Imperial College London

Department of Electrical and Electronic Engineering

Beyond Diagonal Reconfigurable Intelligent Surfaces: Modeling, Architecture Design, and Optimization

Matteo Nerini

2024

Supervised by Prof. Bruno Clerckx

Submitted in part fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Electronic Engineering of Imperial College London and the Diploma of Imperial College London

Statement of Originality

I declare that the content of this thesis is the result of my own work. Information and ideas derived from the work of others have been acknowledged in the text and a list of references is given in the bibliography. The material of this thesis has not been and will not be submitted for another degree at any other university or institution.

Copyright Declaration

To my family

Abstract

Reconfigurable intelligent surface (RIS) is expected to be a key technology in 6G to enhance wireless systems by efficiently and cost-effectively manipulating the propagation environment. In conventional RIS, each RIS element is independently controlled by a tunable load and it is disconnected from the other elements. Thus, conventional RIS is characterized by a diagonal scattering matrix, also known as a phase shift matrix, which has limited passive beamforming capabilities. To enhance the flexibility of RIS, beyond diagonal RIS (BD-RIS) has been introduced as a generalization of conventional RIS, in which the scattering matrix is not restricted to being diagonal. Motivated by the promising benefits unlocked by BD-RIS, in this thesis, we explore the fundamental performance limits of this technology with ideal hardware models and explore how hardware non-idealities impact the performance limits. First, we explore the optimization of the two main existing BD-RIS architectures, i.e., group- and fully-connected RISs. We derive a closed-form global optimal solution to optimize the scattering matrix of group- and fully-connected RISs and optimize the grouping strategy in group-connected RIS based on the channel statistics, showing the benefits of an optimized static grouping strategy. Second, we propose novel BD-RIS architectures to approach the best trade-off between achieved performance and BD-RIS circuit topology complexity. In particular, we propose forest- and tree-connected RISs as novel low-complex BD-RIS architectures and we derive the fundamental limits of the performance-complexity trade-off enabled by BD-RIS. Third, we model and optimize BD-RIS architectures in the presence of hardware impairments. Specifically, we consider BD-RIS with a practical discrete-value scattering matrix, with electromagnetic (EM) mutual coupling between the RIS elements, and with lossy interconnections. In the three technical parts of this thesis, we provide numerical results to corroborate our theoretical insights, confirming the significant superiority of BD-RIS over conventional RIS in EM wave manipulation.

Acknowledgments

This thesis is for my family, who supported me in every decision I took to pursue my dreams, even when this meant spending the last nine years far from them. Thank you for your patience and the sacrifices you made to allow me to study and live abroad. Thank you for always being there for me, and especially for your unconditional love, which gave me the strength to reach this point.

Special thanks to you, Bruno, for believing in me and continuously providing me with opportunities to grow and advance. Thank you for your constructive guidance, for your commitment to my academic development, and for your invaluable insights that led to the completion of this thesis. I feel deeply honored to have had the possibility to work in your research group on such fascinating topics, and I look forward to keep working together.

Thanks to Shanpu and Hongyu, together with whom I shaped most of the ideas and results contained in this thesis. Thank you, Shanpu, for your constant advice, feedback, and mentorship, not only in technical matters but also in navigating the complexities of academic research. Thank you, Hongyu, for supporting me every single day of this journey and for tolerating my humor with minimal complaints, walking through the arduous Ph.D. life has been more enjoyable with you next to me.

Thank you to the most amazing research group one could ever wish to be part of, especially Anup, Hongyu, Jiawei, Kris, Onur, Sibo, Xinze, Yuanwen, Yumeng, and Yunno. You have not just been colleagues, but special friends, with whom I have spent countless joyful moments and learned some (not countless, nor joyful) Chinese words.

Thanks to all the professors, supervisors, and mentors I met during my academic path. In particular, thank you, Golsa, for your continuous advice and support that moved me since my M.Sc. studies and significantly impacted my journey.

Thanks to my friends who made my life in London exciting and full of amazing memories. Special thanks to all my friends of Netherhall House, especially Gianluca, Donato, Leonardo, and Matheus, who are today uniformly distributed around the globe, but with whom I lived unforgettable daily moments. In particular, thanks, Matheus, for being the best friend I have ever had, for your constant presence in my life, and for all the highly majestic memories, jokes, and mathematical proofs we have shared. Thank you, Michele, for your company since before our day zero in London.

Finally, thanks to those who have been part of my life while I was studying towards my Ph.D. I will never forget who was with me the day I began this Ph.D. journey, and accompanied me daily for a significant part of it. A lot of you is in this thesis and I want to thank you for that.

Contents

List of Tables

List of Figures

Acronyms

AWGN additive white Gaussian noise.

BD-RIS beyond diagonal RIS.

BS base station.

CSCG circularly symmetric complex Gaussian.

CSI channel state information.

DFRC dual-function radar-communication.

EM electromagnetic.

i.i.d. independent and identically distributed.

IOS intelligent omni-surface.

IoT internet of things.

IRS intelligent reflecting surface.

LoS line-of-sight.

MEC mobile edge computing.

MIMO multiple-input multiple-output.

MISO multiple-input single-output.

MRT maximum ratio transmission.

NLoS non-line-of-sight.

NOMA non-orthogonal multiple access.

OFDM orthogonal frequency division multiplexing.

PIN positive-intrinsic negative.

RF radio frequency.

RIS reconfigurable intelligent surface.

RSMA rate splitting multiple access.

SDR semidefinite relaxation.

SIM stacked intelligent metasurface.

SIMO single-input multiple-output.

SISO single-input single-output.

SRE smart radio environment.

STAR-RIS simultaneously transmitting and reflecting RIS.

SWIPT simultaneous wireless information and power transfer.

UAV unmanned aerial vehicle.

UE user equipment.

ULA uniform linear array.

UPA uniform planar array.

WPT wireless power transfer.

Chapter 1

Introduction

The ongoing research and development in wireless communications are critical to addressing the expanding demands of our connected society. With the proliferation of mobile devices and emerging technologies such as internet of things (IoT), virtual/augmented reality, and autonomous vehicles, there's an urgent need for more efficient and reliable wireless networks. To meet the growing expectations of users and applications, research endeavors focus on enhancing data rates, extending coverage, and minimizing latency, by facing two main challenges. First, the radio frequency (RF) spectrum available for wireless communication is finite and increasingly congested, which limits the capacity of wireless networks. Second, the wireless channel varies over space and time, often in unpredictable and complex ways, depending on the location, speed, and material of the objects in the propagation environment.

Over the last decades, multiple-input multiple-output (MIMO) emerged as the pivotal technology in advancing the capabilities of wireless systems while addressing these two challenges [1]. Specifically, MIMO technology employs multiple antennas at the transmitter and receiver ends to enable the efficient exploitation of the spatial dimension. The use of multiple antennas was originally proposed at one end of the wireless link, namely at the transmitter in the so-called multiple-input single-output (MISO) systems or at the receiver in single-input multiple-output (SIMO) systems to improve wireless communications through beamforming and spatial diversity. On the one hand, beamforming is a technique used to focus RF energy in a specific direction, enhancing the strength of the received signal. It involves manipulating the phase and amplitude of the transmitted or received signal by an antenna array to create a directive radiation pattern. On the other hand, spatial diversity is used to improve the robustness of the communication link by leveraging independent copies of the signal due to the spatial variations in the wireless channel. The advent of MIMO in the 1990s significantly increased the capacity of wireless links even under limited spectrum usage by turning the multipath characteristic of challenging propagation environments into an advantage [2, 3]. Through signal processing algorithms, MIMO techniques enable the simultaneous transmission of multiple data streams, i.e., spatial multiplexing, further improving the spectral efficiency and enhancing system performance in single- as well as multi-user systems.

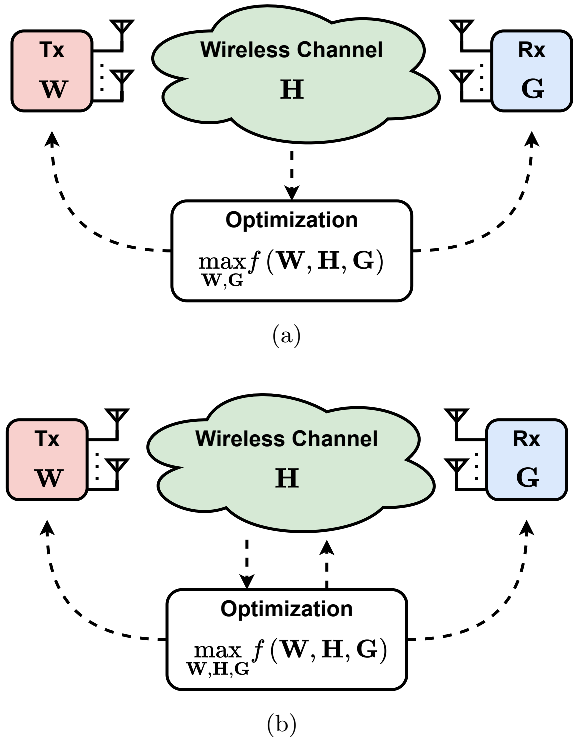


Figure 1.1: (a) Conventional and (b) smart radio environment.

1.1 Smart Radio Environments

MIMO processing has been developed based on the assumption that the wireless channel is given by the propagation environment, and cannot be engineered. Thus, the objective of MIMO processing has been to design precoding strategies at the transmitter side and combining strategies at the receiver side as a function of the wireless channel, often assumed to be known through channel estimation techniques. The channel knowledge, also known as channel state information (CSI), is critical to unlocking the full benefits of MIMO systems. However, while the channel should be known at the transmitter and/or receiver, it remains beyond our control in conventional radio environments, as shown in Fig. 1.1(a).

In the last years, the assumption that the wireless channel cannot be controlled has been overcome by a paradigm shift referred to as smart radio environment (SRE) [4, 5]. In a SRE, the wireless channel can be engineered to improve the quality of wireless communications, as shown in Fig. 1.1(b). This can be achieved by electrically adjusting the reflection coefficients of the objects within the propagation environment such that the EM signal paths interfere constructively at the receiver, with a consequent increase of the received signal power. Besides, the network coverage can be also extended by steering the signal impinging on obstacles toward the direction of the intended receiver, and effectively bypassing blockages.

The history of SREs traces back to the early 2000s. In [6], a "smart surface" with controllable reflective and transmissive capabilities was realized through a metafilm, or metasurface, i.e., the two-dimensional equivalent of a metamaterial. In [7], the concept of "intelligent wall" was introduced as a wall equipped with a frequency-selective surface able to control the radio coverage. Furthermore, electronically tunable metasurfaces have been studied to realize spatial microwave modulators and shape the impinging EM waves [8, 9, 10]. More recently, the number of studies on SRE has increased significantly, and reconfigurable intelligent surface (RIS), or intelligent reflecting surface (IRS), has emerged as the key technology to enable SREs in wireless networks [11, 12, 13, 14].

RIS is a technology that involves deploying arrays consisting of numerous individually controllable elements. The reflection coefficients of these elements can be adjusted in real-time to manipulate electromagnetic waves, allowing for dynamic control over the wireless communication environment and the realization of SREs. RIS-aided communication systems benefit from several advantages. RISs with passive elements are characterized by ultra-low power consumption and do not cause any additive thermal noise or self-interference phenomena. In addition, RIS is a low-profile and cost-effective solution, since it does not include RF chains and expensive active devices [15]. Owing to these benefits, RIS has gained widespread attention.

1.1.1 Diagonal RIS

In a conventional RIS, also known as diagonal RIS or single-connected RIS, the reflection coefficient of each element is independently controlled through a tunable impedance component. This results in conventional RIS having a diagonal scattering matrix, commonly known as a phase shift matrix [16]. Conventional RIS with diagonal scattering matrix has been widely used to enhance different wireless communication systems, such as single-cell [17, 18, 19, 20] and multi-cell systems [21], and systems based on orthogonal frequency division multiplexing (OFDM) [22, 23], non-orthogonal multiple access (NOMA) [24], and rate splitting multiple access (RSMA) [25]. RIS has been also applied to further improve the data rate by modulating and encoding data into the reconfigurable elements [26, 27]. Beyond communication systems, RIS has been used to improve the performance of wireless power transfer (WPT) [28], simultaneous wireless information and power transfer (SWIPT) [29], RF sensing [30], and localization systems [31].

The study of RIS poses several challenges, which have been addressed in recent literature. Path-loss models for RISs under different propagation conditions have been proposed in [32, 33]. In addition, research has been conducted to optimize RIS based on imperfect CSI [34, 35], in multi-RIS scenarios [36, 37], and with discrete phase shifts to account for practical hardware impairments [38, 39]. Channel estimation protocols with reduced pilot overhead for RIS-aided systems have been proposed [40, 41, 42]. Active RIS have been proposed to overcome the multiplicative fading effect limiting the gains of passive RIS by amplifying the reflected signals via amplifiers integrated into their elements [43, 44]. Furthermore, RIS prototypes have been realized in [45, 46, 47], showcasing the benefit of this technology in real-world scenarios.

1.1.2 Beyond Diagonal RIS

Diagonal RIS has shown promising gains but has limited flexibility since each element is independently controlled. To enhance the flexibility of RIS, beyond diagonal RIS (BD-

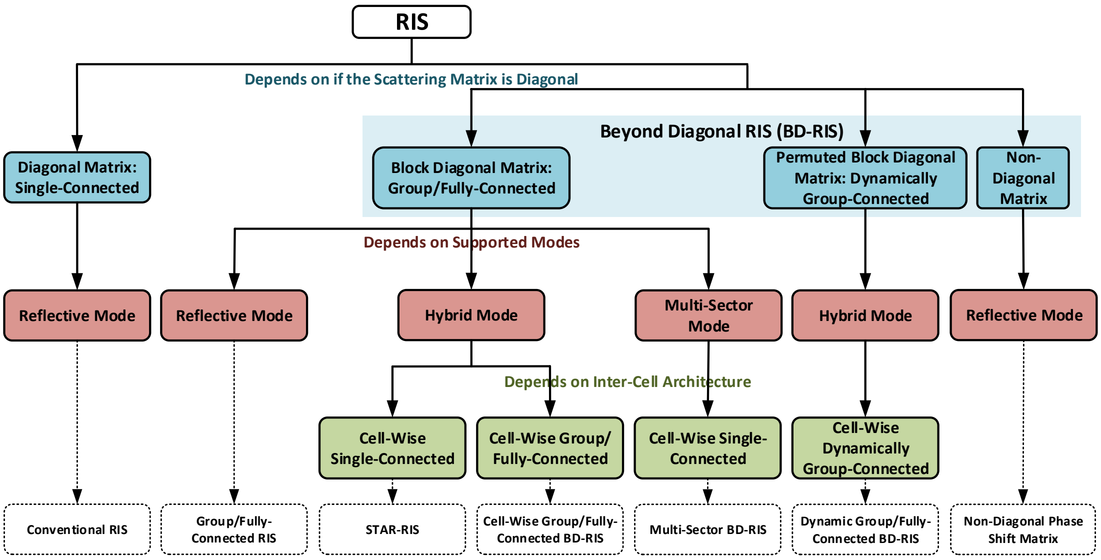


Figure 1.2: RIS classification tree.

RIS) has been proposed as a generalization of diagonal RIS, in which the scattering matrix is not limited to being diagonal [48]. The key novelty of BD-RIS is the presence of tunable impedance components interconnecting the RIS elements, adding further reconfigurability at the expense of additional circuit complexity.

Several BD-RIS architectures have been introduced, as shown in the classification tree in Fig. 1.2. BD-RIS was first introduced in [16], where the fully-connected RIS has been proposed by interconnecting all the RIS elements to each other through tunable impedance components. To bridge between the single- and fully-connected architectures, the group-connected RIS has been also proposed in [16], in which the RIS elements are divided into groups and there are no interconnections between elements in different groups. Group- and fully-connected RISs have been optimized in various wireless systems, showing their superiority over conventional RIS. Specifically, these RIS architectures have been optimized through closed-form globally optimal solutions in single-user systems [49] and through low-complexity optimization algorithms in multi-user systems [50, 51]. A unified approach to optimize power consumption, energy efficiency, and sum-rate in group-connected RIS-aided systems has been proposed in [52]. The capacity of a MIMO link aided by a fully-connected RIS has been characterized in [53], where a solution to reach such a capacity has been proposed. A channel estimation protocol for passive group- and fully-connected RIS-aided systems has been proposed in [54], not requiring expensive and energy-hungry RF chains at the RIS. Group- and fully-connected RISs have been modeled and optimized in the presence of mutual coupling in [55]. Finally, BD-RIS have shown high-performance gains over conventional RIS in RSMA systems [56, 57, 58], massive MIMO systems [59], unmanned aerial vehicle (UAV)-aided systems [60], mobile edge computing (MEC) systems [61], and SWIPT systems [62].

Beyond achieving higher performance gains than conventional RIS, BD-RIS is able to both reflect and transmit the incident signal, allowing us to serve users on both sides of the RIS. Specifically, the concept of simultaneously transmitting and reflecting RIS (STAR-RIS), or intelligent omni-surface (IOS), has been introduced in [63, 64]. This RIS architecture is able to reflect and transmit the incident signal, differently from conventional RISs working only in reflective mode. BD-RIS architectures enabling full-space coverage are denoted to work in the so-called hybrid reflective and transmissive mode [65]. In [65], a BD-RIS model has been developed unifying different BD-RIS working modes (reflective/transmissive/hybrid) and different BD-RIS architectures (single-/group-/fully-connected). In [66], multi-sector BD-RIS has been introduced, where the elements are divided into multiple sectors, each covering a narrow region of space. Owing to the narrow antenna radiation patterns, multi-sector BD-RIS enables highly directional full-space coverage with further gain over BD-RIS working in hybrid mode. BD-RIS with reflective and transmissive capabilities has proved to enlarge the coverage and improve the sum-rate in RSMA systems [67], and to improve communication capacity and sensing precision in dual-function radar-communication (DFRC) systems [68].

Different from the BD-RIS architectures in [16]-[68], which have fixed interconnections between the RIS elements, dynamic BD-RIS have been proposed to further improve the performance by dynamically reconfiguring the interconnections on a per channel realization basis. Dynamically group-connected RIS has been proposed in [69], as a group-connected RIS in which the group size and the grouping strategy can be dynamically reconfigured through a switch network. Besides, a BD-RIS architecture with a non-diagonal phase shift matrix has been proposed in [70, 71], able to achieve a higher rate than conventional RISs. Exploiting its non-reciprocal nature, non-diagonal RIS can enable attacks to break the uplink-downlink channel reciprocity [72].

Motivated by the promising opportunities enabled by BD-RIS, in this thesis, we explore the fundamental limits of this technology under ideal models and analyze the impact of hardware non-idealities on the BD-RIS performance limits. Specifically, we develop our technical contributions in three sequential parts, each constructed upon the foundation laid by the previous. First, we focus on the optimization of the two main existing BD-RIS architectures, i.e., group- and fully-connected RISs. We derive a closed-form global optimal solution to optimize the scattering matrix of group- and fully-connected RISs and optimize the grouping strategy in group-connected RIS based on the channel statistics, showing the benefits of an optimized static grouping strategy. Second, we develop novel BD-RIS architectures to achieve a better trade-off between performance and circuit complexity. In particular, we propose forest- and tree-connected RISs as novel low-complex BD-RIS architectures and we derive the fundamental limits of the performance-complexity trade-off enabled by BD-RIS. Third, we model and optimize BD-RIS in the presence of hardware impairments. Specifically, we consider BD-RIS with a practical discrete-value scattering matrix, with EM mutual coupling between the RIS elements, and with lossy interconnections.

1.2 Contributions

The original contributions of this thesis can be summarized as follows.

1. We derive a closed-form solution for the global optimal scattering matrix of group- and fully-connected RIS. Thus, we prove that their corresponding theoretical perfor-

mance upper bounds can be exactly achieved for any channel realization. We first consider the received signal power maximization in single-user single-input single-output (SISO) systems aided by a BD-RIS working in reflective or transmissive mode. Then, we extend our solution to single-user MIMO and multi-user MIMO systems. We show that our algorithm is less complex than the iterative optimization algorithms employed in the previous literature. The complexity of our algorithm grows linearly (resp. cubically) with the number of RIS elements in the case of group- (resp. fully-) connected architectures.

2. We explore how to optimally group the elements in group-connected RISs to maximize the performance while maintaining a low-complexity circuit. Specifically, we propose and model BD-RIS with a static grouping strategy optimized based on the channel statistics. After formulating the corresponding problems, we design the grouping in single- and multi-user systems. Numerical results reveal the benefits of grouping optimization, i.e., up to 60% sum rate improvement, especially in highly correlated channels.

3. We propose novel modeling, architecture design, and optimization for BD-RIS based on graph theory. This graph theoretical modeling allows us to develop two new efficient BD-RIS architectures, denoted as tree-connected and forest-connected RIS. Tree-connected RIS, whose corresponding graph is a tree, is proven to be the least complex BD-RIS architecture able to achieve the performance upper bound in MISO systems. Besides, forest-connected RIS allows us to strike a balance between performance and complexity, further decreasing the complexity over tree-connected RIS. To optimize tree-connected RIS, we derive a closed-form global optimal solution, while forest-connected RIS is optimized through a low-complexity iterative algorithm. Numerical results confirm that tree-connected (resp. forest-connected) RIS achieves the same performance as fully-connected (resp. group-connected) RIS, while reducing the complexity by up to 16.4 times.

4. We investigate the fundamental limits of the trade-off between performance and circuit complexity enabled by the various BD-RIS architectures. Specifically, we derive the expression of the Pareto frontier for the performance-complexity trade-off in BD-RIS. In addition, we characterize the optimal BD-RIS architectures reaching this Pareto frontier.

5. We propose two solutions to realize discrete-value group- and fully-connected RISs. First, we propose scalar-discrete RISs, in which each entry of the RIS impedance matrix is independently discretized. Second, we propose vector-discrete RISs, where the entries in each group of the RIS impedance matrix are jointly discretized. In both solutions, the codebook is designed offline such as to minimize the distortion caused in the RIS impedance matrix by the discretization operation. Numerical results show that vector-discrete RISs achieve higher performance than scalar-discrete RISs at the cost of increased optimization complexity. Furthermore, fewer resolution bits per impedance are necessary to achieve the performance upper bound as the group size

of the group-connected architecture increases. In particular, only a single resolution bit is sufficient in fully-connected RIS to approximately achieve the performance upper bound.

6. We propose a universal framework for the multiport network analysis of RIS-aided systems. With our framework, we model RIS-aided systems and RIS architectures through impedance, admittance, and scattering parameter analysis. Based on these analyses, three equivalent models are derived accounting for the effects of impedance mismatching and mutual coupling. The three models are then simplified by assuming large transmission distances, perfect matching, and no mutual coupling to understand the role of the RIS in the communication model. The derived simplified models are consistent with the model used in related literature, although we show that an additional approximation is commonly considered in the literature. We discuss the benefits of each analysis in characterizing and optimizing the RIS and how to select the most suitable parameters according to the needs. Numerical results provide additional evidence of the equivalence of the three analyses.

7. We investigate distributed RIS, whose elements are distributed over a wide region, in opposition to localized RIS commonly considered in the literature. The scaling laws of distributed BD-RIS reveal that it offers significant gains over distributed conventional RIS and localized BD-RIS, enabled by its interconnections allowing signal propagation within the BD-RIS. To assess the practical performance of distributed BD-RIS, we model and optimize BD-RIS with lossy interconnections through transmission line theory. Our model accounts for phase changes and losses over the BD-RIS interconnections arising when the interconnection lengths are not much smaller than the wavelength. Numerical results show that the performance of localized BD-RIS is only slightly impacted by losses, given the short interconnection lengths. Besides, distributed BD-RIS can achieve orders of magnitude of gains over conventional RIS, even in the presence of low losses.

1.3 Organization

The rest of this thesis is organized as follows. In Chapter 2, we review the background information required to introduce the concept of BD-RIS. Subsequently, we discuss three technical parts. Part I is composed of Chapters 3 and 4. In Chapter 3, we globally optimize in closed-form group- and fully-connected RISs. In Chapter 4, we optimize the static grouping strategy of group-connected RISs. Part II includes Chapters 5 and 6. In Chapter 5, we propose two novel BD-RIS architectures to achieve the maximum performance with reduced circuit complexity, namely forest- and tree-connected RISs. In Chapter 6, we characterize the Pareto frontier of the performance-complexity trade-off enabled by BD-RIS. Part III contains Chapters 7, 8, and 9, where we investigate BD-RIS with hardware impairments, namely discrete-value scattering matrices, mutual coupling, and lossy interconnections, respectively. Finally, Chapter 10 concludes this thesis. The overview of this thesis is illustrated in Fig. 1.3.

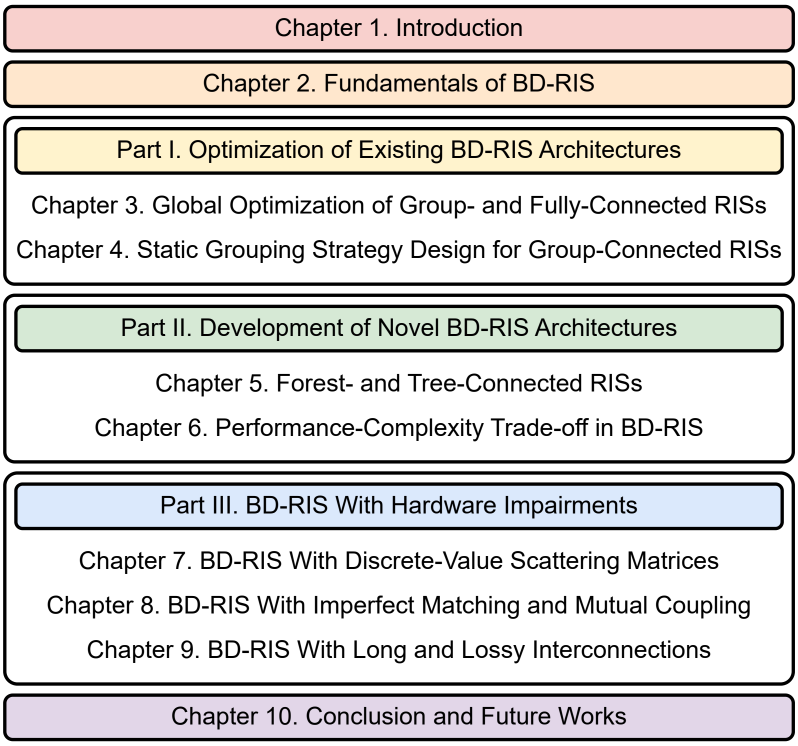


Figure 1.3: Overview of the thesis.

1.4 Notation

Vectors and matrices are denoted with bold lower and bold upper letters, respectively. Scalars are represented with letters not in bold font. $\Re\{a\}$, $\Im\{a\}$, $|a|$, $\arg(a)$, and a^* refer to the real part, imaginary part, modulus, phase, and conjugate of a complex scalar a, respectively. $[\mathbf{a}]_i$ and $\|\mathbf{a}\|$ refer to the ith element and l_2-norm of a vector $\mathbf{a}$, respectively. $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $[\mathbf{A}]_{i,j}$, $\|\mathbf{A}\|$, and $\|\mathbf{A}\|_F$ refer to the conjugate, transpose, conjugate transpose, (i,j)th element, l_2-norm, and Frobenius norm of a matrix $\mathbf{A}$, respectively. $\mathbf{A} \sim \mathbf{B}$ means that the matrices $\mathbf{A}$ and $\mathbf{B}$ are equivalent. $\mathbf{u}_{\max}(\mathbf{A})$ and $\mathbf{v}_{\max}(\mathbf{A})$ denote the dominant left and right singular vectors of a matrix $\mathbf{A}$, respectively. $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{R}$, and $\mathbb{C}$ denote the natural, positive natural, real, and complex number sets, respectively. $j = \sqrt{-1}$ denotes the imaginary unit. $\mathbf{0}$ and $\mathbf{I}$ denote an all-zero matrix and an identity matrix, respectively, with appropriate dimensions. $\mathcal{CN}(0, \sigma^2)$ denotes the distribution of a circularly symmetric complex Gaussian (CSCG) random variable whose real and imaginary parts are independent and normally distributed with mean zero and variance $\sigma^2/2$. $\mathcal{CN}(\mathbf{0}, \mathbf{R})$ denotes the distribution of a CSCG random vector with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{R}$ and $\sim$ stands for "distributed as". $\text{diag}(a_1, \ldots, a_N)$ refers to a diagonal matrix with diagonal elements being $a_1, \ldots, a_N$ and $\text{diag}(\mathbf{a})$ refers to a diagonal matrix with diagonal elements being the element of the vector $\mathbf{a}$. $\text{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_N)$ refers to a block diagonal matrix with blocks being $\mathbf{A}_1, \ldots, \mathbf{A}_N$. $\text{ones}(N)$ refers to an $N \times N$ matrix whose elements are 1.

1.5 Publications

The content of this thesis has appeared in the following publications and preprints, each corresponding to a specific chapter:

1. **M. Nerini**, H. Li, S. Shen, and B. Clerckx, *Beyond diagonal RIS.* Springer, 2024 (to appear as a book chapter). [Chapter 2]

2. **M. Nerini**, S. Shen, and B. Clerckx, "Closed-form global optimization of beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1037-1051, 2024. [Chapter 3]

3. **M. Nerini**, S. Shen, and B. Clerckx, "Static grouping strategy design for beyond diagonal reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, 2024. [Chapter 4]

4. **M. Nerini**, S. Shen, H. Li, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces utilizing graph theory: Modeling, architecture design, and optimization," *IEEE Trans. Wireless Commun.*, 2024. [Chapter 5]

5. **M. Nerini**, and B. Clerckx, "Pareto frontier for the performance-complexity trade-off in beyond diagonal reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2842–2846, 2023. [Chapter 6]

6. **M. Nerini**, S. Shen, and B. Clerckx, "Discrete-value group and fully connected architectures for beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Veh. Tehcnol.*, vol. 72, no. 12, pp. 16354–16368, 2023. [Chapter 7]

7. **M. Nerini**, S. Shen, H. Li, M. Di Renzo, and B. Clerckx, "A universal framework for multiport network analysis of reconfigurable intelligent surfaces," *arXiv preprint arXiv:2311.10561*, 2023 (sub. to *IEEE Trans. Wireless Commun.*). [Chapter 8]

8. **M. Nerini**, G. Ghiaasi, and B. Clerckx, "Localized and distributed beyond diagonal reconfigurable intelligent surfaces with lossy interconnections: Modeling and optimization," *arXiv preprint arXiv:2402.05881*, 2024 (sub. to *IEEE Trans. Wireless Commun.*). [Chapter 9]

In addition, the work carried out during this Ph.D. degree has led to the following publications, which are not included in this thesis:

1. **M. Nerini**, and B. Clerckx, "Physically consistent modeling of stacked intelligent metasurfaces implemented with beyond diagonal RIS," *IEEE Commun. Lett.*, 2024.

2. **M. Nerini**, and B. Clerckx, "Overhead-free blockage detection and precoding through physics-based graph neural networks: LIDAR data meets ray tracing," *IEEE Wireless Commun. Lett.*, vol. 12, no. 3, pp. 565-569, 2023.

3. **M. Nerini**, V. Rizzello, M. Joham, W. Utschick, and B. Clerckx, "Machine learning-based CSI feedback with variable length in FDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 2886-2900, 2023.

Chapter 2

Fundamentals of BD-RIS

In this introductory chapter, we review the fundamental concepts that underpin the work presented in this thesis. We model a single-user SISO communication system aided by a BD-RIS with multiport network theory and show that BD-RIS is a generalization of conventional RIS, as firstly presented in [16].

The content of this chapter has appeared in [73].

2.1 BD-RIS-Aided System Model

Let us consider a single-user SISO scenario in which the communication is aided by an N_I-antenna RIS. The N_I antennas of the RIS are connected to a N_I-port reconfigurable impedance network, with scattering matrix $\boldsymbol{\Theta} \in \mathbb{C}^{N_I \times N_I}$ [74]. As widely adopted in the literature [16], we assume all the RIS antennas perfectly matched and we assume no mutual coupling between the RIS antennas throughout this thesis, unless otherwise specified[1].

Defining $x \in \mathbb{C}$ as the transmitted signal with power $P_T = \mathrm{E}[|x|^2]$ and $y \in \mathbb{C}$ as the received signal, we have $y = hx + n$, where n is the additive white Gaussian noise (AWGN) at the receiver. In particular, the channel h can be written as a function of the scattering matrix $\boldsymbol{\Theta}$ as

$$h = h_{RT} + \mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{h}_{IT}, \tag{2.1}$$

where $h_{RT} \in \mathbb{C}$, $\mathbf{h}_{RI} \in \mathbb{C}^{1 \times N_I}$, and $\mathbf{h}_{IT} \in \mathbb{C}^{N_I \times 1}$ refer to the channels from transmitter to receiver, from RIS to receiver, and from transmitter to RIS, respectively.

In addition to the scattering matrix $\boldsymbol{\Theta}$, the N_I-port reconfigurable impedance network can also be characterized by its impedance matrix $\mathbf{Z}_I \in \mathbb{C}^{N_I \times N_I}$, and its admittance matrix $\mathbf{Y}_I \in \mathbb{C}^{N_I \times N_I}$ [74]. According to microwave theory [74], the scattering matrix $\boldsymbol{\Theta}$ and the impedance matrix $\mathbf{Z}_I$ are related by

$$\boldsymbol{\Theta} = \left(\mathbf{Z}_I + Z_0\mathbf{I}\right)^{-1}\left(\mathbf{Z}_I - Z_0\mathbf{I}\right), \tag{2.2}$$

where Z_0 refers to the reference impedance used to compute the scattering matrix $\boldsymbol{\Theta}$,

[1] In practice, these two assumptions can be achieved by individually matching each RIS antenna to a reference impedance, e.g., $Z_0 = 50\,\Omega$, and by keeping the spacing between the RIS elements larger than half-wavelength. The effects of imperfect matching and mutual coupling are explored in Chapter 8.

usually set as $Z_0 = 50\,\Omega$. Besides, $\boldsymbol{\Theta}$ and $\mathbf{Y}_I$ are related by

$$\boldsymbol{\Theta} = (\mathbf{I} + Z_0\mathbf{Y}_I)^{-1}(\mathbf{I} - Z_0\mathbf{Y}_I), \tag{2.3}$$

since $\mathbf{Y}_I = \mathbf{Z}_I^{-1}$. In the case of an RIS implemented with a reciprocal and lossless reconfigurable impedance network, the matrices $\mathbf{Z}_I$, $\mathbf{Y}_I$, and $\boldsymbol{\Theta}$ are subject to the following constraints. First, assuming the RIS to be reciprocal, $\mathbf{Z}_I$, $\mathbf{Y}_I$, and $\boldsymbol{\Theta}$ are symmetric, i.e., $\mathbf{Z}_I = \mathbf{Z}_I^T$, $\mathbf{Y}_I = \mathbf{Y}_I^T$, and $\boldsymbol{\Theta} = \boldsymbol{\Theta}^T$. Second, assuming the RIS to be lossless, $\mathbf{Z}_I$ is purely reactive and writes as $\mathbf{Z}_I = j\mathbf{X}_I$, where $\mathbf{X}_I \in \mathbb{R}^{N_I \times N_I}$ denotes the reactance matrix of the N_I-port reconfigurable impedance network. Similarly, $\mathbf{Y}_I$ is purely susceptive and writes as $\mathbf{Y}_I = j\mathbf{B}_I$, with $\mathbf{B}_I \in \mathbb{R}^{N_I \times N_I}$ being the susceptance matrix of the N_I-port reconfigurable impedance network. As a consequence, $\boldsymbol{\Theta}$ write as

$$\boldsymbol{\Theta} = (j\mathbf{X}_I + Z_0\mathbf{I})^{-1}(j\mathbf{X}_I - Z_0\mathbf{I}), \tag{2.4}$$

or, equivalently, as

$$\boldsymbol{\Theta} = (\mathbf{I} + jZ_0\mathbf{B}_I)^{-1}(\mathbf{I} - jZ_0\mathbf{B}_I), \tag{2.5}$$

meaning that it is unitary, i.e., $\boldsymbol{\Theta}^H\boldsymbol{\Theta} = \mathbf{I}$. Depending on the circuit topology of the N_I-port reconfigurable impedance network, the matrices $\mathbf{Z}_I$, $\mathbf{Y}_I$, and $\boldsymbol{\Theta}$ satisfy different constraints, resulting in conventional RIS and BD-RIS[2].

2.2 BD-RIS Architectures

In the following, we review the conventional single-connected RIS architecture and present other two BD-RIS architectures, recently proposed in [16].

2.2.1 Single-Connected RIS

The single-connected RIS is the conventional RIS architecture commonly adopted in the RIS literature. In single-connected RIS, the n_Ith port is connected to ground through a tunable impedance denoted as Z_{n_I}, for $n_I = 1, \ldots, N_I$, and is not interconnected to the other ports. Hence, its impedance matrix $\mathbf{Z}_I$ and admittance matrix $\mathbf{Y}_I$ are diagonal matrices, written as

$$\mathbf{Z}_I = \mathrm{diag}(Z_1, Z_2, \ldots, Z_{N_I}), \tag{2.6}$$

$$\mathbf{Y}_I = \mathrm{diag}(Y_1, Y_2, \ldots, Y_{N_I}), \tag{2.7}$$

where $Y_{n_I} = 1/Z_{n_I}$, for $n_I = 1, \ldots, N_I$. Furthermore, according to (2.2), also the scattering matrix $\boldsymbol{\Theta}$ is diagonal, and writes as

$$\boldsymbol{\Theta} = \mathrm{diag}(\Theta_1, \Theta_2, \ldots, \Theta_{N_I}), \tag{2.8}$$

[2]In some of the following chapters, the number of RIS elements N_I is denoted as N and the matrices $\mathbf{Z}_I$, $\mathbf{Y}_I$, $\mathbf{X}_I$, and $\mathbf{B}_I$, are denoted as $\mathbf{Z}$, $\mathbf{Y}$, $\mathbf{X}$, and $\mathbf{B}$, respectively, to ease the notation.

where Θ_{n_I} is the reflection coefficient corresponding to Z_{n_I}, i.e.,

$$\Theta_{n_I} = \frac{Z_{n_I} - Z_0}{Z_{n_I} + Z_0}, \tag{2.9}$$

for $n_I = 1, \ldots, N_I$. In the case of a lossless RIS, we have $Z_{n_I} = jX_{n_I}$, yielding

$$\Theta_{n_I} = \frac{jX_{n_I} - Z_0}{jX_{n_I} + Z_0} = e^{j\theta_{n_I}}, \tag{2.10}$$

where θ_{n_I} is the phase shift for the n_Ith RIS port, for $n_I = 1, \ldots, N_I$. Thus, the scattering matrix assumes in this case the form of the well-known phase shift matrix

$$\boldsymbol{\Theta} = \mathrm{diag}\left(e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_{N_I}}\right), \tag{2.11}$$

which is commonly used in RIS literature [12, 13]. This single-connected circuit topology is the simplest RIS circuit topology, which is characterized by limited flexibility and enables limited performance [16]. In a single-connected RIS, there are N_I tunable admittance components, each connecting an RIS element to ground, yielding a circuit complexity

$$C^{\text{Single}} = N_I. \tag{2.12}$$

As an illustrative example, we report the circuit topology of a 4-element single-connected RIS in Fig. 2.1(a).

2.2.2 Fully-Connected RIS

The first two BD-RIS architectures, namely the fully- and group-connected, have been proposed in [16]. The fully-connected RIS architecture is obtained by connecting every port of the reconfigurable impedance network to all other ports. Therefore, the matrices $\mathbf{Z}_I$ and $\mathbf{Y}_I$ are not restricted to be diagonal and can be any arbitrary symmetric matrices. In the case of a lossless and reciprocal fully-connected RIS, $\boldsymbol{\Theta}$ is a complex symmetric unitary matrix, i.e.,

$$\boldsymbol{\Theta} = \boldsymbol{\Theta}^T,\ \boldsymbol{\Theta}^H\boldsymbol{\Theta} = \mathbf{I}, \tag{2.13}$$

as a consequence of (2.4). Comparing the constraints (2.11) and (2.13), we notice that the fully-connected architecture is a generalization of the single-connected one. More precisely, the scattering matrix $\boldsymbol{\Theta}$ of the fully-connected architecture is not limited to being diagonal with unit-modulus entries because of the presence of reconfigurable impedances connecting the ports to each other. Since every RIS element is connected to ground and to all other elements, the circuit complexity of a fully-connected RIS is given by

$$C^{\text{Fully}} = N_I + \binom{N_I}{2} = \frac{N_I(N_I+1)}{2}. \tag{2.14}$$

The fully-connected RIS has the most general constraint on $\boldsymbol{\Theta}$ and achieves the highest received signal power [16]. This is enabled by the circuit topology, which is characterized by the highest complexity since every RIS port is connected to all other ports. We report

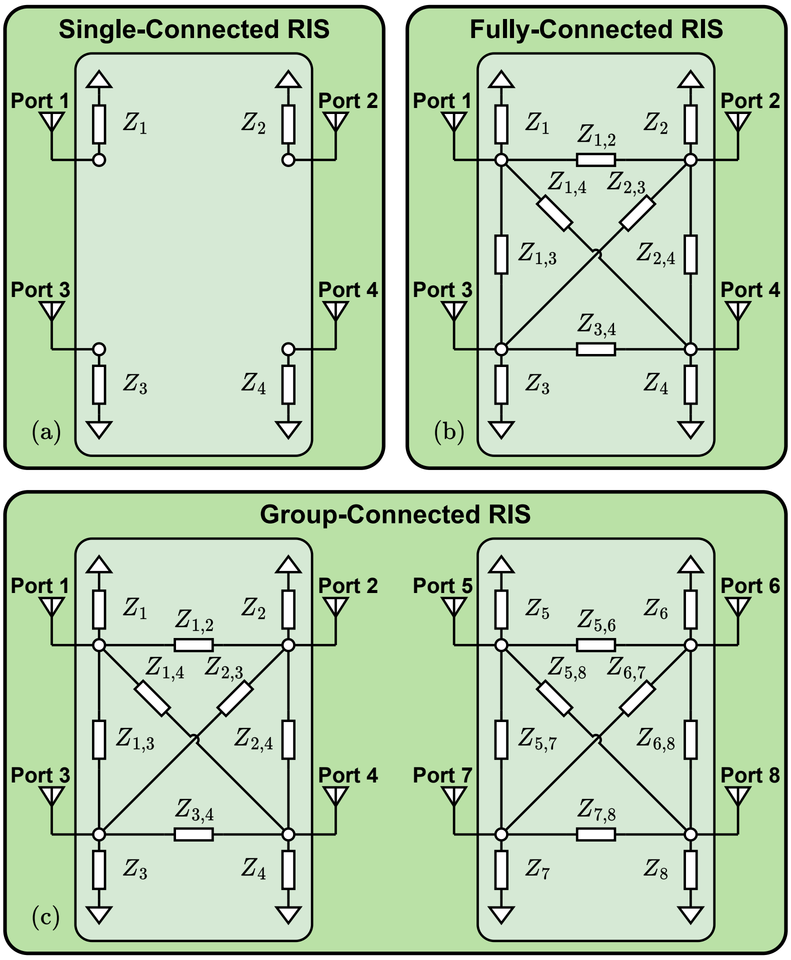


Figure 2.1: (a) 4-element single-connected RIS, (b) 4-element fully-connected RIS, and (c) 8-element group-connected RIS with group size 4.

a 4-element fully-connected RIS in Fig. 2.1(b).

2.2.3 Group-Connected RIS

The group-connected RIS architecture has been proposed as a trade-off between the single-connected and the fully-connected to achieve a favorable trade-off between performance and complexity [16]. In a group-connected RIS, the N_I elements are divided into G groups, each having $N_G = N_I/G$ elements. Each element is connected to all other elements in its group, while there is no connection between elements in different groups. Thus, $\mathbf{Z}_I$ is a block diagonal matrix given by

$$\mathbf{Z}_I = \mathrm{diag}\left(\mathbf{Z}_{I,1}, \mathbf{Z}_{I,2}, \ldots, \mathbf{Z}_{I,G}\right), \; \mathbf{Z}_{I,g} = \mathbf{Z}_{I,g}^T, \; \forall g, \tag{2.15}$$

where $\mathbf{Z}_{I,g} \in \mathbb{C}^{N_G \times N_G}$ is the impedance matrix of the N_G-port fully-connected reconfigurable impedance network for the gth group. In the case of a lossless and reciprocal group-connected RIS, we have $\mathbf{Z}_I = j\mathbf{X}_I$, with $\mathbf{X}_I$ satisfying

$$\mathbf{X}_I = \mathrm{diag}\left(\mathbf{X}_{I,1}, \mathbf{X}_{I,2}, \ldots, \mathbf{X}_{I,G}\right), \; \mathbf{X}_{I,g} = \mathbf{X}_{I,g}^T, \; \forall g, \tag{2.16}$$

where $\mathbf{X}_{I,g} \in \mathbb{R}^{N_G \times N_G}$ is the reactance matrix of the gth group. Similarly, also the matrices $\mathbf{Y}_I$ and $\mathbf{B}_I$ are symmetric and block diagonal. Thus, according to (2.4), the scattering matrix is subject to

$$\mathbf{\Theta} = \mathrm{diag}\left(\mathbf{\Theta}_1, \mathbf{\Theta}_2, \ldots, \mathbf{\Theta}_G\right), \; \mathbf{\Theta}_g = \mathbf{\Theta}_g^T, \; \mathbf{\Theta}_g^H \mathbf{\Theta}_g = \mathbf{I}, \; \forall g, \tag{2.17}$$

which show that $\mathbf{\Theta}$ is a block diagonal matrix with each block $\mathbf{\Theta}_g$ being a complex symmetric unitary matrix, for $g = 1, \ldots, G$. Since the group-connected architecture is composed of G fully-connected architectures with N_G elements each, its circuit complexity is given by

$$C^{\text{Group}} = G \frac{N_G(N_G + 1)}{2} = \frac{N_I(N_G + 1)}{2}, \tag{2.18}$$

depending on the group size N_G. Remarkably, the single- and fully-connected architectures can be seen as two special cases of the group-connected RIS, i.e., with group size $N_G = 1$ and $N_G = N_I$, respectively. As an example, we report an 8-element group-connected RIS with $N_G = 4$ in Fig. 2.1(c).

