathbf{G}_i^{\mathrm{H}} \mathrm{diag}(\boldsymbol{\psi}) \mathbf{u}_i = \mathbf{G}_i^{\mathrm{H}} \mathrm{diag}(\mathbf{u}_i) \boldsymbol{\psi} = \mathbf{H}_i^{\mathrm{H}} \boldsymbol{\psi}, \tag{4.1}$$

where $\mathbf{G}_i \in \mathbb{C}^{M \times N}$ is the BS-RIS channel matrix, $\mathbf{u}_i \in \mathbb{C}^M$ is the RIS-UE channel vector, and $\mathbf{H}_i = \mathrm{diag}(\mathbf{u}_i)^{\mathrm{H}} \mathbf{G}_i \in \mathbb{C}^{M \times N}$ is the RIS reflected channel matrix at the i-th subcarrier. Also, $\boldsymbol{\psi} = [e^{j\psi_1}, \cdots, e^{j\psi_M}]^{\mathrm{T}} \in \mathbb{C}^M$ is the RIS phase shift vector and $\psi_m \in [0, 2\pi)$ is the phase shift of the m-th reflecting element.

As for the channel models of $\mathbf{G}_i$ and $\mathbf{u}_i$, we use the frequency-selective LoS-based THz channel models. The BS-RIS channel $\mathbf{G}_i$ at the i-th subcarrier is expressed as

$$\mathbf{G}_i = \alpha_i e^{-j2\pi f_i \tau_g} \mathbf{a}_M(\phi_i) \mathbf{a}_N^{\mathrm{H}}(\theta_i), \tag{4.2}$$

where α_i is the complex path gain, τ_g is the propagation delay, and $f_i = f_c - \frac{B}{2} + \frac{B}{S-1}(i-1)$ is the i-th subcarrier frequency. Also, $\mathbf{a}_N(\theta_i)$ and $\mathbf{a}_M(\phi_i)$ are the BS and RIS response vectors:

$$\mathbf{a}_N(\theta_i) = [1, e^{j\theta_i}, \cdots, e^{j(N-1)\theta_i}]^{\mathrm{T}}, \tag{4.3}$$

$$\mathbf{a}_M(\phi_i) = [1, e^{j\phi_i}, \cdots, e^{j(M-1)\phi_i}]^{\mathrm{T}}, \tag{4.4}$$

where θ_i and ϕ_i are the spatial directions of the BS and RIS at the i-th subcarrier, respectively, defined as

$$\theta_i = \gamma_i \pi \sin \vartheta, \quad \phi_i = \gamma_i \pi \sin \varphi, \tag{4.5}$$

where $\gamma_i = \frac{f_i}{f_c}$, $\vartheta \in [0, 2\pi)$ and $\varphi \in [0, 2\pi)$ are the AoD at the BS and the AoA at the RIS, respectively. Similarly, the RIS-UE channel $\mathbf{u}_i$ at the i-th subcarrier is expressed as

$$\mathbf{u}_i = \beta_i e^{-j2\pi f_i \tau_u} \mathbf{a}_M(\epsilon_i), \tag{4.6}$$

where β_i is the complex path gain, τ_u is the propagation delay, and $\mathbf{a}_M(\epsilon_i)$ is the RIS response vector, given by

$$\mathbf{a}_M(\epsilon_i) = [1, e^{j\epsilon_i}, \cdots, e^{j(M-1)\epsilon_i}]^{\mathrm{T}}, \tag{4.7}$$

where ϵ_i is the spatial direction of the RIS. Then the RIS reflected channel $\mathbf{H}_i$ can be expressed as

$$\mathbf{H}_i = \text{diag}(\mathbf{u}_i)^{\text{H}}\mathbf{G}_i = \alpha_i\beta_i e^{-j2\pi f_i(\tau_g-\tau_u)}\mathbf{a}_M(\phi_i-\epsilon_i)\mathbf{a}_N^{\text{H}}(\theta_i).$$

Let s_i be the transmitted symbol such that $\mathbb{E}[|s_i|^2]=1$, then the received signal y_i of UE at the i-th subcarrier is given by

$$y_i = \mathbf{h}_i^{\text{H}}\mathbf{f}_i s_i + n_i = (\mathbf{H}_i^{\text{H}}\boldsymbol{\psi})^{\text{H}}\mathbf{f}_i s_i + n_i = \boldsymbol{\psi}^{\text{H}}\mathbf{H}_i\mathbf{f}_i s_i + n_i, \tag{4.8}$$

where $\mathbf{f}_i \in \mathbb{C}^N$ is the beamforming vector and $n_i \sim \mathcal{CN}(0,\sigma_n^2)$ is the additive Gaussian noise at the i-th subcarrier. The corresponding achievable data rate R of the UE is given by

$$R = \frac{1}{S}\sum_{i=1}^{S}\log_2\Big(1+\frac{|\boldsymbol{\psi}^{\text{H}}\mathbf{H}_i\mathbf{f}_i|^2}{\sigma_n^2}\Big). \tag{4.9}$$

One can easily see that when the analog beams and RIS reflected beams align with the spatial directions at the BS and RIS, respectively, the achievable data rate is maximized. In wideband THz systems, however, the beams generated by the frequency-independent analog phase shifters and reflected by the RIS may split into different physical directions at different subcarrier frequencies, which leads to a severe data rate loss.

4.2.2 Conventional True Time Delay-based Phase shifter

To compensate for the data rate loss, the frequency-dependent beamforming techniques realized by the true-time delay (TTD)-based phase shifters have been proposed for wideband THz systems [62–64]. Essence of this approach is to change the phase of the RF signal using multiple TTDs, thereby generating frequency-dependent beams directed to distinct spatial directions at all subcarrier frequencies. As shown in Fig. 4.1, let $s_i(t)$ be the input RF signal, then an output of the n-th TDD for the i-th subcarrier signal will be $x_{i,n}(t,\tau) = s_i(t-(n-1)\tau) = s_i(t)e^{-j2\pi f_i(n-1)\tau}$. By stacking $x_{i,n}(t,\tau)$

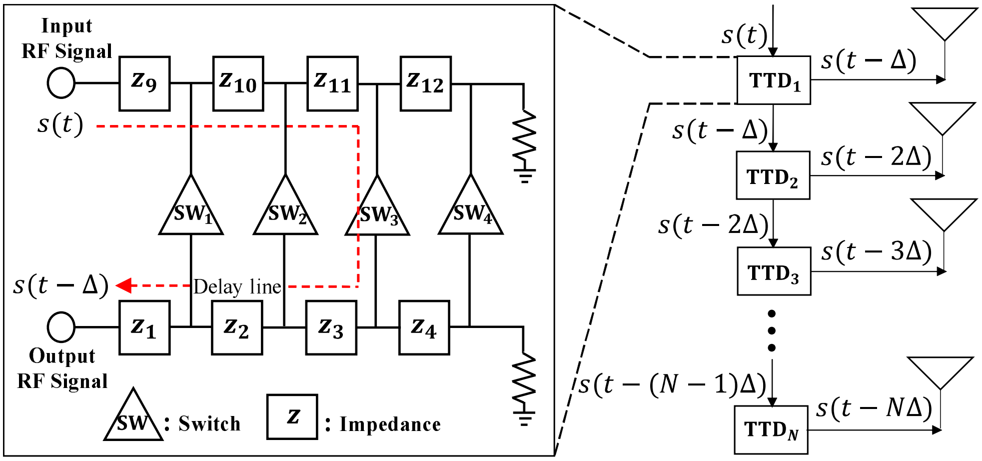

Figure 4.1: Structure of a N-TTD array.

of all N TTDs, the output $\mathbf{x}_i(t)$ of the TTD array for i-th subcarrier signal is

$$\begin{aligned}\mathbf{x}_i(t) &= [x_{i,1}(t,\tau), \cdots, x_{i,N}(t,\tau)]^{\mathrm{T}} \\ &= [s_i(t), \cdots, e^{-j2\pi f_i(N-1)\tau} s_i(t)]^{\mathrm{T}} \\ &= \mathbf{f}_i^{\mathrm{ttd}}(\tau) s_i(t). \end{aligned} \tag{4.10}$$

Note that $\mathbf{f}_i^{\mathrm{ttd}}(\tau) = [1, \cdots, e^{-j2\pi f_i(N-1)\tau}]^{\mathrm{T}} = \mathbf{a}_N(-j2\pi f_i \tau)$ is the TTD beamforming vector. By properly controlling the time delay τ, one can generate the frequency-dependent beamforming vectors $\mathbf{f}_1^{\mathrm{ttd}}(\tau), \cdots, \mathbf{f}_S^{\mathrm{ttd}}(\tau)$ heading towards the spatial directions of different subcarrier frequencies.

4.2.3 Data Rate Maximization Problem Formulation

The achievable data rate maximization problem to optimize the beamforming vector $\{\mathbf{f}_i\}_{i=1}^S$ and the RIS phase shift vector $\boldsymbol{\psi}$ is formulated as

$$\mathcal{P}_1 : \max_{\mathbf{f}_i, \boldsymbol{\psi}} \frac{1}{S} \sum_{i=1}^{S} \log_2\Big(1 + \frac{|\boldsymbol{\psi}^{\mathrm{H}} \mathbf{H}_i \mathbf{f}_i|^2}{\sigma_n^2}\Big) \tag{4.11a}$$

$$\text{s.t. } \|\mathbf{f}_i\| = 1, \quad \forall i = 1, \cdots, S, \tag{4.11b}$$

$$|\psi_m| = 1, \quad \forall m = 1, \cdots, M. \tag{4.11c}$$

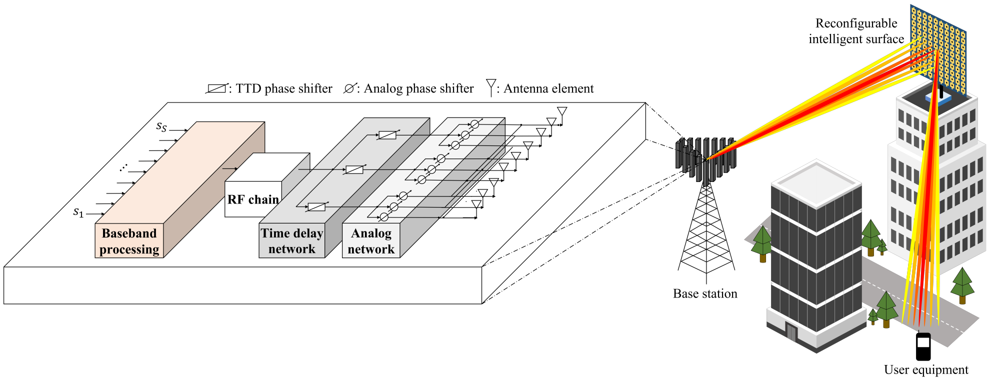

Figure 4.2: Overall structure of the RIS-assisted wideband beamforming (RWB) scheme.

Note that (4.11b) is the unit-norm constraint of BS beamforming vector and (4.11c) is the unit-modulus constraint of RIS phase shift vector. Due to the nonconvexity of the unit-norm and unit-modulus constraints, $\mathcal{P}_1$ is a nonconvex problem. This, together with the quadratic and coupled structure of the rate function in (4.11a), makes $\mathcal{P}_1$ very difficult to solve.

4.3 RIS-Assisted Wideband Beamforming

Main goal of RWB is to find out the beamforming vectors $\{\mathbf{f}_i\}_{i=1}^{S}$ and phase shift vector $\boldsymbol{\psi}$ maximizing the achievable data rate of the RIS-assisted wideband THz systems. Since $\{\mathbf{f}_i\}_{i=1}^{S}$ and $\boldsymbol{\psi}$ are coupled together in the objective function (4.11a), we optimize them in an alternating fashion. We first propose the RWB architecture that generates the frequency-dependent beams aligning with the spatial directions at all subcarriers when $\boldsymbol{\psi}$ is fixed. We then exploit the Riemannian manifold structure of the set of unit-modulus RIS phase shifts and convert $\mathcal{P}_1$ to an unconstrained problem on the Riemannian manifold. By using the Riemannian conjugate gradient (RCG) descent method, we can obtain the optimal RIS phase shift vector maximizing the achievable data rate.

4.3.1 Frequency-Dependent Beamforming Optimization

RWB Structure

The overall structure of RWB includes three main parts (see Fig. 4.2): 1) time delay network generating frequency-dependent beams using the TTD-based phase shifters, 2) analog network changing the spatial directions of the beams generated by the time delay network, and 3) intensifier network suppressing the sidelobes of the subcarrier beams generated by the time delay network and the analog network. Key ingredient of RWB is the intensifier network compensating for the difference between the RWB beam and the desired directional beam, thus improving the beamforming gain.

RWB Beamforming Vector Design

Let $\mathbf{f}_i^{\text{td}}(\tau) \in \mathbb{C}^N$, $\mathbf{f}^{\text{ana}}(\theta_c) \in \mathbb{C}^N$, and $\mathbf{f}_i^{\text{it}}(\eta) \in \mathbb{C}^N$ be the beamforming vectors generated by the time delay network, analog network, and intensifier network, respectively, then the RWB beamforming vector $\mathbf{f}_i \in \mathbb{C}^N$ can be expressed as

$$\mathbf{f}_i = \mathbf{f}_i^{\text{it}}(\eta) \odot \mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau), \quad i = 1, \cdots, S, \tag{4.12}$$

where τ and η are the time delays in the time delay network and intensifier network. By properly controlling the RWB parameters (τ, θ, η), BS can generate the RWB beams $\{\mathbf{f}_i\}_{i=1}^S$ directed to the desired directions $\{\theta_i\}_{i=1}^S$ at different subcarriers. Note that the desired direction area of the generated RWB beams is $[\theta_1, \theta_S]$ with the width $\text{Width}(\{\mathbf{f}_i^{\text{td}}\}_{i=1}^S) = \theta_S - \theta_1$ and the center $\text{Center}(\{\mathbf{f}_i^{\text{td}}\}_{i=1}^S) = \frac{1}{2}(\theta_1 + \theta_S)$.

In the time delay network, we employ $T(< N)$ TTDs each of which is connected to $P = \frac{N}{T}$ phase shifters. The time delay network beamforming vector $\mathbf{f}_i^{\text{td}}(\tau)$ at the i-th subcarrier is

$$\mathbf{f}_i^{\text{td}}(\tau) = [1, \cdots, e^{-j(T-1)2\pi f_i \tau}]^{\text{T}} \otimes \mathbf{1}_P = \mathbf{a}_T(-2\pi f_i \tau) \otimes \mathbf{1}_P. \tag{4.13}$$

Lemma 11. *The spatial direction of the time delay network beamforming vector $\mathbf{f}_i^{\text{td}}(\tau)$ at the i-th subcarrier is $-\frac{2\pi f_i \tau}{P}$.*

Proof. See Appendix A. $\square$

Using Lemma 1, one can see that the spatial directions of the first subcarrier beam $\mathbf{f}_1^{\text{td}}(\tau)$ and the last subcarrier beam $\mathbf{f}_S^{\text{td}}(\tau)$ are $-\frac{2\pi f_1 \tau}{P}$ and $-\frac{2\pi f_S \tau}{P}$, respectively. Since $\text{Width}(\{\mathbf{f}_i^{\text{td}}\}_{i=1}^S) = -\frac{2\pi\tau}{P}(f_S - f_1)$, by adjusting τ as

$$\tau = -\frac{P(\theta_S - \theta_1)}{2\pi(f_S - f_1)} = -\frac{P(\theta_S - \theta_1)}{2\pi B}, \tag{4.14}$$

we can enforce $\text{Width}(\{\mathbf{f}_i^{\text{td}}\}_{i=1}^S) = \theta_S - \theta_1$. Note that the central direction of the generated beams is $\text{Center}(\{\mathbf{f}_i^{\text{td}}\}_{i=1}^S) = -\frac{\pi(f_1+f_S)\tau}{P} = \frac{f_1+f_S}{2B}(\theta_S - \theta_1)$.