Part I

Optimization of Existing BD-RIS Architectures

Chapter 3

Global Optimization of Group- and Fully-Connected RISs

Group- and fully-connected RISs have been proposed in [16] as BD-RIS architectures with more flexibility than the conventional RIS architecture. However, a closed-form solution for the global optimal scattering matrix of group- and fully-connected architectures is not yet available. The scattering matrix has been optimized in recent literature by employing costly iterative optimization algorithms [16]. For this reason, it was possible to show that the theoretical performance upper bounds are tight only numerically.

In this chapter, we provide a closed-form global optimal solution for the scattering matrix of group- and fully-connected RISs. The resulting scattering matrix is proved to exactly achieve the received signal power upper bounds derived in [16] for SISO systems. Thus, we mathematically prove that these upper bounds are tight. Furthermore, we show that our algorithm is less complex than the iterative optimization methods applied to design the scattering matrix in the recent literature [16]. The complexity of our algorithm grows linearly with the number of RIS elements in the case of group-connected architectures, while it grows cubically in fully-connected architectures. Given the non-convexity of the involved optimization problems, this is the first study deriving a closed-form global optimal solution for BD-RISs. Our solution for single-user SISO systems is also proven to be general enough to allow the optimization of multiple problems in RIS-aided multi-antenna systems. The contributions of this chapter are summarized as follows:

First, as the main contribution, we provide a low-complexity closed-form global optimal solution for the scattering matrix of BD-RISs working in reflective mode applied to single-user SISO systems. In these systems, group- and fully-connected RISs designed with our solution exactly achieve their performance upper bounds. The upper bound-achieving property of our solution is valid for any channel realization, with no assumptions on its distribution and correlation.

Second, we consider BD-RISs working in transmissive mode, enabled by the cell-wise group-connected architecture proposed in [65]. We show how our optimal solution can be exploited to also globally optimize these BD-RISs. Also in the case of transmissive mode, the performance upper bounds are always exactly achieved by BD-RISs optimized through our solution.

Third, we exploit our optimal solution to optimize the RIS scattering matrix in single-user MIMO systems, including MISO systems as a special case. For systems aided by a fully-connected RIS and with negligible direct link, we derive a tight upper bound on the received signal power. We show that such an upper bound can be always exactly achieved with our optimal strategy. In addition, we also propose an efficient sub-optimal solution for the case in which the direct link is not negligible. In this case, a tight upper bound on the received signal power is not known.

Fourth, we study the weighted sum power maximization problem in multi-user MISO systems. In the case of systems aided by a fully-connected RIS and with negligible direct links, we provide a tight performance upper bound and an optimal solution to achieve it. Also for multi-user MISO systems, we provide a sub-optimal solution to design the RIS in the case the direct links are not negligible. In fact, a tight performance upper bound is not available in this case.

Organization: In Section 3.1, we derive the upper bound-achieving closed-form solution for the scattering matrix in single-user SISO systems. In Section 3.2, we show that our solution can be also applied to optimally design BD-RISs working in transmissive mode. In Sections 3.3 and 3.4, we extend our closed-form solution to single-user MIMO, and multi-user MISO systems, respectively. In Section 3.5, we assess the obtained performance through numerical simulations. Finally, Section 3.6 contains the concluding remarks.

The content of this chapter has appeared in [75]. For reproducible research, the code is available at `https://github.com/matteonerini/optimization-of-bdris`.

3.1 BD-RIS-Aided Single-User SISO Systems: Reflective Mode

Consider the single-user BD-RIS-aided SISO system model introduced in Section 2.1. In this section, the BD-RIS is assumed to work in reflective mode, as typically considered in the literature. Our goal is to design $\mathbf{\Theta}$ for group- and fully-connected RISs to maximize the received signal power given by $P_R = P_T \left|h_{RT} + \mathbf{h}_{RI}\mathbf{\Theta}\mathbf{h}_{IT}\right|^2$. Thus, the optimization problem for group-connected RISs writes as

$$\max_{\mathbf{\Theta}} \; P_T \left|h_{RT} + \mathbf{h}_{RI}\mathbf{\Theta}\mathbf{h}_{IT}\right|^2 \tag{3.1}$$

$$\text{s.t.} \quad \mathbf{\Theta} = \text{diag}\left(\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_G\right), \; \mathbf{\Theta}_g = \mathbf{\Theta}_g^T, \; \mathbf{\Theta}_g^H\mathbf{\Theta}_g = \mathbf{I}, \, \forall g, \tag{3.2}$$

where $P_T = \text{E}[|x|^2]$ is the transmitted signal power. Since fully-connected RISs can be viewed as a special case of group-connected RISs, the analogous problem for fully-connected RISs can be readily obtained by setting $G = 1$ in (3.2). Remarkably, the term $\mathbf{h}_{RI}\mathbf{\Theta}\mathbf{h}_{IT}$ can be always combined in phase with h_{RT} by applying a suitable phase shift to $\mathbf{\Theta}$. Thus, we first maximize (3.1) by omitting h_{RT} and then we adjust the phase of the resulting $\mathbf{\Theta}$ depending on $\arg\left(h_{RT}\right)$. We assume unit P_T and introduce the normalized channels $\hat{\mathbf{h}}_{RI} = \mathbf{h}_{RI}/\left\|\mathbf{h}_{RI}\right\|$ and $\hat{\mathbf{h}}_{IT} = \mathbf{h}_{IT}/\left\|\mathbf{h}_{IT}\right\|$ such that our problem becomes to maximize

$$\hat{P}_R = |\hat{\mathbf{h}}_{RI}\mathbf{\Theta}\hat{\mathbf{h}}_{IT}|^2. \tag{3.3}$$

For traditional single-connected architectures, it is known that the scattering matrix can be simply optimized in closed-form. The maximum normalized received signal power is

$$\hat{P}_R^{\text{Single}} = \left(\sum_{n_I=1}^{N_I} \left| [\hat{\mathbf{h}}_{RI}]_{n_I} [\hat{\mathbf{h}}_{IT}]_{n_I} \right| \right)^2, \tag{3.4}$$

which is achieved by designing $\boldsymbol{\Theta}$ as in (2.11) with

$$\theta_{n_I} = -\arg\left([\hat{\mathbf{h}}_{RI}]_{n_I} \right) - \arg\left([\hat{\mathbf{h}}_{IT}]_{n_I} \right), \forall n_I. \tag{3.5}$$

However, an exact solution for the optimal $\boldsymbol{\Theta}$ in group- and fully-connected RISs is not known given the non-convexity of problem (3.1)-(3.2). In the following, we consider the global optimization of fully-connected RISs in Section 3.1.1. To this end, we provide an upper bound on the objective function (3.3) and a necessary and sufficient condition to achieve it. This condition is subsequently transformed into an equivalent condition, for which a closed-form solution is available. We generalize our approach to group-connected RISs in Section 3.1.2.

3.1.1 Closed-Form Solution for Optimal Fully-Connected RIS

We begin by observing that constraints (3.2) are equivalent to (2.4) and (2.16). Since $\mathbf{X}_I$ is real symmetric, we can use the eigenvalue decomposition to write $\mathbf{X}_I = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T$, where $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_{N_I}) \in \mathbb{R}^{N_I \times N_I}$ is a diagonal matrix containing the eigenvalues of $\mathbf{X}_I$ ordered in decreasing order and $\mathbf{V} \in \mathbb{R}^{N_I \times N_I}$ is orthonormal. Applying (2.4), the scattering matrix $\boldsymbol{\Theta}$ is given by

$$\boldsymbol{\Theta} = \left(j\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T + Z_0\mathbf{I} \right)^{-1} \left(j\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T - Z_0\mathbf{I} \right) = \mathbf{V}\mathbf{D}\mathbf{V}^T, \tag{3.6}$$

where $\mathbf{D} = \text{diag}\left(e^{jd_1}, \ldots, e^{jd_{N_I}} \right) \in \mathbb{C}^{N_I \times N_I}$ is a diagonal matrix with $e^{jd_{n_I}} = \frac{j\lambda_{n_I} - Z_0}{j\lambda_{n_I} + Z_0}$. Note that the complex diagonal elements of the matrix $\mathbf{D}$ have unit modulus by construction. Using the decomposition of $\boldsymbol{\Theta}$ given by (3.6), the normalized received signal power $\hat{P}_R$ in (3.3) can be expressed as

$$\hat{P}_R = |\hat{\mathbf{h}}_{RI}\mathbf{V}\mathbf{D}\mathbf{V}^T\hat{\mathbf{h}}_{IT}|^2 = \left| \bar{\mathbf{h}}_{RI}\mathbf{D}\bar{\mathbf{h}}_{IT} \right|^2, \tag{3.7}$$

where $\bar{\mathbf{h}}_{RI} = \hat{\mathbf{h}}_{RI}\mathbf{V}$ and $\bar{\mathbf{h}}_{IT} = \mathbf{V}^T\hat{\mathbf{h}}_{IT}$. Note that (3.7) is the squared modulus of the dot product between $\bar{\mathbf{h}}_{RI}$ and $\mathbf{D}\bar{\mathbf{h}}_{IT}$. Thus, using the Cauchy-Schwarz inequality, we have

$$\hat{P}_R \leq \left\| \bar{\mathbf{h}}_{RI} \right\|^2 \left\| \mathbf{D}\bar{\mathbf{h}}_{IT} \right\|^2 = 1, \tag{3.8}$$

where the equality $\hat{P}_R = 1$ is achieved if and only if

$$\left| \left[\bar{\mathbf{h}}_{RI} \right]_{n_I} \right| = \left| \left[\bar{\mathbf{h}}_{IT} \right]_{n_I} \right|, \forall n_I. \tag{3.9}$$

Since we are interested in achieving the received signal power upper bound, our goal is now to find a real orthonormal matrix $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_{N_I}]$ such that condition (3.9) is satisfied.

It is easy to recognize that if the channels $\mathbf{h}_{RI}$ and $\mathbf{h}_{IT}$ are linearly dependent, the optimal $\mathbf{V}$ is $\mathbf{V}=\mathbf{I}$. Consequently, $\mathbf{D}$ can be designed according to (3.5) and the matrix $\mathbf{\Theta}=\mathbf{V}\mathbf{D}\mathbf{V}^T$ is readily obtained. For this reason, in the following discussion, we assume that the channels $\mathbf{h}_{RI}$ and $\mathbf{h}_{IT}$ are linearly independent.

Our objective is now to transform the optimality condition (3.9) into an equivalent condition for which a closed-form solution can be derived. Noting that $\left[\bar{\mathbf{h}}_{RI}\right]_{n_I}=\hat{\mathbf{h}}_{RI}\mathbf{v}_{n_I}$ and $\left[\bar{\mathbf{h}}_{IT}\right]_{n_I}=\mathbf{v}_{n_I}^T\hat{\mathbf{h}}_{IT}$, condition (3.9) becomes equivalent to

$$\left|\hat{\mathbf{h}}_{RI}\mathbf{v}_{n_I}\right|^2=\left|\mathbf{v}_{n_I}^T\hat{\mathbf{h}}_{IT}\right|^2, \tag{3.10}$$

which can be in turn rewritten as

$$\mathbf{v}_{n_I}^T\mathbf{R}_{RI}\mathbf{v}_{n_I}=\mathbf{v}_{n_I}^T\mathbf{R}_{IT}\mathbf{v}_{n_I}, \tag{3.11}$$

where $\mathbf{R}_{RI}=\hat{\mathbf{h}}_{RI}^H\hat{\mathbf{h}}_{RI}\in\mathbb{C}^{N_I\times N_I}$ and $\mathbf{R}_{IT}=\hat{\mathbf{h}}_{IT}\hat{\mathbf{h}}_{IT}^H\in\mathbb{C}^{N_I\times N_I}$. The left- and right-hand sides of (3.11) are quadratic forms. Since $\mathbf{v}_{n_I}^T\mathbf{R}_{RI}\mathbf{v}_{n_I}=\mathbf{v}_{n_I}^T\mathbf{R}_{RI}^T\mathbf{v}_{n_I}$ and $\mathbf{v}_{n_I}^T\mathbf{R}_{IT}\mathbf{v}_{n_I}=\mathbf{v}_{n_I}^T\mathbf{R}_{IT}^T\mathbf{v}_{n_I}$, we can replace in (3.11) the matrices $\mathbf{R}_{RI}$ and $\mathbf{R}_{IT}$ with their symmetric parts $\mathbf{A}_{RI}=1/2\left(\mathbf{R}_{RI}+\mathbf{R}_{RI}^T\right)\in\mathbb{R}^{N_I\times N_I}$ and $\mathbf{A}_{IT}=1/2\left(\mathbf{R}_{IT}+\mathbf{R}_{IT}^T\right)\in\mathbb{R}^{N_I\times N_I}$, respectively, without changing the two quadratic forms. Thus, (3.11) is equivalent to

$$\mathbf{v}_{n_I}^T\mathbf{A}_{RI}\mathbf{v}_{n_I}=\mathbf{v}_{n_I}^T\mathbf{A}_{IT}\mathbf{v}_{n_I}, \tag{3.12}$$

which in turn becomes

$$\mathbf{v}_{n_I}^T\mathbf{A}\mathbf{v}_{n_I}=0, \tag{3.13}$$

where the symmetric matrix $\mathbf{A}=\mathbf{A}_{RI}-\mathbf{A}_{IT}\in\mathbb{R}^{N_I\times N_I}$ has been introduced. To solve (3.13), let us consider the eigenvalue decomposition $\mathbf{A}=\mathbf{U}\mathbf{\Delta}\mathbf{U}^T$, where $\mathbf{\Delta}=\operatorname{diag}\left(\delta_1,\ldots,\delta_{N_I}\right)\in\mathbb{R}^{N_I\times N_I}$ is a diagonal matrix containing the eigenvalues of $\mathbf{A}$ ordered in decreasing order and $\mathbf{U}\in\mathbb{R}^{N_I\times N_I}$ is orthonormal. By introducing the orthonormal vectors $\mathbf{t}_{n_I}=\mathbf{U}^T\mathbf{v}_{n_I}\in\mathbb{R}^{N_I\times 1}$, for $n_I=1,\ldots,N_I$, (3.13) can be reformulated as a diagonal quadratic form

$$\mathbf{t}_{n_I}^T\mathbf{\Delta}\mathbf{t}_{n_I}=0. \tag{3.14}$$

In other words, we need to solve $\mathbf{s}_{n_I}\boldsymbol{\delta}=0$, where $\mathbf{s}_{n_I}=[[\mathbf{t}_{n_I}]_1^2,\ldots,[\mathbf{t}_{n_I}]_{N_I}^2]\in\mathbb{R}^{1\times N_I}$ and $\boldsymbol{\delta}=[\delta_1,\ldots,\delta_{N_I}]^T\in\mathbb{R}^{N_I\times 1}$. Note that we need to find N_I orthonormal vectors $\mathbf{t}_{n_I}$ which are solutions of (3.14). This task is hard in general since the solution space of (3.14) is a non-linear space. However, in our case, we can rely on the special structure of the vector $\boldsymbol{\delta}$. As proved in the following, $\boldsymbol{\delta}$ contains only two, three, and four non-zero elements when $N_I=2$, $N_I=3$, and $N_I\geq 4$, respectively. Thus, we can solve (3.14) in closed-form by separately studying these three cases. The following proposition is introduced to simplify (3.14) in the cases $N_I\in\{2,3\}$.

Proposition 1. *For any linearly independent* $\mathbf{h}_{RI}\in\mathbb{C}^{1\times N_I}$ *and* $\mathbf{h}_{IT}\in\mathbb{C}^{N_I\times 1}$, *with* $N_I\in\{2,3\}$, *the matrix* $\mathbf{A}$ *has rank* $r\left(\mathbf{A}\right)=N_I$ *and trace* $\operatorname{Tr}\left(\mathbf{A}\right)=0$.

Proof. Please refer to Appendix 3.7.1. □

Case $N_I = 2$

In the case of fully-connected RISs with $N_I = 2$, $\mathbf{A}$ has two eigenvalues, both non-zero and one opposite of the other, as a consequence of Proposition 1. Denoting the two eigenvalues of $\mathbf{A}$ as δ_1 and δ_2, we have that the vector $\boldsymbol{\delta}$ writes as $\boldsymbol{\delta} = [\delta_1, \delta_2]^T$, where $\delta_2 = -\delta_1$. Applying Proposition 1, we simplify (3.14) as

$$\delta_1 [\mathbf{t}_{n_I}]_1^2 - \delta_1 [\mathbf{t}_{n_I}]_2^2 = 0. \tag{3.15}$$

Thus, we need to solve

$$\begin{cases} [\mathbf{t}_{n_I}]_1^2 - [\mathbf{t}_{n_I}]_2^2 = 0 \\ [\mathbf{t}_{n_I}]_1^2 + [\mathbf{t}_{n_I}]_2^2 = 1 \end{cases}, \tag{3.16}$$

where the first equation is derived from (3.15) and the second equation is the unitary norm constraint on $\mathbf{t}_{n_I}$, $\forall n_I \in \{1, 2\}$. Solving by substitution, we obtain $[\mathbf{t}_{n_I}]_1^2 = [\mathbf{t}_{n_I}]_2^2 = 1/2$. Finally, we choose $\mathbf{t}_1 = [\sqrt{1/2}, \sqrt{1/2}]^T$ and $\mathbf{t}_2 = [\sqrt{1/2}, -\sqrt{1/2}]^T$ to guarantee orthonormality.

Case $N_I = 3$

Considering fully-connected RISs with $N_I = 3$, we still rely on proposition 1 to simplify (3.14). As a consequence of Proposition 1, $\mathbf{A}$ has three eigenvalues, all non-zero. Denoting the three eigenvalues of $\mathbf{A}$ as δ_1, δ_2, and δ_3, we have that the vector $\boldsymbol{\delta}$ writes as $\boldsymbol{\delta} = [\delta_1, \delta_2, \delta_3]^T$. Applying Proposition 1, we simplify (3.14) as

$$\delta_1 [\mathbf{t}_{n_I}]_1^2 + \delta_2 [\mathbf{t}_{n_I}]_2^2 + \delta_3 [\mathbf{t}_{n_I}]_3^2 = 0. \tag{3.17}$$

We choose the vector $\mathbf{t}_1$ with only the first and the third entries non-zero. Such a vector always exists since Proposition 1 implies $\delta_1 > 0$ and $\delta_3 < 0$. Thus, we need to solve

$$\begin{cases} \delta_1 [\mathbf{t}_1]_1^2 + \delta_3 [\mathbf{t}_1]_3^2 = 0 \\ [\mathbf{t}_1]_1^2 + [\mathbf{t}_1]_3^2 = 1 \end{cases}, \tag{3.18}$$

where the first equation is derived from (3.17) and the second equation is the unitary norm constraint. Solving by substitution, we obtain

$$\begin{cases} [\mathbf{t}_1]_1^2 = \frac{-\delta_3}{\delta_1 - \delta_3} \\ [\mathbf{t}_1]_3^2 = 1 - \frac{-\delta_3}{\delta_1 - \delta_3} = \frac{\delta_1}{\delta_1 - \delta_3} \end{cases}, \tag{3.19}$$

giving $\mathbf{t}_1 = \left[\sqrt{\frac{-\delta_3}{\delta_1 - \delta_3}}, 0, \sqrt{\frac{\delta_1}{\delta_1 - \delta_3}}\right]^T$. Now, we select the two remaining vectors in the form

$$\mathbf{t}_2 = \left[\frac{1}{K}\sqrt{\frac{\delta_1}{\delta_1 - \delta_3}}, \sqrt{1 - \frac{1}{K^2}}, -\frac{1}{K}\sqrt{\frac{-\delta_3}{\delta_1 - \delta_3}}\right]^T, \tag{3.20}$$

$$\mathbf{t}_3 = \left[-\frac{1}{K}\sqrt{\frac{\delta_1}{\delta_1 - \delta_3}}, \sqrt{1 - \frac{1}{K^2}}, \frac{1}{K}\sqrt{\frac{-\delta_3}{\delta_1 - \delta_3}}\right]^T, \tag{3.21}$$

where K is a positive constant. It is easy to recognize that $\mathbf{t}_1$, $\mathbf{t}_2$, and $\mathbf{t}_3$ are an orthonormal basis of $\mathbb{R}^3$ for any $K \neq 1$. Thus, K must be designed such that $\mathbf{t}_2$ and $\mathbf{t}_3$ satisfy (3.17), that is

$$\frac{\delta_1^2}{K^2\left(\delta_1-\delta_3\right)}+\delta_2\left(1-\frac{1}{K^2}\right)-\frac{\delta_3^2}{K^2\left(\delta_1-\delta_3\right)}=0. \tag{3.22}$$

Equation (3.22) can be simplified by substituting $\delta_2 = -\delta_1 - \delta_3$, which is always valid according to Proposition 1. Eventually, (3.22) gives $K = \sqrt{2}$, yielding

$$\mathbf{t}_2=\left[\sqrt{\frac{\delta_1}{2\left(\delta_1-\delta_3\right)}},\sqrt{\frac{1}{2}},-\sqrt{\frac{-\delta_3}{2\left(\delta_1-\delta_3\right)}}\right]^T, \tag{3.23}$$

$$\mathbf{t}_3=\left[-\sqrt{\frac{\delta_1}{2\left(\delta_1-\delta_3\right)}},\sqrt{\frac{1}{2}},\sqrt{\frac{-\delta_3}{2\left(\delta_1-\delta_3\right)}}\right]^T. \tag{3.24}$$

Case $N_I \geq 4$

In the case of fully-connected RISs with $N_I \geq 4$, we introduce the following proposition to simplify (3.14).

Proposition 2. *For any linearly independent* $\mathbf{h}_{RI} \in \mathbb{C}^{1\times N_I}$ *and* $\mathbf{h}_{IT} \in \mathbb{C}^{N_I\times 1}$*, with* $N_I \geq 4$*, the matrix* $\mathbf{A}$ *has rank* $r(\mathbf{A}) = 4$ *and trace* $\mathrm{Tr}(\mathbf{A}) = 0$*. Furthermore, among its four non-zero eigenvalues, two are positive and two are negative.*

Proof. Please refer to Appendix 3.7.2. □

Denoting the first two eigenvalues of $\mathbf{A}$ as δ_1 and δ_2, and the last two as δ_{N_I-1} and δ_{N_I}, we have that the vector $\boldsymbol{\delta}$ writes as $\boldsymbol{\delta} = [\delta_1, \delta_2, 0, \ldots, 0, \delta_{N_I-1}, \delta_{N_I}]^T$. Applying Proposition 2, we simplify (3.14) as

$$\delta_1\left[\mathbf{t}_{n_I}\right]_1^2+\delta_2\left[\mathbf{t}_{n_I}\right]_2^2+\delta_{N_I-1}\left[\mathbf{t}_{n_I}\right]_{N_I-1}^2+\delta_{N_I}\left[\mathbf{t}_{n_I}\right]_{N_I}^2=0. \tag{3.25}$$

We notice that N_I-4 orthonormal solutions to (3.25) are given by the vectors $\mathbf{e}_3, \ldots, \mathbf{e}_{N_I-2}$, where $\mathbf{e}_i \in \mathbb{R}^{N_I\times 1}$ denotes the vector with the ith entry being 1 and the others being 0, for $i = 3, \ldots, N_I-2$. Thus, we now want to find the remaining four orthonormal vectors $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$, $\mathbf{t}_4 \in \mathbb{R}^{N_I}$ solutions of (3.25), all orthogonal to $\mathbf{e}_3, \ldots, \mathbf{e}_{N_I-2}$. To make them orthogonal to $\mathbf{e}_3, \ldots, \mathbf{e}_{N_I-2}$, it is sufficient to set $[\mathbf{t}_i]_{n_I} = 0$ for $i = 1, 2, 3, 4$ and $n_I = 3, \ldots, N_I - 2$.

We choose the first vector $\mathbf{t}_1$ with only the first and the $(N_I - 1)$th entries non-zero. Note that such a vector always exists since $\delta_1 > 0$ and $\delta_{N_I-1} < 0$. Thus, we need to solve

$$\begin{cases}\delta_1\left[\mathbf{t}_1\right]_1^2+\delta_{N_I-1}\left[\mathbf{t}_1\right]_{N_I-1}^2=0\\ \left[\mathbf{t}_1\right]_1^2+\left[\mathbf{t}_1\right]_{N_I-1}^2=1\end{cases}, \tag{3.26}$$

where the first equation is derived from (3.25) and the second equation is the unitary norm constraint. Solving by substitution, we obtain

$$\begin{cases}\left[\mathbf{t}_1\right]_1^2=\frac{-\delta_{N_I-1}}{\delta_1-\delta_{N_I-1}}\\ \left[\mathbf{t}_1\right]_{N_I-1}^2=1-\frac{-\delta_{N_I-1}}{\delta_1-\delta_{N_I-1}}=\frac{\delta_1}{\delta_1-\delta_{N_I-1}}\end{cases}, \tag{3.27}$$

which gives $\mathbf{t}_1 = \left[\sqrt{\frac{-\delta_{N_I-1}}{\delta_1-\delta_{N_I-1}}}, 0, \ldots, \sqrt{\frac{\delta_1}{\delta_1-\delta_{N_I-1}}}, 0\right]^T$. Similarly, we choose the second vector $\mathbf{t}_2$ with only the second and N_Ith entries non-zero. Also this vector always exists since $\delta_2 > 0$ and $\delta_{N_I} < 0$. With a similar procedure, we obtain

$$\begin{cases} [\mathbf{t}_2]_2^2 = \frac{-\delta_{N_I}}{\delta_2-\delta_{N_I}} \\ [\mathbf{t}_2]_{N_I}^2 = 1 - \frac{-\delta_{N_I}}{\delta_2-\delta_{N_I}} = \frac{\delta_2}{\delta_2-\delta_{N_I}} \end{cases}, \tag{3.28}$$

giving $\mathbf{t}_2 = \left[0, \sqrt{\frac{-\delta_{N_I}}{\delta_2-\delta_{N_I}}}, \ldots, 0, \sqrt{\frac{\delta_2}{\delta_2-\delta_{N_I}}}\right]^T$. Note that these first two vectors $\mathbf{t}_1$ and $\mathbf{t}_2$ are orthonormal by construction.

The remaining two vectors $\mathbf{t}_3$ and $\mathbf{t}_4$ must be linear combinations of a basis of the null space of the matrix $M = [\mathbf{t}_1, \mathbf{t}_2]^T$. Such a basis is readily given by two vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ in the form

$$\mathbf{b}_1 = [[\mathbf{t}_1]_{N_I-1}, 0, \ldots, -[\mathbf{t}_1]_1, 0]^T, \tag{3.29}$$

$$\mathbf{b}_2 = [0, [\mathbf{t}_2]_{N_I}, \ldots, 0, -[\mathbf{t}_2]_2]^T, \tag{3.30}$$

whose n_Ith entry is zero, for $n_I = 3, \ldots, N_I - 2$, in addition to the vectors $\mathbf{e}_3, \ldots, \mathbf{e}_{N_I-2}$. Thus, $\mathbf{t}_3$ and $\mathbf{t}_4$ can be expressed as a generic linear combination $\mathbf{c} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2$ given by

$$\mathbf{c} = \left[a_1 [\mathbf{t}_1]_{N_I-1}, a_2 [\mathbf{t}_2]_{N_I}, \ldots, -a_1 [\mathbf{t}_1]_1, -a_2 [\mathbf{t}_2]_2\right]^T. \tag{3.31}$$

Now, our objective is to find a_1 and a_2 satisfying (3.25) and the unitary norm constraint, that is

$$\begin{cases} \delta_1 a_1^2 [\mathbf{t}_1]_{N_I-1}^2 + \delta_2 a_2^2 [\mathbf{t}_2]_{N_I}^2 + \delta_{N_I-1} a_1^2 [\mathbf{t}_1]_1^2 + \delta_{N_I} a_2^2 [\mathbf{t}_2]_2^2 = 0 \\ a_1^2 [\mathbf{t}_1]_{N_I-1}^2 + a_2^2 [\mathbf{t}_2]_{N_I}^2 + a_1^2 [\mathbf{t}_1]_1^2 + a_2^2 [\mathbf{t}_2]_2^2 = 1 \end{cases}. \tag{3.32}$$

Substituting in (3.32) the entries of the vectors $\mathbf{t}_1$ and $\mathbf{t}_2$ given by (3.27) and (3.28), respectively, we obtain

$$\begin{cases} (\delta_1 + \delta_{N_I-1})a_1^2 + (\delta_2 + \delta_{N_I})a_2^2 = 0 \\ a_1^2 + a_2^2 = 1 \end{cases}. \tag{3.33}$$

Solving by substitution, we have

$$\begin{cases} a_1^2 = \frac{-\delta_2-\delta_{N_I}}{\delta_1+\delta_{N_I-1}-\delta_2-\delta_{N_I}} \\ a_2^2 = 1 - \frac{-\delta_2-\delta_{N_I}}{\delta_1+\delta_{N_I-1}-\delta_2-\delta_{N_I}} = \frac{\delta_1+\delta_{N_I-1}}{\delta_1+\delta_{N_I-1}-\delta_2-\delta_{N_I}} \end{cases}. \tag{3.34}$$

Note that this always means $a_1^2 = a_2^2 = 1/2$ since Proposition 2 gives $\delta_1+\delta_{N_I-1} = -\delta_2-\delta_{N_I}$. Finally, we choose $\mathbf{t}_3 = \sqrt{1/2}\mathbf{b}_1 + \sqrt{1/2}\mathbf{b}_2$ and $\mathbf{t}_4 = \sqrt{1/2}\mathbf{b}_1 - \sqrt{1/2}\mathbf{b}_2$ to guarantee orthonormality.

In conclusion, we construct an orthonormal matrix $\mathbf{T} \in \mathbb{R}^{N_I \times N_I}$ depending on the number of RIS elements N_I. If $N_I = 2$, $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2]$; if $N_I = 3$, $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3]$; and if $N_I \geq 4$, $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \mathbf{t}_4, \mathbf{e}_3, \ldots, \mathbf{e}_{N-2}]$. Note that the columns of $\mathbf{T}$ are orthogonal with

Algorithm 1 Optimal fully-connected RIS design for single-user SISO systems

Input: $\mathbf{h}_{RI} \in \mathbb{C}^{1\times N_I}$, $\mathbf{h}_{IT} \in \mathbb{C}^{N_I\times 1}$.

Output: $\bar{\boldsymbol{\Theta}}$.

$\hat{\mathbf{h}}_{RI} = \frac{\mathbf{h}_{RI}}{\|\mathbf{h}_{RI}\|}$, $\hat{\mathbf{h}}_{IT} = \frac{\mathbf{h}_{IT}}{\|\mathbf{h}_{IT}\|}$, $\mathbf{R}_{RI} = \hat{\mathbf{h}}_{RI}^H\hat{\mathbf{h}}_{RI}$, $\mathbf{R}_{IT} = \hat{\mathbf{h}}_{IT}\hat{\mathbf{h}}_{IT}^H$.

$\mathbf{A}_{RI} = \frac{\mathbf{R}_{RI}+\mathbf{R}_{RI}^T}{2}$, $\mathbf{A}_{IT} = \frac{\mathbf{R}_{IT}+\mathbf{R}_{IT}^T}{2}$, $\mathbf{A} \triangleq \mathbf{U}\boldsymbol{\Delta}\mathbf{U}^T = \mathbf{A}_{RI} - \mathbf{A}_{IT}$.

$\boldsymbol{\delta} \triangleq [\delta_1, \ldots, \delta_{N_I}]^T = \operatorname{diag}(\boldsymbol{\Delta})$.

if $N_I == 2$ **then**

$\quad \mathbf{T} = \begin{bmatrix} \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}} \end{bmatrix}$.

else if $N_I == 3$ **then**

$\quad \mathbf{T} = \begin{bmatrix} \sqrt{\frac{-\delta_3}{\delta_1-\delta_3}} & \sqrt{\frac{\delta_1}{2(\delta_1-\delta_3)}} & -\sqrt{\frac{\delta_1}{2(\delta_1-\delta_3)}} \\ 0 & \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \\ \sqrt{\frac{\delta_1}{\delta_1-\delta_3}} & -\sqrt{\frac{-\delta_3}{2(\delta_1-\delta_3)}} & \sqrt{\frac{-\delta_3}{2(\delta_1-\delta_3)}} \end{bmatrix}$.

else

$\quad \mathbf{t}_1 = \left[\sqrt{\frac{-\delta_{N_I-1}}{\delta_1-\delta_{N_I-1}}}, 0, \ldots, \sqrt{\frac{\delta_1}{\delta_1-\delta_{N_I-1}}}, 0\right]^T$, $\mathbf{t}_2 = \left[0, \sqrt{\frac{-\delta_{N_I}}{\delta_2-\delta_{N_I}}}, \ldots, 0, \sqrt{\frac{\delta_2}{\delta_2-\delta_{N_I}}}\right]^T$.

$\quad \mathbf{t}_3 = \frac{1}{\sqrt{2}}\left[[\mathbf{t}_1]_{N_I-1}, [\mathbf{t}_2]_{N_I}, \ldots, -[\mathbf{t}_1]_1, -[\mathbf{t}_2]_2\right]^T$.

$\quad \mathbf{t}_4 = \frac{1}{\sqrt{2}}\left[[\mathbf{t}_1]_{N_I-1}, -[\mathbf{t}_2]_{N_I}, \ldots, -[\mathbf{t}_1]_1, [\mathbf{t}_2]_2\right]^T$.

$\quad \mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \mathbf{t}_4, \mathbf{e}_3, \ldots, \mathbf{e}_{N-2}]$.

end if

$\mathbf{V} = \mathbf{U}\mathbf{T}$.

$d_{n_I} = -\arg\left([\hat{\mathbf{h}}_{RI}\mathbf{V}]_{n_I}\right) - \arg\left([\mathbf{V}^T\hat{\mathbf{h}}_{IT}]_{n_I}\right)$, $\forall n_I$, $\mathbf{D} = \operatorname{diag}\left(e^{jd_1}, \ldots, e^{jd_{N_I}}\right)$.

$\bar{\boldsymbol{\Theta}} = \mathbf{V}\mathbf{D}\mathbf{V}^T$.

Return $\bar{\boldsymbol{\Theta}}$.

each other, have unitary norm, and solve (3.14). At this stage, all the building elements of the optimal scattering matrix, denoted as $\bar{\boldsymbol{\Theta}}$, are available. Applying (3.6), we can write $\bar{\boldsymbol{\Theta}} = \mathbf{V}\mathbf{D}\mathbf{V}^T$, where $\mathbf{V} = \mathbf{U}\mathbf{T}$ by definition of the columns of $\mathbf{T}$, and $\mathbf{D}$ is designed according to (3.5). We summarize the steps necessary to build the optimal $\bar{\boldsymbol{\Theta}}$ in Alg. 1. Note that the solution provided by Alg. 1 is proved to be global optimal by the following two facts. First, it solves (3.14) by construction. Second, since (3.14) is equivalent to (3.9), it allows to exactly achieve the objective upper bound $\hat{P}_R = 1$. To maximize P_R in the presence of the direct link h_{RT}, the scattering matrix can be adjusted as

$$\boldsymbol{\Theta}^\star = e^{j\arg(h_{RT})}\bar{\boldsymbol{\Theta}}, \tag{3.35}$$

such that the term $\mathbf{h}_{RI}\boldsymbol{\Theta}^\star\mathbf{h}_{IT}$ is made in phase with h_{RT}.

3.1.2 Closed-Form Solution for Optimal Group-Connected RIS

Now, we extend our closed-form strategy to design fully-connected architectures to group-connected ones. As previously discussed, we initially omit the direct link h_{RT} and assume unitary transmitted signal power. Thus, the received signal power for group-connected architectures writes as

$$P_R = \left| \sum_{g=1}^{G} \mathbf{h}_{RI,g} \mathbf{\Theta}_g \mathbf{h}_{IT,g} \right|^2 \tag{3.36}$$

where $\mathbf{h}_{RI} = [\mathbf{h}_{RI,1}, \ldots, \mathbf{h}_{RI,G}]$ with $\mathbf{h}_{RI,g} \in \mathbb{C}^{1\times N_G}$ and $\mathbf{h}_{IT} = [\mathbf{h}_{IT,1}, \ldots, \mathbf{h}_{IT,G}]^T$ with $\mathbf{h}_{IT,g} \in \mathbb{C}^{N_G\times 1}$ [16]. It is easy to recognize that (3.36) is maximized when the terms $\mathbf{h}_{RI,g}\mathbf{\Theta}_g\mathbf{h}_{IT,g}$ are all individually maximized in absolute value and they are all co-phased. Recalling the constraint on $\mathbf{\Theta}_g$ given by (2.17), the optimal $\mathbf{\Theta}_g$ that maximizes the absolute value $|\mathbf{h}_{RI,g}\mathbf{\Theta}_g\mathbf{h}_{IT,g}|$ is given by Alg. 1 applied to the truncated channels $\mathbf{h}_{RI,g}$ and $\mathbf{h}_{IT,g}$, $\forall g$. Note that $\mathbf{\Theta}_g$ constructed by Alg. 1 ensures that the complex number $\mathbf{h}_{RI,g}\mathbf{\Theta}_g\mathbf{h}_{IT,g}$ has phase zero. Thus, (3.36) is maximized when the matrices $\mathbf{\Theta}_g$ are constructed by Alg. 1 since all the terms $\mathbf{h}_{RI,g}\mathbf{\Theta}_g\mathbf{h}_{IT,g}$ are co-phased. The block diagonal matrix $\mathbf{\Theta}$ is finally obtained from the matrices $\mathbf{\Theta}_g$ applying (2.17). To maximize P_R in the presence of the direct link h_{RT}, also for group-connected architectures $\mathbf{\Theta}$ can be adjusted as in (3.35).

3.2 BD-RIS-Aided Single-User SISO Systems: Transmissive Mode

In Section 3.1, we assumed the BD-RIS to work in reflective mode. This implies that both the transmitter and the receiver are covered by all the RIS elements. In other words, all the entries of $\mathbf{h}_{RI}$ and $\mathbf{h}_{IT}$ are non-zero in general. In this section, we study the case in which the BD-RIS works in transmissive mode, as modeled in [65]. Following [65], we consider a BD-RIS made of $N_I = 2M_I$ elements, where M_I is the number of RIS cells. Each RIS cell is formed by two RIS elements placed back to back and connected to each other through a reconfigurable impedance. Specifically, we assume that the m_Ith cell is formed by the $(2m_I - 1)$th and $(2m_I)$th RIS elements, for $m_I = 1, \ldots, M_I$. With this notation, the RIS elements can be partitioned into two sectors, where sector 1 is formed by the odd RIS elements and sector 2 is formed by the even RIS elements. Thus, the whole space is divided into two sides, respectively covered by the two sectors. When the BD-RIS is working in transmissive mode, the transmitter and the receiver are located in opposite sectors. In the following, we assume the transmitter to be in sector 1 and the receiver in sector 2.

Denoting as $\tilde{\mathbf{h}}_{RI} \in \mathbb{C}^{1\times 2M_I}$ the channel from the RIS to the receiver, and as $\tilde{\mathbf{h}}_{IT} \in \mathbb{C}^{2M_I\times 1}$ the channel from the transmitter to the RIS, the received signal power maximization problem writes as

$$\max_{\mathbf{\Theta}} \; P_T \left| h_{RT} + \tilde{\mathbf{h}}_{RI}\mathbf{\Theta}\tilde{\mathbf{h}}_{IT} \right|^2 \tag{3.37}$$

$$\text{s.t.} \;\; \mathbf{\Theta} = \mathrm{diag}\left(\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_G\right), \; \mathbf{\Theta}_g = \mathbf{\Theta}_g^T, \; \mathbf{\Theta}_g^H\mathbf{\Theta}_g = \mathbf{I}, \; \forall g, \tag{3.38}$$

where the odd entries of the channel $\tilde{\mathbf{h}}_{RI}$ are zero, as well as the even entries of the channel $\tilde{\mathbf{h}}_{IT}$. To solve (3.37)-(3.38), we can readily apply our optimal solution presented in Section 3.1.2, with the only difference that half of the entries of $\tilde{\mathbf{h}}_{RI}$ and $\tilde{\mathbf{h}}_{IT}$ are zero when the BD-RIS is working in transmissive mode. The reader is referred to [65] for detailed information about the optimization of BD-RISs supporting hybrid transmissive and reflective mode in multi-user scenarios.

3.3 BD-RIS-Aided Single-User MIMO Systems

In this section, we extend our optimal design strategy to single-user MIMO systems. W consider an N_T antenna transmitter and an N_R antenna receiver, whose communication is aided by a BD-RIS working in reflective mode. The equivalent channel writes as $\mathbf{H} = \mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}$, where $\mathbf{H}_{RT} \in \mathbb{C}^{N_R \times N_T}$, $\mathbf{H}_{RI} \in \mathbb{C}^{N_R \times N_I}$, and $\mathbf{H}_{IT} \in \mathbb{C}^{N_I \times N_T}$ are the channels from the transmitter to receiver, from the RIS to the receiver, and from the transmitter to the RIS, respectively. Considering single-stream transmission, we denote as $\mathbf{w} \in \mathbb{C}^{N_T \times 1}$ and $\mathbf{g} \in \mathbb{C}^{1 \times N_R}$ the precoding and combining vectors, respectively, subject to the constraint $\|\mathbf{w}\| = 1$ and $\|\mathbf{g}\| = 1$. Thus, the received signal power is $P_R^{\text{MIMO}} = P_T |\mathbf{g}(\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT})\mathbf{w}|^2$, with corresponding maximization problem

$$\max_{\mathbf{w},\mathbf{g},\mathbf{\Theta}} \; P_T \left|\mathbf{g}\left(\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\right)\mathbf{w}\right|^2 \tag{3.39}$$

$$\text{s.t.} \;\; \mathbf{\Theta} = \text{diag}\left(\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_G\right), \; \mathbf{\Theta}_g = \mathbf{\Theta}_g^T, \; \mathbf{\Theta}_g^H\mathbf{\Theta}_g = \mathbf{I}, \; \forall g, \tag{3.40}$$

$$\|\mathbf{w}\| = 1, \; \|\mathbf{g}\| = 1, \tag{3.41}$$

which is solved by jointly designing $\mathbf{w}$, $\mathbf{g}$, and $\mathbf{\Theta}$.

3.3.1 Fully-Connected RIS-Aided Systems Without Direct Link

We first consider a system aided by a fully-connected RIS, and we assume that the direct channel between transmitter and receiver $\mathbf{H}_{RT} \in \mathbb{C}^{N_R \times N_T}$ is negligible compared to the channel reflected by the RIS. This assumption reflects real scenarios where the direct channel is highly obstructed and significantly weaker than the RIS-aided link. In this case, the optimal precoder and combiner are given by the dominant eigenvectors of the equivalent channel $\mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}$. Thus, the maximum received signal power is given by $P_T \left\|\mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\right\|^2$, which is upper bounded by

$$\bar{P}_R^{\text{MIMO}} = P_T \left\|\mathbf{H}_{RI}\right\|^2 \left\|\mathbf{H}_{IT}\right\|^2, \tag{3.42}$$

following the sub-multiplicativity of the spectral norm. To achieve this upper bound, it is possible to prove that $\mathbf{\Theta}$ must satisfy

$$\mathbf{v}_{RI} = \mathbf{\Theta}\mathbf{u}_{IT}, \tag{3.43}$$

where $\mathbf{v}_{RI}$ is the dominant right singular vector of $\mathbf{H}_{RI}$ and $\mathbf{u}_{IT}$ is the dominant left singular vector of $\mathbf{H}_{IT}$. Note that the equality (3.43) is to be intended up to a phase shift since the complex singular vectors of a matrix are only defined up to a phase shift. Thus, condition (3.43) is satisfied when the cosine similarity

$$\rho = \left|\mathbf{v}_{RI}^H\mathbf{\Theta}\mathbf{u}_{IT}\right|^2 \tag{3.44}$$

Algorithm 2 BD-RISs design for single-user MIMO systems

Input: $\mathbf{H}_{RT} \in \mathbb{C}^{N_R \times N_T}$, $\mathbf{H}_{RI} \in \mathbb{C}^{N_R \times N_I}$, $\mathbf{H}_{IT} \in \mathbb{C}^{N_I \times N_T}$, N_G, ϵ.