The analog network beamforming vector $\mathbf{f}^{\text{ana}}(\theta)$ is

$$\mathbf{f}^{\text{ana}}(\theta) = [1, \cdots, e^{j(N-1)\theta}]^{\text{T}} = \mathbf{a}_N(\theta). \tag{4.15}$$

Then the i-th subcarrier beam $\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau)$ generated by the time delay network and analog network is expressed as

$$\begin{aligned}
\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau) &= \mathbf{a}_N(\theta) \odot (\mathbf{a}_T(-2\pi f_i \tau) \otimes \mathbf{1}_P) \\
&\overset{\text{(a)}}{=} (\mathbf{a}_T(P\theta) \otimes \mathbf{a}_P(\theta)) \odot (\mathbf{a}_T(-2\pi f_i \tau) \otimes \mathbf{1}_P) \\
&\overset{\text{(b)}}{=} (\mathbf{a}_T(P\theta) \odot \mathbf{a}_T(-2\pi f_i \tau)) \otimes (\mathbf{a}_P(\theta) \odot \mathbf{1}_P) \\
&\overset{\text{(c)}}{=} \mathbf{a}_T(P\theta - 2\pi f_i \tau) \otimes \mathbf{a}_P(\theta) \\
&\overset{\text{(d)}}{=} \mathbf{a}_T(P\theta + \frac{P f_i}{B}(\theta_S - \theta_1)) \otimes \mathbf{a}_P(\theta),
\end{aligned} \tag{4.16}$$

where (a), (b), (c), and (d) follow from $\mathbf{a}_N(\theta) = \mathbf{a}_T(P\theta) \otimes \mathbf{a}_P(\theta)$, $(\mathbf{A} \otimes \mathbf{B}) \odot (\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A} \odot \mathbf{C}) \otimes (\mathbf{B} \odot \mathbf{D})$, $\mathbf{a}_N(\theta) \odot \mathbf{a}_N(\phi) = \mathbf{a}_N(\theta + \phi)$, and (4.14), respectively. Using Lemma 1 and (4.16), one can see that the spatial direction of $\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau)$ is $\theta_i' = \theta + \frac{f_i}{B}(\theta_S - \theta_1)$. Since $\text{Center}(\{\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}\}_{i=1}^S) = \theta + \frac{f_1+f_S}{2B}(\theta_S - \theta_1)$, by setting θ as

$$\theta = \frac{1}{2}(\theta_1 + \theta_S) - \frac{f_1 + f_S}{2B}(\theta_S - \theta_1) = \frac{f_S\theta_1 - f_1\theta_S}{B}, \tag{4.17}$$

we can enforce $\text{Center}(\{\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}\}_{i=1}^S) = \frac{1}{2}(\theta_1 + \theta_S)$.

In short, by setting τ and θ as (4.14) and (4.17), we can set the direction area of $\{\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau)\}_{s=1}^S$ to $[\theta_1, \theta_S]$. However, the generated beams suffer from a severe degradation of beamforming gain due to the sidelobe leakage. In the following lemma, we exploit the beamforming gain of the generated beams $\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau)$ to quantify the sidelobe leakage.

Lemma 12. *The beamforming gain G_i of $\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau)$ is*

$$G_i = \left| \frac{1}{N} \mathbf{a}_N^{\text{H}}(\theta_i') \big(\mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau) \big) \right|^2 = \left| \frac{1}{N} \frac{\sin 2\pi f_i \tau}{\sin \frac{2\pi f_i \tau}{P}} \right|^2 .$$

Note that G_i is a function of τ and achieves the maximum value at $\tau = 0$. However, if $\tau = 0$, the spatial directions of beams generated by the time delay network and the analog network will be the same ($\theta_i' = \theta$) so that the data rate loss is still very high. Thus, τ should not be zero, meaning that G_i cannot achieve the maximum value.

Main purpose of the intensifier network is to concentrate the signal power to the mainlobe by compensating for the difference between the RWB beam $\mathbf{f}_i$ and the desired directional beamforming vector $\mathbf{a}_N(\theta_i)$ (θ_i is the spatial direction of $\mathbf{f}_i$). Basically, the intensifier network consists of P TTDs, each of which is connected to $T = \frac{N}{P}$ analog phase shifters. The intensifier beamforming vector $\mathbf{f}_i^{\text{it}}(\eta)$ at the i-th subcarrier is

$$\mathbf{f}_i^{\text{it}}(\eta) = \mathbf{1}_T \otimes [1, \cdots, e^{-j(P-1)2\pi f_i \eta}]^{\text{T}} = \mathbf{1}_T \otimes \mathbf{a}_P(-2\pi f_i \eta).$$

Then the RWB beam $\mathbf{f}_i$ in (4.12) can be re-expressed as

$$\begin{aligned} \mathbf{f}_i &= \mathbf{f}_i^{\text{it}}(\eta) \odot \mathbf{f}^{\text{ana}}(\theta) \odot \mathbf{f}_i^{\text{td}}(\tau) \\ &= (\mathbf{1}_T \otimes \mathbf{a}_P(-2\pi f_i \eta)) \odot \big(\mathbf{a}_T(P\theta + \frac{P f_i}{B}(\theta_S - \theta_1)) \otimes \mathbf{a}_P(\theta) \big) \\ &= \mathbf{a}_T(P\theta - 2\pi f_i \tau) \otimes \mathbf{a}_P(\theta - 2\pi f_i \eta). \end{aligned} \tag{4.18}$$

From Lemma 1, one can see the spatial direction θ_i of $\mathbf{f}_i$ is

$$\theta_i = \theta_1 + \frac{i-1}{S-1}(\theta_S - \theta_1). \tag{4.19}$$

Using θ_i, $\mathbf{f}_i$ can be re-expressed as

$$\mathbf{f}_i = \mathbf{a}_T(P\theta_i) \otimes \mathbf{a}_P(\theta_i - \frac{f_i}{B}(\theta_S - \theta_1) - 2\pi f_i \eta). \tag{4.20}$$

Also, the desired directional beamforming vector $\mathbf{a}_N(\theta_i)$ is

$$\mathbf{a}_N(\theta_i) = \mathbf{a}_T(P\theta_i) \otimes \mathbf{a}_P(\theta_i). \tag{4.21}$$

Comparing (4.20) and (4.21), one can easily obtain that the time delay η satisfying $\mathbf{f}_i = \mathbf{a}_N(\theta_i)$ is given by

$$\eta = \tau/P. \tag{4.22}$$

Finally, by adjusting the RWB parameters (τ, θ, η) as (4.14), (4.17), and (4.22), one can generate the RWB beams $\{\mathbf{f}_i\}_{i=1}^S$ directed to the desired directions $\{\theta_i\}_{i=1}^S$ while achieving the maximum beamforming gain $G_i^{\text{RWB}} = \left|\frac{1}{N}\mathbf{a}_N^{\text{H}}(\theta_i)\mathbf{f}_i\right|^2 = 1$. That is, when $\boldsymbol{\psi}$ is fixed, $\{\mathbf{f}_i\}_{i=1}^S$ are the optimal beamforming vectors such that the achievable data rate is maximized.

4.3.2 RIS Phase Shift Vector Optimization

Once the beamforming vectors are obtained, we convert $\mathcal{P}_1$ to an unconstrained problem on the Riemannian manifold by exploiting the smooth manifold structure of the set of unit-modulus RIS phase shifts [65]. We then design the gradient decent algorithm on the Riemannian manifold and use it to obtain the optimal RIS phase shifts maximizing the achievable data rate of the RIS-assisted wideband THz systems.