Output: $\boldsymbol{\Theta}$.

if $\underline{\mathrm{P}}_R^{\mathrm{dir}} \geq \underline{\mathrm{P}}_R^{\mathrm{refl}}$ **then**

  $\mathbf{w} = \mathbf{v}_{RT}$, $\mathbf{g} = \mathbf{u}_{RT}^H$.

else

  $\mathbf{w} = \mathbf{v}_{IT}$, $\mathbf{g} = \mathbf{u}_{RI}^H$.

end if

while The fractional increase of the objective P_R^{MIMO} is above ϵ **do**

  $h_{RT}^{\mathrm{eff}} = \mathbf{g}\mathbf{H}_{RT}\mathbf{w}$, $\mathbf{h}_{RI}^{\mathrm{eff}} = \mathbf{g}\mathbf{H}_{RI}$, $\mathbf{h}_{IT}^{\mathrm{eff}} = \mathbf{H}_{IT}\mathbf{w}$.

  Compute $\boldsymbol{\Theta}$ by applying Alg. 1 to $\mathbf{h}_{RI}^{\mathrm{eff}}$, $\mathbf{h}_{IT}^{\mathrm{eff}}$.

  $\boldsymbol{\Theta} = e^{j\arg\left(h_{RT}^{\mathrm{eff}}\right)}\boldsymbol{\Theta}$.

  $\mathbf{w} = \mathbf{v}_{\max}\left(\mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\right)$, $\mathbf{g} = \mathbf{u}_{\max}^H\left(\mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\right)$.

end while

Return $\boldsymbol{\Theta}$.

is maximized, i.e., $\rho = 1$. Maximizing (3.44) is similar to the problem of maximizing the normalized received signal power in (3.3), exactly solved for the SISO setting in Section 3.1.1. Thus, the optimal $\boldsymbol{\Theta}$ satisfying (3.43) can be found by applying Alg. 1 to the vectors $\mathbf{v}_{RI}^H$ and $\mathbf{u}_{IT}$ and the upper bound (3.42) is tight. In the MISO setting, the optimal $\boldsymbol{\Theta}$ is given by Alg. 1 applied to the vectors $\mathbf{h}_{RI}$ and $\mathbf{u}_{IT}$.

3.3.2 Fully-/Group-Connected RIS-Aided Systems With Direct Link

In single-user RIS-aided MIMO systems, tight upper bounds on the received signal power are not available in general, i.e., when group-connected RISs are considered or the direct link is not negligible. For these cases, we propose a sub-optimal solution to maximize the received signal power in which the matrix $\boldsymbol{\Theta}$ and the beamforming vectors $\mathbf{w}$ and $\mathbf{g}$ are alternatively optimized, as established in the literature on single-connected RISs [13]. After $\mathbf{w}$ and $\mathbf{g}$ are initialized to feasible values, this optimization process alternates between the two following steps until convergence is reached. With fixed $\mathbf{w}$ and $\mathbf{g}$, we update $\boldsymbol{\Theta}$ by optimally maximizing the objective $|\mathbf{g}\mathbf{H}_{RT}\mathbf{w} + \mathbf{g}\mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\mathbf{w}|^2$ as proposed for SISO systems. The optimal $\boldsymbol{\Theta}$ is obtained by applying the strategy proposed in Section 3.1.2 to the channels $h_{RT}^{\mathrm{eff}} = \mathbf{g}\mathbf{H}_{RT}\mathbf{w}$, $\mathbf{h}_{RI}^{\mathrm{eff}} = \mathbf{g}\mathbf{H}_{RI}$, and $\mathbf{h}_{IT}^{\mathrm{eff}} = \mathbf{H}_{IT}\mathbf{w}$. With fixed $\boldsymbol{\Theta}$, we update $\mathbf{w}$ and $\mathbf{g}$ as the dominant right and left singular vectors of $\mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}$, respectively. The convergence is considered reached when the fractional increase of the objective P_R^{MIMO} in a full iteration is below a certain parameter ϵ.

Depending on the direct link strength, we consider two possible initializations for the beamforming vectors $\mathbf{w}$ and $\mathbf{g}$. In the following, we define the left and right dominant singular vectors of the matrices $\mathbf{H}_{ij}$ as $\mathbf{u}_{ij}$ and $\mathbf{v}_{ij}$, respectively, for $ij \in \{RT, RI, IT\}$. If the direct link is particularly strong, we set $\mathbf{w} = \mathbf{v}_{RT}$ and $\mathbf{g} = \mathbf{u}_{RT}^H$ to capture the energy in the direct channel dominant eigenmode. In this case, at the first iteration of the optimization process, it is possible to design $\boldsymbol{\Theta}$ through our optimal solution to achieve a received signal power

$$P_R^{\text{dir}} = P_T \left(\|\mathbf{H}_{RT}\| + \sum_{g=1}^{G} \|\mathbf{h}_{R,g}\| \|\mathbf{h}_{T,g}\| \right)^2, \tag{3.45}$$

where $\mathbf{h}_R = \mathbf{u}_{RT}^H \mathbf{H}_{RI}$ and $\mathbf{h}_T = \mathbf{H}_{IT} \mathbf{v}_{RT}$. Conversely, if the direct link is weak or a high number of RIS elements is employed, we set $\mathbf{w} = \mathbf{v}_{IT}$ and $\mathbf{g} = \mathbf{u}_{RI}^H$ to capture the energy of the reflected link. In this case, at the first iteration of the optimization process, $\boldsymbol{\Theta}$ can be optimized to achieve a received signal power

$$P_R^{\text{refl}} = P_T \left(\left| \mathbf{u}_{RI}^H \mathbf{H}_{RT} \mathbf{v}_{IT} \right| + \|\mathbf{H}_{RI}\| \|\mathbf{H}_{IT}\| \sum_{g=1}^{G} \left\| \mathbf{v}_{RI,g}^H \right\| \|\mathbf{u}_{IT,g}\| \right)^2. \tag{3.46}$$

Because of the initialization strategy, a lower bound on the received signal power achieved in MIMO systems is given by $\underline{P}_R^{\text{MIMO}} = \max\{P_R^{\text{dir}}, P_R^{\text{refl}}\}$, which is the received signal power obtained after the first iteration of our optimization process. Note that this is a lower bound since the objective function P_R^{MIMO} is non-decreasing over iterations.

We summarize the steps necessary to optimize $\boldsymbol{\Theta}$ in single-user MIMO systems in Alg. 2. The convergence of Alg. 2 is guaranteed by the following two facts. First, at each iteration, the objective P_R^{MIMO} is non-decreasing. Second, the objective function is bounded from above by $P_T \left(\|\mathbf{H}_{RT}\| + \|\mathbf{H}_{RI}\| \|\mathbf{H}_{IT}\| \right)^2$ because of the triangle inequality, and the sub-multiplicativity of the spectral norm. Note that Alg. 2 can be readily applied to the MISO setting, as it is a special case of the MIMO setting, in which $N_R = 1$. Single-user MIMO systems aided by a BD-RIS working in transmissive mode can be similarly optimized by directly applying the discussion made in Section 3.2.

3.4 BD-RIS-Aided Multi-User MISO Systems

In this section, we study the weighted sum power maximization in multi-user MISO systems, which is a problem particularly relevant in WPT applications [76]. Let us consider an N_T antenna transmitter serving K single-antenna receivers through the support of a BD-RIS working in reflective mode. We denote the channel from the transmitter to the kth receiver and from the RIS to the kth receiver as $\mathbf{h}_{RT,k} \in \mathbb{C}^{1 \times N_T}$ and $\mathbf{h}_{RI,k} \in \mathbb{C}^{1 \times N_I}$, respectively. Consequently, the equivalent channel seen by the kth receiver is denoted as $\mathbf{h}_k = \mathbf{h}_{RT,k} + \mathbf{h}_{RI,k} \boldsymbol{\Theta} \mathbf{H}_{IT}$.

In general, the transmitted signal writes as

$$\mathbf{x} = \sqrt{P_T} \sum_{k=1}^{K} \mathbf{w}_k s_k, \tag{3.47}$$

where the precoding vectors $\mathbf{w}_k$ are subject to the constraints $\sum_{k=1}^{K} \|\mathbf{w}_k\|^2 = 1$ and s_k are the energy-carrying signals subject to $\mathrm{E}[|s_k|^2] = 1$. Thus, the received signal power at the

kth receiver writes as

$$P_{R,k} = P_T \sum_{j=1}^{K} |\mathbf{h}_k \mathbf{w}_j|^2 . \tag{3.48}$$

Denoting by $\alpha_k > 0$ the power weight of the kth receiver, the weighted sum power writes as $S_R = \sum_{k=1}^{K} \alpha_k P_{R,k}$, with corresponding maximization problem given by

$$\max_{\mathbf{w}_k, \boldsymbol{\Theta}} \; \sum_{k=1}^{K} \alpha_k P_{R,k} \tag{3.49}$$

$$\text{s.t.} \;\; \boldsymbol{\Theta} = \text{diag}\left(\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_G\right), \; \boldsymbol{\Theta}_g = \boldsymbol{\Theta}_g^T, \; \boldsymbol{\Theta}_g^H \boldsymbol{\Theta}_g = \mathbf{I}, \, \forall g, \tag{3.50}$$

$$\sum_{k=1}^{K} \|\mathbf{w}_k\|^2 = 1. \tag{3.51}$$

Substituting (3.48) into (3.49), we obtain

$$S_R = \sum_{k=1}^{K} \alpha_k P_T \sum_{j=1}^{K} |\mathbf{h}_k \mathbf{w}_j|^2 = \sum_{j=1}^{K} P_T \mathbf{w}_j^H \mathbf{S} \mathbf{w}_j, \tag{3.52}$$

where we introduced $\mathbf{S} = \sum_{k=1}^{K} \alpha_k \mathbf{h}_k^H \mathbf{h}_k$. From (3.52), we notice that the optimal precoding vectors $\mathbf{w}_j$ should be all aligned with the dominant eigenvector of $\mathbf{S}$, denoted as $\mathbf{v}_{\max}(\mathbf{S})$. As in [76], we consider a single-stream precoding given by $\mathbf{w} = \mathbf{v}_{\max}(\mathbf{S})$, with no loss of optimality. With this optimal precoding, the weighted sum power is

$$S_R = P_T \|\mathbf{S}\| = P_T \left\|\mathbf{H}^H \mathbf{H}\right\| = P_T \|\mathbf{H}\|^2 , \tag{3.53}$$

where we introduced $\mathbf{H} = \left[\sqrt{\alpha_1}\mathbf{h}_1^H, \ldots, \sqrt{\alpha_K}\mathbf{h}_K^H\right]^H$. To maximize S_R, it is convenient to rewrite (3.53) by explicitly highlighting the role of $\boldsymbol{\Theta}$. Defining the matrices $\mathbf{G}_{RT} = \left[\sqrt{\alpha_1}\mathbf{h}_{RT,1}^H, \ldots, \sqrt{\alpha_K}\mathbf{h}_{RT,K}^H\right]^H$ and $\mathbf{G}_{RI} = \left[\sqrt{\alpha_1}\mathbf{h}_{RI,1}^H, \ldots, \sqrt{\alpha_K}\mathbf{h}_{RI,K}^H\right]^H$, we can write $\mathbf{H} = \mathbf{G}_{RT} + \mathbf{G}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}$.

3.4.1 Fully-Connected RIS-Aided Systems Without Direct Links

We first consider a system aided by a fully-connected RIS, and we assume that the direct channels between transmitter and receivers are negligible compared to the channels reflected by the RIS. Consequently, the equivalent channel seen by the kth receiver is given by $\mathbf{h}_k = \mathbf{h}_{RI,k}\boldsymbol{\Theta}\mathbf{H}_{IT}$, yielding $\mathbf{H} = \mathbf{G}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}$. Thus, the maximum weighted sum power is given by $P_T \|\mathbf{G}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\|^2$, which is upper bounded by

$$\bar{S}_R = P_T \|\mathbf{G}_{RI}\|^2 \|\mathbf{H}_{IT}\|^2 , \tag{3.54}$$

because of the sub-multiplicativity of the spectral norm. Using the discussion carried out for the single-user MIMO setting, we introduce the vectors $\mathbf{t}_{RI}$ and $\mathbf{u}_{IT}$ as the dominant left singular vectors of $\mathbf{G}_{RI}^H$ and $\mathbf{H}_{IT}$, respectively. Thus, the global optimal $\boldsymbol{\Theta}$ achieving (3.54) can be found by applying Alg. 1 to the vectors $\mathbf{t}_{RI}^H$ and $\mathbf{u}_{IT}$. This proves that the upper bound (3.54) is tight.

3.4.2 Fully-/Group-Connected RIS-Aided Systems With Direct Links

For general systems in which group-connected RISs are considered or the direct links are not negligible, tight performance upper bounds are not available. In this case, we notice that maximizing (3.53) is equivalent to maximizing

$$S_R = P_T \left| \mathbf{z}_1 \left(\mathbf{G}_{RT} + \mathbf{G}_{RI} \boldsymbol{\Theta} \mathbf{H}_{IT} \right) \mathbf{z}_2 \right|^2, \tag{3.55}$$

where $\mathbf{z}_1 \in \mathbb{C}^{1 \times N_R}$ and $\mathbf{z}_2 \in \mathbb{C}^{N_T \times 1}$ are auxiliary variables such that $\|\mathbf{z}_1\| = 1$ and $\|\mathbf{z}_2\| = 1$. Furthermore, maximizing (3.55) is similar to the problem of maximizing the received signal power in (3.39), solved through Alg. 2 in Section 3.3.2. Thus, our sub-optimal strategy provided by Alg. 2 can be readily applied to solve also this maximization problem. Multi-user MISO systems aided by a BD-RIS working in hybrid mode can be similarly optimized by directly applying the discussion made in Section 3.2.

3.5 Performance Evaluation

Let us consider a two-dimensional coordinate system, in which the y-axis represents the height above the ground in meters (m). The transmitter and the receiver are located at $(0, 0)$ and $(52, 0)$, respectively. The RIS is located at $(50, 2)$ and is equipped with N_I antennas. Note that the simulation setting is similar to the setting adopted in [16], with the RIS close to the receiver to maximize the gain brought by the RIS. Nevertheless, the conclusions drawn in this chapter are not impacted by the position of the RIS. The distance-dependent path loss is modeled as $L_{ij}(d_{ij}) = L_0 d_{ij}^{-\alpha_{ij}}$, where L_0 is the reference path loss at distance 1 m, d_{ij} is the distance, and α_{ij} is the path loss exponent for $ij \in \{RT, RI, IT\}$. We set $L_0 = -30$ dB, $\alpha_{RT} = 3.5$, $\alpha_{RI} = 2.8$, $\alpha_{IT} = 2$, and $P_T = 10$ W. For the small-scale fading, the channels are modeled with both Rayleigh and Rician fading, given by

$$\mathbf{h}_{ij} = \sqrt{L_{ij}} \left(\sqrt{\frac{K_F}{1 + K_F}} \mathbf{h}_{ij}^{\mathrm{LoS}} + \sqrt{\frac{1}{1 + K_F}} \mathbf{h}_{ij}^{\mathrm{NLoS}} \right), \tag{3.56}$$

where K_F refers to the Rician factor, while $\mathbf{h}_{ij}^{\mathrm{LoS}}$ and $\mathbf{h}_{ij}^{\mathrm{NLoS}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ represent the small-scale line-of-sight (LoS) and non-line-of-sight (NLoS) (Rayleigh fading) components, respectively, for $ij \in \{RT, RI, IT\}$. To model Rician fading channels, we consider $K_F = 3$ dB.

We start by analyzing the performance of SISO systems aided by single-, group-, and fully-connected RISs working in reflective mode. In Fig. 3.1, we report the average received signal power given in (3.1) obtained by optimizing the scattering matrix $\boldsymbol{\Theta}$ through our optimal strategy proposed in Section 3.1.2, for different group sizes. We compare these results with the average received signal power upper bound given by

$$\bar{P}_R^{\mathrm{Group}} = P_T \left(|h_{RT}| + \sum_{g=1}^{G} \|\mathbf{h}_{RI,g}\| \|\mathbf{h}_{IT,g}\| \right)^2, \tag{3.57}$$

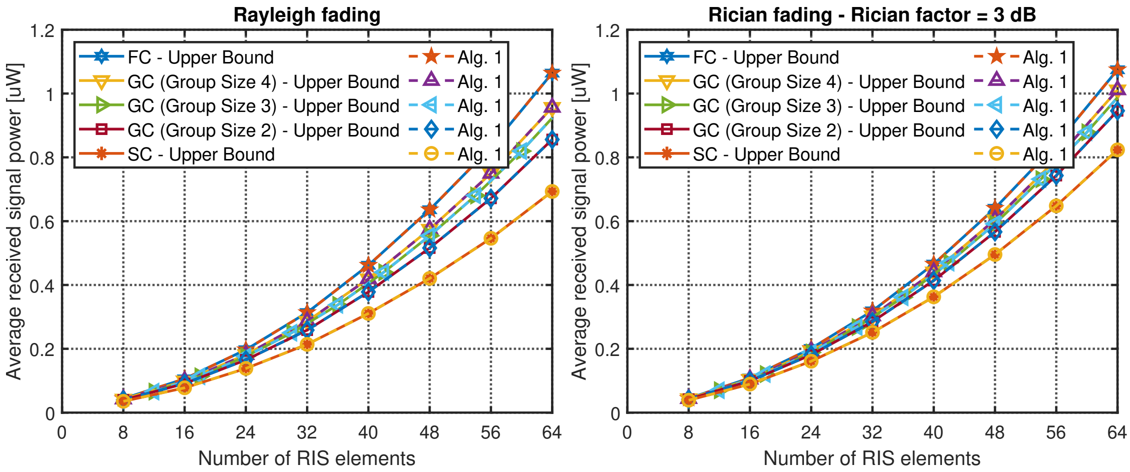


Figure 3.1: Average received signal power in SISO systems aided by single-connected “SC”, group-connected “GC”, and fully-connected “FC” BD-RISs working in reflective mode.

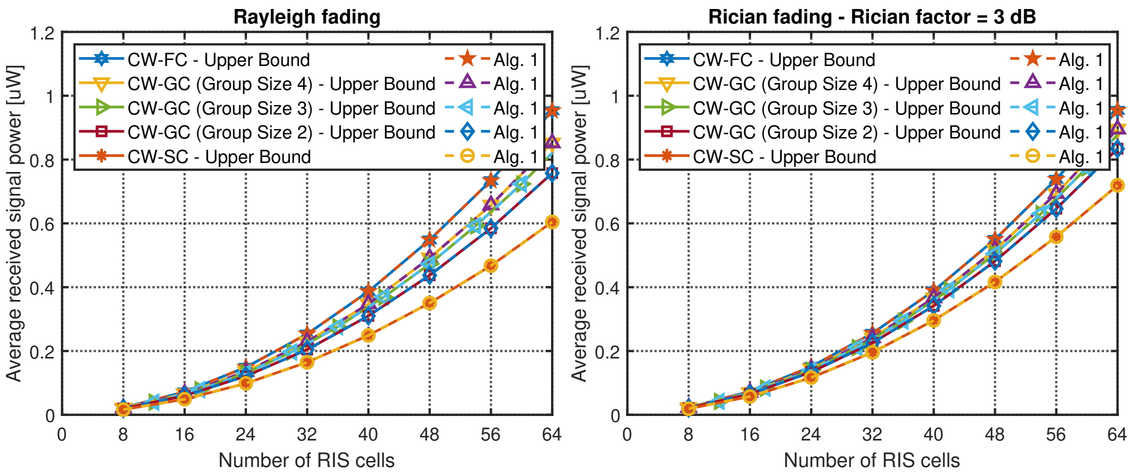


Figure 3.2: Average received signal power in SISO systems aided by cell-wise single-connected “CW-SC”, group-connected “CW-GC”, and fully-connected “CW-FC” BD-RISs working in transmissive mode.

as derived in [16]. The upper bound (3.57), valid for group-connected RISs, boils down to

$$\bar{P}_R^{\text{Single}} = P_T \left(|h_{RT}| + \sum_{n_I=1}^{N_I} \left| [\mathbf{h}_{RI}]_{n_I} [\mathbf{h}_{IT}]_{n_I} \right| \right)^2 \tag{3.58}$$

for single-connected RISs and to

$$\bar{P}_R^{\text{Fully}} = P_T \left(|h_{RT}| + \|\mathbf{h}_{RI}\| \|\mathbf{h}_{IT}\| \right)^2 \tag{3.59}$$

for fully-connected RISs[1]. As expected, we observe that the upper bounds are exactly achieved by our closed-form solution. The upper bounds (3.57), (3.58), and (3.59) can be numerically achieved also by the quasi-Newton method used in [16]. However, the convergence of the quasi-Newton method to a global optimum cannot be mathematically proved.

[1]The upper bounds (3.58) and (3.59) scale as $\mathcal{O}(\Gamma(1.5)^4 N_I^2)$ and $\mathcal{O}(N_I^2)$, respectively, where $\Gamma(\cdot)$ refers to the gamma function, in the presence of Rayleigh fading channels [16]. Thus, fully-connected RISs provide a power gain of $1/\Gamma(1.5)^4 = 16/\pi^2$ over the single-connected RISs when $N_I \to \infty$.

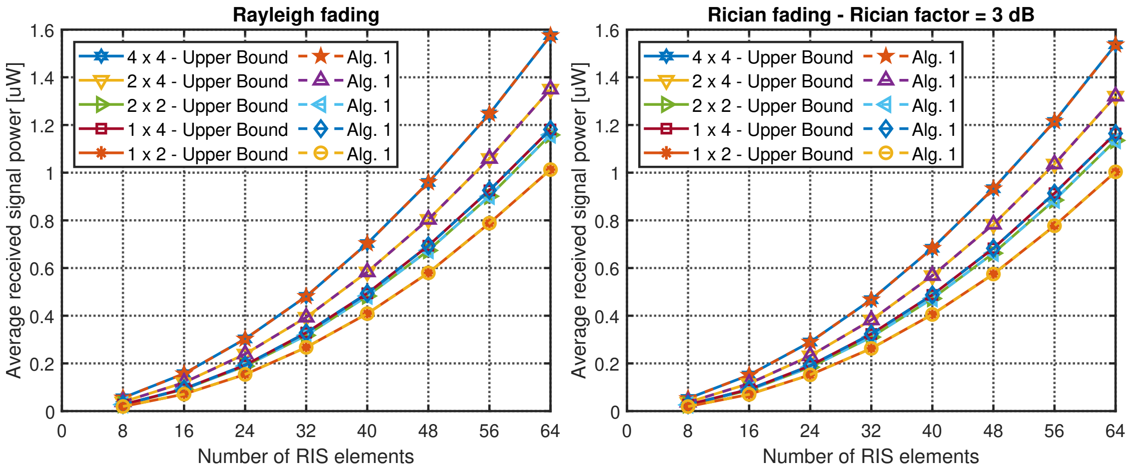


Figure 3.3: Average received signal power in $N_R \times N_T$ systems without direct link aided by fully-connected RISs working in reflective mode.

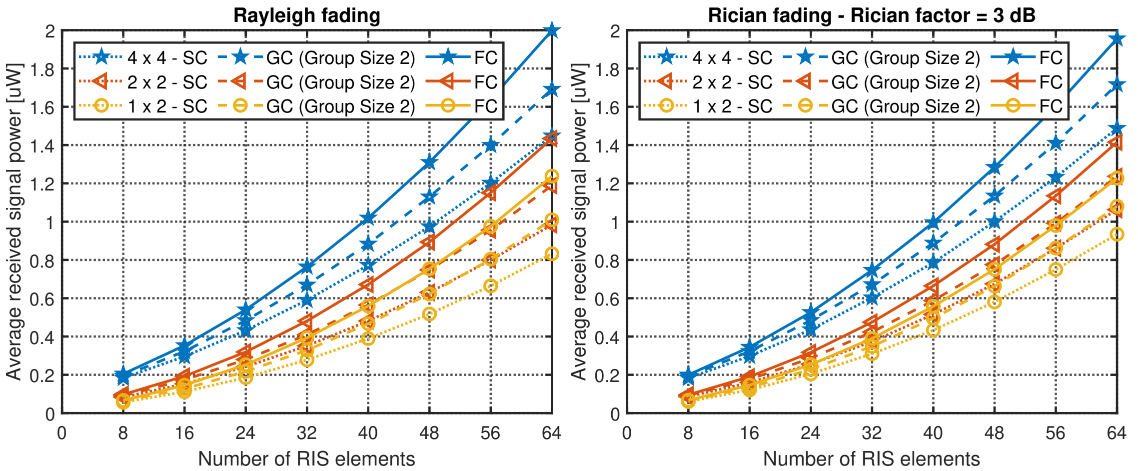


Figure 3.4: Average received signal power in $N_R \times N_T$ systems aided by single-connected "SC", group-connected "GC", and fully-connected "FC" BD-RISs working in reflective mode and optimized through Alg. 2.

Fully-connected RISs achieve the same performance in both Rayleigh and Rician fading conditions. Besides, single and group-connected RISs benefit from the LoS component, achieving higher power with Rician fading, agreeing with [16][2]. Remarkably, the SISO system without RIS can only achieve an average received signal power of 9.88 nW.

We now consider SISO systems aided by a BD-RIS working in transmissive mode. The receiver is now located in $(52, 4)$ and we set $\alpha_{RT} = 4$ to model a weak direct link. As previously discussed, we assume the transmitter to be in sector 1 and the receiver in sector 2. Thus, the odd entries of the channel $\mathbf{h}_{RI}$ and the even entries of the channel $\mathbf{h}_{IT}$ are forced to zero. In Fig. 3.2, we report the received signal power achieved by the optimal design strategy proposed in Section 3.1.2 and its upper bounds. We observe that the received signal power upper bounds are exactly achieved by our solution involving Alg. 1. Thus, our optimal design strategy can be successfully applied to BD-RIS working in both reflective and transmissive modes.

[2] Considering correlated Rayleigh fading distributed according to the exponential correlation model with coefficient $\rho = 0.5$, the performance of single and fully-connected RISs remains unchanged compared to i.i.d. Rayleigh fading. Besides, the performance of group-connected RISs slightly reduces.

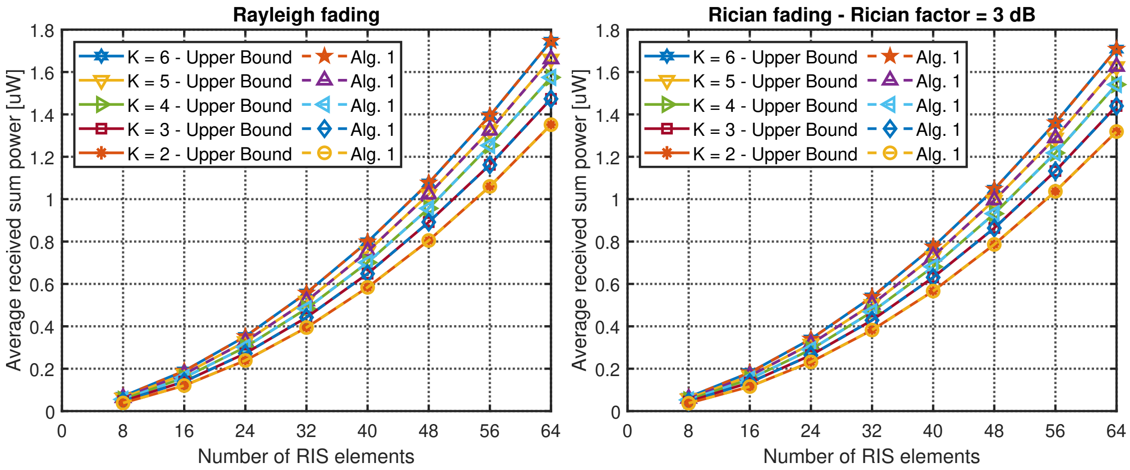


Figure 3.5: Average received sum power in multi-user MISO systems with $N_T = 4$ without direct link aided by fully-connected RISs working in reflective mode.

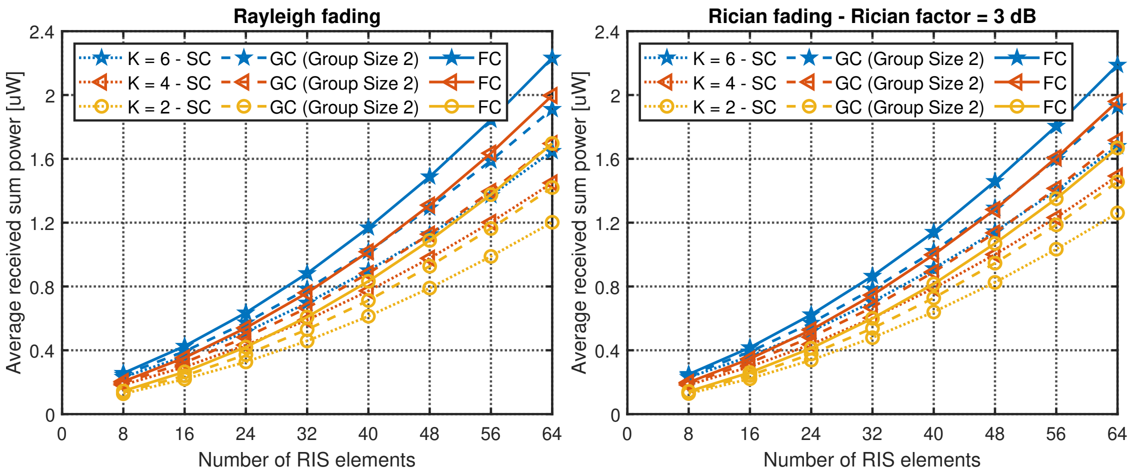


Figure 3.6: Average received sum power in multi-user MISO systems with $N_T = 4$ aided by single-connected "SC", group-connected "GC", and fully-connected "FC" BD-RISs working in reflective mode and optimized through Alg. 2.

We analyze the performance of RIS-aided single-user MIMO and MISO systems, in which the RIS works in reflective mode. In Fig. 3.3, we report the received signal power (3.39) for $N_R \times N_T$ systems aided by a fully-connected RIS, and with negligible direct link. The received signal power obtained by the exact solution provided by Alg. 1 is compared with its upper bound (3.42). We observe that the performance upper bounds are exactly achieved by our solution. Furthermore, higher performance is obtained by increasing the number of antennas N_R and N_T. Rayleigh fading channels allow reaching a slightly higher received signal power since they offer richer scattering. In Fig. 3.4, we consider single-user MIMO and MISO systems aided by a single-, group-, or fully-connected RIS. Alg. 2 is used to maximize the received signal power (3.39) by optimizing the RIS scattering matrix. As expected, fully-connected RISs achieve higher received signal power than group-connected RISs, which in turn outperform single-connected RISs. The performance gap between fully-connected and single-connected architectures is slightly higher in Rayleigh fading conditions.

We now consider the weighted sum power maximization problem in RIS-aided multi-

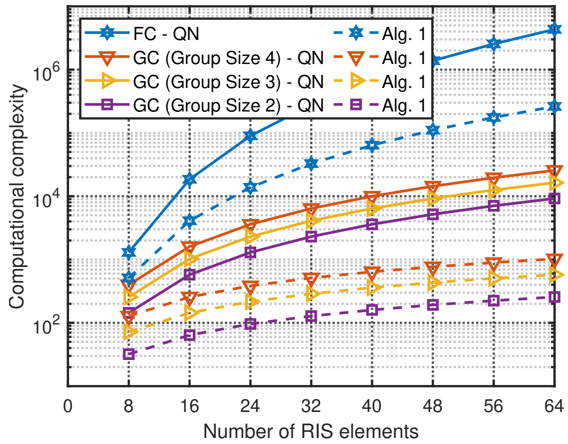


Figure 3.7: Computational complexity versus the number of RIS elements. Alg. 1 is compared with the quasi-Newton method "QN" for group-connected "GC" and fully-connected "FC" BD-RISs.

user MISO systems, with the RIS working in reflective mode. In our simulations, all K receivers are placed in $(52, 0)$, and we set $N_T = 4$ and $\alpha_k = 1\ \forall k$. In Fig. 3.5, we consider multi-user MISO systems aided by a fully-connected RIS, and with negligible direct links between transmitter and receivers. The received sum power (3.49) is maximized by applying the optimal precoding $\mathbf{w} = \mathbf{v}_{\max}(\mathbf{S})$, and by designing $\boldsymbol{\Theta}$ through Alg. 1. We compare this received sum power with its upper bound (3.54). As expected, the solution offered by Alg. 1 is optimal as it exactly achieves the performance upper bounds. In Fig. 3.6, we report the weighted sum power of multi-user MISO systems aided by a single-, group-, or fully-connected RIS, as given in (3.49). In these systems, Alg. 2 is used to optimize the RIS scattering matrix. Fully-connected RISs achieve the highest performance over group and single-connected BD-RISs, while single-connected RISs obtain the lowest weighted sum power. Besides, the weighted sum power increases with the number of receivers K because of the higher diversity offered.

Finally, we assess the computational complexity of our optimal design strategy. In fully-connected architectures, the complexity growth of Alg. 1 as a function of N_I is given by the complexity of eigenvalue decomposition, that is $\mathcal{O}(N_I^3)$. This is less than the complexity of the quasi-Newton optimization adopted in previous literature, which is $\mathcal{O}(N_I^2(N_I+1)^2/4)$ for each iteration [16]. In group-connected architectures with group size N_G, the block diagonal scattering matrix is designed by running $G = N_I/N_G$ times Alg. 1, with complexity $\mathcal{O}(N_G^3)$. Thus, the complexity of our solution is $\mathcal{O}(N_G^2 N_I)$ in this case, less than the complexity of quasi-Newton optimization, given by $\mathcal{O}(N_I^2(N_G+1)^2/4)$ for each iteration [16]. In Fig. 3.7, the computational complexity of Alg. 1 is compared with the complexity of the quasi-Newton method used in [16].

3.6 Conclusion

We provide a low-complexity closed-form solution to design the global optimal scattering matrix in the case of group- and fully-connected RISs. The resulting scattering matrix is proved to achieve exactly the received signal power upper bounds derived in [16]. Our solution is upper bound-achieving for any channel realization since we do not pose assumptions on the channel distribution. We first present a closed-form global optimal design strategy for RIS-aided single-user SISO systems. Subsequently, our strategy is extended to single-user MIMO and multi-user MISO systems. For systems aided by fully-connected RISs and with negligible direct links, we provide tight performance upper bounds. We show that such upper bounds can be exactly achieved with our optimal strategy. Finally, we show that our algorithm is less complex than the iterative optimization methods applied to design BD-RISs in recent literature. The complexity of our algorithm grows linearly (resp. cubically) with the number of RIS elements in the case of group- (resp. fully-) connected RISs.

3.7 Appendix

3.7.1 Proof of Proposition 1

The matrix $\mathbf{A}$ is a linear combination of four outer products matrices, i.e.,

$$\mathbf{A} = \frac{1}{2}\mathbf{R}_{RI} + \frac{1}{2}\mathbf{R}_{RI}^{T} - \frac{1}{2}\mathbf{R}_{IT} - \frac{1}{2}\mathbf{R}_{IT}^{T} = \mathbf{B}\mathbf{B}^{H}, \tag{3.60}$$

where the matrix $\mathbf{B} \in \mathbb{C}^{N_I \times 4}$ is introduced as $\mathbf{B} = 1/\sqrt{2}[\hat{\mathbf{h}}_{RI}^{H}, \hat{\mathbf{h}}_{RI}^{T}, j\hat{\mathbf{h}}_{IT}, j\hat{\mathbf{h}}_{IT}^{*}]$. Since $\mathbf{B}$ is full rank, we have $r(\mathbf{A}) = r(\mathbf{B}) = \min\{4, N_I\}$. Note that we assumed $\mathbf{h}_{RI}$ and $\mathbf{h}_{IT}$ to be independent in this discussion since the linearly dependent case has been trivially addressed. Thus, it holds $r(\mathbf{A}) = N_I$ if $N_I \in \{2, 3\}$. The trace of $\mathbf{A}$ can be readily computed from (3.60) by observing that $\mathrm{Tr}(\mathbf{R}_{RI}) = 1$ and $\mathrm{Tr}(\mathbf{R}_{IT}) = 1$ since $\|\hat{\mathbf{h}}_{RI}\| = 1$ and $\|\hat{\mathbf{h}}_{IT}\| = 1$. By applying the trace linearity property, we have $\mathrm{Tr}(\mathbf{A}) = 0$.

3.7.2 Proof of Proposition 2

To prove that $r(\mathbf{A}) = 4$ and $\mathrm{Tr}(\mathbf{A}) = 0$, we can directly apply the Proof of Proposition 1. To prove that $\mathbf{A}$ has two positive and two negative eigenvalues, we use the fact that $\mathbf{A} = \mathbf{A}_{RI} - \mathbf{A}_{IT}$ is given by the sum of two symmetric matrices. This proof is carried out by treating differently the cases $N_I = 4$ and $N_I > 4$.

In the case $N_I = 4$, $\mathbf{A}$ is a full rank matrix, i.e., $\det(\mathbf{A}) \neq 0$. We denote the decreasingly ordered eigenvalues of $\mathbf{A}_{RI}$ as $\delta_{RI,1}, \ldots, \delta_{RI,N_I}$ and the decreasingly ordered eigenvalues of $\mathbf{A}_{IT}$ as $\delta_{IT,1}, \ldots, \delta_{IT,N_I}$. According to [77], $\det(\mathbf{A})$ can be lower bounded by

$$\min_{P} \prod_{n_I=1}^{N_I} \left(\delta_{RI,n_I} - \delta_{IT,Pn_I}\right) \leq \det(\mathbf{A}), \tag{3.61}$$

where the minimum is taken over all permutations of indices $1, \ldots, N_I$. Since $\mathbf{A}_{RI}$ and $\mathbf{A}_{IT}$ are rank-2, it is always possible to find a permutation P such that

$$\min_{P} \prod_{n_I=1}^{N_I} \left(\delta_{RI,n_I} - \delta_{IT,Pn_I}\right) = 0. \tag{3.62}$$

Thus, recalling that $\det(\mathbf{A}) \neq 0$, we obtain $\det(\mathbf{A}) > 0$. Since $\mathbf{A}$ has four non-zero eigenvalues and $\mathrm{Tr}(\mathbf{A}) = 0$, $\det(\mathbf{A}) > 0$ implies the presence of two positive and two negative eigenvalues. This concludes the proof for $N_I = 4$.

In the case $N_I > 4$, we begin by noticing that $\mathbf{A}_{RI}$ and $\mathbf{A}_{IT}$ have two non-zero eigenvalues, both positives. The reason is that the matrices are rank-2 by construction and positive semi-definite since both are the sum of two positive semi-definite matrices. This means that $\delta_{RI,n_I}, \delta_{IT,n_I} = 0$ if $n_I > 2$. Furthermore, since $\mathbf{A}_{RI}$ and $\mathbf{A}_{IT}$ are Hermitian, we can apply Weyl's inequalities to study the eigenvalues of $\mathbf{A}$. According to Weyl's inequality, we have

$$\delta_i \leq \delta_{RI,i-j} - \delta_{IT,N_I-j}, \tag{3.63}$$

valid for $i = 1, \ldots, N_I$ and $j = 0, \ldots, i-1$ [78]. Considering (3.63) with $i = 3$ and $j = 0$, we have

$$\delta_3 \leq \delta_{RI,3} - \delta_{IT,N_I} = 0, \tag{3.64}$$

since $\delta_{RI,3} = 0$ and $\delta_{IT,N_I} = 0$. Additionally, the dual Weyl's inequality gives

$$\delta_{RI,i+k-1} - \delta_{IT,k} \leq \delta_i, \tag{3.65}$$

valid for $i = 1, \ldots, N_I$ and $k = 1, \ldots, N_I - i + 1$ [78]. Considering (3.65) with $i = 3$ and $k = 3$, we have

$$0 = \delta_{RI,5} + \delta_{IT,3} \leq \delta_3, \tag{3.66}$$

since $\delta_{RI,5} = 0$ and $\delta_{IT,3} = 0$, yielding $\delta_3 = 0$. Now, we consider (3.63) with $i = N_I - 2$ and $j = 0$ to obtain

$$\delta_{N_I-2} \leq \delta_{RI,N_I-2} - \delta_{IT,N_I} = 0, \tag{3.67}$$

since $\delta_{RI,N_I-2} = 0$ and $\delta_{IT,N_I} = 0$. Additionally, (3.65) with $i = N_I - 2$ and $k = 3$ gives

$$0 = \delta_{RI,N_I} + \delta_{IT,3} \leq \delta_{N_I-2}, \tag{3.68}$$

since $\delta_{RI,N_I} = 0$ and $\delta_{IT,3} = 0$, yielding $\delta_{N_I-2} = 0$. Since $\delta_3 = 0$ and $\delta_{N_I-2} = 0$, it has to be $\delta_1, \delta_2 > 0$ and $\delta_{N_I-1}, \delta_{N_I} < 0$. This concludes the proof for $N_I > 4$.

Chapter 4

Static Grouping Strategy Design for Group-Connected RISs

In previous works on group-connected RIS [16, 75], the RIS elements are grouped simply by collecting adjacent RIS elements into a group, without considering the grouping strategy optimization. In [69], a dynamic group strategy for BD-RIS supporting hybrid transmitting and reflecting mode has been proposed, where the groups are dynamically optimized in real-time on a per-channel realization basis. However, to implement the dynamic group strategy optimization, the circuit requires additional switches, which leads to higher circuit complexity compared to conventional group-connected RIS [69]. Therefore, it is worthwhile to consider the design of a static grouping strategy, which does not require additional switches and control overhead, to enhance the performance of group-connected RIS while maintaining the same circuit complexity. To that end, in this chapter, we investigate how to effectively group the RIS elements of group-connected RIS to maximize the performance while not increasing the circuit complexity. Our contributions are outlined as follows.

First, we provide the model of group-connected RIS with an optimized static grouping strategy, designed based on the channel statistics. *Second*, we formalize the corresponding optimization problems for single- and multi-user MISO systems and propose novel and efficient algorithms for the grouping strategy optimization. *Third*, we present numerical results to assess the gains derived from optimizing the grouping strategy, which can be of up to 60% sum rate improvement in highly correlated channels.

Organization: In Section 4.1, we model group-connected RISs with optimized static grouping strategy. In Section 4.2 and 4.3, we optimize the static grouping strategy of group-connected RISs in single- and multi-user systems, respectively. In Section 4.4, we evaluate the performance gains obtained by optimizing the static grouping strategy through numerical simulations. Finally, Section 4.5 contains the concluding remarks.

The content of this chapter has appeared in [79].

4.1 BD-RIS Model

Consider an N-element BD-RIS, modeled as N antenna elements connected to an N-port reconfigurable impedance network, with scattering matrix $\boldsymbol{\Theta} \in \mathbb{C}^{N\times N}$. To maximize the power reflected by the BD-RIS, we consider the N-port reconfigurable impedance network to be lossless. Thus, it can be described through its purely imaginary impedance matrix $\mathbf{Z} = j\mathbf{X} \in \mathbb{C}^{N\times N}$, where $\mathbf{X} \in \mathbb{R}^{N\times N}$ is the reactance matrix. According to network theory [74], $\mathbf{X}$ and $\boldsymbol{\Theta}$ are related by (2.4), where Z_0 denotes the reference impedance according to which the scattering parameters are computed, typically set to 50 Ω.

In a group-connected RIS, the N RIS elements are divided into G groups, each including $N_G = N/G$ elements[1] [16]. Each RIS element is connected to ground and to all other elements in its group through a tunable impedance component, while there is no connection between groups. In previous literature, the RIS elements have been grouped in sequential order according to their indices. In this way, the gth group is composed of the N_G elements indexed by

$$\mathcal{G}_g = \{(g-1)\,N_G + 1, \ldots, gN_G\}, \tag{4.1}$$

for $g = 1, \ldots, G$. As a result of this grouping, $\mathbf{X}$ and $\boldsymbol{\Theta}$ are block diagonal matrices in previous works [16].

Differently from previous literature, in this chapter, we optimize the grouping strategy to additionally enhance the performance of group-connected RIS. The grouping strategy can be described by a permutation π of the N RIS elements such that the gth group is composed of the elements indexed by

$$\mathcal{G}_g = \{\pi\left((g-1)\,N_G + 1\right), \ldots, \pi\left(gN_G\right)\}, \tag{4.2}$$

for $g = 1, \ldots, G$. Accordingly, the reactance matrix is a permuted block diagonal matrix given by

$$\mathbf{X} = \mathbf{P}_\pi \bar{\mathbf{X}} \mathbf{P}_\pi^T, \tag{4.3}$$

where $\mathbf{P}_\pi = [\mathbf{e}_{\pi(1)}, \ldots, \mathbf{e}_{\pi(N)}]$ is the permutation matrix associated to π, with $\mathbf{e}_n \in \mathbb{R}^{N\times 1}$ denoting the vector with the nth entry being 1 and the others being 0, and $\bar{\mathbf{X}} \in \mathbb{R}^{N\times N}$ is a block diagonal matrix fulfilling

$$\bar{\mathbf{X}} = \mathrm{diag}\left(\bar{\mathbf{X}}_1, \ldots, \bar{\mathbf{X}}_G\right),\ \bar{\mathbf{X}}_g = \bar{\mathbf{X}}_g^T,\ \forall g, \tag{4.4}$$

with $\bar{\mathbf{X}}_g \in \mathbb{R}^{N_G\times N_G}$ being the reactance matrix of the N_G-port fully-connected reconfigurable impedance network for the gth group, which is symmetric for a reciprocal network. As a consequence of (2.4) and (4.3), the scattering matrix writes as

$$\boldsymbol{\Theta} = \mathbf{P}_\pi \bar{\boldsymbol{\Theta}} \mathbf{P}_\pi^T, \tag{4.5}$$

[1] Note that in this chapter we focus on equally sized groups to ensure that the number of tunable components does not depend on the grouping strategy. However, grouping strategies allowing non-equally sized groups have been explored in [69] and will be discussed in Chapter 6 of this thesis.

where $\bar{\boldsymbol{\Theta}} \in \mathbb{C}^{N \times N}$ is a block diagonal matrix satisfying

$$\bar{\boldsymbol{\Theta}} = \operatorname{diag}\left(\bar{\boldsymbol{\Theta}}_1, \ldots, \bar{\boldsymbol{\Theta}}_G\right), \; \bar{\boldsymbol{\Theta}}_g = \bar{\boldsymbol{\Theta}}_g^T, \; \bar{\boldsymbol{\Theta}}_g^H \bar{\boldsymbol{\Theta}}_g = \mathbf{I}, \, \forall g, \tag{4.6}$$

with $\bar{\boldsymbol{\Theta}}_g \in \mathbb{C}^{N_G \times N_G}$ being the gth group scattering matrix.

To reconfigure $\boldsymbol{\Theta}$ given by (4.5), we optimize the grouping strategy π, impacting on $\mathbf{P}_\pi$, offline based on the channel statistics, while we optimize $\bar{\boldsymbol{\Theta}}$ online on a per-channel realization basis. In the following, we formulate the corresponding optimization problems for single- and multi-user systems.

4.2 Grouping Strategy Optimization for Single-User Systems

Consider a single-user RIS-aided MISO system where the transmitter is equipped with M antennas. The channel $\mathbf{h} \in \mathbb{C}^{1 \times M}$ between the transmitter and the receiver is expressed as $\mathbf{h} = \mathbf{h}_R \boldsymbol{\Theta} \mathbf{H}_T$, where $\mathbf{h}_R \in \mathbb{C}^{1 \times N}$ and $\mathbf{H}_T \in \mathbb{C}^{N \times M}$ denote the channels from the RIS to the receiver and from the transmitter to the RIS, respectively. The transmitted signal is $\mathbf{x} = \mathbf{w}s$, where $\mathbf{w} \in \mathbb{C}^{M \times 1}$ and $s \in \mathbb{C}$ are the precoding vector and data symbol. The precoding vector is subject to $\|\mathbf{w}\|_2^2 = 1$ while the data symbol is subject to the transmit power constraint $\mathrm{E}[|s|^2] = P_T$. Thus, the received signal is given by $y = \mathbf{h}_R \boldsymbol{\Theta} \mathbf{H}_T \mathbf{w} s + n$, where $n \sim \mathcal{CN}\left(0, \sigma_n^2\right)$ is the AWGN, with received signal power given by

$$P_R = P_T \left|\mathbf{h}_R \boldsymbol{\Theta} \mathbf{H}_T \mathbf{w}\right|^2, \tag{4.7}$$

which we want to maximize by optimizing $\boldsymbol{\Theta}$ and $\mathbf{w}$. Observing that (4.7) is maximized when $\mathbf{w}$ is given by maximum ratio transmission (MRT), i.e., $\mathbf{w} = \mathbf{h}^H / \|\mathbf{h}\|_2$, our problem reduces to maximize the channel gain $\|\mathbf{h}\|_2^2$ by optimizing $\boldsymbol{\Theta}$.

The resulting optimization problem can be formalized as a bi-level problem composed of a lower- and upper-level problem. The lower-level problem, solved on a per-channel realization basis, is given by

$$\max_{\bar{\boldsymbol{\Theta}}} \; \left\|\mathbf{h}_R \mathbf{P}_\pi \bar{\boldsymbol{\Theta}} \mathbf{P}_\pi^T \mathbf{H}_T\right\|_2^2 \tag{4.8}$$

$$\text{s.t.} \; \bar{\boldsymbol{\Theta}} = \operatorname{diag}\left(\bar{\boldsymbol{\Theta}}_1, \ldots, \bar{\boldsymbol{\Theta}}_G\right), \; \bar{\boldsymbol{\Theta}}_g = \bar{\boldsymbol{\Theta}}_g^T, \; \bar{\boldsymbol{\Theta}}_g^H \bar{\boldsymbol{\Theta}}_g = \mathbf{I}, \, \forall g, \tag{4.9}$$

according to which the matrix $\bar{\boldsymbol{\Theta}}$ is optimized given a fixed permutation matrix $\mathbf{P}_\pi$ depending on the static grouping strategy π. Besides, the upper-level problem, solved offline, optimizes π based on a training set of C channel realizations representative of the channel statistics. This problem is formalized as

$$\max_{\pi} \; \frac{1}{C} \sum_{c=1}^{C} \left\|\mathbf{h}_R^{(c)} \mathbf{P}_\pi \bar{\boldsymbol{\Theta}}^{(c)} \mathbf{P}_\pi^T \mathbf{H}_T^{(c)}\right\|_2^2 \tag{4.10}$$

$$\text{s.t.} \; \mathbf{P}_\pi = [\mathbf{e}_{\pi(1)}, \ldots, \mathbf{e}_{\pi(N)}], \tag{4.11}$$

$$\bar{\boldsymbol{\Theta}}^{(c)} \text{ solves (4.8)-(4.9)}, \, \forall c, \tag{4.12}$$

where $\mathbf{h}_R^{(c)}$ and $\mathbf{H}_T^{(c)}$ are the cth channel realizations in the training set, for $c = 1, \dots, C$. Remarkably, the bi-level problem (4.8)-(4.12) is hard to solve since it involves two nested optimization problems.

4.2.1 Optimizing the Static Grouping Strategy Offline

To efficiently optimize π by solving (4.10)-(4.12), we first remove the dependence on the variables $\{\bar{\boldsymbol{\Theta}}^{(c)}\}_{c=1}^{C}$ from the objective (4.10). By introducing $\bar{\mathbf{h}}_R^{(c)} = \mathbf{h}_R^{(c)}\mathbf{P}_\pi$ and $\bar{\mathbf{H}}_T^{(c)} = \mathbf{P}_\pi^T\mathbf{H}_T^{(c)}$, the objective $\|\mathbf{h}_R^{(c)}\mathbf{P}_\pi\bar{\boldsymbol{\Theta}}^{(c)}\mathbf{P}_\pi^T\mathbf{H}_T^{(c)}\|_2^2$ can be lower bounded by

$$\left\|\bar{\mathbf{h}}_R^{(c)}\bar{\boldsymbol{\Theta}}^{(c)}\bar{\mathbf{H}}_T^{(c)}\right\|_2^2 \geq \left|\bar{\mathbf{h}}_R^{(c)}\bar{\boldsymbol{\Theta}}^{(c)}\bar{\mathbf{H}}_T^{(c)}\mathbf{v}_T^{(c)}\right|^2 \tag{4.13}$$

$$= \left\|\mathbf{H}_T^{(c)}\right\|_2^2\left|\bar{\mathbf{h}}_R^{(c)}\bar{\boldsymbol{\Theta}}^{(c)}\mathbf{u}_T^{(c)}\right|^2, \tag{4.14}$$

where $\mathbf{v}_T^{(c)} \in \mathbb{C}^{M\times 1}$ and $\mathbf{u}_T^{(c)} \in \mathbb{C}^{N\times 1}$ are the dominant right and left singular vectors of $\bar{\mathbf{H}}_T^{(c)}$, respectively. According to [75], we can find in closed-form the optimal $\bar{\boldsymbol{\Theta}}^{(c)}$ maximizing the terms $|\bar{\mathbf{h}}_R^{(c)}\bar{\boldsymbol{\Theta}}^{(c)}\mathbf{u}_T^{(c)}|^2$, denoted as $\bar{\boldsymbol{\Theta}}^{\star(c)}$, yielding

$$\left|\bar{\mathbf{h}}_R^{(c)}\bar{\boldsymbol{\Theta}}^{\star(c)}\mathbf{u}_T^{(c)}\right| = \sum_{g=1}^{G}\left\|\left[\bar{\mathbf{h}}_R^{(c)}\right]_{\mathcal{G}_g}\right\|_2\left\|\left[\mathbf{u}_T^{(c)}\right]_{\mathcal{G}_g}\right\|_2, \tag{4.15}$$

where $\mathcal{G}_g = \{(g-1)N_G+1, \dots, gN_G\}$, for $g = 1, \dots, G$. Thus, problem (4.10)-(4.12) can be simplified into

$$\max_{\pi}\ \frac{1}{C}\sum_{c=1}^{C}\left(\sum_{g=1}^{G}\left\|\left[\bar{\mathbf{h}}_R^{(c)}\right]_{\mathcal{G}_g}\right\|_2\left\|\left[\mathbf{u}_T^{(c)}\right]_{\mathcal{G}_g}\right\|_2\right)^2 \tag{4.16}$$

$$\text{s.t.}\quad \mathbf{P}_\pi = [\mathbf{e}_{\pi(1)}, \dots, \mathbf{e}_{\pi(N)}], \tag{4.17}$$

which no longer contains a nested optimization problem as the objective solely depends on the permutation matrix $\mathbf{P}_\pi$.

To solve problem (4.16)-(4.17) through exhaustive search has prohibitive complexity because of the high number of possible grouping strategies π. Specifically, in an RIS with N elements grouped into G groups, each containing N_G elements, the number of possible grouping strategies N_S is given by

$$N_S = \frac{1}{G!}\binom{N}{N_G}\binom{N-N_G}{N_G}\cdots\binom{N_G}{N_G} = \frac{1}{G!}\frac{N!}{(N_G!)^G}, \tag{4.18}$$

growing with $N!$. To decrease the search space, we solve problem (4.16)-(4.17) through a local search process, as shown in Alg. 3. Specifically, we find the optimal grouping strategy $\pi^\star$ maximizing the objective (4.16) as follows. The grouping strategy is initialized to the trivial permutation. Thus, denoting as π_0 the initial grouping strategy, we have $\pi_0(n) = n$, for $n = 1, \dots, N$. At the ith iteration, we generate all possible grouping strategies obtainable by swapping two elements in π_{i-1}, where the swapping operation consists of selecting two elements belonging to two different groups and assigning each of

Algorithm 3 Optimizing the grouping strategy

Input: $N_G, \{\mathbf{h}_R^{(c)}, \mathbf{H}_T^{(c)}\}_{c=1}^{C}$.

Output: The optimized grouping strategy $\pi^{\star}$.

1: $i \leftarrow 0$.
2: $\pi_0(n) = n$, $\forall n$.
3: **while** $\pi_i \neq \pi_{i-1}$ **do**
4: $\quad i \leftarrow i + 1$.
5: $\quad$ Update Π_i as the set of all the grouping strategies obtainable by swapping two elements in π_{i-1}.
6: $\quad$ Update π_i by solving problem (4.19)-(4.21).
7: **end while**
8: Return $\pi^{\star} \leftarrow \pi_i$.

them to the group of the other. The resulting set of grouping strategies is denoted as Π_i and it is possible to show that its cardinality is $N(N - N_G)/2$. Then, the objective (4.16) is computed for each grouping strategy in Π_i. Finally, π_i is given by the grouping strategy in $\{\Pi_i \cup \pi_{i-1}\}$ maximizing (4.16), namely

$$\max_{\pi} \ \frac{1}{C} \sum_{c=1}^{C} \left(\sum_{g=1}^{G} \left\| \left[\bar{\mathbf{h}}_R^{(c)} \right]_{\mathcal{G}_g} \right\|_2 \left\| \left[\mathbf{u}_T^{(c)} \right]_{\mathcal{G}_g} \right\|_2 \right)^2 \tag{4.19}$$

$$\text{s.t.} \quad \mathbf{P}_{\pi} = [\mathbf{e}_{\pi(1)}, \ldots, \mathbf{e}_{\pi(N)}], \tag{4.20}$$

$$\pi \in \{\Pi_i \cup \pi_{i-1}\}, \tag{4.21}$$

which can be solved by an exhaustive search given the limited search space. We update π iteratively until the convergence is reached, i.e., when $\pi_i = \pi_{i-1}$. Note that Alg. 3 is ensured to converge by the following two properties. First, the same grouping strategy is never selected in multiple iterations since (4.16) is strictly increasing over iterations. Second, the total number of grouping strategies is limited.

4.2.2 Optimizing the Scattering Matrix Online

When π is fixed, we optimize $\bar{\boldsymbol{\Theta}}$ by solving problem (4.8)-(4.9) on a per-channel realization basis. Introducing the equivalent channels $\bar{\mathbf{h}}_R = \mathbf{h}_R \mathbf{P}_{\pi}$ and $\bar{\mathbf{H}}_T = \mathbf{P}_{\pi}^T \mathbf{H}_T$, we solve (4.8)-(4.9) by alternatively optimizing $\bar{\boldsymbol{\Theta}}$ and the auxiliary variable $\mathbf{w}$ subject to $\|\mathbf{w}\|_2 = 1$ to maximize $|\bar{\mathbf{h}}_R \bar{\boldsymbol{\Theta}} \bar{\mathbf{H}}_T \mathbf{w}|^2$. First, with fixed $\bar{\boldsymbol{\Theta}}$, we update $\mathbf{w}$ as $\mathbf{w} = (\bar{\mathbf{h}}_R \bar{\boldsymbol{\Theta}} \bar{\mathbf{H}}_T)^H / \|\bar{\mathbf{h}}_R \bar{\boldsymbol{\Theta}} \bar{\mathbf{H}}_T\|_2$. Second, with fixed $\mathbf{w}$, we update $\bar{\boldsymbol{\Theta}}$ through the global optimal solution provided in [75]. These two steps are iterated until the objective (4.8) converges.

4.3 Grouping Strategy Optimization for Multi-User Systems

Consider a multi-user RIS-aided MISO system where the transmitter is equipped with M antennas and there are K single-antenna receivers. The channel $\mathbf{h}_k \in \mathbb{C}^{1\times M}$ between the transmitter and the kth receiver can be expressed as $\mathbf{h}_k = \mathbf{h}_{R,k}\boldsymbol{\Theta}\mathbf{H}_T$, where $\mathbf{h}_{R,k} \in \mathbb{C}^{1\times N}$ and $\mathbf{H}_T \in \mathbb{C}^{N\times M}$ are the channels from the RIS to the kth receiver and from the transmitter to the RIS, respectively. The transmitted signal is $\mathbf{x} = \sum_{k=1}^{K} \mathbf{w}_k s_k$, where $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ and $s_k \in \mathbb{C}$ are the precoding vector and data symbol for the kth receiver, which are subject to $\sum_{k=1}^{K} \|\mathbf{w}_k\|_2^2 = 1$ and $\mathrm{E}[|s_k|^2] = P_T$. Thus, the signal at the kth receiver can be expressed as $y_k = \mathbf{h}_k \sum_i \mathbf{w}_i s_i + n_k$, where $n_k \sim \mathcal{CN}\left(0, \sigma_n^2\right)$ is the AWGN at the kth receiver, yielding a sum rate given by

$$S_R = \sum_{k=1}^{K} \log\left(1 + \frac{|\mathbf{h}_k\mathbf{w}_k|^2}{\sum_{i\neq k} |\mathbf{h}_k\mathbf{w}_i|^2 + \sigma_n^2}\right). \tag{4.22}$$

To maximize (4.22) by jointly optimizing $\boldsymbol{\Theta}$ and $\mathbf{w}_1, \ldots, \mathbf{w}_K$ is a hard problem given the non-convex objective function. Thus, we optimize the BD-RIS and the precoding vectors in two different stages, as also adopted in [50]. First, the BD-RIS is optimized to maximize the sum of the channel gains of all users $\sum_k \|\mathbf{h}_k\|_2^2 = \|\mathbf{H}\|_F^2$, where $\mathbf{H} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_K^T]^T \in \mathbb{C}^{K\times M}$. Second, the precoding vectors $\mathbf{w}_1, \ldots, \mathbf{w}_K$ are designed through zero-forcing beamforming.

When optimizing the grouping strategy and the tunable impedance components of the BD-RIS, the resulting optimization problem can be formalized as follows. The lower-level optimization problem is given by

$$\max_{\bar{\boldsymbol{\Theta}}} \quad \left\|\mathbf{H}_R\mathbf{P}_\pi\bar{\boldsymbol{\Theta}}\mathbf{P}_\pi^T\mathbf{H}_T\right\|_F^2 \tag{4.23}$$

$$\text{s.t.} \quad \bar{\boldsymbol{\Theta}} = \mathrm{diag}\left(\bar{\boldsymbol{\Theta}}_1, \ldots, \bar{\boldsymbol{\Theta}}_G\right), \; \bar{\boldsymbol{\Theta}}_g = \bar{\boldsymbol{\Theta}}_g^T, \; \bar{\boldsymbol{\Theta}}_g^H\bar{\boldsymbol{\Theta}}_g = \mathbf{I}, \, \forall g, \tag{4.24}$$

where we introduced $\mathbf{H}_R = [\mathbf{h}_{R,1}^T, \ldots, \mathbf{h}_{R,K}^T]^T \in \mathbb{C}^{K\times N}$ and $\mathbf{P}_\pi$ is fixed. In addition, the upper-level optimization problem writes as

$$\max_{\pi} \quad \frac{1}{C}\sum_{c=1}^{C} \left\|\mathbf{H}_R^{(c)}\mathbf{P}_\pi\bar{\boldsymbol{\Theta}}^{(c)}\mathbf{P}_\pi^T\mathbf{H}_T^{(c)}\right\|_F^2 \tag{4.25}$$

$$\text{s.t.} \quad \mathbf{P}_\pi = [\mathbf{e}_{\pi(1)}, \ldots, \mathbf{e}_{\pi(N)}], \tag{4.26}$$

$$\bar{\boldsymbol{\Theta}}^{(c)} \text{ solves (4.23)-(4.24)}, \, \forall c, \tag{4.27}$$

which results in a challenging bi-level problem.

4.3.1 Optimizing the Static Grouping Strategy Offline

Similarly to the single-user case, we optimize π through (4.25)-(4.27) by removing the dependence on the variables $\{\bar{\boldsymbol{\Theta}}^{(c)}\}_{c=1}^{C}$ from the objective (4.25). To this end, introducing

$\bar{\mathbf{H}}_R^{(c)} = \mathbf{H}_R^{(c)}\mathbf{P}_\pi$, we can lower bound the objective $\|\mathbf{H}_R^{(c)}\mathbf{P}_\pi\bar{\mathbf{\Theta}}^{(c)}\mathbf{P}_\pi^T\mathbf{H}_T^{(c)}\|_F^2$ as

$$\left\|\bar{\mathbf{H}}_R^{(c)}\bar{\mathbf{\Theta}}^{(c)}\bar{\mathbf{H}}_T^{(c)}\right\|_F^2 \geq \left\|\bar{\mathbf{H}}_R^{(c)}\bar{\mathbf{\Theta}}^{(c)}\bar{\mathbf{H}}_T^{(c)}\right\|_2^2 \tag{4.28}$$

$$\geq \left|\mathbf{u}_R^{(c)H}\bar{\mathbf{H}}_R^{(c)}\bar{\mathbf{\Theta}}^{(c)}\bar{\mathbf{H}}_T^{(c)}\mathbf{v}_T^{(c)}\right|^2 \tag{4.29}$$

$$= \left\|\mathbf{H}_R^{(c)}\right\|_2^2\left\|\mathbf{H}_T^{(c)}\right\|_2^2\left|\mathbf{v}_R^{(c)H}\bar{\mathbf{\Theta}}^{(c)}\mathbf{u}_T^{(c)}\right|^2, \tag{4.30}$$

where $\mathbf{u}_R^{(c)} \in \mathbb{C}^{K\times 1}$ and $\mathbf{v}_R^{(c)} \in \mathbb{C}^{N\times 1}$ denote the dominant left and right singular vectors of $\bar{\mathbf{H}}_R^{(c)}$, respectively. Remarkably, the terms $|\mathbf{v}_R^{(c)H}\bar{\mathbf{\Theta}}^{(c)}\mathbf{u}_T^{(c)}|^2$ can be maximized through a closed-form expression for the optimal matrix $\bar{\mathbf{\Theta}}^{\star(c)}$ [75], giving

$$\left|\mathbf{v}_R^{(c)H}\bar{\mathbf{\Theta}}^{\star(c)}\mathbf{u}_T^{(c)}\right| = \sum_{g=1}^{G}\left\|\left[\mathbf{v}_R^{(c)}\right]_{\mathcal{G}_g}\right\|_2\left\|\left[\mathbf{u}_T^{(c)}\right]_{\mathcal{G}_g}\right\|_2, \tag{4.31}$$

where $\mathcal{G}_g = \{(g-1)N_G+1,\dots,gN_G\}$, for $g = 1,\dots,G$. Consequently, problem (4.25)-(4.27) can be rewritten as

$$\max_{\pi}\ \frac{1}{C}\sum_{c=1}^{C}\left(\sum_{g=1}^{G}\left\|\left[\mathbf{v}_R^{(c)}\right]_{\mathcal{G}_g}\right\|_2\left\|\left[\mathbf{u}_T^{(c)}\right]_{\mathcal{G}_g}\right\|_2\right)^2 \tag{4.32}$$

$$\text{s.t.}\quad \mathbf{P}_\pi = [\mathbf{e}_{\pi(1)},\dots,\mathbf{e}_{\pi(N)}], \tag{4.33}$$

which the dependence on $\{\bar{\mathbf{\Theta}}^{(c)}\}_{c=1}^C$ has been removed. Since the objective in (4.32)-(4.33) solely depends on the permutation matrix $\mathbf{P}_\pi$, it can be solved through Alg. 3 by replacing (4.32) into the objective of problem (4.19)-(4.21).

4.3.2 Optimizing the Scattering Matrix Online

Once π has been fixed offline, $\bar{\mathbf{\Theta}}$ is reconfigured by solving problem (4.23)-(4.24) for each given channel realization. Constraint (4.24) indicates that $\bar{\mathbf{\Theta}}$ is a block diagonal matrix with each block being a complex symmetric unitary matrix, which complicates the optimization. Thus, the relationship between $\bar{\mathbf{\Theta}}$ and $\bar{\mathbf{X}}$ deriving from (2.4) can be exploited to equivalently reformulate (4.23)-(4.24) as

$$\max_{\bar{\mathbf{X}}_g}\ \left\|\bar{\mathbf{H}}_R\bar{\mathbf{\Theta}}\bar{\mathbf{H}}_T\right\|_F^2 \tag{4.34}$$

$$\text{s.t.}\quad \bar{\mathbf{\Theta}} = \mathrm{diag}\left(\bar{\mathbf{\Theta}}_1,\dots,\bar{\mathbf{\Theta}}_G\right), \tag{4.35}$$

$$\bar{\mathbf{\Theta}}_g = \left(j\bar{\mathbf{X}}_g + Z_0\mathbf{I}\right)^{-1}\left(j\bar{\mathbf{X}}_g - Z_0\mathbf{I}\right),\ \forall g, \tag{4.36}$$

$$\bar{\mathbf{X}}_g = \bar{\mathbf{X}}_g^T,\ \forall g, \tag{4.37}$$

where we introduced the equivalent channels $\bar{\mathbf{H}}_R = \mathbf{H}_R\mathbf{P}_\pi$ and $\bar{\mathbf{H}}_T = \mathbf{P}_\pi^T\mathbf{H}_T$, which can be directly transformed into an unconstrained problem. More precisely, exploiting the constraints (4.35)-(4.37), the objective (4.34) can be expressed as a function of $\bar{\mathbf{X}}_1,\dots,\bar{\mathbf{X}}_G$. Since $\bar{\mathbf{X}}_g$ is an arbitrary $N_G \times N_G$ real symmetric matrix, $\bar{\mathbf{X}}_g$ is an unconstrained function of the $N_G(N_G+1)/2$ entries in its upper triangular part. Thus, problem (4.34)-(4.37) is

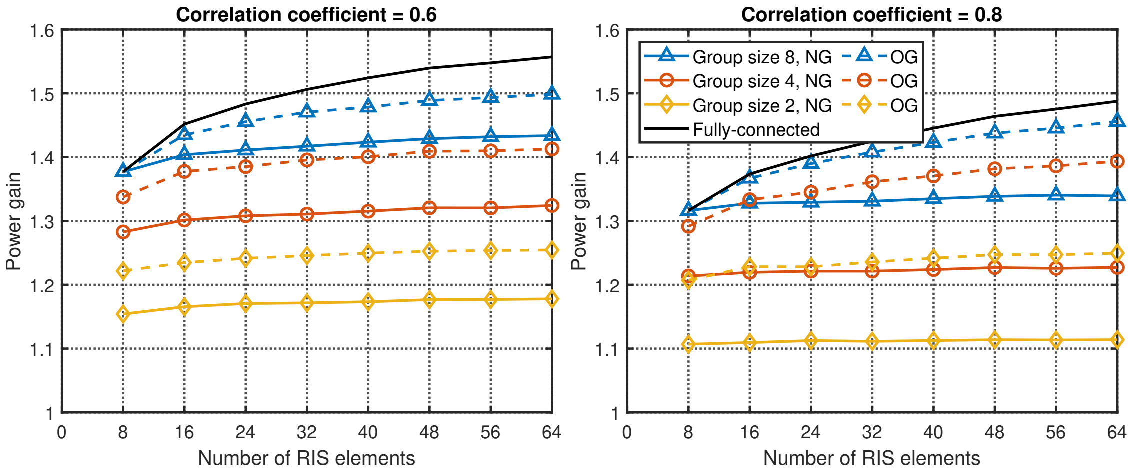


Figure 4.1: Power gain in single-user systems aided by fully- and group-connected RISs with non-optimized grouping "NG" and optimized grouping "OG".

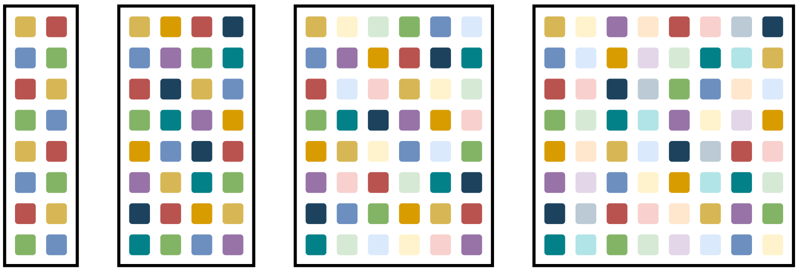


Figure 4.2: Optimized grouping for RISs with 2×8, 4×8, 6×8, and 8×8 elements and group size 4.

an unconstrained problem in the variables $[\bar{\mathbf{X}}_g]_{i,j}$, with $i \leq j$, $\forall g$, and can be solved by using the quasi-Newton method to find the optimal upper triangular part of each block $\bar{\mathbf{X}}_g$ without any constraints.

4.4 Performance Evaluation

We now evaluate the performance of single- and multi-user MISO systems aided by a group-connected RIS with an optimized grouping strategy. The transmitter, the RIS, and the receiver(s) are located at $(0, 0)$, $(50, 2)$, and $(52, 0)$ in meters (m), respectively. The transmitter is equipped with an uniform linear array (ULA) composed of $M = 4$ antennas, while the RIS is an uniform planar array (UPA) composed of $N_H \times N_V$ antennas, with $N_V = 8$ and $N_H N_V = N$. The path loss of the channels is given by the distance-dependent model $L_i(d_i) = L_0(d_i)^{-\alpha_i}$, where L_0 is the reference path loss at distance 1 m, d_i is the distance, and α_i is the path loss exponent for $i \in \{R, T\}$. We set $L_0 = -30$ dB, $\alpha_R = 2.8$, and $\alpha_T = 2$. The small-scale fading effects are modeled as correlated Rayleigh fading, i.e., $\mathbf{h}_{R,k} \sim \mathcal{CN}(\mathbf{0}, L_R\mathbf{R}_{RIS})$, for $k = 1, \ldots, K$, and $\text{vec}(\mathbf{H}_T) \sim \mathcal{CN}(\mathbf{0}, L_T\mathbf{R}_T)$. The covariance matrix $\mathbf{R}_{RIS} \in \mathbb{R}^{N \times N}$ is given by the Kronecker correlation model for UPAs $\mathbf{R}_{RIS} = \mathbf{R}_H \otimes \mathbf{R}_V$, where $\mathbf{R}_H \in \mathbb{R}^{N_H \times N_H}$ and $\mathbf{R}_V \in \mathbb{R}^{N_V \times N_V}$ are defined through the exponential correlation model as $[\mathbf{R}_H]_{i,j} = \rho^{|i-j|}$ and $[\mathbf{R}_V]_{i,j} = \rho^{|i-j|}$, with ρ being the correlation coefficient. Besides, $\mathbf{R}_T$ is given by the Kronecker correlation model for MIMO

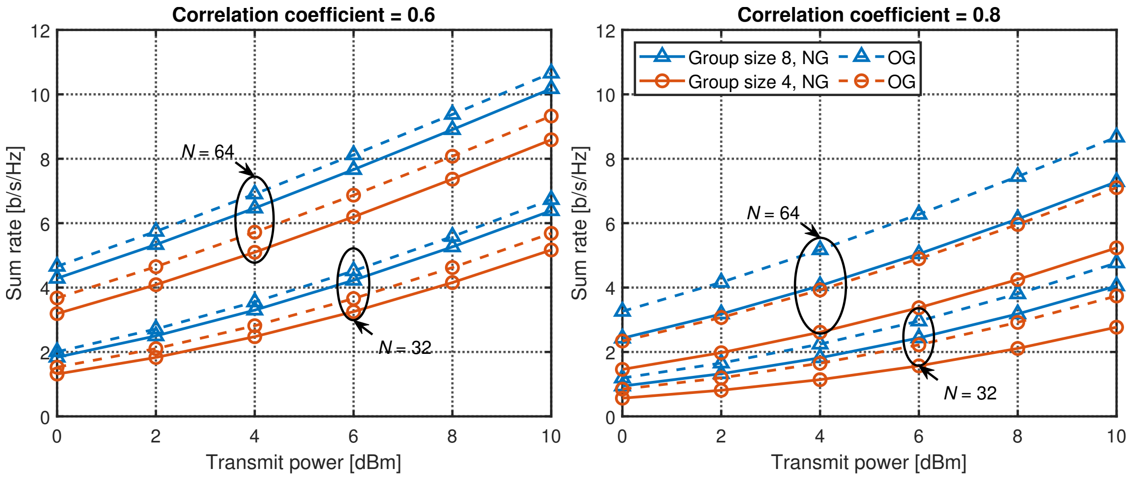


Figure 4.3: Sum rate in multi-user systems aided by group-connected RISs with non-optimized grouping "NG" and optimized grouping "OG".

channels $\mathbf{R}_T = \mathbf{R}_{RIS} \otimes \mathbf{R}_{TX}$, where $\mathbf{R}_{TX} \in \mathbb{R}^{M\times M}$ is defined as $[\mathbf{R}_{TX}]_{i,j} = \rho^{|i-j|}$. We consider mildly correlated channels by setting $\rho = 0.6$ and highly correlated channels with $\rho = 0.8$.

In Fig. 4.1, we investigate the impact of the grouping strategy optimization in single-user systems. To this end, we report the power gains of fully- and group-connected RISs over single-connected RIS, defined as $\mathcal{G}^{\text{Fully}} = P_R^{\text{Fully}}/P_R^{\text{Single}}$ and $\mathcal{G}^{\text{Group}} = P_R^{\text{Group}}/P_R^{\text{Single}}$, respectively, where P_R^{Fully}, P_R^{Single}, and P_R^{Group} are the received signal power of fully-, single-, and group-connected RISs, respectively. As a baseline, we consider the non-optimized grouping strategy given by (4.1), obtained by grouping adjacent RIS elements, as in previous works [16]. We observe that an optimized grouping strategy can significantly improve the performance of group-connected RIS, for any group size, particularly in the presence of highly correlated channels.

In Fig. 4.2, we illustrate the resulting optimized grouping strategy in single-user systems with $\rho = 0.8$, for RISs with group size $N_G = 4$, where RIS elements in the same group have the same color. Interestingly, we notice that the optimized grouping tends to maximize the distance between RIS elements grouped together. Note that an optimized static grouping strategy does not require additional hardware and online optimization complexity, different from an optimized dynamic grouping strategy [69].

In Fig. 4.3, we report the sum rate obtained by group-connected RISs in multi-user systems with $K = 2$ users and $\sigma_n^2 = -80$ dBm. As expected, the sum rate increases with the number of RIS elements N and the group size N_G. Besides, we notice that optimizing the grouping strategy can also visibly contribute to improving the sum rate by maintaining the same circuit complexity. Also in multi-user systems, the grouping strategy especially impacts the performance in highly correlated channels. Specifically, with $\rho = 0.8$, the sum rate is improved by up to 60% in the case $N = 64$ and $N_G = 4$.

4.5 Conclusion

We address the problem of optimally grouping the RIS elements in group-connected RIS based on the channel statistics, to improve the performance of BD-RIS while not increasing the circuit complexity. To this end, we model and optimize group-connected RIS with an optimized static grouping strategy. Specifically, we show how to optimize the grouping based on the channel statistics (offline) to maximize the ergodic performance. Besides, we optimize the tunable impedance components on a per-channel realization basis (online) to maximize the received signal power and sum rate in single- and multi-user systems, respectively. Numerical results show the gain of optimizing the grouping strategy offline, especially in the presence of highly correlated channels. Remarkably, this gain is achieved without increasing the circuit complexity or the online optimization computational complexity.

Part II

Development of Novel BD-RIS Architectures

Chapter 5

Forest- and Tree-Connected RISs

The fully-connected RIS architecture provides the highest flexibility and best performance, which is however at the expense of high circuit complexity since it has a large number of tunable impedance components [16]. Although the group-connected architecture can simplify the circuit complexity, it still remains an open problem to develop a systematic approach to design BD-RIS architecture achieving the optimal trade-off between performance and circuit complexity, that is achieving the best performance with the least complexity. To address this issue, in this chapter, we propose novel modeling, architecture design, and optimization for BD-RIS by utilizing graph theory. The contributions of this chapter are summarized as follows.

First, we propose a novel modeling of BD-RIS architectures utilizing graph theory. To the authors' best knowledge, it is the first time that BD-RIS (or any other RIS) architectures are modeled through graph theory. Specifically, we model each BD-RIS architecture by a graph, whose vertices characterize the RIS ports and whose edges characterize the tunable impedance components connecting the RIS ports. This graph theoretical modeling lays down the foundations for developing new efficient BD-RIS architectures.

Second, we derive a necessary and sufficient condition for a BD-RIS architecture to achieve the performance upper bound in MISO systems. Consequently, we characterize the least complex BD-RISs achieving such an upper bound. The resulting least-complexity BD-RIS architectures are referred to as tree-connected RIS since their corresponding graph is a tree. In addition, two specific examples of tree-connected RIS are proposed, namely tridiagonal RIS and arrowhead RIS. Tree-connected RIS achieves the same performance in MISO systems as fully-connected RIS with a circuit complexity significantly reduced by 16.4 times, when 64 RIS elements are considered.

Third, we prove that there always exists one and only one global optimal solution to optimize tree-connected RIS. To find this solution, we derive a closed-form algorithm valid for any tree-connected architecture. With this algorithm, tree-connected RIS achieves a performance improvement of 51.7% over single-connected RIS. We show that the proposed algorithm is optimal in single-user as well as multi-user systems.

Fourth, we propose forest-connected RIS as an additional BD-RIS family to reach a trade-off between performance and circuit complexity. Forest-connected RIS serves as a bridge between single- and tree-connected RIS. We propose a low-complexity iterative algorithm to optimize forest-connected RIS. Numerical results show that forest-connected RIS achieves the same performance as group-connected RIS, while reducing the circuit complexity by 2.4 times.

Organization: In Section 5.1, we introduce the graph theoretical modeling of BD-RIS. In Section 5.2, we derive tree- and forest-connected RIS architectures and provide two examples of tree-connected RIS. In Section 5.3 we provide a global optimal closed-form solution to optimize tree-connected RIS and an iterative approach to optimize forest-connected RIS. In Section 5.4, we evaluate the performance of the proposed BD-RISs. Finally, Section 5.5 concludes this chapter.

The content of this chapter has appeared in [80]. For reproducible research, the code is available at `https://github.com/matteonerini/bdris-utilizing-graph-theory`.

5.1 BD-RIS Modeling Utilizing Graph Theory

Consider a MISO system with an M-antenna transmitter and a single-antenna receiver, which is aided by an N-element RIS. The N-element RIS can be modeled as N antennas connected to an N-port reconfigurable impedance network that is characterized by the scattering matrix $\mathbf{\Theta} \in \mathbb{C}^{N \times N}$ [16]. We assume the direct channel between the transmitter and receiver is blocked and thus is negligible compared to the indirect channel provided by RIS. Therefore, the overall channel $\mathbf{h} \in \mathbb{C}^{1 \times M}$ between the transmitter and receiver can be written as

$$\mathbf{h} = \mathbf{h}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}, \tag{5.1}$$

where $\mathbf{h}_{RI} \in \mathbb{C}^{1 \times N}$ and $\mathbf{H}_{IT} \in \mathbb{C}^{N \times M}$ are the channel matrices from the RIS to the receiver, and from the transmitter to the RIS, respectively. We denote the transmit signal as $\mathbf{x} = \mathbf{w}s$, where $\mathbf{w} \in \mathbb{C}^{M \times 1}$ is the precoder satisfying $\|\mathbf{w}\| = 1$ and $s \in \mathbb{C}$ is the transmit symbol with power $P_T = \mathrm{E}[|s|^2]$. Hence, the received signal is given by $y = \mathbf{h}\mathbf{x} + n$, where n is the AWGN.

In BD-RIS, the RIS ports can also be connected to each other through additional tunable admittance components. We denote the tunable admittance connecting the nth port to the mth port as $Y_{n,m}$. According to [81], given the admittance components Y_n and $Y_{n,m}$, the (n,m)th entry of the admittance matrix $\mathbf{Y}$ is given by

$$[\mathbf{Y}]_{n,m} = \begin{cases} -Y_{n,m} & n \neq m \\ Y_n + \sum_{k \neq n} Y_{n,k} & n = m \end{cases}. \tag{5.2}$$

Thus, by selecting $Y_{n,m} = -[\mathbf{Y}]_{n,m}$ and $Y_n = \sum_k [\mathbf{Y}]_{n,k}$, we can implement an arbitrary symmetric $\mathbf{Y}$ so that the admittance matrix $\mathbf{Y}$ and the scattering matrix $\mathbf{\Theta}$ are beyond diagonal. Particularly, there are two special categories of BD-RIS, namely fully-connected and group-connected RIS [16]. Due to the enhanced flexibility in $\mathbf{Y}$ and $\mathbf{\Theta}$, BD-RIS is proven to outperform single-connected RIS [16].

To model the general circuit topology of BD-RIS, we resort for the first time to graph theoretical tools [82]. Specifically, we represent the circuit topology of BD-RIS through a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the *vertex set* of $\mathcal{G}$ and is given by the set of indices of RIS ports, i.e.

$$\mathcal{V} = \{1, 2, \ldots, N\}, \tag{5.3}$$

and $\mathcal{E}$ represents the *edge set* of $\mathcal{G}$ and is given by

$$\mathcal{E} = \{(n_\ell, m_\ell) \mid n_\ell, m_\ell \in \mathcal{V},\ Y_{n_\ell, m_\ell} \neq 0,\ n_\ell \neq m_\ell\}. \tag{5.4}$$

Thus, there exists an edge between vertex n_ℓ and vertex m_ℓ if and only if there is a tunable admittance connecting port n_ℓ and port m_ℓ, namely the ℓth admittance. Accordingly, given a graph $\mathcal{G}$ with edge set $\mathcal{E} = \{(n_\ell, m_\ell)\}_{\ell=1}^{L}$, the corresponding BD-RIS admittance matrix has $2L$ non-zero off-diagonal entries, i.e. $[\mathbf{Y}]_{n_\ell, m_\ell}$ and $[\mathbf{Y}]_{m_\ell, n_\ell}$ for $\ell = 1, \ldots, L$, while the other off-diagonal entries are zero. In the following, we define the circuit complexity of a BD-RIS architecture as the number of tunable admittance components in its circuit topology. Thus, the circuit complexity of a BD-RIS represented by a graph with L edges is given by $N + L$ since it includes N tunable admittance components connecting each element to ground, and L tunable admittance components interconnecting the elements to each other.

To better clarify our graph theoretical modeling of BD-RIS, we show a 4-element single-connected RIS with its associated graph in Fig. 5.1(a), and four examples of 4-element BD-RISs with their associated graphs in Fig. 5.1(b)-(e). We also use Fig. 5.1 to illustrate the graph theoretical definitions [82] that are used in this chapter, as briefly summarized in the following.

- A graph is *empty* when there is no edge.
- A graph is *connected* when there is a *path*, i.e., a finite sequence of distinct edges joining a sequence of distinct vertices, from any vertex to any other vertex.
- A graph that is not connected is called *disconnected*.
- A *cyclic* graph is a graph containing at least one *cycle*, i.e., a finite sequence of distinct edges joining a sequence of vertices, where only the first and last vertices are equal.
- A graph that is not cyclic is called *acyclic*, or a *forest*.
- A *tree* is defined as a connected and acyclic graph.
- A graph is *complete* when each pair of distinct vertices is joined by an edge.
- A graph is *simple* when there is no edge with identical ends (called *loops*) and no multiple edges joining the same pair of vertices.

Given these definitions, we recognize that the graph for the single-connected RIS in Fig. 5.1(a) is empty, the graph in Fig. 5.1(b) is disconnected and acyclic, the graph in Fig. 5.1(c) is connected and acyclic, i.e., a tree, the graph in Fig. 5.1(d) is connected and cyclic as it presents the cycle $1-2-3$, and the graph for the fully-connected RIS

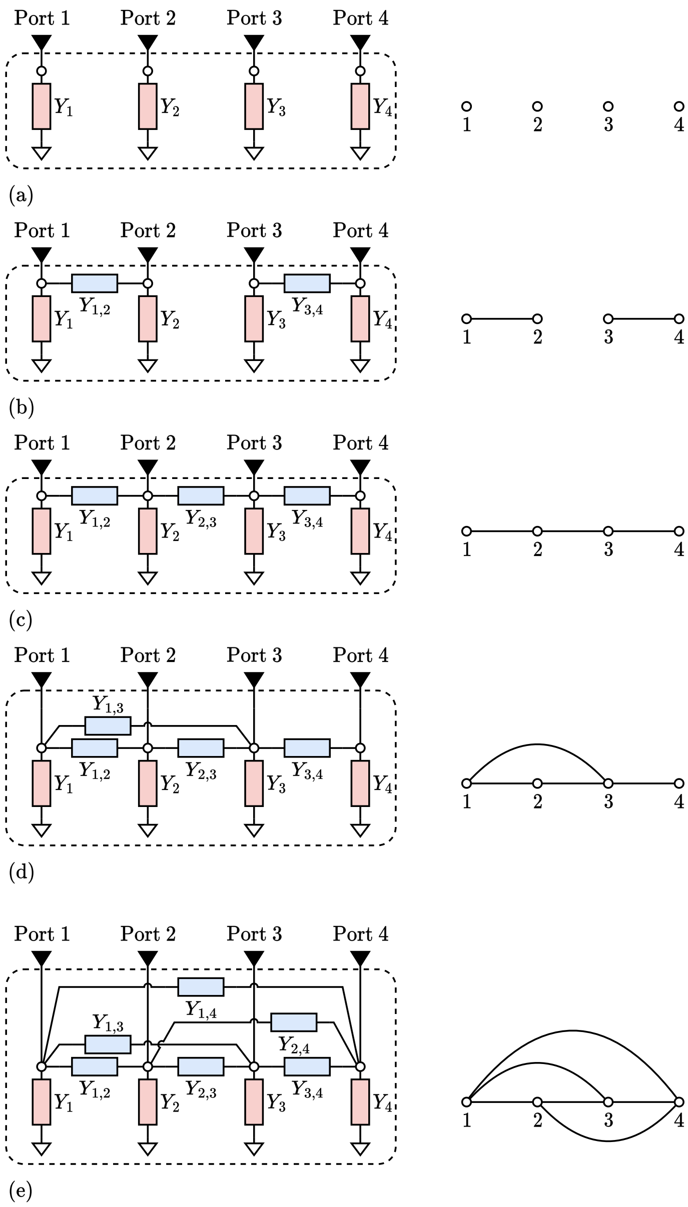


Figure 5.1: Examples of 4-port BD-RIS architectures (left) and their corresponding graphs $\mathcal{G}$ (right), where $\mathcal{G}$ is (a) empty, (b) disconnected and acyclic, (c) connected and acyclic, i.e., a tree, (d) connected and cyclic, (e) complete.

in Fig. 5.1(e) is complete. In addition, graphs associated with BD-RIS are always simple[1]. Therefore, graph theory is effective to characterize and model the circuit topology of BD-RIS and the corresponding admittance matrix $\mathbf{Y}$.