Specifically, for the given $\{\mathbf{f}_i\}_{i=1}^S$, $\mathcal{P}_1$ is reduced to the unconstrained problem on the complex circle manifold $\mathcal{M}$:

$$\mathcal{P}_{\boldsymbol{\psi}} : \min_{\boldsymbol{\psi}\in\mathcal{M}} R(\boldsymbol{\psi}) = \frac{1}{S}\sum_{i=1}^{S} \log_2\Big(1 + \frac{|\boldsymbol{\psi}^{\text{H}}\mathbf{H}_i\mathbf{f}_i|^2}{\sigma_n^2}\Big), \tag{4.23}$$

where $\mathcal{M}=\{\boldsymbol{\psi}\in\mathbb{C}^M : |\psi_m|=1\}$ is the complex circle manifold with the inner product $g(\mathbf{z}_1,\mathbf{z}_2)=\text{Re}\{\mathbf{z}_1^{\text{H}}\mathbf{z}_2\}$ and the tangent space $\mathcal{T}_{\boldsymbol{\psi}}\mathcal{M}=\{\mathbf{z}\in\mathbb{C}^M : \text{Re}\{\mathbf{z}^* \odot \boldsymbol{\psi}\}=\mathbf{0}_M\}$ of $\mathcal{M}$ at $\boldsymbol{\psi}$.

To minimize the objective function $R(\boldsymbol{\psi})$, we need a Riemannian gradient which is obtained by projecting the Euclidean gradient of $R(\boldsymbol{\psi})$ at $\boldsymbol{\psi}$ onto $\mathcal{T}_{\boldsymbol{\psi}}\mathcal{M}$. That is,

$$\begin{aligned}\mathrm{grad}_{\mathcal{M}} R(\boldsymbol{\psi}) &= \mathbf{P}_{\mathcal{T}_{\boldsymbol{\psi}}\mathcal{M}}(\nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi})) \\ &= \nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi}) - \mathrm{Re}\{\boldsymbol{\psi}^* \odot \nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi})\} \odot \boldsymbol{\psi},\end{aligned} \tag{4.24}$$

where $\nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi}) = \frac{1}{S \ln 2} \sum_{s=1}^{S} \frac{\sigma_n^2 \mathbf{H}_s \mathbf{f}_s \mathbf{f}_s^{\mathrm{H}} \mathbf{H}_s^{\mathrm{H}} \boldsymbol{\psi}}{\sigma_n^2 + \boldsymbol{\psi}^{\mathrm{H}} \mathbf{H}_s \mathbf{f}_s \mathbf{f}_s^{\mathrm{H}} \mathbf{H}_s^{\mathrm{H}} \boldsymbol{\psi}}$ is the Euclidean gradient of $R(\boldsymbol{\psi})$ and $\mathbf{P}_{\mathcal{T}_{\boldsymbol{\psi}}\mathcal{M}}(\nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi})) = \nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi}) - \mathrm{Re}\{\boldsymbol{\psi}^* \odot \nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi})\} \odot \boldsymbol{\psi}$ is the orthogonal projection of $\nabla_{\boldsymbol{\psi}} R(\boldsymbol{\psi})$ onto $\mathcal{T}_{\boldsymbol{\psi}}\mathcal{M}$. Then the conjugate direction $\mathbf{D}$ is updated as

$$\mathbf{D}^t = -\mathrm{grad}_{\mathcal{M}} R(\boldsymbol{\psi}^t) + \beta^t \mathbf{P}_{\mathcal{T}_{\boldsymbol{\psi}^t}\mathcal{M}}(\mathbf{D}^{t-1}), \tag{4.25}$$

where β^t is the Fletcher-Reeves conjugate gradient parameter.

Finally, the optimal RIS phase shift vector $\boldsymbol{\psi}^{t+1}$ is

$$\boldsymbol{\psi}^{t+1} = \mathbf{R}_{\boldsymbol{\psi}^t}(\alpha^t \mathbf{D}^t), \tag{4.26}$$

where $\mathbf{R}_{\boldsymbol{\psi}^t}(\alpha^t \mathbf{D}^t) = (\boldsymbol{\psi}^t + \alpha^t \mathbf{D}^t) \odot \frac{1}{|\boldsymbol{\psi}^t + \alpha^t \mathbf{D}^t|}$ is the retraction operation ensuring that the point $\boldsymbol{\psi}$ is updated in the direction of the tangent space $\mathcal{T}_{\boldsymbol{\psi}}\mathcal{M}$ while staying on $\mathcal{M}$ and α^t is the step size [49].

4.4 Numerical Results

4.4.1 Simulation Setup

In this section, we present numerical results to validate the effectiveness of the proposed RWB technique. We consider the RIS-assisted wideband THz system where a BS equipped with $N = 256$ antennas serves a single-antenna UE with the aid of an RIS equipped with $M = 128$ passive reflecting elements. The RIS and UE are located randomly around the BS within the cell radius of $R = 20\,\mathrm{m}$. For both BS-RIS and UE-RIS channels, we use the wideband THz LoS channel model where the carrier frequency is $f_c = 0.1\,\mathrm{THz}$, the bandwidth is $B = 40\,\mathrm{GHz}$, and the number

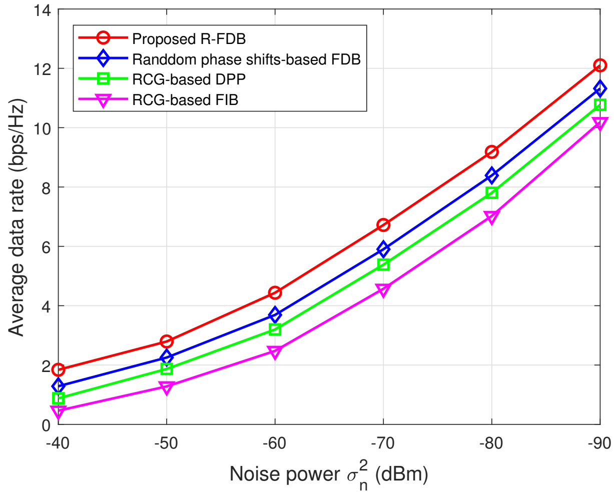

Figure 4.3: Achievable data rate vs. noise power σ_n^2.

of subcarriers is $S = 70$. The number of TTDs used in the time delay network and intensifier network are set to $T = P = \sqrt{N}$. The noise power is $\sigma_n^2 = -90\,\text{dBm}$. At each point of the figures, we test $1,000$ randomly generated RIS-assisted wideband THz systems.

For comparison, we use three benchmark schemes: 1) random phase shifts-based FDB scheme exploiting the proposed frequency-dependent beamforming scheme and randomly generated phase shifts, 2) RCG-based delay phase precoding (DPP) scheme exploiting the DPP scheme and the RCG-based phase shift design [66], 3) RCG-based frequency-independent beamforming (FIB) scheme exploiting the frequency-invariant beams and the RCG-based phase shift design.