Utilizing graph theory, we propose two novel families of BD-RIS with remarkable properties, which are referred to as tree-connected RIS and forest-connected RIS, in the following section.

[1]In this chapter, we consider graphs where the edges are not oriented since we assume the RIS to have a reciprocal reconfigurable impedance network. BD-RIS architectures with non-reciprocal reconfigurable impedance networks could be modeled through *directed* graphs, i.e., graphs with oriented edges, but are beyond the scope of this chapter.

5.2 BD-RIS Architecture Design Utilizing Graph Theory

5.2.1 Tree-Connected RIS

As a starting point, we investigate how the graph of BD-RIS influences the system performance. To that end, we formulate the BD-RIS optimization problem for MISO systems utilizing graph theory. Specifically, we assume the BD-RIS is lossless and purely susceptive and aim to maximize the received signal power $P_R = P_T \left|\mathbf{h}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\mathbf{w}\right|^2$ by jointly optimizing $\mathbf{w}$ and $\mathbf{\Theta}$[2]. Thus, we can formulate the problem as

$$\max_{\mathbf{\Theta},\mathbf{w}} \; P_T \left|\mathbf{h}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\mathbf{w}\right|^2 \tag{5.5}$$

$$\text{s.t.} \;\; \mathbf{\Theta} = \left(\mathbf{I} + jZ_0\mathbf{B}\right)^{-1}\left(\mathbf{I} - jZ_0\mathbf{B}\right), \tag{5.6}$$

$$\mathbf{B} = \mathbf{B}^T, \; \mathbf{B} \in \mathcal{B}_{\mathcal{G}}, \; \|\mathbf{w}\| = 1, \tag{5.7}$$

where $\mathcal{B}_{\mathcal{G}}$ represents the set of possible susceptance matrices of BD-RISs characterized by the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and is written as

$$\mathcal{B}_{\mathcal{G}} = \{\mathbf{B} | \left[\mathbf{B}\right]_{n,m} = 0, n \neq m, n, m \in \mathcal{V}, (n, m) \notin \mathcal{E}\} \tag{5.8}$$

which means that an off-diagonal entry is zero if there exists no corresponding edge in the graph. The graph $\mathcal{G}$ is fixed in (5.5)-(5.7) since the interconnections between the RIS elements are fixed to avoid additional circuit complexity. Note that problem (5.5)-(5.7) considers the optimization of the RIS based on the instantaneous channel, as typically performed in related literature [13].

From (5.5)-(5.7), we can deduce that the maximum received signal power depends on the graph of the BD-RIS. Specifically, when the graph $\mathcal{G}$ has more edges, the susceptance matrix $\mathbf{B}$ is more flexible so that the maximum received signal power is higher. However, the enhanced performance is achieved at the expense of higher circuit complexity (more connections from the increased number of edges). Thus, it is worth designing the graph of BD-RIS to achieve the best performance with the least complexity.

To that end, we first consider the case that the graph $\mathcal{G}$ is complete, that is the fully-connected RIS. As shown in [16] and [75], maximum ratio transmission (MRT) gives the optimal $\mathbf{w}$ and subsequently the fully-connected RIS can achieve the upper bound of the received signal power given by

$$\bar{P}_R = P_T \left\|\mathbf{h}_{RI}\right\|^2 \left\|\mathbf{H}_{IT}\right\|^2, \tag{5.9}$$

following the sub-multiplicativity of the l_2-norm and that $\mathbf{\Theta}^H\mathbf{\Theta} = \mathbf{I}$. Leveraging this upper bound, herein we define a BD-RIS that achieves the performance upper bound $\bar{P}_R$ for any channel realization as MISO optimal. Thus, the fully-connected RIS is MISO optimal. More generally, the following lemma provides a sufficient and necessary condition for a BD-RIS to be MISO optimal.

[2]Maximizing the received signal power or the achievable rate are equivalent problems in single-user systems. Thus, we consider the received signal power since it is independent of the noise power.

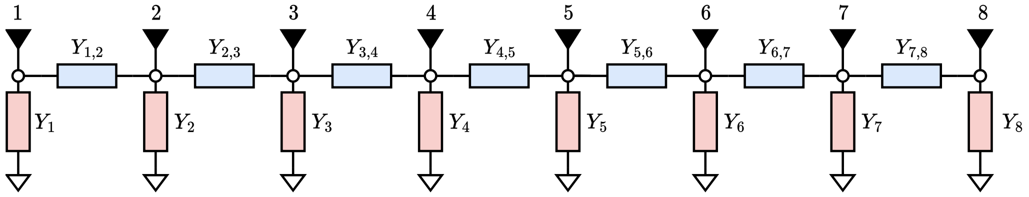


Figure 5.2: Tridiagonal RIS with $N = 8$.

Lemma 1. *A BD-RIS with associated graph $\mathcal{G}$ is MISO optimal if and only if $\mathcal{G}$ is a connected graph.*

Proof. Please refer to Appendix 5.6.1. □

Using Lemma 1, we have the following proposition providing the least complexity for MISO optimal BD-RIS.

Proposition 3. *The minimum number of edges in the graph of a MISO optimal BD-RIS is $N - 1$, with N denoting the number of RIS elements.*

Proof. Please refer to Appendix 5.6.2. □

Remarkably, a connected graph $\mathcal{G}$ on N vertices with $N - 1$ edges is a tree. To show this, we observe that removing any edge from $\mathcal{G}$ makes the graph disconnected because of Proposition 3. Thus, $\mathcal{G}$ cannot have a cycle and must be a tree (connected and acyclic). Therefore, to minimize the BD-RIS circuit complexity while maintaining optimal performance, we propose tree-connected RIS as BD-RIS whose corresponding graph is a tree. According to this definition, tree-connected RIS includes N admittance components connecting each port to ground and $N - 1$ admittance components interconnecting the ports to each other, yielding a total of $2N - 1$ admittance components, which is much less than the $N(N + 1)/2$ admittance components in fully-connected RIS.

5.2.2 Two Examples of Tree-Connected RIS

Tree-connected RIS refers to a family of BD-RISs, including multiple possible architectures. To clarify the concept of tree-connected RIS, in this subsection we provide two specific examples of tree-connected RIS, referred to as tridiagonal and arrowhead RIS.

Tridiagonal RIS

We define tridiagonal RIS as tree-connected RIS whose graph $\mathcal{G}$ is a *path graph.* In graph theory, the degree of a vertex is defined as the number of edges incident with it, and a path graph is defined as a graph where two vertices have *degree* one and all the others (if any) have degree two [82]. The vertex set of the graph associated with tridiagonal RIS is defined as

$$\mathcal{V}^{\text{Tri}} = \{n_1, n_2, \ldots, n_N\}, \tag{5.10}$$

and the edge set can be expressed accordingly as

$$\mathcal{E}^{\text{Tri}} = \{(n_\ell, n_{\ell+1}) \,|\, \ell = 1, \ldots, N - 1\}. \tag{5.11}$$

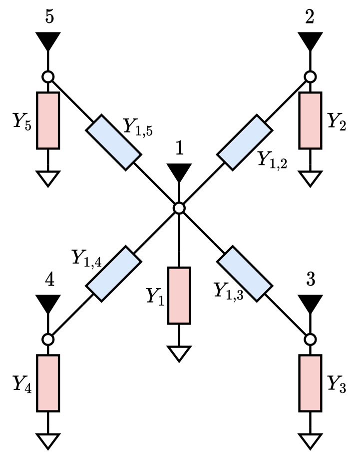


Figure 5.3: Arrowhead RIS with $N = 5$ where port 1 is the central vertex.

Interestingly, the tridiagonal architectures have also been proposed to realize low-complexity impedance matching networks for multiple antenna systems [81]. In Fig. 5.2, we provide an illustrative example of tridiagonal RIS with $N = 8$. As a consequence of such architecture, the susceptance matrix of tridiagonal RIS is symmetric and tridiagonal ($[\mathbf{B}]_{i,j} = 0$ if $|i - j| > 1$), written as

$$\mathbf{B} = \begin{bmatrix} [\mathbf{B}]_{1,1} & [\mathbf{B}]_{1,2} & \cdots & 0 \\ [\mathbf{B}]_{1,2} & [\mathbf{B}]_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & [\mathbf{B}]_{N-1,N} \\ 0 & \cdots & [\mathbf{B}]_{N-1,N} & [\mathbf{B}]_{N,N} \end{bmatrix}. \tag{5.12}$$

Note that if the port indexes in $\mathcal{V}^{\mathrm{Tri}}$ are permuted, $\mathbf{B}$ is a permuted tridiagonal matrix.

Arrowhead RIS

We define arrowhead RIS as tree-connected RIS whose graph $\mathcal{G}$ is a *star graph*. In graph theory, a star graph is defined as a graph where at most one vertex, called *central*, has degree greater than one and all the others have degree one [82]. Accordingly, the vertex set of the graph associated with arrowhead RIS is defined as

$$\mathcal{V}^{\mathrm{Arrow}} = \{c, n_1, n_2, \ldots, n_{N-1}\}, \tag{5.13}$$

where c denotes the central vertex, while the edge set is consequently expressed as

$$\mathcal{E}^{\mathrm{Arrow}} = \{(c, n_\ell) \mid \ell = 1, \ldots, N-1\}. \tag{5.14}$$

In Fig. 5.3, we provide an illustrative example of arrowhead RIS with $N = 5$ where port 1 is the central vertex. As a consequence, the susceptance matrix of arrowhead RIS is symmetric and arrowhead ($[\mathbf{B}]_{i,j} = 0$ if $i \neq c$, $j \neq c$, and $i \neq j$) [83]. Assuming port 1 is the central vertex, the susceptance matrix of arrowhead RIS can be written as

$$\mathbf{B} = \begin{bmatrix} [\mathbf{B}]_{1,1} & [\mathbf{B}]_{1,2} & \cdots & [\mathbf{B}]_{1,N} \\ [\mathbf{B}]_{1,2} & [\mathbf{B}]_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ [\mathbf{B}]_{1,N} & 0 & \cdots & [\mathbf{B}]_{N,N} \end{bmatrix}. \tag{5.15}$$

Note that if the central vertex is not port 1, $\mathbf{B}$ is a permuted arrowhead matrix.

It should be noted that all tree-connected RISs satisfy the necessary condition to be MISO optimal and contain the same number of tunable admittance components. Thus, tridiagonal and arrowhead RIS have the same circuit complexity and MISO optimality. In practice, we can select the suitable tree-connected RIS architecture according to the use case. On the one hand, tridiagonal RIS can be easily implemented with cables or tapes, thanks to the path graph circuit topology, so it enables the practical development of tree-connected RIS with ULAs. In addition, tridiagonal RIS is the tree-connected architecture that minimizes the length of the interconnections since it connects only adjacent RIS elements. On the other hand, arrowhead BD-RIS is more suitable to develop tree-connected RIS with UPAs.

5.2.3 Forest-Connected RIS

To further reduce the circuit complexity of tree-connected RIS, especially for the large-scale case, we propose forest-connected RIS. A forest-connected RIS is defined as a BD-RIS where the N RIS elements are divided into $G = N/N_G$ groups with group size N_G and each group utilizes the tree-connected architecture (i.e. the graph for each group is a tree on N_G vertices)[3]. Therefore, the graph $\mathcal{G}$ associated with the forest-connected RIS is a forest, and the corresponding vertex and edge sets can be represented as

$$\mathcal{V} = \cup_{g=1}^{G} \mathcal{V}_g,\ \mathcal{E} = \cup_{g=1}^{G} \mathcal{E}_g, \tag{5.16}$$

where $\mathcal{V}_g$ and $\mathcal{E}_g$ denote the vertex and edge sets of the graph associated with the gth group for $g = 1, \ldots, G$, with $\mathcal{V}_{g_1} \cap \mathcal{V}_{g_2} = \emptyset$ for all $g_1 \neq g_2$.

More specifically, the tree-connected architecture for each group in the forest-connected RIS can be either tridiagonal or arrowhead. When the tridiagonal architecture is used, the vertex and edge sets of the graph associated with forest-connected RIS are defined as

$$\mathcal{V}_{\text{Forest}}^{\text{Tri}} = \cup_{g=1}^{G} \left\{ n_{g,1}, n_{g,2}, \ldots, n_{g,N_G} \right\}, \tag{5.17}$$

$$\mathcal{E}_{\text{Forest}}^{\text{Tri}} = \cup_{g=1}^{G} \left\{ (n_{g,\ell}, n_{g,\ell+1}) \mid \ell = 1, \ldots, N_G - 1 \right\}, \tag{5.18}$$

similarly to (5.10) and (5.11). Accordingly, the susceptance matrix of forest-connected RIS with tridiagonal architecture for each group is block diagonal with each block being symmetric and tridiagonal. On the other hand, when the arrowhead architecture is used,

[3]Note that the tree-connected RIS is a special case of the forest-connected RIS, i.e., with $G = 1$. This is consistent with the fact that in graph theory a tree is a special case of forest.

the vertex and edge sets of the graph associated with forest-connected RIS are defined as

$$\mathcal{V}_{\text{Forest}}^{\text{Arrow}} = \cup_{g=1}^{G} \{c_g, n_{g,1}, n_{g,2}, \ldots, n_{g,N_G-1}\}, \tag{5.19}$$

$$\mathcal{E}_{\text{Forest}}^{\text{Arrow}} = \cup_{g=1}^{G} \{(c_g, n_{g,\ell}) \mid \ell = 1, \ldots, N_G - 1\}, \tag{5.20}$$

where c_g is the index of the central vertex for the gth group, similarly to (5.13) and (5.14). Accordingly, the susceptance matrix of forest-connected RIS with arrowhead architecture for each group is block diagonal with each block being symmetric and arrowhead.

The forest-connected RIS achieves a good performance-complexity trade-off between the single-connected and the tree-connected RIS. Note that the single-connected and tree-connected RIS can be viewed as two special cases of the forest-connected RIS, with $N_G = 1$ and $N_G = N$, respectively. In addition, forest-connected RIS is equivalent to group-connected RIS when $N_G = 2$ [16]. The number of tunable admittance components in forest-connected RIS with group size N_G is $N(2 - 1/N_G)$.

5.3 Optimization of Tree- and Forest-Connected RIS for MISO Systems

In this section, we optimize tree-connected and forest-connected RIS to maximize the received signal power in MISO systems. For tree-connected RIS, we show that any tree-connected RIS is MISO optimal and provide a closed-form global optimal solution. For forest-connected RIS, we provide a low-complexity iterative solution.

5.3.1 Tree-Connected RIS Optimization

The tree-connected RIS optimization to maximize the received signal power in MISO systems can be formulated as (5.5)-(5.7) plus a constraint that graph $\mathcal{G}$ is a tree. Recall that the received signal power in BD-RIS-aided MISO systems is upper-bounded by $\bar{P}_R = P_T \|\mathbf{h}_{RI}\|^2 \|\mathbf{H}_{IT}\|^2$. In the following, we prove that there is always one and only one solution for the optimal susceptance matrix $\mathbf{B}$ of tree-connected RIS to achieve the upper bound $\bar{P}_R$, and propose an algorithm to find such optimal $\mathbf{B}$.

Considering an RIS implemented through a reciprocal and lossless circuit, the key to achieving the upper bound $\bar{P}_R$ given by (5.9) is to find a complex symmetric unitary matrix $\boldsymbol{\Theta}$ satisfying

$$\hat{\mathbf{h}}_{RI}^H = \boldsymbol{\Theta}\mathbf{u}_{IT}, \tag{5.21}$$

where $\hat{\mathbf{h}}_{RI} = \mathbf{h}_{RI} / \|\mathbf{h}_{RI}\|$ and $\mathbf{u}_{IT} = \mathbf{u}_{\max}(\mathbf{H}_{IT})$ is the dominant left singular vector of $\mathbf{H}_{IT}$[4]. Substituting (2.5) into (5.21), we can equivalently rewrite (5.21) as

$$(\mathbf{I} + jZ_0\mathbf{B})\,\hat{\mathbf{h}}_{RI}^H = (\mathbf{I} - jZ_0\mathbf{B})\,\mathbf{u}_{IT}, \tag{5.22}$$

[4]Condition (5.21) is derived by considering $P_R \leq P_T \max_{\|\mathbf{x}\|=1} \|\mathbf{h}_{RI}\|^2 \|\boldsymbol{\Theta}\mathbf{H}_{IT}\mathbf{x}\|^2 \leq P_T \max_{\|\mathbf{x}\|=1} \|\mathbf{h}_{RI}\|^2 \|\boldsymbol{\Theta}\mathbf{H}_{IT}\|^2 \|\mathbf{x}\|^2$. Note that the equality holds in the two inequalities when $\hat{\mathbf{h}}_{RI}^H$ is equal to the dominant left singular vector of $\boldsymbol{\Theta}\mathbf{H}_{IT}$, i.e., $\boldsymbol{\Theta}\mathbf{u}_{IT}$.

which can be expressed in a compact form as

$$\mathbf{B}\boldsymbol{\alpha} = \boldsymbol{\beta}, \tag{5.23}$$

where $\boldsymbol{\alpha} = jZ_0 \left(\mathbf{u}_{IT} + \hat{\mathbf{h}}_{RI}^H\right) \in \mathbb{C}^{N\times 1}$ and $\boldsymbol{\beta} = \mathbf{u}_{IT} - \hat{\mathbf{h}}_{RI}^H \in \mathbb{C}^{N\times 1}$. The equation in (5.23) is composed of N linear equations with $2N-1$ unknowns. The linear equation coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are complex vectors, while the $2N-1$ unknowns are real since they are the entries of the susceptance matrix $\mathbf{B}$ of the tree-connected RIS, which are not constrained to be zero as shown in (5.12) and (5.15). We aim to solve (5.23) as it is a sufficient and necessary condition to achieve the performance upper bound $\bar{P}_R$. To that end, we rewrite (5.23) as $2N$ linear equations with real coefficients and $2N-1$ real unknowns, namely

$$\begin{bmatrix} \mathbf{B} \\ \hdashline \mathbf{B} \end{bmatrix} \mathbf{a} = \mathbf{b}, \tag{5.24}$$

where $\mathbf{a} \in \mathbb{R}^{2N\times 1}$ and $\mathbf{b} \in \mathbb{R}^{2N\times 1}$ are defined as

$$\mathbf{a} = \begin{bmatrix} \Re\{\boldsymbol{\alpha}\} \\ \hdashline \Im\{\boldsymbol{\alpha}\} \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} \Re\{\boldsymbol{\beta}\} \\ \hdashline \Im\{\boldsymbol{\beta}\} \end{bmatrix}. \tag{5.25}$$

To solve (5.24), we further rewrite the equation such that the $2N-1$ real unknowns are explicitly collected in a vector $\mathbf{x} \in \mathbb{R}^{(2N-1)\times 1}$. Specifically, given a tree-connected RIS whose graph has an edge set $\mathcal{E} = \{(n_\ell, m_\ell)\}_{\ell=1}^{N-1}$ such as (5.11) or (5.14), the $2N-1$ real unknowns are collected as

$$\mathbf{x} = \left[[\mathbf{B}]_{1,1}, \ldots, [\mathbf{B}]_{N,N}, [\mathbf{B}]_{n_1,m_1}, \ldots, [\mathbf{B}]_{n_{N-1},m_{N-1}}\right]^T, \tag{5.26}$$

where $[\mathbf{B}]_{n,n}$ for $n = 1, \ldots, N$ denote the N diagonal entries of $\mathbf{B}$ and $[\mathbf{B}]_{n_\ell,m_\ell}$ for $\ell = 1, \ldots, N-1$ denote the $N-1$ off-diagonal entries of $\mathbf{B}$ specified by the edge set of graph. Accordingly, (5.24) can be equivalently rewritten as

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{5.27}$$

where $\mathbf{A} \in \mathbb{R}^{2N\times(2N-1)}$ is the coefficient matrix given by

$$\mathbf{A} = \begin{bmatrix} \Re\{\mathbf{A}_1\} & \Re\{\mathbf{A}_2\} \\ \hdashline \Im\{\mathbf{A}_1\} & \Im\{\mathbf{A}_2\} \end{bmatrix}, \tag{5.28}$$

where $\mathbf{A}_1 \in \mathbb{C}^{N\times N}$ and $\mathbf{A}_2 \in \mathbb{C}^{N\times(N-1)}$. Because the real and imaginary parts of $\mathbf{A}_1$ are multiplied by the N diagonal entries of $\mathbf{B}$ in (5.27), to build the equivalence between (5.24) and (5.27), $\mathbf{A}_1$ should be given by

$$\mathbf{A}_1 = \mathrm{diag}\left(\boldsymbol{\alpha}\right). \tag{5.29}$$

On the other hand, the real and imaginary parts of $\mathbf{A}_2$ are multiplied by the $N-1$ off-diagonal entries of $\mathbf{B}$ including $[\mathbf{B}]_{n_1,m_1}, \ldots, [\mathbf{B}]_{n_{N-1},m_{N-1}}$. For this reason, the structure

Algorithm 4 Closed-form global optimization of tree-connected RIS for MISO systems

Input: $\mathbf{h}_{RI} \in \mathbb{C}^{1\times N}$, $\mathbf{H}_{IT} \in \mathbb{C}^{N\times M}$.

Output: $\mathbf{B}$.

1: Obtain $\hat{\mathbf{h}}_{RI} = \frac{\mathbf{h}_{RI}}{\|\mathbf{h}_{RI}\|}$, $\mathbf{u}_{IT} = \mathbf{u}_{\max}(\mathbf{H}_{IT})$.

2: Obtain $\boldsymbol{\alpha} = jZ_0\left(\mathbf{u}_{IT} + \hat{\mathbf{h}}_{RI}^H\right)$, $\boldsymbol{\beta} = \mathbf{u}_{IT} - \hat{\mathbf{h}}_{RI}^H$.

3: Obtain $\mathbf{A}_1$ and $\mathbf{A}_2$ by (5.29) and (5.30), respectively.

4: Obtain $\mathbf{A}$ and $\mathbf{b}$ by (5.28) and (5.25), respectively.

5: Obtain $\mathbf{x} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$.

6: Obtain $\mathbf{B}$ by (5.26).

7: Return $\mathbf{B}$.

of $\mathbf{A}_2$ depends on how the RIS ports are connected to each other, that is the specific graph associated with the tree-connected RIS. Accordingly, the entry of the ℓth column of $\mathbf{A}_2$ for $\ell = 1, \ldots, N-1$ is given by

$$[\mathbf{A}_2]_{k,\ell} = \begin{cases} [\boldsymbol{\alpha}]_{n_\ell} & \text{if } k = m_\ell \\ [\boldsymbol{\alpha}]_{m_\ell} & \text{if } k = n_\ell \\ 0 & \text{otherwise} \end{cases}. \tag{5.30}$$

Remarkably, the constraint on the tree-connected architecture, i.e., the position of the non-zero elements of $\mathbf{B}$, is dealt with by properly constructing $\mathbf{A}_2$ as in (5.30).

The following proposition provides a useful property of tree-connected RIS, which we can leverage to solve the linear equations (5.27).

Proposition 4. *For any tree-connected RIS with N elements, the coefficient matrix $\mathbf{A} \in \mathbb{R}^{2N\times(2N-1)}$ has full column rank, i.e., $r(\mathbf{A}) = 2N-1$.*

Proof. Please refer to Appendix 5.6.3. □

According to Proposition 4, the linear equations (5.27) has one solution or no solution [84]. Furthermore, we show that it has exactly one solution by using the following proposition.

Proposition 5. *For any tree-connected RIS with N elements, the augmented matrix $[\mathbf{A}|\mathbf{b}] \in \mathbb{R}^{2N\times 2N}$ has rank $r([\mathbf{A}|\mathbf{b}]) = 2N-1$.*

Proof. Please refer to Appendix 5.6.4. □

According to the Rouche-Capelli theorem [84], a system of linear equations is consistent if the rank of the coefficient matrix $\mathbf{A}$ is equal to the rank of the augmented matrix $[\mathbf{A}|\mathbf{b}]$. Thus, according to Proposition 5, the system of linear equations (5.27) is consistent and has exactly one solution. As a result, we can find the solution of (5.27) as $\mathbf{x} = \mathbf{A}^\dagger\mathbf{b}$, where $\mathbf{A}^\dagger = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$ is the Moore-Penrose pseudo-inverse of $\mathbf{A}$, and subsequently the optimal susceptance matrix $\mathbf{B}$ of the tree-connected RIS can be found through (5.26).

To conclude, we prove that for any tree-connected RIS there is always one and only one global optimal solution for the susceptance matrix $\mathbf{B}$ to achieve the upper bound $\bar{P}_R$

for any channel realization. In Alg. 4, we summarize the steps required to optimize tree-connected RIS for MISO systems. Interestingly, Alg. 4 can be used to globally optimize tree-connected RISs in single-user MIMO and multi-user MISO systems. More precisely, in single-user MIMO systems using single-stream transmission, the received signal power is maximized by applying Alg. 4 to the vectors $\mathbf{v}_{RI}$ and $\mathbf{u}_{IT}$, with $\mathbf{v}_{RI}$ being the dominant right singular vector of $\mathbf{H}_{RI}$. In addition, in multi-user MISO systems, the sum power is maximized by applying Alg. 4 to $\mathbf{t}_{RI}$ and $\mathbf{u}_{IT}$, with $\mathbf{t}_{RI}$ being the dominant right singular vector of $\mathbf{G}_{RI} = \left[\mathbf{h}_{RI,1}^H, \ldots, \mathbf{h}_{RI,U}^H\right]^H$, where $\mathbf{h}_{RI,u} \in \mathbb{C}^{1\times N}$ denotes the channel from the RIS to the uth receiver. The computational complexity of Alg. 4 is driven by the complexity of Step 5, i.e., it is the complexity of computing the inverse of $\mathbf{A}^T\mathbf{A}$. Since $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{(2N-1)\times(2N-1)}$, the computational complexity of Alg. 4 is $\mathcal{O}(8N^3)$.

5.3.2 Forest-Connected RIS Optimization

The forest-connected RIS optimization to maximize the received signal power in MISO systems is formulated as

$$\max_{\boldsymbol{\Theta},\mathbf{w}} \quad P_T \left|\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\mathbf{w}\right|^2 \tag{5.31}$$

$$\text{s.t.} \quad \boldsymbol{\Theta} = \left(\mathbf{I} + jZ_0\mathbf{B}\right)^{-1}\left(\mathbf{I} - jZ_0\mathbf{B}\right), \tag{5.32}$$

$$\mathbf{B} = \mathrm{diag}\left(\mathbf{B}_1, \ldots, \mathbf{B}_G\right), \tag{5.33}$$

$$\mathbf{B}_g = \mathbf{B}_g^T,\ \mathbf{B}_g \in \mathcal{B}_{\mathcal{G},g},\ \forall g,\ \|\mathbf{w}\| = 1, \tag{5.34}$$

where the susceptance matrix $\mathbf{B}$ of the forest-connected RIS is a block diagonal matrix with $\mathbf{B}_g$ being the gth block and $\mathcal{B}_{\mathcal{G},g}$ represents the set of the gth block written as

$$\mathcal{B}_{\mathcal{G},g} = \left\{\mathbf{B}_g \middle| \left[\mathbf{B}_g\right]_{n,m} = 0, n \neq m, n, m \in \mathcal{V}_g, (n,m) \notin \mathcal{E}_g\right\}, \tag{5.35}$$

where $\mathcal{V}_g$ and $\mathcal{E}_g$ are the vertex and edge sets of the graph for the gth group for $g = 1, \ldots, G$. Specifically, $\mathcal{V}_g$ and $\mathcal{E}_g$ can be found by (5.17) and (5.18) when the tridiagonal architecture is used for the gth group, or by (5.19) and (5.20) when the arrowhead architecture is used.

Different from optimizing the tree-connected RIS, it is hard to find a tight upper bound on the received signal power in forest-connected RIS-aided MISO systems due to block diagonal characteristics of the susceptance matrix, which makes it hard to derive a closed-form global optimal solution. To solve the challenging forest-connected RIS optimization problem, we propose an iterative optimization approach, where the precoder $\mathbf{w}$ and the RIS susceptance matrix $\mathbf{B}$ are alternatively optimized.

Optimizing B with Fixed w

When the precoder $\mathbf{w}$ is fixed, we introduce $\mathbf{h}_{IT}^{\text{eff}} = \mathbf{H}_{IT}\mathbf{w}$ as the effective channel between the transmitter and RIS so that the forest-connected RIS-aided MISO system is equivalently converted to a forest-connected RIS-aided SISO system. Thus, we can equivalently

Algorithm 5 Optimization of forest-connected RIS for MISO systems

Input: $\mathbf{h}_{RI} \in \mathbb{C}^{1\times N}$, $\mathbf{H}_{IT} \in \mathbb{C}^{N\times M}$.

Output: $\mathbf{B}$ and $\mathbf{w}$.

Initialize $\mathbf{w}$.
while no convergence of objective (5.31) **do**
  Update $\mathbf{B}_g$ by Alg. 1 with input $\mathbf{h}_{RI,g}$, $\mathbf{h}_{IT,g}^{\text{eff}}$, $\forall g$.
  Obtain $\boldsymbol{\Theta}$ from $\mathbf{B} = \text{diag}\left(\mathbf{B}_1, \ldots, \mathbf{B}_G\right)$ by (2.5).
  Update $\mathbf{w} = \left(\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\right)^H / \|\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\|$.
end while
Return $\mathbf{B}$ and $\mathbf{w}$.

rewrite the problem (5.31)-(5.34) as

$$\max_{\boldsymbol{\Theta}_g} \ P_T \left| \sum_{g=1}^{G} \mathbf{h}_{RI,g}\boldsymbol{\Theta}_g\mathbf{h}_{IT,g}^{\text{eff}} \right|^2 \tag{5.36}$$

$$\text{s.t.} \quad \boldsymbol{\Theta}_g = \left(\mathbf{I} + jZ_0\mathbf{B}_g\right)^{-1}\left(\mathbf{I} - jZ_0\mathbf{B}_g\right), \tag{5.37}$$

$$\mathbf{B}_g = \mathbf{B}_g^T, \ \mathbf{B}_g \in \mathcal{B}_{\mathcal{G},g}, \ \forall g, \ \|\mathbf{w}\| = 1, \tag{5.38}$$

where $\mathbf{h}_{RI,g} \in \mathbb{C}^{1\times N_G}$ and $\mathbf{h}_{IT,g}^{\text{eff}} \in \mathbb{C}^{N_G\times 1}$ consist of the N_G entries of $\mathbf{h}_{RI}$ and $\mathbf{h}_{IT}^{\text{eff}}$ corresponding to the RIS elements of the gth group, respectively [16]. To solve the problem (5.36)-(5.38), we first consider the case that the graph for each group is complete, that is the group-connected RIS where $\mathbf{B}_g$ can be arbitrary symmetric matrix. As shown in [16] and [75], the received signal power in the group-connected RIS-aided SISO system, i.e. $P_T \left| \sum_{g=1}^{G} \mathbf{h}_{RI,g}\boldsymbol{\Theta}_g\mathbf{h}_{IT,g}^{\text{eff}} \right|^2$, is upper bounded by

$$\bar{P}_R^{\text{eff}} = P_T \left(\sum_{g=1}^{G} \|\mathbf{h}_{RI,g}\| \left\| \mathbf{h}_{IT,g}^{\text{eff}} \right\| \right)^2. \tag{5.39}$$

The key to achieve such upper bound (5.39) is to find complex symmetric unitary matrices $\boldsymbol{\Theta}_g$ for $g = 1, \ldots, G$ satisfying

$$\hat{\mathbf{h}}_{RI,g}^H = \boldsymbol{\Theta}_g\hat{\mathbf{h}}_{IT,g}^{\text{eff}}, \ \forall g, \tag{5.40}$$

where $\hat{\mathbf{h}}_{RI,g} = \mathbf{h}_{RI,g}/\|\mathbf{h}_{RI,g}\|$ and $\hat{\mathbf{h}}_{IT,g}^{\text{eff}} = \mathbf{h}_{IT,g}^{\text{eff}}/\|\mathbf{h}_{IT,g}^{\text{eff}}\|$. Recall that each group in the forest-connected RIS utilizes the tree-connected architecture. Thus, referring to the tree-connected RIS optimization shown in Section 5.3.1, it is straightforward to show that for each group of the forest-connected RIS there is always one and only one solution for $\mathbf{B}_g$ whose corresponding $\boldsymbol{\Theta}_g$ satisfying (5.40), so that the upper bound (5.39) can be achieved. Moreover, the optimal $\mathbf{B}_g$ for each group can be found by Alg. 4 with the input of the channels $\mathbf{h}_{RI,g}$ and $\mathbf{h}_{IT,g}^{\text{eff}}$. Therefore, the susceptance matrix $\mathbf{B}$ of the forest-connected RIS can be globally optimized to achieve the upper bound (5.39) when $\mathbf{w}$ is fixed.

Optimizing w with Fixed B

When the RIS susceptance matrix $\mathbf{B}$ is fixed, it is obvious that the optimal $\mathbf{w}$ is the MRT given by $\mathbf{w} = (\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT})^H / \|\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\|$, where $\boldsymbol{\Theta}$ can be computed by (2.5) as a function of $\mathbf{B}$.

The precoder $\mathbf{w}$ and the RIS susceptance matrix $\mathbf{B}$ are alternatively optimized until convergence of the objective (5.31). We summarize the steps to optimize forest-connected RIS in MISO systems in Alg. 5. The iterative algorithm is initialized by randomly setting $\mathbf{w}$ to a feasible value. In addition, the convergence of Alg. 5 is guaranteed by the following two facts. First, at each iteration, the objective given by the received signal power P_R is non-decreasing. Second, the objective function is upper bounded by $P_T \|\mathbf{h}_{RI}\|^2 \|\mathbf{H}_{IT}\|^2$ because of the sub-multiplicativity of the spectral norm. The computational complexity per iteration of Alg. 5 is driven by the complexity of Step 3, i.e., it is the complexity of applying N/N_G times Alg. 4 to an N_G-element BD-RIS. Thus, the computational complexity per iteration of Alg. 5 is $\mathcal{O}(8NN_G^2)$.

5.4 Performance Evaluation

In this section, we provide the performance evaluation for the proposed tree- and forest-connected RIS. We consider a two-dimensional coordinate system, where the transmitter and receiver are located at $(0,0)$ and $(52,0)$ in meters (m), respectively. The RIS is located at $(50,2)$ and it is equipped with N elements. For the large-scale path loss, we use the distance-dependent path loss modeled as $L_{ij}(d_{ij}) = L_0(d_{ij}/D_0)^{-\alpha_{ij}}$, where L_0 is the reference path loss at distance $D_0 = 1$ m, d_{ij} is the distance, and α_{ij} is the path loss exponent for $ij \in \{RI, IT\}$. We set $L_0 = -30$ dB, $\alpha_{RI} = 2.8$, $\alpha_{IT} = 2$, and $P_T = 10$ mW. For the small-scale fading, we assume that the channel from the RIS to the receiver is Rayleigh fading. The channel from the transmitter to the RIS is modeled with Rician fading, given by

$$\mathbf{H}_{IT} = \sqrt{L_{IT}}\left(\sqrt{\frac{K}{1+K}}\mathbf{H}_{IT}^{\mathrm{LoS}} + \sqrt{\frac{1}{1+K}}\mathbf{H}_{IT}^{\mathrm{NLoS}}\right), \tag{5.41}$$

where K refers to the Rician factor, while $\mathbf{H}_{IT}^{\mathrm{LoS}}$ and $\mathrm{vec}(\mathbf{H}_{IT}^{\mathrm{NLoS}}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ represent the small-scale LoS and NLoS (Rayleigh fading) components, respectively. We consider two scenarios, with Rician factor $K = 0$ dB and $K = 10$ dB.

We first evaluate the performance of the tree-connected RIS to verify that it is MISO optimal. Specifically, the tree-connected RIS is optimized through Alg. 4 and compared with the performance upper bound (5.9) achieved by the fully-connected RIS. We recall that in the fully-connected RIS, all the RIS ports are connected with each other, so that $\mathbf{B}$ is an arbitrary symmetric matrix and $\boldsymbol{\Theta}$ satisfies (2.13). In Fig. 5.4, we provide the received signal power achieved in MISO systems aided by fully- and tree-connected RIS. We can make the following observations. *First*, as expected, the tree-connected RIS can always achieve the performance upper bound. However, it has much lower circuit complexity than the fully-connected RIS, which will be quantitatively shown in the following.

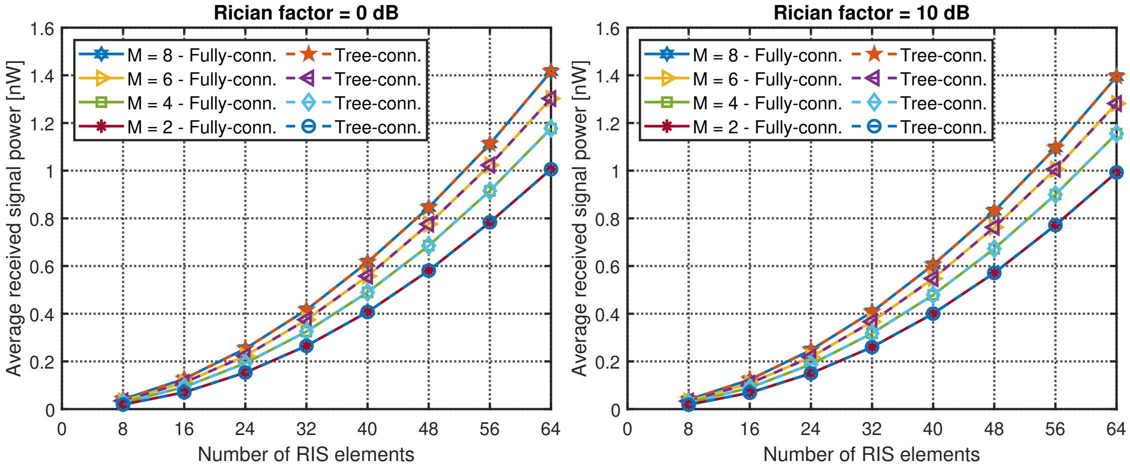


Figure 5.4: Received signal power in MISO systems aided by fully- and tree-connected RIS.

Second, higher received signal power can be obtained by increasing the number of transmit antennas M for both fully- and tree-connected RIS. *Last*, Rician fading channels with a lower Rician factor offer richer scattering, allowing to reach a slightly higher performance.

We next evaluate the performance of the forest-connected RIS. Specifically, the forest-connected RIS is optimized through Alg. 5 and it is compared with tree-, group-, and single-connected RIS. For the group-connected RIS, we recall that the N elements are divided into G groups and all the RIS ports within the same group are connected with each other, so that $\mathbf{B}$ is a symmetric block diagonal matrix and $\boldsymbol{\Theta}$ satisfies (2.17). In Figs. 5.5 and 5.6, we provide the received signal power in MISO systems aided by tree-, group-, forest-, and single-connected RIS, with $M = 2$ and $M = 8$, respectively. Group-connected and single-connected RIS are optimized as proposed in [75, Alg. 2]. We can make the following observations. *First*, the forest-connected RIS can always achieve the same received signal power as the group-connected RIS with the same group size. However, it has a much lower circuit complexity than the group-connected RIS with the same group size, as quantitatively shown in the following. *Second*, the forest-connected RIS achieves a higher received signal power than the single-connected RIS. For example, forest-connected RIS with group size 8 brings an improvement in the received signal power of 44.6%, when $M = 2$, $N = 64$, and $K = 0$ dB. On the other hand, it achieves a lower received signal power than the tree-connected RIS due to the simplified circuit complexity. Therefore, it is shown that the forest-connected RIS achieves a good performance-complexity trade-off between the single- and tree-connected RIS. *Third*, the received signal power increases with the group size in the forest-connected RIS, which is because increasing the group size can provide more flexibility to the BD-RIS. *Fourth*, higher received signal power can be obtained by increasing the number of transmit antennas M. *Last*, BD-RIS is particularly beneficial over single-connected RIS in the presence of fading channels with lower Rician factors, in agreement with [16].

We finally evaluate the circuit topology complexity of the proposed tree- and forest-connected RIS. As analyzed in Section 5.2, the circuit topology complexity in terms of the number of the tunable admittance components of the tree- and forest-connected RIS is

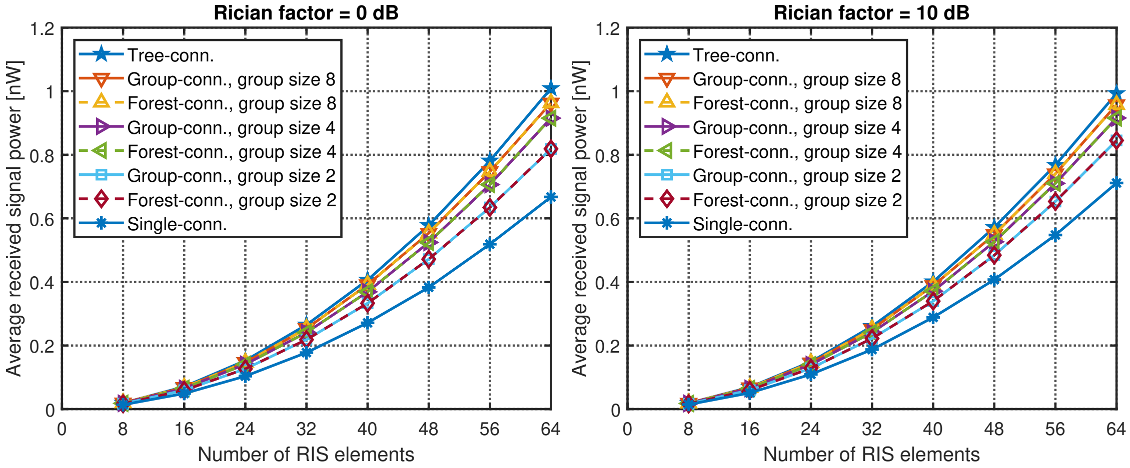


Figure 5.5: Received signal power in MISO systems aided by tree-, group-, forest-, and single-connected RIS, with $M = 2$.

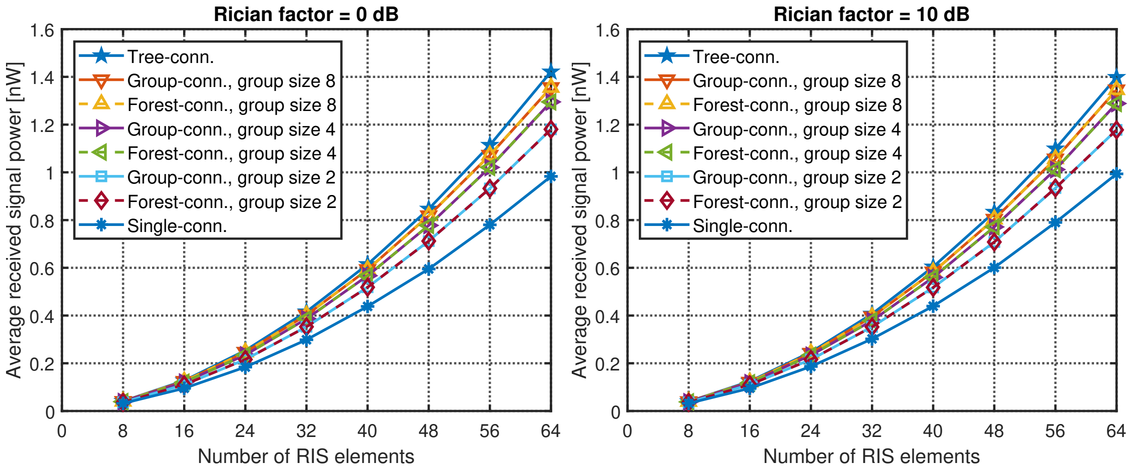


Figure 5.6: Received signal power in MISO systems aided by tree-, group-, forest-, and single-connected RIS, with $M = 8$.

$2N-1$ and $N(2-1/N_G)$, respectively. For comparison, we also consider the circuit topology complexity of fully-, group-, and single-connected RIS, given by $N(N+1)/2$, $N(N_G+1)/2$, and N, respectively [16]. In Fig. 5.7, we provide the number of tunable admittance components in the fully-, group, tree-, forest-, and single-connected RIS. We can make the following observations. *First*, compared with the fully-connected RIS, the tree-connected RIS has much lower circuit topology complexity. For example, the number of tunable admittance components is decreased by 16.4 times when $N = 64$ in the tree-connected RIS compared to the fully-connected RIS, but they achieve the same performance. *Second*, compared with the group-connected RIS with same group size, the forest-connected RIS has much lower circuit topology complexity. For example, the number of tunable admittance components is reduced by 2.4 times when $N = 64$ in the forest-connected RIS compared to the group-connected RIS with group size 8, but they achieve the same performance. *Third*, increasing the group size in the forest-connected RIS will increase the number of tunable admittance components while also enhancing the performance. *Fourth*, compared with the single-connected RIS, the tree- and forest-connected RIS introduce appropriate complexity but achieves an improvement in the received power.

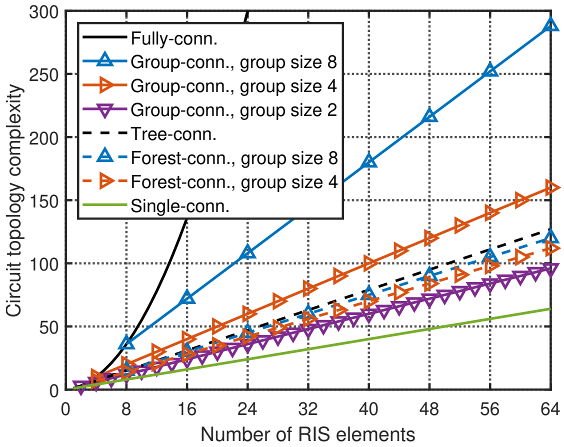


Figure 5.7: Circuit topology complexity, i.e., the number of tunable admittance components, of fully-, group-, tree-, forest-, and single-connected RIS.

To conclude, we demonstrate that the benefit of the tree-connected (resp. forest-connected) RIS over the fully-connected (resp. group-connected) RIS lies in their highly simplified circuit complexity while maintaining optimal performance. Therefore, the proposed tree- and forest-connected architectures significantly improve the performance-complexity trade-off over existing BD-RIS architectures.

5.5 Conclusion

We propose novel modeling, architecture design, and optimization for BD-RIS by utilizing graph theory. In particular, we model BD-RIS architectures as graphs, capturing the presence of interconnections between the RIS elements. Through this modeling, we prove that a BD-RIS achieves the performance upper bound in MISO systems if and only if its associated graph is connected. This remarkable result allows us to characterize the least complex BD-RIS architectures able to achieve the performance upper bound in MISO systems, denoted as tree-connected RIS. We also propose forest-connected RIS to bridge between the single-connected and the tree-connected architectures. To optimize these novel BD-RIS architectures, we derive a closed-form global optimal solution for tree-connected RIS, and an iterative algorithm for forest-connected RIS. Numerical results confirm that tree-connected (resp. forest-connected) RIS achieve the same performance as fully-connected (resp. group-connected) RIS, with a significantly reduced circuit complexity by up to 16.4 times. We leave the optimization and performance evaluation of the proposed tree- and forest-connected RISs in other scenarios such as multi-user systems as the object of future research. In addition, the study of the optimal graph topologies in the presence of hardware impairments, such as mutual coupling, also constitutes an interesting future direction.

5.6 Appendix

5.6.1 Proof of Lemma 1

We prove the necessary condition by showing that if $\mathcal{G}$ is not a connected graph, the corresponding BD-RIS is not MISO optimal. If $\mathcal{G}$ is disconnected, it has $C \geq 2$ connected components, where the cth component includes $N^{(c)}$ ports, for $c = 1, \ldots, C$, with $\sum_{c=1}^{C} N^{(c)} = N$. With no loss of generality, we assume that each component includes adjacent ports. Thus, the admittance matrix $\mathbf{Y}$ given by (5.2) is block diagonal, with the cth block having dimensions $N^{(c)} \times N^{(c)}$, for $c = 1, \ldots, C$. As a consequence of (2.3), also the scattering matrix $\boldsymbol{\Theta}$ is block diagonal, and writes as $\boldsymbol{\Theta} = \text{diag}(\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_C)$, where $\boldsymbol{\Theta}_c \in \mathbb{C}^{N^{(c)} \times N^{(c)}}$, for $c = 1, \ldots, C$. Thus, the received signal power (5.5) in the case of a disconnected graph $\mathcal{G}$ is

$$P_R^{\text{DisC}} = P_T \left| \sum_{c=1}^{C} \mathbf{h}_{RI,c} \boldsymbol{\Theta}_c \mathbf{H}_{IT,c} \mathbf{w} \right|^2, \tag{5.42}$$

where $\mathbf{h}_{RI,c} \in \mathbb{C}^{1 \times N^{(c)}}$ and $\mathbf{H}_{IT,c} \in \mathbb{C}^{N^{(c)} \times M}$ contain the $N^{(c)}$ elements of $\mathbf{h}_{RI}$ and rows of $\mathbf{H}_{IT}$ corresponding to the $N^{(c)}$ RIS elements grouped into the cth component, respectively. The received signal power P_R^{DisC} is upper bounded by

$$\bar{P}_R^{\text{DisC}} = P_T \left(\sum_{c=1}^{C} \left\| \mathbf{h}_{RI,c} \right\| \left\| \mathbf{H}_{IT,c} \mathbf{w} \right\| \right)^2, \tag{5.43}$$

following the Cauchy-Schwarz inequality and that $\boldsymbol{\Theta}_c^H \boldsymbol{\Theta}_c = \mathbf{I}$, for $c = 1, \ldots, C$. Furthermore, it holds that

$$\bar{P}_R^{\text{DisC}} \overset{(a)}{\leq} P_T \left\| \mathbf{h}_{RI} \right\|^2 \left\| \mathbf{H}_{IT} \mathbf{w} \right\|^2 \overset{(b)}{\leq} \bar{P}_R, \tag{5.44}$$

where (a) follows by the Cauchy-Schwarz inequality and (b) follows by the definition of the spectral norm. Note that (b) is an equality if and only if $\mathbf{w}$ is the dominant right singular vector of $\mathbf{H}_{IT}$. However, (a) is in general a strict inequality if $\mathbf{w}$ is the dominant right singular vector of $\mathbf{H}_{IT}$, which thus proves the necessary condition of the lemma.

On the other hand, we prove the sufficient condition by showing that if $\mathcal{G}$ is a connected graph, the corresponding BD-RIS is MISO optimal. This is straightforward to prove because 1) for a connected graph, we can always remove some edges to make it a tree, i.e. setting the corresponding off-diagonal entries of $\mathbf{B}$ zero, 2) there exist one and only one solution for any RIS whose associated graph is a tree to achieve the performance upper bound (i.e., to be MISO optimal) as shown in Section 5.3.1.

5.6.2 Proof of Proposition 3

Applying Lemma 1, we need to prove that a connected graph with N vertices has at least $N-1$ edges. This is achieved by induction. The base case is easily verified: a connected graph with a single vertex has at least zero edges. As the induction step, we consider a connected graph with N vertices. From this graph, we remove edges (at least one) until we obtain a disconnected graph with two connected components. Assuming the two components have K and $N-K$ vertices, they have at least $K-1$ and $N-K-1$ edges by the induction hypothesis, respectively. Since we removed at least one edge to disconnect the graph, it had originally at least $(K-1)+(N-K-1)+1=N-1$ edges.

5.6.3 Proof of Proposition 4

This proof is conducted by induction. As the base case, we consider the only tree-connected RIS that can be realized with $N=2$ ports. This RIS includes two tunable admittance components connecting the two ports to ground and a further tunable admittance connecting the two ports to each other.

Based on the susceptance matrix $\mathbf{B}\in\mathbb{R}^{2\times 2}$, the left-hand side of the system (5.27) is built by setting $\mathbf{x}=[[\mathbf{B}]_{1,1},[\mathbf{B}]_{2,2},[\mathbf{B}]_{1,2}]^T$. Accordingly, $\mathbf{A}\in\mathbb{R}^{4\times 3}$ is given by (5.28) with $\mathbf{A}_1\in\mathbb{C}^{2\times 2}$ as in (5.29) and $\mathbf{A}_2\in\mathbb{C}^{2\times 1}$ given by $\mathbf{A}_2=[[\boldsymbol{\alpha}]_2,[\boldsymbol{\alpha}]_1]^T$. Thus, it is easy to recognize that $\mathbf{A}$ has in general full column rank, i.e., $r(\mathbf{A})=3$. The proposition is hence verified for the case $N=2$.

As the induction step, we prove that if the proposition is valid for RISs with $N-1$ ports, it also holds for RISs with N ports. We consider a tree-connected RIS with $N-1$ elements, whose coefficient matrix is given by

$$\mathbf{A}^{(N-1)}=\left[\begin{array}{cc}\Re\left\{\mathbf{A}_1^{(N-1)}\right\} & \Re\left\{\mathbf{A}_2^{(N-1)}\right\}\\ \hdashline \Im\left\{\mathbf{A}_1^{(N-1)}\right\} & \Im\left\{\mathbf{A}_2^{(N-1)}\right\}\end{array}\right]. \tag{5.45}$$

To this $(N-1)$-port tree-connected RIS, we connect an additional port creating an N-port tree-connected RIS. With no loss of generality, we assume that the additional port is connected to the $(N-1)$th port through a tunable admittance. In this way, the resulting N-port BD-RIS has the coefficient matrix

$$\mathbf{A}^{(N)}=\left[\begin{array}{cc}\Re\left\{\mathbf{A}_1^{(N)}\right\} & \Re\left\{\mathbf{A}_2^{(N)}\right\}\\ \hdashline \Im\left\{\mathbf{A}_1^{(N)}\right\} & \Im\left\{\mathbf{A}_2^{(N)}\right\}\end{array}\right], \tag{5.46}$$

where $\mathbf{A}_1^{(N)}\in\mathbb{C}^{N\times N}$ and $\mathbf{A}_2^{(N)}\in\mathbb{C}^{N\times(N-1)}$ are given by

$$\mathbf{A}_1^{(N)}=\left[\begin{array}{cc}\mathbf{A}_1^{(N-1)} & \mathbf{0}_{(N-1)\times 1}\\ \hdashline \mathbf{0}_{1\times(N-1)} & [\boldsymbol{\alpha}]_N\end{array}\right],\ \mathbf{A}_2^{(N)}=\left[\begin{array}{cc}\mathbf{A}_2^{(N-1)} & \begin{array}{c}\mathbf{0}_{(N-2)\times 1}\\ [\boldsymbol{\alpha}]_N\end{array}\\ \hdashline \mathbf{0}_{1\times(N-2)} & [\boldsymbol{\alpha}]_{N-1}\end{array}\right], \tag{5.47}$$

with $\mathbf{0}_{R\times C}$ denoting an $R\times C$ all-zero matrix. To prove the induction step, we need to show that $r(\mathbf{A}^{(N-1)})=2(N-1)-1$ implies $r(\mathbf{A}^{(N)})=2N-1$. To this end, $\mathbf{A}^{(N)}$ is rewritten as

$$\mathbf{A}^{(N)}\sim\left[\begin{array}{ccc}\mathbf{A}^{(N-1)} & \mathbf{0}_{(2N-2)\times 1} & \begin{array}{c}\mathbf{0}_{(N-2)\times 1}\\ \Re\left\{[\boldsymbol{\alpha}]_N\right\}\\ \mathbf{0}_{(N-2)\times 1}\\ \Im\left\{[\boldsymbol{\alpha}]_N\right\}\end{array}\\ \hdashline \mathbf{0}_{2\times(2N-3)} & \begin{array}{c}\Re\{[\boldsymbol{\alpha}]_N\}\\ \Im\{[\boldsymbol{\alpha}]_N\}\end{array} & \begin{array}{c}\Re\left\{[\boldsymbol{\alpha}]_{N-1}\right\}\\ \Im\left\{[\boldsymbol{\alpha}]_{N-1}\right\}\end{array}\end{array}\right], \tag{5.48}$$

by applying appropriate row and column swapping operations. From (5.48), we notice that if $\mathbf{A}^{(N-1)}$ has full column rank and the vectors $[\Re\{[\boldsymbol{\alpha}]_N\}, \Im\{[\boldsymbol{\alpha}]_N\}]$ and $\left[\Re\left\{[\boldsymbol{\alpha}]_{N-1}\right\}, \Im\left\{[\boldsymbol{\alpha}]_{N-1}\right\}\right]$ are linearly independent, then $\mathbf{A}^{(N)}$ has full column rank. Since $[\Re\{[\boldsymbol{\alpha}]_N\}, \Im\{[\boldsymbol{\alpha}]_N\}]$ and $\left[\Re\left\{[\boldsymbol{\alpha}]_{N-1}\right\}, \Im\left\{[\boldsymbol{\alpha}]_{N-1}\right\}\right]$ are in practice linearly independent with probability 1, the induction step is proven and the proof is concluded.

5.6.4 Proof of Proposition 5

To prove that $r([\mathbf{A}|\mathbf{b}]) = 2N-1$, we show that it is always possible to obtain a zero row in $[\mathbf{A}|\mathbf{b}]$ by applying appropriate row operations. Firstly, we execute on $[\mathbf{A}|\mathbf{b}]$ the N row operations

$$\mathbf{r}_{N+n} = \mathbf{r}_{N+n} - \frac{\Im\{[\boldsymbol{\alpha}]_n\}}{\Re\{[\boldsymbol{\alpha}]_n\}}\mathbf{r}_n, \tag{5.49}$$

for $n = 1, \ldots, N$, where $\mathbf{r}_n$ denotes the nth row of $[\mathbf{A}|\mathbf{b}]$. The resulting matrix is given by

$$[\mathbf{A}|\mathbf{b}] \sim \left[\begin{array}{ccc} \Re\{\mathbf{A}_1\} & \Re\{\mathbf{A}_2\} & \Re\{\boldsymbol{\beta}\} \\ \hdashline \mathbf{0}_{N\times N} & \mathbf{A}' & \mathbf{b}' \end{array}\right], \tag{5.50}$$

where we introduced $\mathbf{A}' \in \mathbb{R}^{N\times(N-1)}$ and $\mathbf{b}' \in \mathbb{R}^{N\times 1}$. The matrix $\mathbf{A}'$ has exactly two non-zero elements in each column, in the same positions as the non-zero elements in $\mathbf{A}_2$. Specifically,

$$\begin{cases} [\mathbf{A}']_{n_\ell,\ell} & = \Im\left\{[\boldsymbol{\alpha}]_{m_\ell}\right\} - \frac{\Im\left\{[\boldsymbol{\alpha}]_{n_\ell}\right\}}{\Re\left\{[\boldsymbol{\alpha}]_{n_\ell}\right\}}\Re\left\{[\boldsymbol{\alpha}]_{m_\ell}\right\}, \\ [\mathbf{A}']_{m_\ell,\ell} & = \Im\left\{[\boldsymbol{\alpha}]_{n_\ell}\right\} - \frac{\Im\left\{[\boldsymbol{\alpha}]_{m_\ell}\right\}}{\Re\left\{[\boldsymbol{\alpha}]_{m_\ell}\right\}}\Re\left\{[\boldsymbol{\alpha}]_{n_\ell}\right\}, \\ [\mathbf{A}']_{p,\ell} & = 0, \forall p \neq m_\ell, n_\ell, \end{cases} \tag{5.51}$$

for $\ell = 1, \ldots, N-1$, with n_ℓ and m_ℓ being the row indexes of the two non-zero elements in the ℓth column of $\mathbf{A}_2$. The vector $\mathbf{b}'$ writes as

$$\left[\mathbf{b}'\right]_n = \Im\{[\boldsymbol{\beta}]_n\} - \frac{\Im\{[\boldsymbol{\alpha}]_n\}}{\Re\{[\boldsymbol{\alpha}]_n\}}\Re\{[\boldsymbol{\beta}]_n\}, \tag{5.52}$$

for $n = 1, \ldots, N$. Secondly, we execute on the resulting $[\mathbf{A}|\mathbf{b}]$ the N row operations

$$\mathbf{r}_{N+n} = \Re\{[\boldsymbol{\alpha}]_n\}\mathbf{r}_{N+n}, \tag{5.53}$$

for $n = 1, \ldots, N$, followed by the $N-1$ row operations

$$\mathbf{r}_{2N} = \mathbf{r}_{2N} + \mathbf{r}_{N+n}, \tag{5.54}$$

for $n = 1, \ldots, N-1$. The resulting matrix is given by

$$[\mathbf{A}|\mathbf{b}] \sim \left[\begin{array}{ccc} \Re\{\mathbf{A}_1\} & \Re\{\mathbf{A}_2\} & \Re\{\boldsymbol{\beta}\} \\ \hdashline \mathbf{0}_{N\times N} & \mathbf{A}'' & \mathbf{b}'' \end{array}\right], \tag{5.55}$$

where we introduced $\mathbf{A}'' \in \mathbb{R}^{N\times N-1}$ and $\mathbf{b}'' \in \mathbb{R}^{N\times 1}$.

We now show that (5.55) has rank $2N-1$ since its $2N$th row is all-zero. To this end, we need to prove that $[\mathbf{A}'']_{N,\ell} = 0$, for $\ell = 1, \dots, N-1$, and that $[\mathbf{b}'']_N = 0$. Indeed, we have

$$\left[\mathbf{A}''\right]_{N,\ell} = \Re\left\{[\boldsymbol{\alpha}]_{n_\ell}\right\}\left[\mathbf{A}'\right]_{n_\ell,\ell} + \Re\left\{[\boldsymbol{\alpha}]_{m_\ell}\right\}\left[\mathbf{A}'\right]_{m_\ell,\ell} = 0 \tag{5.56}$$

for $\ell = 1, \dots, N-1$, with n_ℓ and m_ℓ being the row indexes of the two non-zero elements in the ℓth column of $\mathbf{A}_2$. Furthermore, the Nth entry of vector $\mathbf{b}''$ is given by

$$\left[\mathbf{b}''\right]_N = \sum_{n=1}^{N} \Re\left\{[\boldsymbol{\alpha}]_n\right\}\left[\mathbf{b}'\right]_n = \sum_{n=1}^{N} \Re\left\{[\boldsymbol{\alpha}]_n\right\}\Im\left\{[\boldsymbol{\beta}]_n\right\} - \Im\left\{[\boldsymbol{\alpha}]_n\right\}\Re\left\{[\boldsymbol{\beta}]_n\right\}, \tag{5.57}$$

where

$$\Re\left\{[\boldsymbol{\alpha}]_n\right\}\Im\left\{[\boldsymbol{\beta}]_n\right\} - \Im\left\{[\boldsymbol{\alpha}]_n\right\}\Re\left\{[\boldsymbol{\beta}]_n\right\} \tag{5.58}$$

$$= jZ_0\left(\Im\left\{\left[\hat{\mathbf{h}}_{RI}^H\right]_n\right\}^2 - \Im\left\{[\mathbf{u}_{IT}]_n\right\}^2 + \Re\left\{\left[\hat{\mathbf{h}}_{RI}^H\right]_n\right\}^2 - \Re\left\{[\mathbf{u}_{IT}]_n\right\}^2\right) \tag{5.59}$$

$$= jZ_0\left(\left|\left[\hat{\mathbf{h}}_{RI}^H\right]_n\right|^2 - \left|[\mathbf{u}_{IT}]_n\right|^2\right). \tag{5.60}$$

Thus, recalling that $\|\hat{\mathbf{h}}_{RI}^H\|^2 = 1$ and $\|\mathbf{u}_{IT}\|^2 = 1$, we have

$$\left[\mathbf{b}''\right]_N = jZ_0\left(\left\|\hat{\mathbf{h}}_{RI}^H\right\|^2 - \|\mathbf{u}_{IT}\|^2\right) = 0, \tag{5.61}$$

proving that the $2N$th row of (5.55) is all-zero.

Chapter 6

Performance-Complexity Trade-off in BD-RIS

When designing new BD-RIS architectures, the critical issue is the trade-off between performance and circuit complexity, given by the number of tunable impedance components in the BD-RIS architecture [16, 80]. On the one hand, the single-connected RIS is the architecture with the lowest circuit complexity since there are no interconnections among the RIS elements. Due to its limited flexibility, the single-connected RIS can only achieve a reduced performance. On the other hand, the fully-connected RIS has the highest circuit complexity since each RIS element is connected to all others through a tunable impedance component, enabling the highest performance. Several BD-RIS architectures have been proposed to trade performance and complexity. However, the fundamental limits of this trade-off are still unexplored. To fill this gap, we investigate how to optimally trade the achievable performance and the circuit complexity in BD-RIS architectures.

The contribution of this chapter is twofold. *First*, we derive the Pareto frontier for the performance-complexity trade-off offered by BD-RIS in SISO systems. *Second*, we characterize the optimal BD-RIS architectures allowing to achieve this Pareto frontier.

Organization: In Section 6.1, we formulate the problem of finding the Pareto frontier of the performance-complexity trade-off enabled by BD-RIS. In Section 6.2, we derive the expression of the Pareto frontier. In Section 6.3, we show the numerical results. Finally, Section 6.4 concludes this chapter.

The content of this chapter has appeared in [85]. For reproducible research, the code is available at `https://github.com/matteonerini/pareto-frontier-for-bdris`.

6.1 Problem Formulation

Consider a SISO communication system aided by an N-element RIS. The N elements of the RIS are connected to a N-port reconfigurable impedance network, with scattering matrix $\boldsymbol{\Theta} \in \mathbb{C}^{N \times N}$. Defining $x \in \mathbb{C}$ as the transmitted signal and $y \in \mathbb{C}$ as the received signal, we have $y = hx + n$, where $h \in \mathbb{C}$ is the wireless channel and $n \in \mathbb{C}$ is the AWGN at the receiver. Assuming that the direct link between the transmitter and the receiver is negligible compared to the RIS-aided link, the channel h writes as $h = \mathbf{h}_R \boldsymbol{\Theta} \mathbf{h}_T$, where

$\mathbf{h}_R \in \mathbb{C}^{1\times N}$ and $\mathbf{h}_T \in \mathbb{C}^{N\times 1}$ refer to the channels from the RIS to the receiver and from the transmitter to the RIS, respectively [16][1]. We assume independent and identically distributed (i.i.d.) Rayleigh channels to obtain fundamental insights, having unit channel gains with no loss of generality, i.e., $\mathbf{h}_R \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ and $\mathbf{h}_T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$.

When reconfiguring an RIS, the scattering matrix $\boldsymbol{\Theta}$ is typically optimized to maximize the performance given by the received signal power

$$P_R = P_T \left|\mathbf{h}_R \boldsymbol{\Theta} \mathbf{h}_T\right|^2, \tag{6.1}$$

where $P_T = \mathrm{E}[|x|^2]$ is the transmitted signal power. Considering passive RISs with lossless and reciprocal impedance networks, the matrix $\boldsymbol{\Theta}$ is in general subject to the constraints $\boldsymbol{\Theta}^H \boldsymbol{\Theta} = \mathbf{I}$ and $\boldsymbol{\Theta} = \boldsymbol{\Theta}^T$ [74]. Furthermore, additional constraints on $\boldsymbol{\Theta}$, limiting the received signal power, are present depending on the BD-RIS architecture [16, 80].

Conventional RIS, also known as single-connected RIS, is the least complex architecture achieving the lowest performance, given by

$$\bar{P}_R^{\text{Single}} = P_T \left(\sum_{n=1}^{N} \left| [\mathbf{h}_R]_n [\mathbf{h}_T]_n \right| \right)^2, \tag{6.2}$$

since it includes only N tunable impedance components [16]. In contrast, tree-connected RIS is proved to be the least complex architecture achieving the performance upper bound

$$\bar{P}_R^{\text{Tree}} = P_T \left\|\mathbf{h}_R\right\|^2 \left\|\mathbf{h}_T\right\|^2, \tag{6.3}$$

with $2N-1$ tunable impedance components [80]. In this chapter, our goal is to determine the maximum performance achievable by BD-RIS architectures with circuit complexity $C \in [N, 2N-1]$, representing the number of tunable components[2]. In other words, we want to characterize the Pareto frontier of the performance-complexity trade-off enabled by BD-RIS. Furthermore, we are interested in which BD-RIS architectures allow us to reach such a frontier, denoted as "optimal" BD-RIS architectures in the following.

We begin by characterizing the maximum received signal power achievable by a given BD-RIS architecture. To this end, we consider the modeling of BD-RIS based on graph theory developed in [80]. According to [80], each BD-RIS architecture can be described through a graph $\mathcal{G}$ capturing the presence of tunable impedance components between its RIS elements. We denote as G the number of connected components of such a graph $\mathcal{G}$, where a connected component of a graph is defined as a connected subgraph that is not part of any larger connected subgraph [82]. Besides, $N_g \geq 1$ is the number of RIS elements included in the gth component, with $\sum_{g=1}^{G} N_g = N$. In agreement with previous work on BD-RIS [16], we refer to the connected components of $\mathcal{G}$ as the "groups" of the corresponding BD-RIS architecture. According to [80], the maximum received signal

[1]Since $\mathbf{h}_R \boldsymbol{\Theta} \mathbf{h}_T$ can always be co-phased with the direct link, our conclusions are not impacted by the direct link. In the case of a non-negligible direct link, the performance would merely be scaled up, depending on its strength. Thus, we neglect the direct link to gain fundamental insights not depending on its strength.

[2]In our analysis, we preclude BD-RISs with dynamic interconnections, since they require switches and hence additional circuit complexity.

power obtained by the BD-RIS associated with $\mathcal{G}$ is given by

$$\bar{P}_R = P_T \left(\sum_{g=1}^{G} \|\mathbf{h}_{R,g}\| \, \|\mathbf{h}_{T,g}\| \right)^2 , \tag{6.4}$$

where $\mathbf{h}_{R,g} \in \mathbb{C}^{1\times N_g}$ and $\mathbf{h}_{T,g} \in \mathbb{C}^{N_g\times 1}$ contain the N_g elements of $\mathbf{h}_R$ and $\mathbf{h}_T$ corresponding to the N_g RIS elements included into the gth group, respectively. In the case of i.i.d. fading channels, we can assume that each group includes adjacent RIS elements with no loss of generality, such that $\mathbf{h}_R = [\mathbf{h}_{R,1}, \ldots, \mathbf{h}_{R,G}]$ and $\mathbf{h}_T = [\mathbf{h}_{T,1}^T, \ldots, \mathbf{h}_{T,G}^T]^T$. Thus, the maximum received signal power $\bar{P}_R$ achievable by a given BD-RIS solely depends on G and the group sizes $N_1, \ldots, N_G$.

To express the maximum received signal power $\bar{P}_R$ achievable with a circuit complexity C as a function of C, we introduce the following three results. First, we characterize the optimal BD-RIS architectures through the following lemma.

Lemma 2. *All the optimal BD-RIS architectures have a corresponding graph being acyclic, also known as a forest.*

Proof. Please refer to Appendix 6.5.1. □

In other words, a BD-RIS architecture can be optimal only if its graph does not contain any cycle, i.e., a finite sequence of distinct edges joining a sequence of vertices, where only the first and last vertices are equal [82]. Second, we use the following result from graph theory [82].

Lemma 3. *If a graph $\mathcal{G}$ is a forest, then it has $G = N - L$ connected components, where N is the number of vertices and L is the number of edges.*

Proof. Please refer to Appendix 6.5.2. □

Third, by using Lemma 2 and Lemma 3, we can derive the following proposition.

Proposition 6. *An optimal BD-RIS architecture with N elements and circuit complexity C, with $C \in [N, 2N-1]$, has a corresponding graph with $G = 2N - C$ connected components.*

Proof. Please refer to Appendix 6.5.3. □

According to Proposition 6, given a circuit complexity C, the number of groups G in the corresponding optimal BD-RIS is fixed. Thus, our problem is to find the group sizes $N_1, \ldots, N_G$ of the BD-RIS architecture that maximize the performance $\mathrm{E}\left[\bar{P}_R\right]$, with fixed G. The corresponding optimization problem is given by

$$\max_{N_1,\ldots,N_G} \; \mathrm{E}\left[\bar{P}_R\right] \tag{6.5}$$

$$\text{s.t.} \quad N_g \geq 1, \, \forall g, \; \sum_{g=1}^{G} N_g = N, \tag{6.6}$$

where $G = 2N - C$ is fixed depending on the complexity C.

6.2 Pareto Frontier

We now solve problem (6.5)-(6.6) by rewriting and simplifying the objective (6.5), then we use the obtained optimal group sizes $N_1, \ldots, N_G$ to derive the desired Pareto frontier. To solve problem (6.5)-(6.6), we assume $P_T = 1$ with no loss of generality and we write the average received signal power (6.4) as

$$\mathrm{E}\left[\bar{P}_R\right] = \sum_{g=1}^{G} \mathrm{E}\left[\|\mathbf{h}_{R,g}\|^2\right]^2 + \sum_{g_1 \neq g_2} \mathrm{E}\left[\|\mathbf{h}_{R,g_1}\|\right]^2 \mathrm{E}\left[\|\mathbf{h}_{R,g_2}\|\right]^2, \tag{6.7}$$

where we exploited the i.i.d. channels assumption and the fact that $\mathbf{h}_R$ and $\mathbf{h}_T$ are identically distributed. Using the moments of the chi distribution with $2N_g$ degrees of freedom, we have that $\mathrm{E}[\|\mathbf{h}_{R,g}\|] = \Gamma(N_g + 1/2)/\Gamma(N_g)$ and $\mathrm{E}[\|\mathbf{h}_{R,g}\|^2] = N_g$, $\forall g$, where $\Gamma(\cdot)$ is the gamma function. Thus, we can write

$$\mathrm{E}\left[\bar{P}_R\right] = \sum_{g=1}^{G} N_g^2 + \sum_{g_1 \neq g_2} \left(\frac{\Gamma\left(N_{g_1} + 1/2\right)}{\Gamma\left(N_{g_1}\right)}\right)^2 \left(\frac{\Gamma\left(N_{g_2} + 1/2\right)}{\Gamma\left(N_{g_2}\right)}\right)^2. \tag{6.8}$$

The expression of $\mathrm{E}\left[\bar{P}_R\right]$ in (6.8) can be now simplified by using the relationship

$$\left(\frac{\Gamma\left(M + 1/2\right)}{\Gamma\left(M\right)}\right)^2 = M - \frac{1}{4} + \frac{1}{32M} + \mathcal{O}\left(\left(\frac{1}{M}\right)^2\right), \tag{6.9}$$

given by the Laurent series expansion at $M = \infty$ [86]. Remarkably, the function $(\Gamma(M + 1/2)/\Gamma(M))^2$ is well approximated by $M - 1/4 + 1/(32M)$ for any positive integer M, despite the series being computed at $M = \infty$. Thus, we can approximate (6.8) as

$$\mathrm{E}\left[\bar{P}_R\right] = \sum_{g=1}^{G} N_g^2 + \sum_{g_1 \neq g_2} \left(N_{g_1} - \frac{1}{4} + \frac{1}{32N_{g_1}}\right)\left(N_{g_2} - \frac{1}{4} + \frac{1}{32N_{g_2}}\right). \tag{6.10}$$

By developing the product and reorganizing the terms, we get

$$\begin{aligned}\mathrm{E}\left[\bar{P}_R\right] = \sum_{g=1}^{G} N_g^2 + \sum_{g_1 \neq g_2} \Bigg(& N_{g_1}N_{g_2} + \frac{1}{16} - \frac{N_{g_1}}{4} - \frac{N_{g_2}}{4} \\ &+ \frac{N_{g_1}}{32N_{g_2}} + \frac{N_{g_2}}{32N_{g_1}} - \frac{1}{128N_{g_1}} - \frac{1}{128N_{g_2}} + \frac{1}{1024N_{g_1}N_{g_2}}\Bigg).\end{aligned} \tag{6.11}$$

Recalling that $N_g \geq 1$, $\forall g$, the term $1/(1024N_{g_1}N_{g_2})$ in (6.11) is negligible since it is at least 1024 times smaller than the term $N_{g_1}N_{g_2}$. Thus, omitting $1/(1024N_{g_1}N_{g_2})$ and completing the computations, we obtain

$$\begin{aligned}\mathrm{E}\left[\bar{P}_R\right] = N^2 + \frac{G\left(G-1\right)}{16} - \frac{N\left(G-1\right)}{2}& \\ + \sum_{g_1 \neq g_2} \left(\frac{N_{g_1}}{32N_{g_2}} + \frac{N_{g_2}}{32N_{g_1}} - \frac{1}{128N_{g_1}} - \frac{1}{128N_{g_2}}\right)&.\end{aligned} \tag{6.12}$$

Observing that

$$\sum_{g_1 \neq g_2} \frac{N_{g_1}}{32N_{g_2}} = \sum_{g_1 \neq g_2} \frac{N_{g_2}}{32N_{g_1}} = \sum_{g=1}^{G} \frac{N - N_g}{32N_g}, \tag{6.13}$$

$$\sum_{g_1 \neq g_2} \frac{1}{128N_{g_1}} = \sum_{g_1 \neq g_2} \frac{1}{128N_{g_2}} = \sum_{g=1}^{G} \frac{G-1}{128N_g}, \tag{6.14}$$

we can eventually rewrite $\mathrm{E}\left[\bar{P}_R\right]$ as

$$\mathrm{E}\left[\bar{P}_R\right] = N^2 + \frac{G-1}{16}(G - 8N) - \frac{G}{16} + \frac{4N - G + 1}{64} \sum_{g=1}^{G} \frac{1}{N_g}. \tag{6.15}$$

We notice from (6.15) that maximizing $\mathrm{E}\left[\bar{P}_R\right]$ is equivalent to maximize $\sum_{g=1}^{G} 1/N_g$. Thus, problem (6.5)-(6.6) can be equivalently expressed as

$$\max_{N_1,\ldots,N_G} \sum_{g=1}^{G} \frac{1}{N_g} \tag{6.16}$$

$$\text{s.t.} \quad N_g \geq 1,\, \forall g,\, \sum_{g=1}^{G} N_g = N, \tag{6.17}$$

which is solved in the following proposition.

Proposition 7. *The solution to problem* (6.16)-(6.17) *is given by*

$$N_1 = N_2 = \ldots = N_{G-1} = 1, \tag{6.18}$$

$$N_G = N - G + 1, \tag{6.19}$$

up to a permutation of the group sizes.

Proof. Please refer to Appendix 6.5.4. □

Given the optimal group sizes $N_1, \ldots, N_G$ provided by Proposition 7, we can derive in closed form the expression of the desired Pareto frontier. Specifically, plugging (6.18) and (6.19) into (6.8), we obtain

$$\begin{aligned}\mathrm{E}\left[\bar{P}_R\right] = G - 1 + (N - G + 1)^2 + (G-1)(G-2)\Gamma(3/2)^4 \\ + 2(G-1)\left(\frac{\Gamma(N - G + 3/2)\Gamma(3/2)}{\Gamma(N - G + 1)}\right)^2,\end{aligned} \tag{6.20}$$

giving the maximum performance achievable with a BD-RIS architecture having G groups. Finally, recalling that $G = 2N - C$, the expression of the maximum performance achievable with a circuit complexity $C \in [N, 2N-1]$ is given by

$$\begin{aligned}\mathrm{E}\left[\bar{P}_R\right] = (C - N)^2 + C + (2N - C - 1)(2N - C - 2)\Gamma(3/2)^4 \\ + 2(2N - C - 1)\left(\frac{\Gamma(C - N + 3/2)\Gamma(3/2)}{\Gamma(C - N + 1)}\right)^2,\end{aligned} \tag{6.21}$$

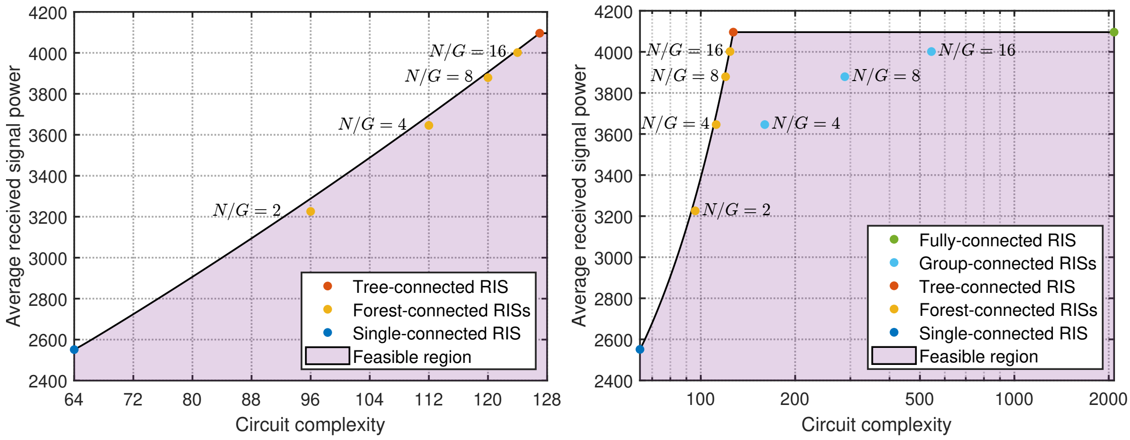


Figure 6.1: Pareto frontier for the performance-complexity trade-off achieved by BD-RISs, with $N = 64$.

representing the Pareto frontier of the performance-complexity trade-off offered by BD-RISs.

6.3 Numerical Results

In Fig. 6.1, we report the Pareto frontier given by (6.21), delimiting the region of feasible BD-RIS architectures[3]. Specifically, the feasible region is delimited by (6.21) when $C \in [N, 2N-1]$ since (6.21) gives the maximum performance achievable with complexity C, and by the horizontal line $\mathrm{E}[\bar{P}_R] = N^2$ when $C > 2N-1$ since N^2 is the performance upper bound [16]. We fix $N = 64$ in Fig. 6.1 to obtain a fair comparison in terms of the space occupied by the RIS.

The derived frontier is compared with the performance-complexity trade-off achieved by the BD-RIS architectures recently proposed in [16, 80]. Since each BD-RIS is characterized by its circuit complexity C and average received signal power $\mathrm{E}[\bar{P}_R]$, each BD-RIS is represented as a point in Fig. 6.1 with coordinates $(C, \mathrm{E}[\bar{P}_R])$. More precisely, we report group- and forest-connected RISs, both achieving a performance

$$\mathrm{E}\left[\bar{P}_R^{\mathrm{Group}}\right] = \frac{N^2}{G} + G(G-1)\left(\frac{\Gamma(N/G + 1/2)}{\Gamma(N/G)}\right)^4, \tag{6.22}$$

and with complexity $C^{\mathrm{Group}} = N(N/G+1)/2$ and $C^{\mathrm{Forest}} = 2N-G$, respectively, depending on the group size N/G [16, 80]. Note that (6.22) is obtained by setting $N_g = N/G$, $\forall g$, in (6.8). The four forest-connected RISs in Fig. 6.1 have group sizes 2, 4, 8, and 16, while the three group-connected RISs have group sizes 4, 8, and 16 since forest- and group-connected RISs with group sizes 2 are equivalent [80]. Besides, we report fully- and tree-connected RISs, both achieving

$$\mathrm{E}\left[\bar{P}_R^{\mathrm{Fully}}\right] = N^2, \tag{6.23}$$

and with complexity $C^{\mathrm{Fully}} = N(N+1)/2$ and $C^{\mathrm{Tree}} = 2N-1$, respectively [16, 80], where

[3]Note that the average received signal power has no units since is computed with unit transmit power and unit channel gains with no loss of generality.

(6.23) is derived by setting $G = 1$ in (6.22). Finally, we report the single-connected RIS architecture, achieving a performance given by

$$\mathrm{E}\left[\bar{P}_R^{\mathrm{Single}}\right] = N + N\left(N-1\right)\Gamma\left(3/2\right)^4, \tag{6.24}$$

and with complexity $C^{\mathrm{Single}} = N$ [16], where (6.24) is derived by setting $G = N$ in (6.22).

We make the following remarks. *First*, on the one hand, the single-connected RIS is the least complex architecture, achieving the lowest performance due to its limited architecture. On the other hand, the tree-connected RIS allows us to reach the performance upper bound, with the lowest possible complexity. *Second*, forest-connected RISs approach the Pareto frontier, but they are slightly suboptimal. This is because forest-connected RISs have equally sized groups, i.e., they all have group size N/G. However, the optimal group sizes are not all equal, as given by (6.18)-(6.19). *Third*, the fully-connected (resp. group-connected) RIS achieves the same performance as the tree-connected (resp. forest-connected) RIS, but with higher circuit complexity. Thus, in SISO systems, fully- and group-connected RISs are highly suboptimal. Note that the exact shape of the Pareto frontier depends on the channel distribution. Specifically, with Rician or correlated channels, the gain of the tree-connected over the single-connected RIS decreases, and less complex BD-RIS architectures are expected to approach the performance upper bound [16, 80].

6.4 Conclusion

We derive the Pareto frontier for the performance-complexity trade-off in BD-RISs. This frontier provides the BD-RIS architectures that can be optimally used to bridge between the single-connected RIS and the tree-connected RIS. The presented fundamental results are expected to drive the development of novel BD-RIS architectures and prototypes. Remarkably, a multi-dimensional Pareto frontier could be derived accounting for multiple parameters in addition to the circuit complexity, such as the number of RIS elements, the channel estimation overhead, and the optimization complexity, thus representing a future research direction.

6.5 Appendix

6.5.1 Proof of Lemma 2

To prove Lemma 2, we show that any BD-RIS whose corresponding graph is not a forest, is not optimal. Consider a BD-RIS with circuit complexity C with a corresponding graph $\mathcal{G}$ that is not a forest, i.e., it has at least one cycle [82]. According to (6.4), the received signal power achievable by a BD-RIS solely depends on its number of groups G, and the group sizes $N_1, \ldots, N_G$. Thus, by removing one edge from a cycle in $\mathcal{G}$, the resulting BD-RIS has complexity $C-1$ and achieves the same performance as the original one since its graph still has G connected components with sizes $N_1, \ldots, N_G$. Since this resulting BD-RIS achieves the same performance as the original one with reduced complexity, the original BD-RIS is not optimal, and Lemma 2 is proved.

6.5.2 Proof of Lemma 3

Consider a forest $\mathcal{G}$ with G connected components, where the gth component has N_g vertices, with $\sum_{g=1}^{G} N_g = N$. Since $\mathcal{G}$ is a forest, each component is a tree, i.e., is a connected forest, and the gth component includes $N_g - 1$ edges [82, Theorem 2.2]. Thus, the number of edges in $\mathcal{G}$ is

$$L = \sum_{g=1}^{G} (N_g - 1) = \sum_{g=1}^{G} N_g - G = N - G, \tag{6.25}$$

proving that $G = N - L$.

6.5.3 Proof of Proposition 6

According to [80], a BD-RIS with N elements and C tunable impedance components has a corresponding graph with $L = C - N$ edges since N tunable impedance components are used to connect each RIS element to ground. Besides, we know from Lemma 2 and Lemma 3 that the graph of an optimal BD-RIS with N elements and complexity C is a forest, having $G = N - L$ connected components. By plugging $L = C - N$ into $G = N - L$, we obtain $G = 2N - C$.

6.5.4 Proof of Proposition 7

The proof is conducted by induction on the number of groups G. As the base case, we consider $G = 2$, where problem (6.16)-(6.17) boils down to

$$\min_{N_1, N_2} \quad N_1 N_2 \tag{6.26}$$

$$\text{s.t.} \quad N_1 \geq 1, \; N_2 \geq 1, \; N_1 + N_2 = N. \tag{6.27}$$

The solution to this problem is clearly given by $N_1 = 1$ and $N_2 = N - 1$, or vice versa. The proposition is consequently verified for the case $G = 2$.

As the induction step, we prove that if the proposition is valid for $G-1$ groups, it also holds for G groups. To this end, we rewrite problem (6.16)-(6.17) as

$$\max_{N_1, \ldots, N_G} \quad \sum_{g=1}^{G-1} \frac{1}{N_g} + \frac{1}{N_G} \tag{6.28}$$

$$\text{s.t.} \quad N_g \geq 1, \; \forall g, \; \sum_{g=1}^{G-1} N_g = N - N_G. \tag{6.29}$$

By the induction hypothesis, we have

$$N_1 = N_2 = \ldots = N_{G-2} = 1, \tag{6.30}$$

$$N_{G-1} = N - N_G - G + 2. \tag{6.31}$$

Using (6.30) and (6.31), problem (6.28)-(6.29) can be simplified as

$$\min_{N_{G-1},N_G} \quad N_{G-1}N_G \tag{6.32}$$

$$\text{s.t.} \quad N_{G-1} \geq 1,\ N_G \geq 1, \tag{6.33}$$

$$N_{G-1} + N_G = N - G + 2, \tag{6.34}$$

where the only unknown are N_{G-1} and N_G. By solving this problem as done for the base case, we obtain $N_{G-1} = 1$ and $N_G = N - G + 1$, proving the induction step.

Part III

BD-RIS With Hardware Impairments

Chapter 7

BD-RIS With Discrete-Value Scattering Matrices

The group- and fully-connected RISs proposed in [16] have been studied assuming continuous-value scattering matrices. However, the scattering matrix of a RIS can be reconfigured by choosing among a finite number of discrete values in practical implementation, and it is not clear how to design these BD-RIS architectures with discrete-value scattering matrices. Discrete-value single-connected RISs have been typically realized by considering uniform discretization of the phase shifts in the interval $[0, 2\pi)$. This is made possible by the diagonal structure of the scattering matrix of these RISs, which is completely described by its phase shifts. Thus, this discretization design cannot be applied to group- and fully-connected RISs, whose scattering matrices are not diagonal. This gap is addressed in this chapter, where we propose a design strategy for discrete-value group- and fully-connected RISs. The contributions of this chapter are summarized as follows.

First, we propose a strategy to design discrete-value group- and fully-connected RISs that assigns a finite number of bits to each entry of the RIS impedance matrix. The resulting RISs are denoted as scalar-discrete RISs, since they are obtained with a scalar quantization approach. Firstly, a suitable codebook is designed in the offline learning stage. Secondly, the discrete elements of the impedance matrix are optimized based on the resulting codebook in the online deployment stage.