4.4.2 Simulation Results

In Fig. 4.3, we plot the achievable data rate as a function of the noise power σ_n^2. We observe that the proposed RWB scheme outperforms the conventional beamforming schemes in all regimes under test. For example, when $\sigma_n^2 = -40\,\text{dBm}$, RWB achieves a significant rate gain (more than 290% data rate improvement) over the RCG-based FIB

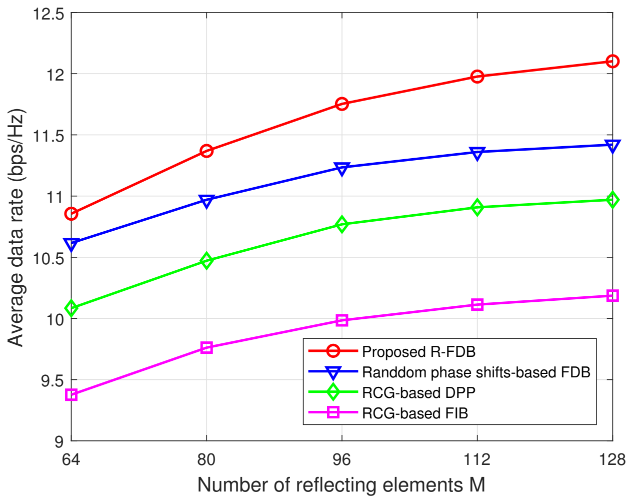

Figure 4.4: Achievable data rate vs. number of reflecting elements M.

scheme. This is because that a phase shift of the RCG-based FIB scheme relying on the analog phase shifters is invariant to the frequency so the beams for all subcarriers are directed towards the same spatial direction. Since the directions of THz subcarrier channels are distinct due to the beam split effect, a mismatch between the analog beam and the subcarrier channels is unavoidable, resulting in a considerable loss of the data rate. Whereas, in the proposed scheme, multiple frequency-dependent beams are generated using the TTD-based phase shifters so that the achievable data rate loss caused by the beam split effect can be effectively mitigated.

In Fig. 4.4, we plot the achievable data rate as a function of the number of RIS reflecting elements M. We observe that the proposed RWB achieves a significant achievable data rate gain over the benchmark schemes. For example, when $M = 128$, RWB achieves more than 2 bps/Hz data rate gain over the RCG-based FIB scheme. We also observe that the achievable data rate gain of RWB over the benchmark schemes increases with the number of reflecting elements. When $M = 64$, RWB shows around 3 bps/Hz achievable data rate gain over the random phase shifts-based FDB scheme but it increases up to 5 bps/Hz when $M = 128$. This is because when the number of

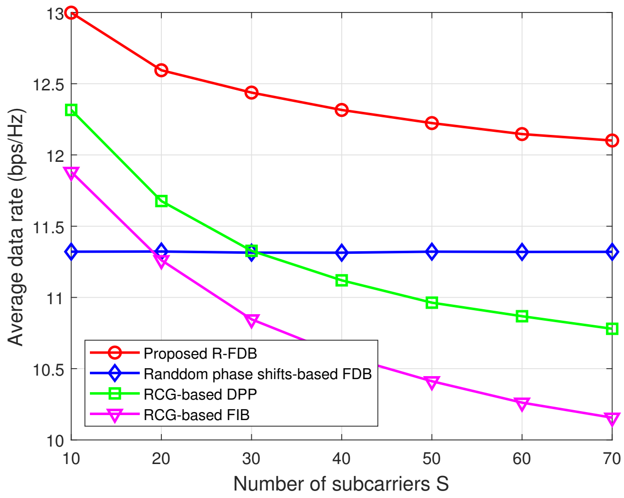

Figure 4.5: Achievable data rate vs. number of subcarriers S.

reflecting elements increases, the RIS reflected beams become sharper so that the loss of beamforming gain caused by the misalignment between the RIS reflected beams and the spatial directions at the RIS also increases.

In Fig. 4.5, we plot the achievable data rate as a function of the number of subcarriers S. We observe that the performance gain of RWB increases with the number of subcarriers. Specifically, when the number of subcarriers increases from $S = 10$ to $S = 70$, the data rate gain of RWB over the RCG-based DPP scheme increases from 6% to 12%. The reason is that when the number of subcarriers increases, the sidelobe leakage of DPP beam also increases (see Lemma 2) but such is not the case for RWB due to the suppression of the sidelobe leakage using the intensifier network.

4.5 Summary

In this chapter, we proposed a novel RIS-assisted wideband beamforming (RWB) scheme for the THz system to improve the achievable data rate. Key idea of the proposed RWB scheme is to alternately optimize the analog beamforming vector and the RIS phase shift vector by properly designing the time delays, analog phase shifts, and RIS phase shifts of the RIS-assisted frequency-dependent beamforming network such that the achievable data rate is maximized. To do so, we first exploit a small number of TTD-based phase shifters and analog phase shifters to simultaneously generate frequency-dependent beams aligning with the spatial directions at different subcarriers. We then exploit the Riemannian conjugate gradient method to optimize the phase shifts that maximize the achievable data rate of the RIS-assisted wideband THz systems. From the simulation results, we demonstrated that RWB outperforms the conventional phase shift control and beamforming schemes by a large margin.

4.6 Proofs

4.6.1 Proof of Lemma 1

The beamforming gain of the delay network beamforming vector $\mathbf{f}_i^{\mathrm{td}}(\tau)$ on the spatial direction θ_i is

$$\begin{aligned}
|\mathbf{a}_N^{\mathrm{H}}(\theta_i)\mathbf{f}_i^{\mathrm{td}}(\tau)|^2 &= \left|\sum_{t=1}^{T} e^{j(t-1)(-P\theta_i-2\pi f_i\tau)} \sum_{p=1}^{P} e^{-j(p-1)\theta_i}\right|^2 \\
&= \left|\frac{\sin\frac{T}{2}(-P\theta_i-2\pi f_i\tau)}{\sin\frac{1}{2}(-P\theta_i-2\pi f_i\tau)} \times \frac{\sin\frac{P}{2}(-\theta_i)}{\sin\frac{1}{2}(-\theta_i)}\right|^2 \\
&= \left|\Delta_T(-P\theta_i-2\pi f_i\tau)\right|^2 \times \left|\Delta_P(-\theta_i)\right|^2.
\end{aligned}$$

We can see that the beamforming gain of $\mathbf{f}_i^{\mathrm{td}}(\tau)$ is the product of two Dirichlet sinc functions (i.e., $\Delta_N(x) = \frac{\sin\frac{Nx}{2}}{\sin\frac{x}{2}}$). For $\left|\Delta_T(-P\theta_i-2\pi f_i\tau)\right|^2$, the maximum point is $\theta_{T,i}^{\max} = -\frac{2\pi f_i\tau}{P}$ by setting $P\theta_c - P\theta_i - 2\pi f_i\tau = 0$, and the mainlobe width is $\frac{4\pi}{N}$. Similarly for $\left|\Delta_P(\theta_i)\right|^2$, the maximum point is $\theta_{P,i}^{\max} = 0$ and the mainlobe width is $\frac{4\pi}{P}$. Considering that $-2\pi f_i\tau \in [-\pi,\pi]$, we obtain $\theta_{T,i}^{\max} \in [\theta_c - \frac{\pi}{P}, \theta_c + \frac{\pi}{P}]$, meaning that the maximum point of $\left|\Delta_T(-P\theta_i-2\pi f_i\tau)\right|^2$ locates in the mainlobe of $\left|\Delta_P(\theta_i)\right|^2$ with the range of $[-\frac{2\pi}{P}, \frac{2\pi}{P}]$. Then, considering that the mainlobe width of $\left|\Delta_P(\theta_i)\right|^2$ is T times wider than $\left|\Delta_T(-P\theta_i-2\pi f_i\tau)\right|^2$, the variation range of $\left|\Delta_P(\theta_i)\right|^2$ in the mainlobe of $\left|\Delta_T(-P\theta_i - 2\pi f_i\tau)\right|^2$ is much smaller than the variation range of $\left|\Delta_T(-P\theta_i - 2\pi f_i\tau)\right|^2$. Therefore, the maximum value of $|\mathbf{a}_N^{\mathrm{H}}(\theta_i)\mathbf{f}_i^{\mathrm{td}}(\tau)|^2$ can be approximated as

$$\theta_{T,i}^{\max} = \arg\max_{\theta_i} |\mathbf{a}_N^{\mathrm{H}}(\theta_i)\mathbf{f}_i^{\mathrm{td}}(\tau)|^2 = -\frac{2\pi f_i \tau}{P}. \tag{4.27}$$

Chapter 5

Conclusion

In this dissertation, THz channel acquisition schemes in the context of 6G wireless networks have been extensively studied. Specifically, we have made the following contributions.