Second, we prove that our codebook always achieves a received signal power growth of $\mathcal{O}(N_I^2)$ in group-connected architectures, where N_I is the number of RIS elements. Thus, the proposed discretization strategy does not degrade the growth of the received signal power as a function of N_I.

Third, we develop a different discretization strategy for group- and fully-connected RISs, exploiting the block diagonal structure of their impedance matrices. According to this strategy, we jointly discretize and optimize the entries in each block of the RIS impedance matrix through vector quantization. RISs designed with this strategy, denoted as vector-discrete RISs, achieve higher performance than scalar-discrete RISs at the cost of increased optimization computational complexity. Vector discretization has never been considered for conventional RISs since their impedance matrix is diagonal.

Fourth, we assess the performance in terms of received signal power of single-, group-

, and fully-connected RISs employing different numbers of resolution bits. We verify that the more connections are present in the reconfigurable impedance network, the less resolution is needed to achieve the optimal performance of continuous-value RISs. Because of this property, in the fully-connected architecture, a single resolution bit allocated to each reconfigurable impedance is sufficient to achieve the performance of ideal RISs.

Organization: In Section 7.1, we introduce the system model and the problem formulation. In Sections 7.2 and 7.3, we present our novel discrete group- and fully-connected RIS design based on scalar and vector quantization, respectively. In Section 7.4, we assess the performance of MIMO systems aided by discrete RISs in terms of received signal power. Finally, Section 7.5 contains the concluding remarks.

The content of this chapter has appeared in [87].

7.1 System Model

Let us consider an RIS-aided MIMO system, where we denote as N_T the number of antennas at the transmitter, N_R the number of antennas at the receiver, and N_I the number of antennas at the RIS. The N_I antennas of the RIS are connected to a N_I-port reconfigurable impedance network, with scattering matrix $\mathbf{\Theta} \in \mathbb{C}^{N_I \times N_I}$. By assuming perfect matching and no mutual coupling at the RIS antennas, the channel matrix seen by the receiver $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ can be written as

$$\mathbf{H} = \mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}, \tag{7.1}$$

where $\mathbf{H}_{RT} \in \mathbb{C}^{N_R \times N_T}$, $\mathbf{H}_{RI} \in \mathbb{C}^{N_R \times N_I}$, and $\mathbf{H}_{IT} \in \mathbb{C}^{N_I \times N_T}$ are the channels from transmitter to receiver, RIS to receiver, and transmitter to RIS, respectively [16].

We consider a single-stream MIMO transmission scheme, to exploit the diversity gain offered by the multiple antennas at the transmitter and receiver, and by the reflecting elements at the RIS. We denote the transmit signal as $\mathbf{x} = \mathbf{w}s \in \mathbb{C}^{N_T \times 1}$, where $\mathbf{w} \in \mathbb{C}^{N_T \times 1}$ is the precoding vector subject to the constraint $\|\mathbf{w}\| = 1$, and $s \in \mathbb{C}$ is the transmit symbol with average power $P_T = \mathrm{E}[|s|^2]$. Denoting the receive signal as $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$, we have $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{n} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_n^2\mathbf{I}\right)$ is the AWGN at the receiver with variance σ_n^2. Thus, the signal used for detection is given by $z = \mathbf{g}\mathbf{y} \in \mathbb{C}$, where $\mathbf{g} \in \mathbb{C}^{1 \times N_R}$ is the combining vector subject to the constraint $\|\mathbf{g}\| = 1$. Finally, we can express z as $z = \mathbf{g}\mathbf{H}\mathbf{w}s + \tilde{n}$, where $\tilde{n} = \mathbf{g}\mathbf{n}$ is the AWGN with variance σ_n^2.

In the following, we propose two novel strategies to design discrete-value group- and fully-connected RISs. In the first, scalar-discrete RISs, B resolution bits are allocated to each reactance matrix entry $[\mathbf{X}_I]_{i,j}$. In the second, vector-discrete RISs, B_V bits are set to each reactance matrix block $\mathbf{X}_{I,g}$.

7.2 Scalar-Discrete Group-/Fully-Connected RIS

Our goal is to design the discrete-value matrix $\mathbf{\Theta}$ and the vectors $\mathbf{g}$ and $\mathbf{w}$ to maximize the received signal power, given by $P_R = P_T \left|\mathbf{g}\left(\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\right)\mathbf{w}\right|^2$. In the case of single-stream transmission, the optimal precoding and combining vectors are given by the dominant eigenmode transmission. Thus, maximizing the received signal power is equivalent to maximize $\|\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\|^2$. To investigate the design of discrete-value group- and fully-connected RISs, it is necessary to first consider the design of group- and fully-connected RISs with continuous-value $\mathbf{\Theta}$. Then, the continuous-value optimization is used to optimize $\mathbf{\Theta}$ with discrete values.

7.2.1 Continuous-Value Group-/Fully-Connected RIS

The received signal power maximization problem for continuous-value $\mathbf{\Theta}$ in MIMO systems can be formulated as

$$\max_{\mathbf{X}_{I,g}} \ \|\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\|^2 \tag{7.2}$$

$$\text{s.t.} \ \ \mathbf{\Theta} = \text{diag}\left(\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_G\right), \ \mathbf{\Theta}_g = \mathbf{\Theta}_g^T, \ \mathbf{\Theta}_g^H\mathbf{\Theta}_g = \mathbf{I}, \ \forall g, \tag{7.3}$$

where the constraints derived from (2.17) indicate that the scattering matrix $\mathbf{\Theta}$ is a block diagonal matrix with each block being a complex symmetric unitary matrix. These constraints complicate the optimization problem, which needs to be reformulated. Thus, the relationship between $\mathbf{\Theta}$ and the reactance matrix $\mathbf{X}_I$ given by (2.4) is exploited to equivalently rewrite problem (7.2)-(7.3) as

$$\max_{\mathbf{X}_{I,g}} \ \|\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\|^2 \tag{7.4}$$

$$\text{s.t.} \ \ \mathbf{\Theta} = \text{diag}\left(\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_G\right), \tag{7.5}$$

$$\mathbf{\Theta}_g = \left(j\mathbf{X}_{I,g} + Z_0\mathbf{I}\right)^{-1}\left(j\mathbf{X}_{I,g} - Z_0\mathbf{I}\right), \ \mathbf{X}_{I,g} = \mathbf{X}_{I,g}^T, \ \forall g, \tag{7.6}$$

which can be transformed into an unconstrained problem. More precisely, exploiting the constraints (7.5) and (7.6), the objective $\|\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\|^2$ can be expressed as a function of $\mathbf{X}_{I,g}$, $\forall g$. Since $\mathbf{X}_{I,g}$ is an arbitrary $N_G \times N_G$ real symmetric matrix, $\mathbf{X}_{I,g}$ is an unconstrained function of the $N_G\left(N_G+1\right)/2$ entries in its upper triangular part. Thus, the obtained problem is an unconstrained optimization problem in the variables $\left[\mathbf{X}_{I,g}\right]_{i,j}$, with $i \leq j$ and $\forall g$.

Problem (7.4)-(7.6) has been solved for RIS-aided SISO systems in [16] by using the quasi-Newton method to find the optimal upper triangular part of each $\mathbf{X}_{I,g}$ without any constraints. For RIS-aided MIMO systems, we propose to solve it by alternatively optimizing $\mathbf{X}_I$ and the beamforming vectors $\mathbf{g}$ and $\mathbf{w}$, as established in the literature on single-connected RISs [13]. After $\mathbf{g}$ and $\mathbf{w}$ are initialized to feasible values, this optimization process alternates between the two following steps until convergence is reached. With fixed $\mathbf{g}$ and $\mathbf{w}$, we update $\mathbf{X}_I$ by maximizing the objective $|\mathbf{g}\mathbf{H}_{RT}\mathbf{w} + \mathbf{g}\mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\mathbf{w}|$ with the quasi-Newton method as in [16], where $\mathbf{\Theta}$ is a function of $\mathbf{X}_I$ given by (2.4). With fixed $\mathbf{X}_I$, and consequently fixed $\mathbf{\Theta}$, we update $\mathbf{g}$ and $\mathbf{w}$ as the dominant left and right singular vectors of the matrix $\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}$, respectively. Note that this method is proven to converge to a stationary point of the objective.

7.2.2 Problem Formulation for Discrete-Value RIS

In (7.4)-(7.6), the reactance matrix entries are allowed to assume arbitrary real values, which is hard to realize in practice. Thus, we are interested in an RIS design strategy in which each reactance matrix entry is selected from a codebook. We consider a discrete reactance codebook $\mathcal{C}$ symmetric around zero, written as

$$\mathcal{C} = \left\{\pm c_1, \pm c_2, \ldots, \pm c_{2^{B-1}}\right\}, \tag{7.7}$$

where $c_b > 0$, for $b = 1, \ldots, 2^{B-1}$. When selecting each reactance matrix entry from a codebook, the RIS optimization problem cannot be simply solved as an unconstrained problem. Thus, instead of (7.4)-(7.6), the following optimization problem can be considered to optimize $\mathbf{X}_I$

$$\max_{\mathbf{X}_{I,g}} \;\; \left\|\mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\right\|^2 \tag{7.8}$$

$$\text{s.t.} \;\; \boldsymbol{\Theta} = \operatorname{diag}\left(\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_G\right), \tag{7.9}$$

$$\boldsymbol{\Theta}_g = \left(j\mathbf{X}_{I,g} + Z_0\mathbf{I}\right)^{-1}\left(j\mathbf{X}_{I,g} - Z_0\mathbf{I}\right), \; \mathbf{X}_{I,g} = \mathbf{X}_{I,g}^T, \; \left[\mathbf{X}_{I,g}\right]_{i,j} \in \mathcal{C}, \; \forall g. \tag{7.10}$$

Since the entries of $\mathbf{X}_I$ are not distributed in a finite interval, to define the optimal codebook $\mathcal{C}$ is not as straightforward as in the single-connected case [13]. Thus, we introduce a second optimization problem to design $\mathcal{C}$ as

$$\max_{\mathcal{C}} \;\; \mathrm{E}\left[\left\|\mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\right\|^2\right] \tag{7.11}$$

$$\text{s.t.} \;\; \boldsymbol{\Theta} = \operatorname{diag}\left(\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_G\right), \tag{7.12}$$

$$\boldsymbol{\Theta}_g = \left(j\mathbf{X}_{I,g} + Z_0\mathbf{I}\right)^{-1}\left(j\mathbf{X}_{I,g} - Z_0\mathbf{I}\right), \; \mathbf{X}_{I,g} \text{ solves (7.8)-(7.10)}, \; \forall g, \tag{7.13}$$

where the objective is the ergodic received signal power and the expectation operator $\mathrm{E}[\cdot]$ represents the average over the channel realizations[1]. Note that in (7.11)-(7.13), $\mathbf{X}_{I,g}$ implicitly depends on $\mathcal{C}$ because of constraint (7.13). These two nested optimization problems form a bilevel programming problem, in which (7.11)-(7.13) is the upper-level problem, and (7.8)-(7.10) is the lower-level problem. This bilevel problem is difficult to solve since in the upper-level problem we do not have an expression for the constraint (7.13), and we cannot explicitly write $\mathbf{X}_{I,g}$ as a function of $\mathcal{C}$.

For this reason, when considering the optimization of discrete group- and fully-connected RISs, two problems arise. Firstly, during the offline learning stage, a suitable discrete reactance codebook has to be defined for the elements of $\mathbf{X}_I$. Secondly, during the online deployment stage, the discrete values of $\mathbf{X}_I$ are optimized by choosing among the values of the fixed codebook. In the following, these two stages are described. The offline learning stage is performed differently for binary-level ($B = 1$) and multi-level ($B > 1$) codebooks since a lower complexity solution is available when $B = 1$.

[1] Note that the received signal power in (7.8)-(7.13) can be replaced by any objective function depending on the CSI and the RIS scattering matrix. Thus, the proposed codebook design and optimization framework is general enough to accommodate any objective function, including the metrics typically considered in multi-user systems, e.g., the sum rate.

7.2.3 Offline Learning for Binary-Level Codebook

During offline learning, we assume that the codebook design can exploit a training set of N_0 channel realization triplets. Such a training set can be practically collected through a sampling campaign to be conducted offline. During this offline sampling campaign, channel realizations are collected and stored with dedicated transceiver devices. The n_0th triplet is denoted as $\mathcal{H}^{[n_0]} = (\mathbf{H}_{RT}^{[n_0]}, \mathbf{H}_{RI}^{[n_0]}, \mathbf{H}_{IT}^{[n_0]})$ and includes the channels from transmitter to receiver, RIS to receiver, and transmitter to RIS, respectively. Thus, the training set is a set of triplets defined as

$$\mathcal{H}_0 = \left\{\mathcal{H}^{[1]}, \mathcal{H}^{[2]}, \ldots, \mathcal{H}^{[N_0]}\right\}. \tag{7.14}$$

When the number of resolution bits per reconfigurable reactance is $B = 1$, the entries of the reactance matrix $\mathbf{X}_I$ can assume only the two values in $\mathcal{C} = \{-c_1, +c_1\}$, where c_1 is a positive real number. Exploiting this special structure of binary-level codebooks, we can obtain the optimal codebook value c_1 and binary-value reactance matrix $\bar{\mathbf{X}}_I$ for each channel triplet in the training set by solving the bilevel optimization problem (7.8)-(7.13).

Firstly, problem (7.8)-(7.10) can be solved given a fixed value of c_1 through alternating optimization. With this method, each non-zero entry of the reactance matrix is optimized individually by searching among the two possible values in $\mathcal{C}$, while fixing the other $N_I\left(N_G+1\right)/2-1$ entries. This procedure is repeated for all the reactance matrix entries, iterating multiple times until the convergence is reached. In this way, we can evaluate the optimal $\mathbf{X}_{I,g}$ as function of c_1 in constraint (7.13). Secondly, the upper-level problem (7.11)-(7.13) is solvable with one-dimensional pattern search since its objective can be readily evaluated as a function of c_1. We employ pattern search since it is an optimization algorithm that only requires the evaluation of the objective function in a series of points until the convergence is reached, without using derivatives [88]. Thus, we build the two sets

$$\mathcal{C}_1 = \left\{c_1^{[1]}, c_1^{[2]}, \ldots, c_1^{[N_0]}\right\},\ \bar{\mathcal{X}}_I = \left\{\bar{\mathbf{X}}_I^{[1]}, \bar{\mathbf{X}}_I^{[2]}, \ldots, \bar{\mathbf{X}}_I^{[N_0]}\right\}, \tag{7.15}$$

where $c_1^{[n_0]}$ and $\bar{\mathbf{X}}_I^{[n_0]}$ are the optimal codebook value c_1 and binary-value reactance matrix $\bar{\mathbf{X}}_I$ associated to the channel triplet $\mathcal{H}^{[n_0]}$, respectively.

Given the two sets $\mathcal{C}_1$ and $\bar{\mathcal{X}}_I$, we design the optimal value for c_1, denoted as $c_1^\star$. We obtain $c_1^\star$ as the value that minimizes the distortion caused in the reactance matrix

$$c_1^\star = \underset{c_1}{\arg\min}\ \frac{1}{N_0}\sum_{n_0=1}^{N_0}\left\|\bar{\mathbf{X}}_I^{[n_0]} - \mathbf{X}_I\right\|_F^2 \tag{7.16}$$

$$\text{s.t.}\quad \mathbf{X}_I = \operatorname{diag}\left(\mathbf{X}_{I,1}, \ldots, \mathbf{X}_{I,G}\right),\ \left[\mathbf{X}_{I,g}\right]_{i,j} \in \left\{-c_1, +c_1\right\},\ \forall g, \tag{7.17}$$

where (7.16) is a sample average computed over the channel triplets in the training set. In (7.16), each non-zero squared entry of the matrix $\bar{\mathbf{X}}_I^{[n_0]} - \mathbf{X}_I$ is given by $(c_1^{[n_0]} - c_1)^2$ because of the symmetry of the codebook. Thus, problem (7.16) can be reformulated as

$$c_1^\star = \underset{c_1}{\arg\min}\ \frac{N_I N_G}{N_0}\sum_{n_0=1}^{N_0}\left(c_1^{[n_0]} - c_1\right)^2. \tag{7.18}$$

By setting the derivative of the objective function in (7.18) with respect to c_1 to zero, we obtain

$$c_1^\star = \frac{1}{N_0} \sum_{n_0=1}^{N_0} c_1^{[n_0]}. \tag{7.19}$$

Finally, the binary-value codebook is given by $\mathcal{C} = \{-c_1^\star, +c_1^\star\}$.

The complexity of the offline learning for binary-level codebook is driven by the complexity of solving the bilevel problem (7.8)-(7.13) N_0 times to find $\mathcal{C}_1$. Since one-dimensional pattern search requires computing twice the objective function per iteration, the complexity per iteration is given by $\mathcal{O}(2N_0N_I(N_G+1))$.

7.2.4 Offline Learning for Multi-Level Codebook

Pattern search can be used to solve the upper-level problem (7.11)-(7.13) only when the optimization involves a one-dimensional search, that is only when $B = 1$. When more resolution bits are considered, pattern search becomes highly sub-optimal since it is likely to converge to a local maximum in a multi-dimensional search space. For this reason, we rely on a different method to define multi-level codebooks. The main idea is to learn the codebook from the distribution of the optimal continuous-value $\mathbf{X}_I$ which solve (7.4)-(7.6). More precisely, we find the optimal set of values $\{c_1^\star, \ldots, c_K^\star\}$, with $K = 2^{B-1}$, such that the distortion on the optimal continuous-value reactance matrix is minimized. The process leading to the multi-level codebook $\mathcal{C} = \{\pm c_1^\star, \ldots, \pm c_{2^{B-1}}^\star\}$ is detailed in the following.

We firstly solve (7.4)-(7.6) for each training set triplet by alternatively optimizing $\mathbf{X}_I$ and the beamforming vectors $\mathbf{g}$ and $\mathbf{w}$. As a result, we construct the set

$$\mathcal{X}_I = \left\{\mathbf{X}_I^{[1]}, \mathbf{X}_I^{[2]}, \ldots, \mathbf{X}_I^{[N_0]}\right\}, \tag{7.20}$$

where $\mathbf{X}_I^{[n_0]}$ is the optimal continuous-value reactance matrix associated to the channel triplet $\mathcal{H}^{[n_0]}$. Given the set $\mathcal{X}_I$, the optimal values $\{c_1^\star, \ldots, c_K^\star\}$ are determined by solving

$$\{c_1^\star, \ldots, c_K^\star\} = \underset{\{c_1,\ldots,c_K\}}{\arg\min}\ \frac{1}{N_0} \sum_{n_0=1}^{N_0} \left\|\mathbf{X}_I^{[n_0]} - \mathbf{X}_I\right\|_F^2 \tag{7.21}$$

$$\text{s.t.} \quad \mathbf{X}_I = \operatorname{diag}\left(\mathbf{X}_{I,1}, \ldots, \mathbf{X}_{I,G}\right), \ \left[\mathbf{X}_{I,g}\right]_{i,j} \in \{\pm c_1, \ldots, \pm c_K\}, \ \forall g. \tag{7.22}$$

To simplify problem (7.21), we introduce the vectors $\mathbf{x}^{[n_0]} \in \mathbb{R}^{1\times N_I N_G}$ containing the non-zero entries of $\mathbf{X}_I^{[n_0]}$. Then, the N_0 vectors $\mathbf{x}^{[n_0]}$ are concatenated in $\mathbf{x}_0 \in \mathbb{R}^{1\times N_0 N_I N_G}$ as $\mathbf{x}_0 = [\mathbf{x}^{[1]}, \ldots, \mathbf{x}^{[N_0]}]$. In this way, problem (7.21) becomes

$$\{c_1^\star, \ldots, c_K^\star\} = \underset{\{c_1,\ldots,c_K\}}{\arg\min}\ \frac{1}{N_0} \sum_{n=1}^{N_0 N_I N_G} \left|[\mathbf{x}_0]_n - [\mathbf{x}_I]_n\right|^2 \tag{7.23}$$

$$\text{s.t.} \quad [\mathbf{x}_I]_n \in \{\pm c_1, \ldots, \pm c_K\}, \ \forall n. \tag{7.24}$$

where the vector $\mathbf{x}_I \in \mathbb{R}^{1\times N_0 N_I N_G}$ has been introduced as an auxiliary variable. Exploiting the symmetry of the codebook $\mathcal{C}$, (7.23) can be further reformulated as

$$\{c_1^\star,\ldots,c_K^\star\} = \underset{\{c_1,\ldots,c_K\}}{\arg\min}\ \frac{1}{N_0}\sum_{n=1}^{N_0N_IN_G}\left|\left|[\mathbf{x}_0]_n\right|-[\mathbf{x}_I]_n\right|^2 \tag{7.25}$$

$$\text{s.t.}\quad [\mathbf{x}_I]_n\in\{c_1,\ldots,c_K\}\,,\ \forall n, \tag{7.26}$$

which is a K-means clustering problem in a one-dimensional space, with $K=2^{B-1}$ [89]. For each data point $x_n=\left|[\mathbf{x}_0]_n\right|$ we introduce the binary indicator $r_{nk}\in\{0,1\}$ such that $r_{nk}=1$ if x_n is assigned to c_k, and $r_{ni}=0$ otherwise. The resulting K-means clustering problem is formalized as

$$\{c_1^\star,\ldots,c_K^\star\} = \underset{\{c_1,\ldots,c_K\}}{\arg\min}\ \sum_{n=1}^{N_0N_IN_G}\sum_{k=1}^{2^{B-1}} r_{nk}\left|x_n-c_k\right|^2 \tag{7.27}$$

To solve (7.27), we iteratively optimize the sets $\{r_{nk}|n=1,\ldots,N_0N_IN_G,k=1,\ldots,K\}$ and $\{c_1,\ldots,c_K\}$. Firstly, when the values $\{c_1,\ldots,c_K\}$ are fixed, we set $\{r_{nk}\}$ by assigning the nth data point x_n to the closest cluster, i.e.,

$$r_{nk}=\begin{cases}1 & \text{if } k=\arg\min_j\left|x_n-c_j\right|^2\\ 0 & \text{otherwise}\end{cases}, \tag{7.28}$$

for $n=1,\ldots,N_0N_IN_G$ and $k=1,\ldots,K$. Secondly, when $\{r_{nk}\}$ are fixed, the cluster centers $\{c_1,\ldots,c_K\}$ are updated. This can be done in close form by setting the derivative of the objective function in (7.27) with respect to c_k to zero

$$c_k=\frac{\sum_n r_{nk}x_n}{\sum_n r_{nk}}, \tag{7.29}$$

for $k=1,\ldots,K$. The two steps of assigning data points to clusters and updating the cluster centers are repeated alternatively until there is no further change in the cluster assignment. Finally, the multi-value codebook is given by $\mathcal{C}=\{\pm c_1^\star,\ldots,\pm c_{2^{B-1}}^\star\}$.

The complexity of the offline learning for multi-level codebook is given by $\mathcal{O}(N_0(N_I(N_G+1)/2)^2+N_0N_IN_G2^{B-1})$, where the first term is due to solving (7.4)-(7.6) N_0 times to find $\mathcal{X}_I$, and the second term is due to the K-means clustering algorithm used to solve (7.27).

7.2.5 Online Deployment

Once the codebook has been learned offline, the discrete $\mathbf{X}_I$ entries are optimized during the online deployment stage. This is achieved by solving (7.8)-(7.10) depending on the instantaneous CSI $\mathbf{H}_{RT}$, $\mathbf{H}_{RI}$, and $\mathbf{H}_{IT}$ and on the designed codebook $\mathcal{C}$. To solve this problem, an exhaustive search could be employed since each reactance matrix entry can assume a finite number of values. However, to reduce the search space, we employ an alternating optimization method. In this sub-optimal method, each non-zero entry of the reactance matrix is optimized individually by searching among all possible 2^B values, while fixing the other $N_I\left(N_G+1\right)/2-1$ entries. This procedure is repeated for all $N_I\left(N_G+1\right)/2$ reactance matrix entries, iterating multiple times until the convergence

Algorithm 6 Scalar-Discrete RISs Design

Input: B, $\mathcal{H}_0$, $\mathbf{H}_{RT}$, $\mathbf{H}_{RI}$, $\mathbf{H}_{IT}$, N_G.
Output: $\mathcal{C}$, $\mathbf{X}_I$.

Offline Learning
if $B == 1$ **then**
  Calculate $\mathcal{C}_1$ by the bilevel problem (7.8)-(7.13).
  Calculate $c_1^\star$ by (7.19).
  $\mathcal{C} \leftarrow \{-c_1^\star, +c_1^\star\}$.
else
  Calculate $\mathcal{X}_I$ by (7.4)-(7.6).
  Calculate $\{c_1^\star, \ldots, c_K^\star\}$ by iteratively solving (7.27), with $K = 2^{B-1}$.
  $\mathcal{C} \leftarrow \{\pm c_1^\star, \ldots, \pm c_K^\star\}$.
end if
Online Deployment
Initialize $\mathbf{X}_{I,g} = \mathbf{0}$, $\forall g$.
while no convergence of objective (7.8) **do**
  for $g \leftarrow 1$ to G **do**
    for $i \leftarrow 1$ to N_G **do**
      for $j \leftarrow i$ to N_G **do**
        Update $[\mathbf{X}_{I,g}]_{i,j} \in \mathcal{C}$ by searching among the 2^B possible values.
        $[\mathbf{X}_{I,g}]_{j,i} \leftarrow [\mathbf{X}_{I,g}]_{i,j}$.
      end for
    end for
  end for
end while

is reached. The convergence is considered reached when the fractional increase of the objective value in a full iteration is below a certain parameter ϵ. The convergence is guaranteed by the following two facts. First, at each iteration, the objective function $\|\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\|^2$ is non-decreasing. This holds since each reactance matrix entry is optimized by exhaustively searching among the possible values in the codebook $\mathcal{C}$. Second, the objective function is bounded from above by $(\|\mathbf{H}_{RT}\| + \|\mathbf{H}_{RI}\|\|\mathbf{H}_{IT}\|)^2$. The design of scalar-discrete RISs is summarized in Alg. 6, which solves the problem (7.8)-(7.13).

Note that we assume that the statistical properties of the channels are unchanged during the offline learning and online deployment stages. This is necessary since we approximate the expectation of the received signal power with its sample average, computed over the training set. However, numerical simulations showed that learning the codebook based on outdated channel statistics only slightly degrades the performance. Besides, perfect instantaneous CSI is assumed during the online deployment stage, which can be obtained with the semi-passive channel estimation strategy proposed in [13].

7.2.6 Scaling Law

To analyze the optimality of the codebook in (7.7), we characterize the scaling law of the received signal power as a function of the number of RIS elements N_I, when $\mathcal{C}$ is used to discretize the reactance matrix. Note that this analysis is necessary since the scaling law of the proposed discrete RIS architectures is not known and not straightforward given the novel constraints of BD-RISs. To obtain fundamental insights, we consider a SISO system, i.e., $N_T = 1$ and $N_R = 1$. Furthermore, the direct link h_{RT} is negligible in comparison with the reflected signal power for asymptotically large N_I. Considering unitary transmit power, the received signal power is given by $P_R = |\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{h}_{IT}|^2$, where $\boldsymbol{\Theta}$ is obtained by discretizing the reactance matrix $\mathbf{X}_I$. Such a scaling law is given in the following proposition.

Proposition 8. *Assume i.i.d. Rayleigh fading channels, i.e.,* $\mathbf{h}_{RI} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ *and* $\mathbf{h}_{IT} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. *A group-connected RIS whose reactance matrix entries are chosen from* $\mathcal{C}$ *can achieve an average received signal power in the order of* $\mathcal{O}(N_I^2)$ *as* N_I *increases, even with one resolution bit per each reactance.*

Proof. Let us consider the worst case $B = 1$, and assume that the matrix block $\mathbf{X}_{I,g}$ can only assume two values $\mathbf{X}_{I,g} = \pm c_1 \text{ones}(N_G)$. Note that proving Proposition 1 for the case $\mathcal{C} = \{\pm c_1\}$ also establishes its validity for all codebooks $\mathcal{C} = \{\pm c_1, \pm c_2, \ldots, \pm c_{2^{B-1}}\}$. Through eigenmode decomposition, we write $\mathbf{X}_{I,g} = \mathbf{V}_g\boldsymbol{\Lambda}_g\mathbf{V}_g^T$, where $\boldsymbol{\Lambda}_g = \text{diag}(\boldsymbol{\lambda}_g)$ is diagonal containing the eigenvalues of $\mathbf{X}_{I,g}$ and $\mathbf{V}_g$ is orthonormal. Among the eigenvalues of $\mathbf{X}_{I,g}$, only $[\boldsymbol{\lambda}_g]_1$ is non zero since $\mathbf{X}_{I,g}$ is rank one. By applying (2.4), the scattering matrix $\boldsymbol{\Theta}_g$ can be expressed as $\boldsymbol{\Theta}_g = \mathbf{V}_g\mathbf{D}_g\mathbf{V}_g^T$, where $\mathbf{D}_g = \text{diag}(\mathbf{d}_g)$ is a diagonal matrix with $[\mathbf{d}_g]_{n_G} = \frac{j[\boldsymbol{\lambda}_g]_{n_G} - Z_0}{j[\boldsymbol{\lambda}_g]_{n_G} + Z_0}$. According to [16], the average received signal power for group-connected RISs is written as

$$\mathrm{E}[P_R] = \mathrm{E}\left[\left|\sum_{g=1}^{G}\mathbf{h}_{RI,g}\boldsymbol{\Theta}_g\mathbf{h}_{IT,g}\right|^2\right] = \mathrm{E}\left[\left|\sum_{g=1}^{G}\sum_{n_G=1}^{N_G}\left[\bar{\mathbf{h}}_{RI,g}\right]_{n_G}[\mathbf{d}_g]_{n_G}\left[\bar{\mathbf{h}}_{IT,g}\right]_{n_G}\right|^2\right], \quad (7.30)$$

where we considered $\boldsymbol{\Theta}_g = \mathbf{V}_g\mathbf{D}_g\mathbf{V}_g^T$ and introduced $\bar{\mathbf{h}}_{RI,g} = \mathbf{h}_{RI,g}\mathbf{V}_g$ and $\bar{\mathbf{h}}_{IT,g} = \mathbf{V}_g^T\mathbf{h}_{IT,g}$. Recalling that $[\mathbf{d}_g]_{n_G} = -1$ if $n_G \geq 2$, we have

$$\begin{aligned}
\mathrm{E}[P_R] &= \mathrm{E}\left[\left|\left(\sum_{g=1}^{G}\left[\bar{\mathbf{h}}_{RI,g}\right]_1[\mathbf{d}_g]_1\left[\bar{\mathbf{h}}_{IT,g}\right]_1\right) - \left(\sum_{g=1}^{G}\sum_{n_G=2}^{N_G}\left[\bar{\mathbf{h}}_{RI,g}\right]_{n_G}\left[\bar{\mathbf{h}}_{IT,g}\right]_{n_G}\right)\right|^2\right] && (7.31)\\
&> \mathrm{E}\left[\left|\sum_{g=1}^{G}\left[\bar{\mathbf{h}}_{RI,g}\right]_1[\mathbf{d}_g]_1\left[\bar{\mathbf{h}}_{IT,g}\right]_1\right|^2\right] && (7.32)\\
&= \mathrm{E}\left[\left|\sum_{g=1}^{G}\left|\left[\bar{\mathbf{h}}_{RI,g}\right]_1\right|\left|\left[\bar{\mathbf{h}}_{IT,g}\right]_1\right|e^{j\left[\arg([\mathbf{d}_g]_1) - \theta_g^\star\right]}\right|^2\right], && (7.33)
\end{aligned}$$

where we introduced $\theta_g^\star = -\arg([\bar{\mathbf{h}}_{RI,g}]_1) - \arg([\bar{\mathbf{h}}_{IT,g}]_1)$. We denote as $\theta_1 = \arg\left(\frac{j[\boldsymbol{\lambda}_g^+]_1 - Z_0}{j[\boldsymbol{\lambda}_g^+]_1 + Z_0}\right) \in (0, \pi)$, where $[\boldsymbol{\lambda}_g^+]_1$ is the non zero eigenvalue of $\mathbf{X}_{I,g}^+ = +c_1\text{ones}(N_G)$. Thus, we have $-\theta_1 = \arg\left(\frac{j[\boldsymbol{\lambda}_g^-]_1 - Z_0}{j[\boldsymbol{\lambda}_g^-]_1 + Z_0}\right) \in (-\pi, 0)$, where $[\boldsymbol{\lambda}_g^-]_1 = -[\boldsymbol{\lambda}_g^+]_1$ is the non zero eigenvalue of $\mathbf{X}_{I,g}^- = -c_1\text{ones}(N_G)$. We consider a quantization approach giving $\arg([\mathbf{d}_g]_1) = \theta_1$ if $\theta_g^\star \in [0, \pi)$, or $\arg([\mathbf{d}_g]_1) = -\theta_1$ otherwise. The quantization error is denoted as

$\bar{\theta}_g = \arg([\mathbf{d}_g]_1) - \theta_g^\star$. Thus, (7.33) can be rewritten as

$$\mathrm{E}[P_R] > \mathrm{E}\Bigg[\sum_{g=1}^{G} \left|\left[\bar{\mathbf{h}}_{RI,g}\right]_1\right|^2 \left|\left[\bar{\mathbf{h}}_{IT,g}\right]_1\right|^2 + \sum_{g\neq f} \left|\left[\bar{\mathbf{h}}_{RI,g}\right]_1\right| \left|\left[\bar{\mathbf{h}}_{IT,g}\right]_1\right| \left|\left[\bar{\mathbf{h}}_{RI,f}\right]_1\right| \left|\left[\bar{\mathbf{h}}_{IT,f}\right]_1\right| e^{j\bar{\theta}_g - j\bar{\theta}_f}\Bigg]. \quad (7.34)$$

Noting that $|[\bar{\mathbf{h}}_{RI,g}]_1|$, $|[\bar{\mathbf{h}}_{IT,g}]_1|$, and $e^{j\bar{\theta}_g}$ are independent with each other, with $\mathrm{E}\left[|[\bar{\mathbf{h}}_{RI,g}]_1|^2\right] = 1$, $\mathrm{E}\left[|[\bar{\mathbf{h}}_{RI,g}]_1|\right] = \sqrt{\pi}/2$, and $\mathrm{E}\left[e^{j\bar{\theta}_g}\right] = \mathrm{E}\left[e^{-j\bar{\theta}_g}\right] = \frac{2}{\pi}\sin(\theta_1)$, we have

$$\mathrm{E}[P_R] > G + G(G-1)\left(\frac{\sqrt{\pi}}{2}\right)^4 \left(\frac{2}{\pi}\sin(\theta_1)\right)^2 > N_I^2 \left(\frac{\sin(\theta_1)}{2N_G^2}\right)^2. \quad (7.35)$$

Since $\sin(\theta_1) > 0$, (7.35) proves that $\mathcal{C}$ can achieve an average received signal power growth of $\mathcal{O}\left(N_I^2\right)$ as N_I increases. □

According to Proposition 8, the entries of an RIS reactance matrix can be chosen from a codebook $\mathcal{C}$ with no degradation in the received signal power growth. This justifies the use of a symmetric codebook for our discrete-value design strategy.

7.3 Vector-Discrete Group-/Fully-Connected RIS

In this section, we propose a second discretization strategy based on vector quantization. To realize vector-discrete RISs, we assign B_V resolution bits to each reactance matrix block $\mathbf{X}_{I,g}$. Thus, we introduce the codebook $\mathcal{C}_V$ as a set of vectors

$$\mathcal{C}_V = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{2^{B_V}}\}. \quad (7.36)$$

where $\mathbf{c}_k \in \mathbb{R}^{\frac{N_G(N_G+1)}{2}}$, for $k = 1, \ldots, 2^{B_V}$. Here, $\mathbf{c}_k$ is a possible value that the upper triangular part of each block $\mathbf{X}_{I,g}$ can assume.

7.3.1 Offline Learning

Similarly to the scalar discretization case, $\mathcal{C}_V$ can be obtained with K-means clustering through an offline learning phase, with the difference that a multi-dimensional feature space with dimension $N_G(N_G+1)/2$ is now considered, and $K = 2^{B_V}$. To construct the codebook $\mathcal{C}_V$, let us consider the set $\mathcal{X}_I$ containing the optimal reactance matrices associated with the N_0 training channel triplets. To formalize the K-means clustering problem, we introduce the matrices $\mathbf{X}^{[n_0]} \in \mathbb{R}^{N_G(N_G+1)/2\times G}$ containing in their gth column the entries of the upper triangular part of $\mathbf{X}_{I,g}^{[n_0]}$, for $g = 1, \ldots, G$. Then, the N_0 matrices $\mathbf{X}^{[n_0]}$ are concatenated in $\mathbf{X}_0 \in \mathbb{R}^{N_G(N_G+1)/2\times N_0 G}$ as $\mathbf{X}_0 = [\mathbf{X}^{[1]}, \ldots, \mathbf{X}^{[N_0]}]$. In this way, the resulting K-means clustering problem writes as

$$\{\mathbf{c}_1^\star, \ldots, \mathbf{c}_K^\star\} = \underset{\{\mathbf{c}_1,\ldots,\mathbf{c}_K\}}{\arg\min} \sum_{n=1}^{N_0 G} \sum_{k=1}^{2^{B_V}} r_{nk} \left|\mathbf{x}_n - \mathbf{c}_k\right|^2, \quad (7.37)$$

where the nth data point $\mathbf{x}_n \in \mathbb{R}^{N_G(N_G+1)/2\times 1}$ is the nth column of the matrix $\mathbf{X}_0$. Remarkably, (7.37) is in the form of a K-means clustering problem. Thus, we solve (7.37) by alternatively optimizing the sets $\{r_{nk}\}$ and $\{\mathbf{c}_1,\ldots,\mathbf{c}_K\}$, as described in the following. Firstly, with fixed $\{\mathbf{c}_1,\ldots,\mathbf{c}_K\}$ are fixed, we set $\{r_{nk}\}$ by assigning the nth data point $\mathbf{x}_n$ to the closest cluster center $\mathbf{c}_k$, yielding

$$r_{nk} = \begin{cases} 1 & \text{if } k = \arg\min_j \|\mathbf{x}_n - \mathbf{c}_j\|^2 \\ 0 & \text{otherwise} \end{cases}, \tag{7.38}$$

for $n = 1,\ldots,N_0G$ and $k = 1,\ldots,K$. Secondly, with fixed $\{r_{nk}\}$, the cluster centers $\{\mathbf{c}_1,\ldots,\mathbf{c}_K\}$ are optimally updated in closed-form as

$$\mathbf{c}_k = \frac{\sum_n r_{nk}\mathbf{x}_n}{\sum_n r_{nk}}, \tag{7.39}$$

for $k = 1,\ldots,K$. These two steps are alternatively repeated until convergence. Finally, the vector-quantization codebook is given by $\mathcal{C}_V = \{\mathbf{c}_1^\star,\ldots,\mathbf{c}_K^\star\}$.

The complexity of the offline learning for vector-discrete group- and fully-connected RISs is given by $\mathcal{O}(N_0(N_I(N_G+1)/2)^2 + N_0N_I(N_G+1)2^{B_V-1})$, where the first term is due to solving (7.4)-(7.6) N_0 times to find $\mathcal{X}_I$, and the second term is due to the K-means clustering algorithm used to solve (7.37).

7.3.2 Online Deployment

Once the codebook has been learned offline, the online optimization problem is given by

$$\max_{\mathbf{X}_{I,g}} \ \|\mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\|^2 \tag{7.40}$$

$$\text{s.t.} \ \boldsymbol{\Theta} = \mathrm{diag}\left(\boldsymbol{\Theta}_1,\ldots,\boldsymbol{\Theta}_G\right), \tag{7.41}$$

$$\boldsymbol{\Theta}_g = \left(j\mathbf{X}_{I,g} + Z_0\mathbf{I}\right)^{-1}\left(j\mathbf{X}_{I,g} - Z_0\mathbf{I}\right), \ \mathbf{X}_{I,g} = \mathbf{X}_{I,g}^T, \ \forall g, \tag{7.42}$$

$$\mathbf{X}_{I,g} \in \left\{\mathbf{c}_1^\star,\ldots,\mathbf{c}_{2^{B_V}}^\star\right\}, \ \forall g. \tag{7.43}$$

Alternating optimization is the selected strategy to solve this problem, in which the G blocks $\mathbf{X}_{I,g}$ are iteratively optimized by searching among their possible 2^{B_V} values contained in $\mathcal{C}_V$. The convergence of this iterative process can be proved as done for scalar-discrete RISs in Section III E. The proposed design of vector-discrete RISs is summarized in Alg. 7, where problem (7.37) is solved offline to design the codebook while (7.40)-(7.43) is solved online to optimize the RIS reactance matrix.

The cost of vector-discrete RISs relies on their optimization complexity. For example, if B bits are allocated to each reactance element, the number of clusters needed to quantize a vector of $N_G\left(N_G+1\right)/2$ elements is $k = 2^{B\frac{N_G(N_G+1)}{2}}$, growing exponentially with the square of N_G. Consequently, the number of search times in a complete iteration of the alternating optimization algorithm becomes $G2^{B\frac{N_G(N_G+1)}{2}}$. However, this discretization strategy based on vector quantization brings two benefits. First, for the same number of total resolution bits, vector quantization is inherently more efficient than scalar quanti-

Algorithm 7 Vector-Discrete RISs Design

Input: B_V, $\mathcal{H}_0$, $\mathbf{H}_{RT}$, $\mathbf{H}_{RI}$, $\mathbf{H}_{IT}$, N_G.
Output: $\mathcal{C}$, $\mathbf{X}_I$.

```
Offline Learning
Calculate 𝒳_I by (7.4)-(7.6).
Calculate {c*_1, ..., c*_K} by iteratively solving (7.37), with K = 2^{B_V}.
𝒞 ← {c*_1, c*_2, ..., c*_K}.
Online Deployment
Initialize X_{I,g} = 0, ∀g.
while no convergence of objective (7.40) do
  for g ← 1 to G do
    Update X_{I,g} ∈ 𝒞_V by searching among the 2^{B_V} possible values.
  end for
end while
```

zation. Second, the number of total resolution bits can be chosen with more degrees of freedom, since it is no longer limited to B bits for each reactance element.

7.4 Performance Evaluation

In this section, we evaluate the performance of an RIS-aided MIMO system with different discrete RIS configurations, each characterized by its number of elements N_I, number of resolution bits, and group size N_G. The performance is measured in terms of received signal power, given by $P_R = P_T \|\mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}\|^2$.

Let us consider a three-dimensional coordinate system (x, y, z), in which the z-axis represents the height above the ground in meters (m). The transmitter, the RIS, and the receiver are ULAs composed of N_T, N_I, and N_R antennas, respectively, with half wavelength antenna spacing[2]. We set $N_T = 4$ and $N_R = 2$, while considering several values of N_I as specified in the following. The transmitter is located at $(5, -250, 25)$ and its antennas are arranged along the x-axis. The RIS is located at $(0, 0, 5)$ and its antennas are arranged along the y-axis. Finally, the receiver is located at $(5, 5, 1.5)$ with random orientation. The path loss is modeled as $L_{ij}(d_{ij}) = L_0 d_{ij}^{-\alpha_{ij}}$, where L_0 is the path loss at distance 1 m, d_{ij} is the distance, and α_{ij} is the path loss exponent for $ij \in \{RT, RI, IT\}$. We set $L_0 = -30$ dB, $\alpha_{RT} = 4$, $\alpha_{RI} = 2.8$, $\alpha_{IT} = 2$, and $P_T = 10$ W.

We analyze two scenarios with different small-scale fading effects: i.i.d. Rayleigh fading and correlated fading. In the former scenario, the small-scale fading of all channels is assumed to be i.i.d. Rayleigh distributed. Thus, we have $\mathbf{H}_{RT} \sim \mathcal{CN}(\mathbf{0}, L_{RT}\mathbf{I})$, $\mathbf{H}_{RI} \sim \mathcal{CN}(\mathbf{0}, L_{RI}\mathbf{I})$ and $\mathbf{H}_{IT} \sim \mathcal{CN}(\mathbf{0}, L_{IT}\mathbf{I})$. In the latter scenario, we generate the small-scale effects with QuaDRiGa version 2.4, a MATLAB based statistical ray-tracing channel simulator [90]. An urban macrocell propagation environment is simulated, with LoS in the transmitter-RIS link, and with NLoS in the transmitter-receiver and RIS-receiver links. The channel models "3GPP_38.901_UMa_LOS" and "3GPP_38.901_UMa_NLOS"

[2] Our discrete design for BD-RIS is valid for any RIS array shape.