- In Chapter 2, we introduced an efficient channel estimation technique tailored for near-field RIS-assisted wideband THz systems. The proposed PF-RCE scheme estimates the multipath components (angles, distances, and path gains) of the near-field RIS-assisted THz channel by exploiting the polar-domain sparsity and common support properties. Since the number of multipath components is much smaller than that of the RIS reflecting elements, the pilot overhead can be reduced significantly. In PF-RCE, by exploiting the polar-domain sparsity, the multipath components estimation problem is converted into the sparse recovery problem in the polar-domain. Then using the common support property, the multipath components of all subcarriers are jointly estimated via the block-sparse recovery algorithm. We demonstrated from numerical evaluations that PF-RCE can accurately estimate the near-field RIS-assisted wideband THz channel with low pilot overhead.

- In Chapter 3, we proposed an RIS phase shift and BS beamforming optimization technique to minimize the uplink transmit power of an RIS-assisted IoT network. Key idea of the proposed RCG-JO algorithm is to jointly optimize the RIS phase shifts and BS beamforming vectors using the Riemannian conjugate gradient method. By leveraging the product Riemannian manifold structure of the sets of unit-modulus RIS phase shift and unit-norm BS beamforming vector, we converted the uplink power minimization problem to an unconstrained problem on the Riemannian manifold. Then, we employed the Riemannian conjugate gradient method to find out the optimal RIS phase shifts and the BS beamforming vectors simultaneously. We demonstrated from the performance analysis and numerical evaluations that the proposed RCG-JO algorithm is effective in saving the uplink transmit power of RIS-aided IoT networks.

- In Chapter 4, we proposed a novel RIS-assisted wideband beamforming scheme for wideband THz systems, aiming at enhancing the achievable data rate. Key idea of the proposed RWB scheme is to alternately optimize the analog beamforming vector and the RIS phase shift vector by properly designing the time delays, analog phase shifts, and RIS phase shifts of the RIS-assisted frequency-dependent beamforming network such that the achievable data rate is maximized. To do so, we first exploit a small number of TTD-based phase shifters and analog phase shifters to simultaneously generate frequency-dependent beams aligning with the spatial directions at different subcarrier frequencies. We then exploit the Riemannian conjugate gradient method to optimize the phase shifts that maximize the achievable data rate of the RIS-assisted wideband THz systems. From the simulation results, we demonstrated that RWB outperforms the conventional phase shift control and beamforming schemes by a large margin.

Bibliography

[1] M. Cui and L. Dai, “Channel estimation for extremely large-scale MIMO: Far-field or near-field?” *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, 2022.

[2] Z. Chen, B. Ning, C. Han, Z. Tian, and S. Li, “Intelligent Reflecting Surface Assisted Terahertz Communications Toward 6G,” *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 110–117, 2021.

[3] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2019.

[4] J. Wu and B. Shim, “Power minimization of intelligent reflecting surface-aided uplink iot networks,” in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.

[5] S. Kim, J. Wu, and B. Shim, “Efficient Channel Probing and Phase Shift Control for mmWave Reconfigurable Intelligent Surface-Aided Communications,” *submitted to IEEE Trans. Wireless Commun.*, 2022.

[6] Y. Lin, S. Jin, M. Matthaiou, and X. You, “Tensor-based algebraic channel estimation for hybrid IRS-assisted MIMO-OFDM,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3770–3784, 2021.

[7] G. Zhou, C. Pan, H. Ren, P. Popovski, and A. L. Swindlehurst, "Channel estimation for RIS-aided multiuser millimeter-wave systems," *IEEE Trans. Sig. Process.*, vol. 70, pp. 1478–1492, 2022.

[8] S. Ma, W. Shen, J. An, and L. Hanzo, "Wideband channel estimation for IRS-aided systems in the face of beam squint," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6240–6253, 2021.

[9] X. Ma, Z. Chen, W. Chen, Z. Li, Y. Chi, C. Han, and S. Li, "Joint channel estimation and data rate maximization for intelligent reflecting surface assisted terahertz MIMO communication systems," *IEEE Access*, vol. 8, pp. 99 565–99 581, 2020.

[10] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel Estimation for Reconfigurable Intelligent Surface Aided Multi-User mmWave MIMO Systems," *IEEE Trans. Wireless Commun.*, 2023.

[11] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for near-field multi-user MIMO communications," *IEEE Trans. Wireless Commun.*, 2022.

[12] B. Wang, F. Gao, S. Jin, H. Lin, G. Y. Li, S. Sun, and T. S. Rappaport, "Spatial-wideband effect in massive MIMO with application in mmWave systems," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 134–141, 2018.

[13] Z. Wang, Z. Liu, Y. Shen, A. Conti, and M. Z. Win, "Wideband Localization with Reconfigurable Intelligent Surfaces," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Jun. 2022, pp. 1–6.

[14] A. Guerra, F. Guidi, D. Dardari, and P. M. Djurić, "Near-field tracking with large antenna arrays: Fundamental limits and practical algorithms," *IEEE Trans. Sig. Process.*, vol. 69, pp. 5723–5738, 2021.

[15] X. Zhang, H. Zhang, and Y. C. Eldar, “Near-Field Sparse Channel Representation and Estimation in 6G Wireless Communications,” *arXiv preprint arXiv:2212.13527*, 2022.

[16] Z. Abu-Shaban, K. Keykhosravi, M. F. Keskin, G. C. Alexandropoulos, G. Seco-Granados, and H. Wymeersch, “Near-field localization with a reconfigurable intelligent surface acting as lens,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[17] F. Guidi and D. Dardari, “Radio positioning with EM processing of the spherical wavefront,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3571–3586, 2021.

[18] H. Zhang, N. Shlezinger, G. C. Alexandropoulos, I. Alamzadeh, M. F. Imani, and Y. C. Eldar, “Channel estimation with hybrid reconfigurable intelligent metasurfaces,” *arXiv preprint arXiv:2206.03913*, 2022.

[19] J. Sherman, “Properties of focused apertures in the Fresnel region,” *IRE Trans. Antennas Propag.*, vol. 10, no. 4, pp. 399–408, 1962.

[20] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, “Channel estimation and hybrid combining for wideband terahertz massive MIMO systems,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604–1620, 2021.

[21] P. Wang, J. Fang, H. Duan, and H. Li, “Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems,” *IEEE Sig. Process. Lett.*, vol. 27, pp. 905–909, 2020.

[22] S. Kim, P. Jungjae, M. Jihoon, and B. Shim, “Fast and Accurate Terahertz Beam Management via Frequency-dependent Beamforming,” *submitted to IEEE Trans. Wireless Commun.*, 2022.

[23] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1527–1550, 2017.

[24] J. Kim, J. Wang, and B. Shim, "Optimal restricted isometry condition of normalized sampling matrices for exact sparse recovery with orthogonal least squares," *IEEE Trans. Sig. Process.*, vol. 69, pp. 1521–1536, 2021.

[25] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3—a MATLAB software package for semidefinite programming, version 1.3," *Optim. Methods Softw.*, vol. 11, no. 1-4, pp. 545–581, 1999.

[26] J. Wu, S. Kim, and B. Shim, "Energy-Efficient Power Control and Beamforming for Reconfigurable Intelligent Surface-Aided Uplink IoT Networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10 162–10 176, 2022.

[27] L. Dong, H.-M. Wang, J. Bai, and H. Xiao, "Double intelligent reflecting surface for secure transmission with inter-surface signal reflection," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2912–2916, 2021.

[28] L. T. Nguyen, J. Kim, S. Kim, and B. Shim, "Localization of IoT networks via low-rank matrix completion," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5833–5847, 2019.

[29] B. Houska, J. Frasch, and M. Diehl, "An augmented Lagrangian based algorithm for distributed nonconvex optimization," *SIAM J. Opt.*, vol. 26, no. 2, pp. 1101–1127, 2016.

[30] M. Stoyanova, Y. Nikoloudakis, S. Panagiotakis, E. Pallis, and E. K. Markakis, "A survey on the internet of things (IoT) forensics: challenges, approaches, and open issues," *IEEE Commun. Surv. Tut.*, vol. 22, no. 2, pp. 1191–1221, 2020.

[31] Q. Wu and R. Zhang, "Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2020.

[32] X. Cao, B. Yang, H. Zhang, C. Huang, C. Yuen, and Z. Han, "Reconfigurable intelligent surface-assisted MAC for wireless networks: protocol design, analysis, and optimization," *IEEE Internet Things J.*, vol. 8, no. 18, pp. 14 171–14 186, 2021.

[33] X. Cao, B. Yang, C. Huang, C. Yuen, M. Di Renzo, Z. Han, D. Niyato, H. V. Poor, and L. Hanzo, "AI-Assisted MAC for Reconfigurable Intelligent Surface-Aided Wireless Networks: Challenges and Opportunities," *arXiv preprint arXiv:2105.00437*, 2021.

[34] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Commun. Surv. Tut.*, vol. 22, no. 4, pp. 2283–2314, 2020.

[35] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, 2019.

[36] S. Jia, X. Yuan, and Y.-C. Liang, "Reconfigurable Intelligent Surfaces for Energy Efficiency in D2D Communication Network," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 683–687, 2020.

[37] Q. Wu, Y. Zhang, C. Huang, Y. Chau, Z. Yang, and M. Shikh-Bahaei, "Energy Efficient Intelligent Reflecting Surface Assisted Terahertz Communications," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021.

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

[39] L. You, J. Xiong, D. W. K. Ng, C. Yuen, W. Wang, and X. Gao, “Energy efficiency and spectral efficiency tradeoff in RIS-aided multiuser MIMO uplink transmission,” *IEEE Trans. Sig. Process.*, vol. 69, pp. 1407–1421, 2020.

[40] Z. Chu, P. Xiao, D. Mi, W. Hao, M. Khalily, and L.-L. Yang, “A Novel Transmission Policy for Intelligent Reflecting Surface Assisted Wireless Powered Sensor Networks,” *IEEE J. Sel. Top. Sig. Process.*, 2021.

[41] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, “On beamforming with finite rate feedback in multiple-antenna systems,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, 2003.

[42] L. T. Nguyen, J. Kim, S. Kim, and B. Shim, “Localization of IoT Networks via Low-Rank Matrix Completion,” *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5833–5847, 2019.

[43] X. Yu, D. Xu, and R. Schober, “MISO wireless communication systems via intelligent reflecting surfaces,” in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2019.

[44] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, “Multicell MIMO communications relying on intelligent reflecting surfaces,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, 2020.

[45] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, “Intelligent Reflecting Surface Aided Wireless Communications: A Tutorial,” *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, 2021.

[46] W. Kim, H. Ji, and B. Shim, "Channel Aware Sparse Transmission for Ultra Low-Latency Communications in TDD Systems," *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 1175–1186, 2020.

[47] L. T. Nguyen, J. Kim, and B. Shim, "Low-rank matrix completion: A contemporary survey," *IEEE Access*, vol. 7, pp. 94 215–94 237, 2019.

[48] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming," 2016. [Online]. Available: http://cvxr.com/cvx

[49] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[50] Y.-H. Dai and Y. Yuan, "Nonlinear conjugate gradient methods," *Shanghai Science and Technology Publisher, Shanghai*, 2000.

[51] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," *lecture notes of EE392o, Stanford University, Autumn Quarter*, vol. 2004, pp. 2004–2005, 2003.

[52] W. Ring and B. Wirth, "Optimization methods on Riemannian manifolds and their application to shape space," *SIAM J. Opt.*, vol. 22, no. 2, pp. 596–627, 2012.

[53] M. Fu, Y. Zhou, and Y. Shi, "Intelligent reflecting surface for downlink non-orthogonal multiple access networks," in *Proc. IEEE Global Commun. Conf. (Globecom) Workshops*, Mar. 2019.

[54] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, 2020.

[55] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, 2020.

[56] J. Wu, S. Kim, and B. Shim, "Near-Field Channel Estimation for RIS-Assisted Wideband Terahertz Systems," *Proc. of IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022.

[57] C. Han and Y. Chen, "Propagation modeling for wireless communications in the terahertz band," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 96–101, 2018.

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

[59] C. Huang, Z. Yang, G. C. Alexandropoulos, K. Xiong, L. Wei, C. Yuen, Z. Zhang, and M. Debbah, "Multi-hop RIS-empowered terahertz communications: A DRL-based hybrid beamforming design," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1663–1677, 2021.

[60] Z. Wan, Z. Gao, F. Gao, M. Di Renzo, and M.-S. Alouini, "Terahertz massive MIMO with holographic reconfigurable intelligent surfaces," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4732–4750, 2021.

[61] Y. Lu, M. Hao, and R. Mackenzie, "Reconfigurable intelligent surface based hybrid precoding for THz communications," *Intelligent and Converged Networks*, vol. 3, no. 1, pp. 103–118, 2022.

[62] J. Tan and L. Dai, "Wideband beam tracking in THz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1693–1710, 2021.

[63] E. Ghaderi, A. S. Ramani, A. A. Rahimi, D. Heo, S. Shekhar, and S. Gupta, "An integrated discrete-time delay-compensating technique for large-array beamformers," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 66, no. 9, pp. 3296–3306, 2019.

[64] V. Boljanovic, H. Yan, C.-C. Lin, S. Mohapatra, D. Heo, S. Gupta, and D. Cabric, "Fast beam training with true-time-delay arrays in wideband millimeter-wave systems," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 68, no. 4, pp. 1727–1739, 2021.

[65] J. Wu, S. Kim, and B. Shim, "Energy-Efficient Power Control and Beamforming for Reconfigurable Intelligent Surface-Aided Uplink IoT Networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10 162–10 176, 2022.

[66] L. Dai, J. Tan, Z. Chen, and H. V. Poor, "Delay-phase precoding for wideband THz massive MIMO," *IEEE Trans. Wireless Commun.*, 2022.

초 록

전 세계적으로 5세대(5G) 시스템이 구축됨에 따라 2030년 이후를 위한 6세대 (6G) 시스템 연구가 본격화되었으며, 6G는 센싱, 컴퓨팅, 인공지능(AI) 등 비통신 기술의 결합을 구상하고 있다, 5G 애플리케이션 시나리오의 기능을 계승하고 개발하는 것 외에도. 주요 동인은 증가하는 모바일 트래픽과 홀로그램 텔레프레젠스, 확장 현실(XR), 디지털 트윈 및 자율 시스템과 같은 새로운 애플리케이션이다. 이러한 새로운 애플리케이션은 6G의 주요 성능 지표(KPI)를 일반적으로 속도, 신뢰성, 대기 시간, 이동성 및 에너지 소비 측면에서 10 ~ 100 배 더 높게 만든다. 최대 데이터 속도는 1 테라비트/초(Tbps)에 도달하고 대기 시간은 밀리초 미만 수준으로 감소할 것으로 예상된다. 현재의 메커니즘과 기존의 접근 방식은 이러한 엄격한 요구 사항을 지원할 수 없기 때문에, 6G 비전을 실현하기 위한 새로운 기술이 필요하다. 논문의 첫 번째 부분에서, 우리는 근거리 재구성 가능 지능형 표면(RIS) 지원 테라헤르츠(THz) 시스템을 위한 채널 추정 프레임워크를 연구한다. 6G 무선 네트워크에서 매우 높은 데이터 속도를 지원하기 위해 재구성 가능한 지능형 표면(RIS) 지원 테라헤르츠(THz) 통신이 최근 몇 년 동안 많은 관심을 받고 있다. RIS는 반사 요소의 위상 이동을 조작함으로써 THz 시스템의 무선 전파 환경을 능동적으로 조정하여 전체 처리량을 크게 향상시킬 수 있다. RIS 지원 THz 시스템의 잠재력을 최대한 실현하려면 정확한 채널 정보를 획득하는 것이 매우 중요하다. 그러나 THz 전자기 신호의 파면은 구형이기 때문에 평면 파면 가정을 사용하는 기존의 채널 추정 기법은 근거리 RIS 지원 THz 시스템에서 심각한 성능 저하를 겪는다. 이 연구의 목표는 근거리 RIS 지원 광대역 THz 시스템에 대한 효율적인 채널 추정 기법을 제안하는

것이다. 제안된 극성 도메인 주파수 종속 RIS 지원 채널 추정(PF-RCE) 체계의 핵심 아이디어는 극성 도메인 희소성과 공통 지원 속성을 활용하여 근거리 THz 채널의 희소 다중 경로 구성 요소(즉, 각도, 거리 및 경로 이득)를 추정하는 것이다.

논문의 두 번째 부분에서, 우리는 RIS 지원 사물인터넷(IoT) 네트워크를 위한 에너지 효율적인 전력 제어 및 빔포밍 체계를 연구한다. 많은 수의 저가 반사 요소로 구성된 평면 메타 표면인 RIS는 무선 전파 환경을 재구성하여 스펙트럼과 에너지 효율을 모두 향상시킬 수 있는 능력으로 인해 많은 관심을 받아왔다. 본 연구에서는 RIS 지원 IoT 네트워크의 업링크 전송 전력을 최소화하는 RIS 위상 편이 및 BS 빔포밍 최적화 기술을 제안한다. 리만 공역 그레이디언트 기반 조인트 최적화(RCG-JO)라고 하는 제안된 체계의 핵심 아이디어는 리만 공역 그레이디언트 기법을 사용하여 RIS 위상 이동과 BS 빔포밍 벡터를 공동으로 최적화하는 것이다. 단위-모듈러스 위상 이동 및 단위-노름 빔포밍 벡터 세트의 제품 리만 다양체 구조를 활용하여 비볼록 업링크 전력 최소화 문제를 제약 없는 문제로 변환한 다음 제품 리만 다양체에 대한 최적의 솔루션을 찾는다.

논문의 세 번째 부분에서, 우리는 RIS 지원 광대역 THz 시스템을 위한 주파수 의존적 빔 형성 체계를 연구한다. 광대역 THz 통신의 주요 과제 중 하나는 경로 구성요소가 서로 다른 서브캐리어 주파수에서 서로 다른 공간 방향으로 분할되는 빔 분할 효과로 인한 심각한 어레이 이득 손실이다. 따라서, 종래의 위상 편이 제어 및 빔포밍 기법은 광대역 THz 시스템에 직접 적용될 수 없다. 본 연구에서는 RIS 지원 광대역 THz 시스템의 평균 데이터 속도를 최대화하는 RIS 지원 주파수 종속 빔포밍(RWB) 기술을 제안한다. RWB의 핵심 아이디어는 광대역 THz 시스템의 평균 데이터 속도가 최대화되도록 RWB 네트워크의 매개 변수를 적절하게 설계하여 아날로그 빔포밍 벡터와 RIS 위상 편이 벡터를 교대로 최적화하는 것이다.

주요어: 6G 무선 통신, 지능형 반사평면, 테라헤르츠 통신, 채널 추정, 주파수 종속적 빔포밍

학번: 2019-36083

Acknowledgement

The five years spent pursuing my PhD in South Korea have flown by like the blink of an eye, fleeting yet profound. However, at this moment, memories of these five years flicker in my mind like a carousel, vivid yet distant, deeply etched and unforgettable, yet slipping through my grasp like sand.

Looking back to the spring of 2018, I distinctly remember the moment when Professor Shim responded to my email and offered me the opportunity for joining ISLAB. In our initial interaction, I felt his warmth and kindness, brimming with anticipation for the upcoming journey of doctoral research, ignited by an overwhelming passion. Without the opportunity provided by Professor Shim, I might have ventured into the corporate world, pursued studies in another country, or delved into a different field. As a consequence, I would have missed encountering the best supervisor, would not have assimilated within the nurturing family of ISLAB, and would have forfeited these invaluable five years that hold great significance to me now.

Over these five years of doctoral studies, Professor Shim has not only provided guidance in my research, but has also extended tremendous support and encouragement in my daily life. I still vividly remember the time when I received my first paper rejection, he promptly consoled and encouraged me, dedicating weekends to assist me in refining the problems in my paper. I express heartfelt gratitude to my supervisor, Professor Shim, for the continuous encouragement and trust, which have enabled me to grow stronger and more confident in research and life, embarking on the journey toward the future with greater courage.

As an international student, from the moment I stepped into ISLAB, I felt an instant sense of belonging. It is this group of friendly friends who have offered unwavering assistance in my daily life, allowing me to embrace the most wonderful and unforgettable five years in a foreign land. They not only helped me manage non-academic tasks, but also tailored meal choices to my preferences, introduced me to Korean culture, taught me the language, embraced my mistakes, and understood even my less-amusing jokes... How fortunate I am! Wonjun has been my mentor and my first friend in Korea. I wholeheartedly appreciate him for his patient guidance in helping me to understand papers, teaching me coding, and assisting me in settling down on the scorching summer days of 2018. Seungnyun has been an incomparable partner; without any doubt, he is the best! I extend sincere thanks to Seungnyun for his generous help and trust, and for the treasured moments spent together revising papers. I am grateful to Hyunsoo, who not only shared many meals with me together but also taught me skiing and assisted with countless documentations. While direct in words, Hyunkyu has offered me invaluable advice, and I appreciate his help and understanding. Jihoon, standing tall at 187cm yet exuding undeniable cuteness, has endured my inexplicable jokes. I thank him for the rabbit doll he gave me on my birthday. The lovely younger sisters, Yiying and Jiajie, they deserve of every positive description. I wish their dreams to come true. My gratitude extends to Khoa, Dr. Luong, Sangtae, Yongseok, Junhan, Jinhoon, Yongjun, Sunwoo, Jiseob, Yoonsung, Jungjae, Ahnho, Yunseo, Taeyup, Seokhyun, and Sangmok, for their assistance during my stay in Korea.

Upon joining ISLAB, Dr. Tang and Dr. Di were still working in the lab. They were not only my admirable seniors but also role models for my studies. I sincerely thank them for their help, and wish them success in their research endeavors. Additionally, I appreciate Chen Dawei and Dr. Ruoyu Zhang, for making my lab life more enjoyable and relaxed when we were working together.

From attending the local elementary school near my home since the age of four to now at the age of 30, my family has consistently been the unwavering force driving me

forward, offering selfless love without reservation. I thank my father for encouraging me to pursue knowledge and dreams. His diligent and upright character serves as my spiritual role model. I am grateful to my mother, the most beautiful person in the world, who nurtured me through growth. My sister, who manages household matters, has afforded me the freedom to focus on studying abroad. She is not only the best sister but also my closest friend from childhood to adulthood. My gratitude extends to her and her loving husband, as well as to my adorable little nephew, who recently joined our world. My younger brother has grown up and matured, I am thankful for his genuine and sincere love. My grandfather was the greatest person in my life; he held my hand as I entered school, sheltered me all the time, and guided me through life's questions. My perspectives and values have been shaped by his teachings. While my grandfather is no longer with us, his presence remains alive within me. Despite her advanced age, my grandmother is still the same as the one my sister described in her childhood composition: ”My grandma is adorable, with a face as round as a red apple.” I love my grandmother dearly and wish her good health. From this point onward, I vow to dedicate all my strength to safeguarding my beloved family.

August 7th, 2023

WU JIAO