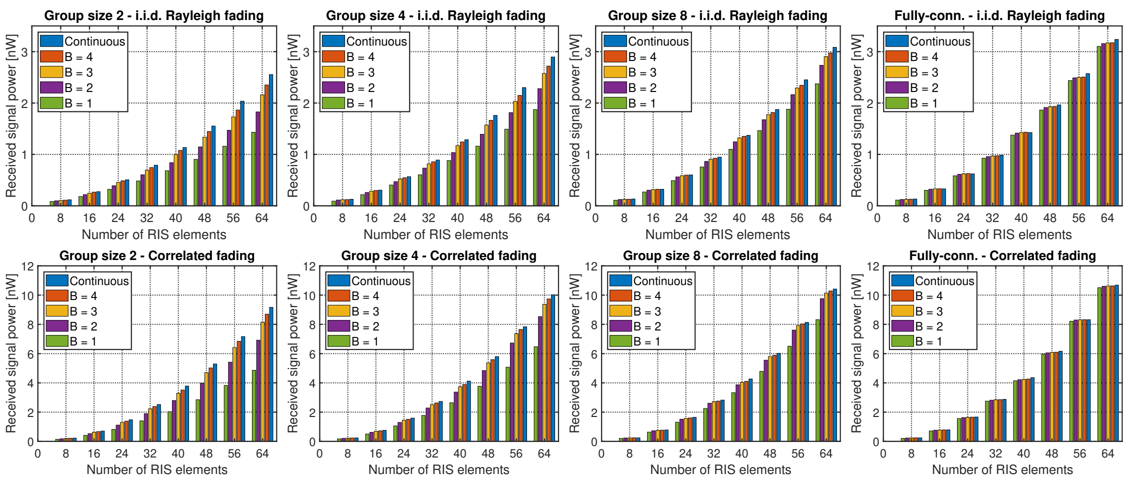


Figure 7.1: Average received signal power versus the number of RIS elements for different values of group size.

have been used to simulate the LoS and NLoS channels, respectively. The ULAs at the transmitter, the RIS, and the receiver are modeled as “3gpp-3d” array and the carrier frequency is set to $f_c = 2.5$ GHz.

For both scenarios, offline learning stages have been carried out to define the optimal codebooks. We employ a training set $\mathcal{H}_0$ composed of $N_0 = 100$ channel realization triplets for scalar-discrete RISs, and $N_0 = 500$ triplets for vector-discrete RISs. These values of N_0 have been set such that in the K-means clustering problems involved in Alg. 6 and Alg. 7, the number of training samples is always higher than the number of clusters. In the alternating optimizations, convergence is considered reached when the fractional increase of the objective value is below $\epsilon = 10^{-3}$. Finally, the average received signal power has been computed for each RIS configuration using the Monte Carlo method.

Before comparing discrete-value group- and fully-connected RISs with discrete-value single-connected RISs, we briefly review how single-connected RISs have been discretized in the literature. In single-connected RISs, $\mathbf{\Theta}$ is completely described by the phase shifts θ_{n_I}. The phase shifts have been typically discretized uniformly in $[0, 2\pi)$. By assigning B resolution bits per phase shift, they are chosen in the codebook

$$\mathcal{C}^{\text{Single}} = \left\{0, \delta, \ldots, \left(2^B - 1\right)\delta\right\}, \tag{7.44}$$

where $\delta = \frac{2\pi}{2^B}$ is the quantization step. Thus, the received signal power maximization problem is formulated in the discrete case as

$$\max_{\theta_{n_I}} \quad \left\|\mathbf{H}_{RT} + \mathbf{H}_{RI}\mathbf{\Theta}\mathbf{H}_{IT}\right\|^2 \tag{7.45}$$

$$\text{s.t.} \quad \mathbf{\Theta} = \text{diag}\left(e^{j\theta_1}, \ldots, e^{j\theta_{N_I}}\right), \; \theta_{n_I} \in \left\{0, \delta, \ldots, \left(2^B - 1\right)\delta\right\}, \; \forall n_I. \tag{7.46}$$

To solve this problem, we employ the widely adopted alternating optimization method [38]. In this method, each of the N_I phase shifts is optimized individually by searching among all the possible 2^B values, while fixing the other $N_I - 1$ phase shifts. This procedure is repeated for all N_I phase shifts, iterating multiple times until the convergence is reached,

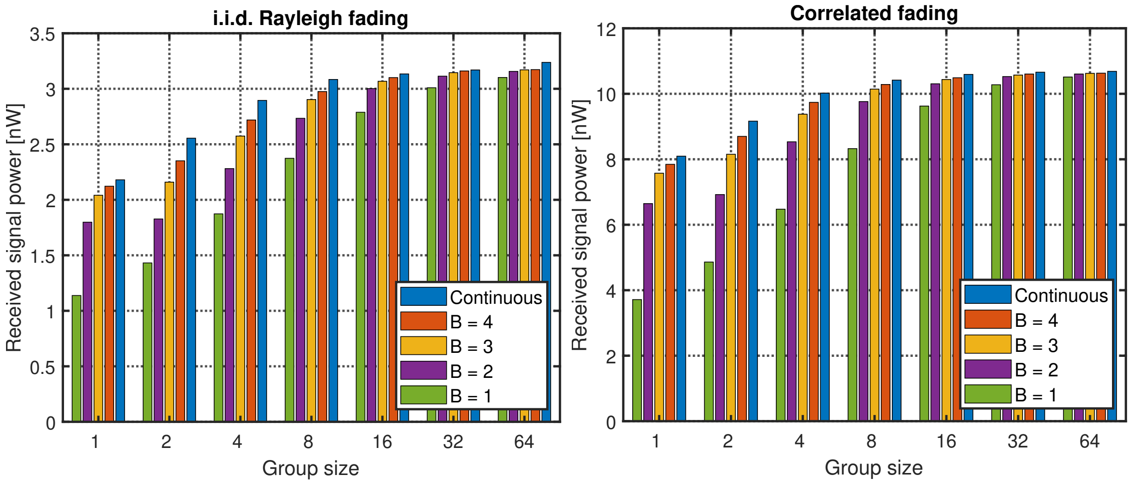


Figure 7.2: Average received signal power versus the group size, with $N_I = 64$.

i.e., until the fractional increase of the objective value is below a certain parameter ϵ.

We first evaluate the received signal power in the case of scalar-discrete RISs. In Fig. 7.1, we show the received signal power when group-connected (with group size $N_G \in \{2, 4, 8\}$) and fully-connected RISs are used in the presence of i.i.d. Rayleigh fading channels and correlated channels. The performance of scalar-discrete RISs is compared with the performance of continuous-value RISs, optimized by solving (7.4)-(7.6). We observe that the higher the group size, the fewer resolution bits are necessary to reach the performance of continuous-value RISs. The rationale behind this behavior is given by the necessary and sufficient condition for $\boldsymbol{\Theta}$ to achieve the maximum P_R with no direct link. By extending the analysis in [16] to MIMO settings, it is possible to prove that this condition is given by

$$\mathbf{u}_{RI,g} = \boldsymbol{\Theta}_g \mathbf{u}_{IT,g}, \forall g, \tag{7.47}$$

where $\mathbf{u}_{RI} = [\mathbf{u}_{RI,1}, \mathbf{u}_{RI,2}, \ldots, \mathbf{u}_{RI,G}]$ with $\mathbf{u}_{RI,g} \in \mathbb{C}^{N_G \times 1}$, and $\mathbf{u}_{IT} = [\mathbf{u}_{IT,1}, \mathbf{u}_{IT,2}, \ldots, \mathbf{u}_{IT,G}]^T$ with $\mathbf{u}_{IT,g} \in \mathbb{C}^{N_G \times 1}$, are the dominant left singular vectors of $\mathbf{H}_{RI}^H$ and $\mathbf{H}_{IT}$, respectively[3]. This condition consists of an underdetermined system of N_I equations in $N_I (N_G + 1)/2$ unknowns. Thus, a higher group size N_G implies more degrees of freedom, and fewer resolution bits are required to satisfy it. In particular, in the fully-connected architecture, a single resolution bit is sufficient to obtain approximately the same performance as the optimized continuous-value RISs.

We now compare the performance obtained with different group sizes when the number of RIS elements N_I is fixed. To this end, we optimize RISs with $N_I = 64$ elements considering all the possible group sizes, spanning from the single-connected architecture ($N_G = 1$) to the fully-connected ($N_G = 64$). In Fig. 7.2, the upper bound of the received signal power is reported together with the power maximized with different resolutions. We notice that the received signal power increases with the group size for every B con-

[3]Condition (7.47) is derived by considering $P_R \leq P_T \max_{\|\mathbf{x}\|=1} \|\mathbf{H}_{RI}\|^2 \|\boldsymbol{\Theta}\mathbf{H}_{IT}\mathbf{x}\|^2 \leq P_T \max_{\|\mathbf{x}\|=1} \|\mathbf{H}_{RI}\|^2 \|\boldsymbol{\Theta}\mathbf{H}_{IT}\|^2 \|\mathbf{x}\|^2$. Note that the equality holds in the two inequalities when $\mathbf{u}_{RI}$ is equal to the dominant left singular vector of $\boldsymbol{\Theta}\mathbf{H}_{IT}$, i.e., $\boldsymbol{\Theta}\mathbf{u}_{IT}$, which is equivalent to (7.47) in the case of block diagonal $\boldsymbol{\Theta}$.

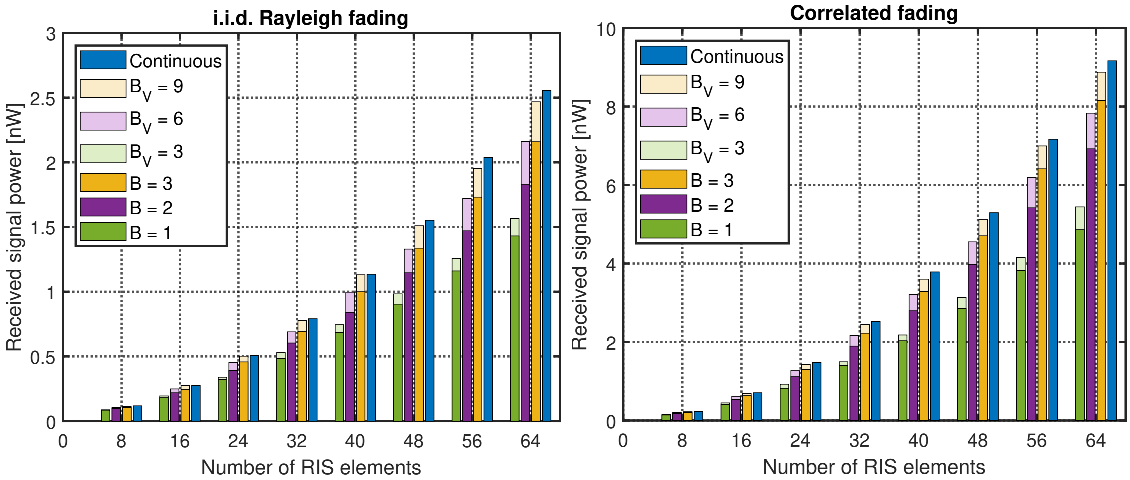


Figure 7.3: Average received signal power versus the number of RIS elements in the scalar-discrete and vector-discrete group-connected architectures, with $N_G = 2$.

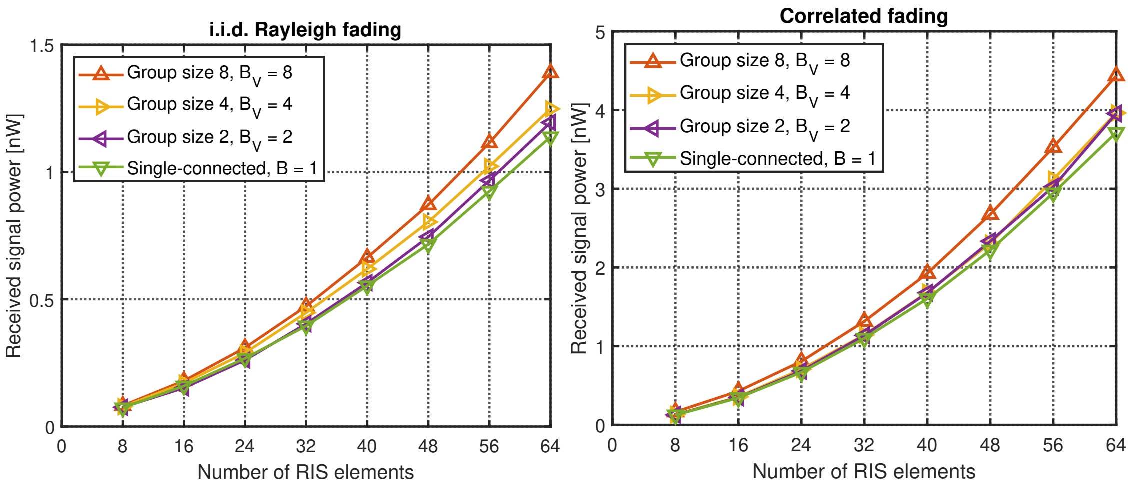


Figure 7.4: Average received signal power versus the number of RIS elements for different group sizes, with N_I total resolution bits.

sidered. Furthermore, the performance of discrete-value RISs with $B = 1$ approaches the performance of continuous-value RISs as the group size increases, converging to it in the fully-connected architecture.

Let us analyze the performance of vector-discrete RISs, in comparison with scalar-discrete RISs. To this end, we set the number of resolution bits assigned to each block $\mathbf{X}_{I,g}$ to $B_V = B\frac{N_G(N_G+1)}{2}$. This is equivalent, in terms of total bits employed, to assigning B bits to each reactance element, allowing a comparison between scalar-discrete and vector-discrete RISs. Fig. 7.3 shows the performance achieved by vector-discrete group-connected RISs with group size $N_G = 2$, when the number of resolution bits considered is $B_V \in \{3, 6, 9\}$. In terms of total resolution bits employed, $B_V \in \{3, 6, 9\}$ bits per impedance network block are equivalent to $B \in \{1, 2, 3\}$ bits per impedance element, respectively. Thus, for the sake of comparison, we report in Fig. 7.3 also the performance of scalar-discrete group-connected RISs with $B \in \{1, 2, 3\}$. We notice that vector discretization brings a significant performance improvement over scalar discretization. In particular, $B_V = 9$ resolution bits per impedance network block are sufficient to approach

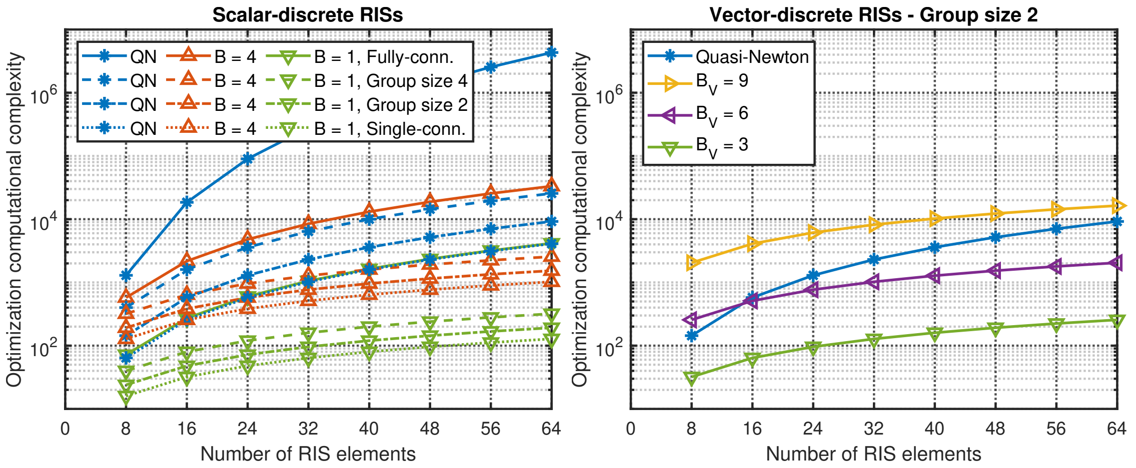


Figure 7.5: Optimization computational complexity versus the number of RIS elements. The quasi-Newton method "QN" is compared with alternating optimization for scalar-discrete RISs (on the left) and vector-discrete RISs (on the right).

the performance of continuous-value RISs. However, $B = 3$ resolution bits per impedance element are not sufficient to reach the performance of continuous-value RISs when scalar discretization is considered.

We now compare the performance of RISs with different group sizes when the same number of total bits is employed. To this end, we fix the number of total bits used to N_I, which are required by the single-connected architecture with $B = 1$. To use N_I total bits, the feature space having $N_G(N_G+1)/2$ dimensions is quantized with $B_V = N_G$ resolution bits. Fig. 7.4 shows the received signal power achieved with different values of group size. Here, we can notice that higher group sizes achieve better performance than lower group sizes. This is due to the optimality condition (7.47), which has more degrees of freedom as the group size increases. Thus, it can be more easily satisfied when higher group sizes are considered.

To conclude our discussion, we compare continuous-value RISs with discrete-value RISs in terms of optimization computational complexity. On the one hand, when the quasi-Newton method is used to optimize the continuous values of $\mathbf{X}_I$, the computational complexity of each iteration is $\mathcal{O}((N_I(N_G+1)/2)^2)$ [16]. On the other hand, when alternating optimization is used to optimize the discrete values of $\mathbf{X}_I$, we distinguish two cases, namely scalar-discrete RISs and vector-discrete RISs. In the case of scalar-discrete RISs, the total number of search times in a complete iteration of the alternating optimization algorithm is $G\frac{N_G(N_G+1)}{2}2^B$ in group-connected architectures. This number boils down to $N_I 2^B$ in single-connected and to $\frac{N_I(N_I+1)}{2}2^B$ in fully-connected architectures. Instead, in the case of vector-discrete RISs, the number of search times per iteration becomes $\frac{N_I}{N_G}2^{B_V}$. The values of these computational complexities are reported in Fig. 7.5. In the case of scalar-discrete RISs, we observe that our RIS design strategy based on discrete-value impedance networks is beneficial also in terms of optimization computational complexity. Even with $B = 4$, the complexity is far less than the complexity of the quasi-Newton method for medium-high numbers of RIS elements.

7.5 Conclusion

We propose a novel BD-RIS design based on discrete-value group- and fully-connected architectures. Our strategy is composed of two stages. Firstly, through the offline learning stage, we build the codebook containing the possible values for the reconfigurable impedances. Secondly, during the online deployment stage, we optimize the reconfigurable impedances with alternating optimization on the basis of the designed codebook. Two different approaches are developed, namely scalar-discrete and vector-discrete RISs, based on scalar and vector quantization, respectively. Vector-discrete RISs achieve higher performance than scalar-discrete RISs at the cost of increased optimization computational complexity. Numerical results show that a single resolution bit per reconfigurable impedance is sufficient to achieve optimality in fully-connected RISs. In a practical scenario, this simplifies significantly the hardware complexity of the fully-connected RIS.

Chapter 8

BD-RIS With Imperfect Matching and Mutual Coupling

While most of the research on conventional RIS and BD-RIS has primarily focused on system-level RIS optimization and on RIS architecture development, only a few works have been devoted to model and analyze RIS-aided communication systems accounting for the EM properties of the RIS [91]. With a focus on RIS modeled as an antenna array connected to a reconfigurable impedance network, three equivalent analyses are available to show the role of an RIS in wireless networks from different perspectives. First, in [16], an RIS-aided communication model has been derived using multiport network analysis based on scattering parameters. Second, in [92], another RIS-aided communication model has been developed by using the impedance parameters. Interestingly, it has been shown in [93] that the scattering parameters analysis of [16] and the impedance parameters analysis of [92] lead to the same conclusion under the assumptions of perfect matching and no mutual coupling. Third, another equivalent analysis based on admittance parameters has been proposed in [80] to model BD-RIS with sparse interconnections among the RIS elements. However, there is currently no universal framework for modeling RIS-aided systems according to impedance, admittance, and scattering parameters and clarifying the meaning of the different components in each model. Such a universal framework would improve our understanding of the different EM-compliant models for RIS-aided systems, and would provide additional tools to perform system level RIS optimization. Furthermore, it would enable a thorough analysis of the impact of each approximation and assumption necessary to simplify existing RIS-aided communication models based on the different parameters.

Motivated by the above considerations, in this chapter, we depart from existing works on EM-compliant RIS-aided communication modeling [16, 92, 93] and derive a universal framework to analyze RIS-aided communication systems. This universal framework can be used to carry out a rigorous multiport network analysis of RIS-aided communication systems based on impedance, admittance, and scattering parameters, also denoted as Z-, Y-, and S-parameters, respectively. As a result of these analyses, we can derive three RIS-aided communication models accounting for mutual coupling effects and impedance mismatching at the transmitter, RIS, and receiver. Our universal framework shows the equivalence between channel models based on Z-, Y-, and S-parameters also under different assumptions and approximations. This is achieved by providing the mappings between the terms of the models based on the three different parameters. The contributions of this chapter are summarized as follows.

First, we propose a universal framework to derive the communication model of an RIS-aided system. We exploit the proposed universal framework to analyze RIS-aided communication systems based on the Z-, Y-, and S-parameters. Through these three independent analyses, we first derive three general channel models, accounting for the effects of impedance mismatching and mutual coupling at the transmitter, RIS, and receiver.

Second, since it is difficult to interpret the derived general channel models, we approximate them by considering the so-called unilateral approximation, applicable in the case of large transmission distances[1]. Interestingly, the derived simplified models have tractable expressions yet capture the effect of impedance mismatching and mutual coupling.

Third, we further simplify the channel models to get engineering insights into the role of the RIS by assuming that all the antennas are perfectly matched and there is no mutual coupling. We show that these models are consistent with the model widely accepted in related literature, although an additional approximation is commonly considered in the literature. We numerically quantify the impact of such an approximation and show that it vanishes as the number of RIS elements increases.

Fourth, we characterize conventional RIS and BD-RIS architectures by using Z-, Y-, and S-parameters. Since the three descriptions are equivalent, we discuss how to select the most suitable parameters to describe the RIS architecture according to the needs. While Z- and S-parameters have been widely used to characterize conventional RIS with and without mutual coupling, respectively, we show that Y-parameters are particularly useful for BD-RIS. In addition, we illustrate the advantages of the three parameters in optimizing conventional RIS and BD-RIS in different scenarios by studying three optimization problems.

Organization: In Section 8.1, we provide the multiport network analysis. In Section 8.2, we derive the general RIS-aided communication models based on Z-, Y-, and S-parameters. In Section 8.3, we simplify the derived models by using the unilateral approximation. In Section 8.4, we further simplify the models assuming perfect matching and no mutual coupling, and compare them with the widely used RIS-aided channel model. In Section 8.5, we discuss the advantages of using the different parameters in characterizing and optimizing an RIS. In Section 8.6, we evaluate the performance of an RIS-aided system obtained by solving the presented case studies. Finally, Section 8.7 concludes this chapter.

The content of this chapter has appeared in [95].

[1] The unilateral approximation consists in setting to zero the feedback channel between a transmitter and a receiver, i.e., the channel from the receiver to the transmitter, and holds when the electrical properties at the transmitter are independent of the electrical properties at the receiver [94].

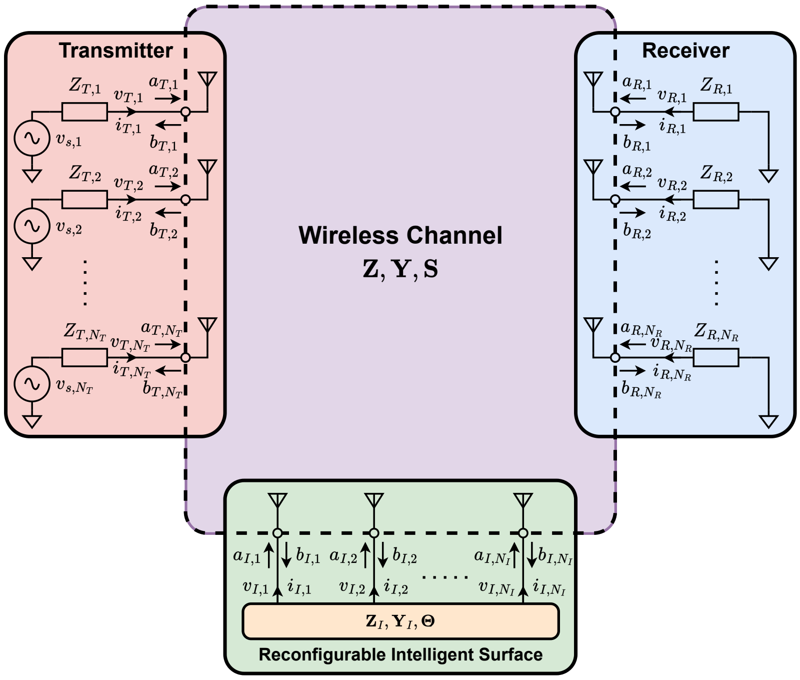


Figure 8.1: Multiport model of an RIS-aided communication system.

8.1 Multiport Network Analysis

We consider a MIMO communication system aided by an RIS, where there are N_T antennas at the transmitter, N_R antennas at the receiver, and N_I antennas at the RIS, as represented in Fig. 8.1. As in [16], we model the wireless channel as an N-port network, with $N = N_T + N_I + N_R$. According to the multiport network analysis [74, Chapter 4], this N-port network can be characterized by using its impedance, admittance, or scattering matrix, as described in the following.

8.1.1 Modeling Based on the Impedance Parameters

The N-port network modeling the wireless channel can be characterized by its impedance matrix $\mathbf{Z} \in \mathbb{C}^{N\times N}$, such that

$$\mathbf{v} = \mathbf{Z}\mathbf{i}, \tag{8.1}$$

where $\mathbf{v} \in \mathbb{C}^{N\times 1}$ and $\mathbf{i} \in \mathbb{C}^{N\times 1}$ are the voltages and currents at the N ports, respectively. We can partition $\mathbf{v}$, $\mathbf{i}$, and $\mathbf{Z}$ as

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_T \\ \mathbf{v}_I \\ \mathbf{v}_R \end{bmatrix}, \; \mathbf{i} = \begin{bmatrix} \mathbf{i}_T \\ \mathbf{i}_I \\ \mathbf{i}_R \end{bmatrix}, \; \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{TT} & \mathbf{Z}_{TI} & \mathbf{Z}_{TR} \\ \mathbf{Z}_{IT} & \mathbf{Z}_{II} & \mathbf{Z}_{IR} \\ \mathbf{Z}_{RT} & \mathbf{Z}_{RI} & \mathbf{Z}_{RR} \end{bmatrix}, \tag{8.2}$$

where $\mathbf{v}_i \in \mathbb{C}^{N_i\times 1}$ and $\mathbf{i}_i \in \mathbb{C}^{N_i\times 1}$ for $i \in \{T, I, R\}$ refer to the voltages and currents at the antennas of the transmitter, RIS, and receiver, respectively. Accordingly, $\mathbf{Z}_{TT} \in \mathbb{C}^{N_T\times N_T}$, $\mathbf{Z}_{II} \in \mathbb{C}^{N_I\times N_I}$, and $\mathbf{Z}_{RR} \in \mathbb{C}^{N_R\times N_R}$ refer to the impedance matrices of the antenna arrays at the transmitter, RIS, and receiver, respectively. The diagonal entries of $\mathbf{Z}_{TT}$, $\mathbf{Z}_{II}$, and $\mathbf{Z}_{RR}$ refer to the antenna self-impedance while the off-diagonal entries refer to antenna mutual coupling. $\mathbf{Z}_{RT} \in \mathbb{C}^{N_R \times N_T}$, $\mathbf{Z}_{IT} \in \mathbb{C}^{N_I \times N_T}$, and $\mathbf{Z}_{RI} \in \mathbb{C}^{N_R \times N_I}$ refer to the transmission impedance matrices from the transmitter to receiver, from the transmitter to the RIS, and from the RIS to the receiver, respectively.

At the transmitter, the n_Tth antenna is connected in series with a source voltage v_{s,n_T} and a source impedance Z_{T,n_T}, for $n_T = 1, \ldots, N_T$. Therefore, $\mathbf{v}_T$ and $\mathbf{i}_T$ are related by

$$\mathbf{v}_T = \mathbf{v}_{s,T} - \mathbf{Z}_T \mathbf{i}_T, \tag{8.3}$$

where $\mathbf{v}_{s,T} = [v_{s,1}, v_{s,2}, \ldots, v_{s,N_T}]^T \in \mathbb{C}^{N_T \times 1}$ refers to the source voltage vector and $\mathbf{Z}_T \in \mathbb{C}^{N_T \times N_T}$ is a diagonal matrix given by $\mathbf{Z}_T = \mathrm{diag}(Z_{T,1}, Z_{T,2}, \ldots, Z_{T,N_T})$. At the RIS, the N_I antennas are connected to an N_I-port reconfigurable impedance network and $\mathbf{v}_I$ is related to $\mathbf{i}_I$ by

$$\mathbf{v}_I = -\mathbf{Z}_I \mathbf{i}_I, \tag{8.4}$$

where $\mathbf{Z}_I \in \mathbb{C}^{N_I \times N_I}$ is the impedance matrix of the N_I-port reconfigurable impedance network. At the receiver, the n_Rth antenna is connected in series with a load impedance Z_{R,n_R}, for $n_R = 1, \ldots, N_R$. Therefore, $\mathbf{v}_R$ and $\mathbf{i}_R$ are related by

$$\mathbf{v}_R = -\mathbf{Z}_R \mathbf{i}_R, \tag{8.5}$$

where $\mathbf{Z}_R \in \mathbb{C}^{N_R \times N_R}$ is a diagonal matrix given by $\mathbf{Z}_R = \mathrm{diag}(Z_{R,1}, Z_{R,2}, \ldots, Z_{R,N_R})$.

8.1.2 Modeling Based on the Admittance Parameters

The wireless channel can also be characterized by its admittance matrix $\mathbf{Y} \in \mathbb{C}^{N \times N}$, with $\mathbf{Y} = \mathbf{Z}^{-1}$, such that

$$\mathbf{i} = \mathbf{Y}\mathbf{v}. \tag{8.6}$$

Similar to $\mathbf{Z}$, also $\mathbf{Y}$ can be partitioned as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{TT} & \mathbf{Y}_{TI} & \mathbf{Y}_{TR} \\ \mathbf{Y}_{IT} & \mathbf{Y}_{II} & \mathbf{Y}_{IR} \\ \mathbf{Y}_{RT} & \mathbf{Y}_{RI} & \mathbf{Y}_{RR} \end{bmatrix}, \tag{8.7}$$

where $\mathbf{Y}_{ij} \in \mathbb{C}^{N_i \times N_j}$ for $i, j \in \{T, I, R\}$.

At the transmitter, $\mathbf{v}_T$ and $\mathbf{i}_T$ are related by

$$\mathbf{i}_T = \mathbf{i}_{s,T} - \mathbf{Y}_T \mathbf{v}_T, \tag{8.8}$$

where $\mathbf{i}_{s,T} \in \mathbb{C}^{N_T \times 1}$ refers to the source current vector and $\mathbf{Y}_T \in \mathbb{C}^{N_T \times N_T}$ is a diagonal matrix given by $\mathbf{Y}_T = \mathbf{Z}_T^{-1}$. At the RIS, $\mathbf{v}_I$ and $\mathbf{i}_I$ are related by

$$\mathbf{i}_I = -\mathbf{Y}_I \mathbf{v}_I, \tag{8.9}$$

where $\mathbf{Y}_I \in \mathbb{C}^{N_I \times N_I}$ is the admittance matrix of the N_I-port reconfigurable impedance network, given by $\mathbf{Y}_I = \mathbf{Z}_I^{-1}$. At the receiver, $\mathbf{v}_R$ and $\mathbf{i}_R$ are related by

$$\mathbf{i}_R = -\mathbf{Y}_R\mathbf{v}_R, \tag{8.10}$$

where $\mathbf{Y}_R \in \mathbb{C}^{N_R\times N_R}$ is diagonal given by $\mathbf{Y}_R = \mathbf{Z}_R^{-1}$.

8.1.3 Modeling Based on the Scattering Parameters

Finally, the wireless channel can be characterized by its scattering matrix $\mathbf{S} \in \mathbb{C}^{N\times N}$, so that we have

$$\mathbf{b} = \mathbf{S}\mathbf{a}, \tag{8.11}$$

where $\mathbf{a} \in \mathbb{C}^{N\times 1}$ and $\mathbf{b} \in \mathbb{C}^{N\times 1}$ are the incident and reflected waves at the ports, respectively. We partition $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{S}$ as

$$\mathbf{a} = \begin{bmatrix}\mathbf{a}_T\\ \mathbf{a}_I\\ \mathbf{a}_R\end{bmatrix},\ \mathbf{b} = \begin{bmatrix}\mathbf{b}_T\\ \mathbf{b}_I\\ \mathbf{b}_R\end{bmatrix},\ \mathbf{S} = \begin{bmatrix}\mathbf{S}_{TT} & \mathbf{S}_{TI} & \mathbf{S}_{TR}\\ \mathbf{S}_{IT} & \mathbf{S}_{II} & \mathbf{S}_{IR}\\ \mathbf{S}_{RT} & \mathbf{S}_{RI} & \mathbf{S}_{RR}\end{bmatrix}, \tag{8.12}$$

where $\mathbf{a}_i \in \mathbb{C}^{N_i\times 1}$ and $\mathbf{b}_i \in \mathbb{C}^{N_i\times 1}$ for $i \in \{T, I, R\}$ refer to the incident and reflected waves at the antennas of the transmitter, RIS, and receiver, respectively, and $\mathbf{S}_{ij} \in \mathbb{C}^{N_i\times N_j}$ for $i, j \in \{T, I, R\}$. Remarkably, according to [74, Chapter 4], the vectors $\mathbf{v}$ and $\mathbf{i}$ are related to $\mathbf{a}$ and $\mathbf{b}$ though

$$\mathbf{v} = \mathbf{a} + \mathbf{b},\ \mathbf{i} = \frac{\mathbf{a} - \mathbf{b}}{Z_0} = Y_0\left(\mathbf{a} - \mathbf{b}\right), \tag{8.13}$$

respectively, where Z_0 is the characteristic impedance used to compute the S-parameters, typically set as $Z_0 = 50\ \Omega$, and $Y_0 = 1/Z_0$ is the characteristic admittance. In addition, the matrix $\mathbf{S}$ can be expressed as a function of $\mathbf{Z}$ as

$$\mathbf{S} = \left(\mathbf{Z} + Z_0\mathbf{I}\right)^{-1}\left(\mathbf{Z} - Z_0\mathbf{I}\right). \tag{8.14}$$

At the transmitter, $\mathbf{a}_T$ and $\mathbf{b}_T$ are related by

$$\mathbf{a}_T = \mathbf{b}_{s,T} + \mathbf{\Gamma}_T\mathbf{b}_T, \tag{8.15}$$

where $\mathbf{b}_{s,T} \in \mathbb{C}^{N_T\times 1}$ refers to the source wave vector and $\mathbf{\Gamma}_T \in \mathbb{C}^{N_T\times N_T}$ is diagonal with its (n_T, n_T)th element being the reflection coefficient of the n_Tth source impedance, i.e.,

$$\mathbf{\Gamma}_T = \left(\mathbf{Z}_T + Z_0\mathbf{I}\right)^{-1}\left(\mathbf{Z}_T - Z_0\mathbf{I}\right). \tag{8.16}$$

At the RIS, $\mathbf{a}_I$ and $\mathbf{b}_I$ are related by

$$\mathbf{a}_I = \mathbf{\Theta}\mathbf{b}_I, \tag{8.17}$$

where $\mathbf{\Theta} \in \mathbb{C}^{N_I\times N_I}$ denotes the scattering matrix of the N_I-port reconfigurable impedance network, written as

$$\mathbf{\Theta} = \left(\mathbf{Z}_I + Z_0\mathbf{I}\right)^{-1}\left(\mathbf{Z}_I - Z_0\mathbf{I}\right). \tag{8.18}$$

At the receiver, $\mathbf{a}_R$ and $\mathbf{b}_R$ are related by

$$\mathbf{a}_R = \mathbf{\Gamma}_R \mathbf{b}_R, \tag{8.19}$$

where $\mathbf{\Gamma}_R \in \mathbb{C}^{N_R \times N_R}$ is diagonal with its (n_R, n_R)th element being the reflection coefficient of the n_Rth load, i.e.,

$$\mathbf{\Gamma}_R = \left(\mathbf{Z}_R + Z_0 \mathbf{I}\right)^{-1} \left(\mathbf{Z}_R - Z_0 \mathbf{I}\right). \tag{8.20}$$

8.2 General RIS-Aided Communication Model

We have characterized the relationships between the electrical properties at the transmitter, RIS, and receiver of an RIS-aided communication system. In this section, we determine the channel $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ relating the voltage vector at the transmitter (transmitted signal) $\mathbf{v}_T$ and the voltage vector at the receiver (received signal) $\mathbf{v}_R$ through $\mathbf{v}_R = \mathbf{H}\mathbf{v}_T$. To this end, we introduce a universal framework enabling three equivalent analyses, based on Z-, Y-, and S-parameters.

8.2.1 Universal Framework

Consider the following mathematical problem. Given two variable vectors $\mathbf{x} \in \mathbb{C}^{N \times 1}$ and $\mathbf{y} \in \mathbb{C}^{N \times 1}$, and a constant matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$, respectively partitioned as

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{bmatrix}, \ \mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}, \tag{8.21}$$

where $\mathbf{x}_i \in \mathbb{C}^{N_i \times 1}$, $\mathbf{y}_i \in \mathbb{C}^{N_i \times 1}$, $\mathbf{A}_{ij} \in \mathbb{C}^{N_i \times N_j}$ for $i, j \in \{1, 2, 3\}$, and $N_1 + N_2 + N_3 = N$, our goal is to solve

$$\begin{cases} \mathbf{y} = \mathbf{A}\mathbf{x}, \\ \mathbf{x}_1 = \mathbf{c}_1 + \mathbf{A}_1 \mathbf{y}_1, \\ \mathbf{x}_2 = \mathbf{A}_2 \mathbf{y}_2, \\ \mathbf{x}_3 = \mathbf{A}_3 \mathbf{y}_3, \end{cases} \tag{8.22}$$

where $\mathbf{c}_1 \in \mathbb{C}^{N_1 \times 1}$ and $\mathbf{A}_i \in \mathbb{C}^{N_i \times N_i}$ for $i \in \{1, 2, 3\}$ are constant. In other words, we want to characterize the two variable vectors $\mathbf{x}$ and $\mathbf{y}$ as functions of the constants $\mathbf{A}$, $\mathbf{c}_1$, $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$. Remarkably, system (8.22) provides a universal framework that can be used to describe the equations of the multiport network analysis based on Z-, Y-, and S-parameters. In Tab. 8.1, we report the relationship between the variables in this framework and the quantities in the three different parameters introduced in Section 8.1.

To solve system (8.22), we rewrite it in a compact form

$$\begin{cases} \mathbf{y} = \mathbf{A}\mathbf{x}, \\ \mathbf{x} = \mathbf{c} + \overline{\mathbf{A}}\mathbf{y}, \end{cases} \tag{8.23}$$

Table 8.1: Relationship between the universal framework and the three parameters.

Framework	Z-parameters	Y-parameters	S-parameters
$\mathbf{x}$	$\mathbf{v}$	$\mathbf{i}$	$\mathbf{a}$
$\mathbf{y}$	$\mathbf{i}$	$\mathbf{v}$	$\mathbf{b}$
$\mathbf{A}$	$\mathbf{Z}^{-1}$	$\mathbf{Y}^{-1}$	$\mathbf{S}$
$\mathbf{c}_1$	$\mathbf{v}_{s,T}$	$\mathbf{i}_{s,T}$	$\mathbf{b}_{s,T}$
$\mathbf{A}_1$	$-\mathbf{Z}_T$	$-\mathbf{Y}_T$	$\mathbf{\Gamma}_T$
$\mathbf{A}_2$	$-\mathbf{Z}_I$	$-\mathbf{Y}_I$	$\mathbf{\Theta}$
$\mathbf{A}_3$	$-\mathbf{Z}_R$	$-\mathbf{Y}_R$	$\mathbf{\Gamma}_R$

where we introduce $\mathbf{c} \in \mathbb{C}^{N \times 1}$ and $\overline{\mathbf{A}} \in \mathbb{C}^{N \times N}$ as

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \ \overline{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_3 \end{bmatrix}. \tag{8.24}$$

System (8.23) can be solved by noting that

$$\mathbf{x} = \mathbf{c} + \overline{\mathbf{A}}\mathbf{A}\mathbf{x}, \tag{8.25}$$

gives

$$\mathbf{x} = \left(\mathbf{I} - \overline{\mathbf{A}}\mathbf{A}\right)^{-1} \mathbf{c}. \tag{8.26}$$

Thus, by introducing $\widetilde{\mathbf{A}} = \left(\mathbf{I} - \overline{\mathbf{A}}\mathbf{A}\right)^{-1}$, partitioned as

$$\widetilde{\mathbf{A}} = \begin{bmatrix} \widetilde{\mathbf{A}}_{11} & \widetilde{\mathbf{A}}_{12} & \widetilde{\mathbf{A}}_{13} \\ \widetilde{\mathbf{A}}_{21} & \widetilde{\mathbf{A}}_{22} & \widetilde{\mathbf{A}}_{23} \\ \widetilde{\mathbf{A}}_{31} & \widetilde{\mathbf{A}}_{32} & \widetilde{\mathbf{A}}_{33} \end{bmatrix}, \tag{8.27}$$

we have

$$\mathbf{x}_1 = \widetilde{\mathbf{A}}_{11}\mathbf{c}_1, \ \mathbf{x}_2 = \widetilde{\mathbf{A}}_{21}\mathbf{c}_1, \ \mathbf{x}_3 = \widetilde{\mathbf{A}}_{31}\mathbf{c}_1, \tag{8.28}$$

providing an expression for the vector $\mathbf{x}$. Given $\mathbf{x}$, the vector $\mathbf{y}$ is determined by applying $\mathbf{y} = \mathbf{A}\mathbf{x}$ and system (8.22) is solved.

8.2.2 Z-, Y-, and S-Parameters Analysis

With the Z-parameters, the channel $\mathbf{H}$ is derived by solving the system of the four linear equations (8.1), (8.3), (8.4), and (8.5). According to Section 8.2.1, this system gives

$$\mathbf{v} = \left(\mathbf{I} + \overline{\mathbf{Z}}\mathbf{Z}^{-1}\right)^{-1} \mathbf{v}_s, \tag{8.29}$$

where we define $\mathbf{v}_s \in \mathbb{C}^{N\times 1}$ and $\overline{\mathbf{Z}} \in \mathbb{C}^{N\times N}$ as

$$\mathbf{v}_s = \begin{bmatrix} \mathbf{v}_{s,T} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \; \overline{\mathbf{Z}} = \begin{bmatrix} \mathbf{Z}_T & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_R \end{bmatrix}. \tag{8.30}$$

Thus, by introducing $\widetilde{\mathbf{Z}} = \left(\mathbf{I} + \overline{\mathbf{Z}}\mathbf{Z}^{-1}\right)^{-1}$, partitioned as

$$\widetilde{\mathbf{Z}} = \begin{bmatrix} \widetilde{\mathbf{Z}}_{TT} & \widetilde{\mathbf{Z}}_{TI} & \widetilde{\mathbf{Z}}_{TR} \\ \widetilde{\mathbf{Z}}_{IT} & \widetilde{\mathbf{Z}}_{II} & \widetilde{\mathbf{Z}}_{IR} \\ \widetilde{\mathbf{Z}}_{RT} & \widetilde{\mathbf{Z}}_{RI} & \widetilde{\mathbf{Z}}_{RR} \end{bmatrix}, \tag{8.31}$$

we have

$$\mathbf{v}_T = \widetilde{\mathbf{Z}}_{TT}\mathbf{v}_{s,T}, \; \mathbf{v}_R = \widetilde{\mathbf{Z}}_{RT}\mathbf{v}_{s,T}. \tag{8.32}$$

As a consequence of (8.32), we obtain

$$\mathbf{v}_R = \widetilde{\mathbf{Z}}_{RT}\mathbf{v}_{s,T} = \widetilde{\mathbf{Z}}_{RT}\widetilde{\mathbf{Z}}_{TT}^{-1}\mathbf{v}_T, \tag{8.33}$$

yielding

$$\mathbf{H} = \widetilde{\mathbf{Z}}_{RT}\widetilde{\mathbf{Z}}_{TT}^{-1}, \tag{8.34}$$

which is the general channel model based on the Z-parameters.

With the Y-parameters, the channel $\mathbf{H}$ is derived by solving the system of the four linear equations (8.6), (8.8), (8.9), and (8.10). According to Section 8.2.1, this system gives

$$\mathbf{i} = \left(\mathbf{I} + \overline{\mathbf{Y}}\mathbf{Y}^{-1}\right)^{-1}\mathbf{i}_s, \tag{8.35}$$

where we define $\mathbf{i}_s \in \mathbb{C}^{N\times 1}$ and $\overline{\mathbf{Y}} \in \mathbb{C}^{N\times N}$ as

$$\mathbf{i}_s = \begin{bmatrix} \mathbf{i}_{s,T} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \; \overline{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y}_T & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_R \end{bmatrix}. \tag{8.36}$$

Thus, by introducing $\widetilde{\mathbf{Y}} = \left(\mathbf{I} + \overline{\mathbf{Y}}\mathbf{Y}^{-1}\right)^{-1}$, partitioned as

$$\widetilde{\mathbf{Y}} = \begin{bmatrix} \widetilde{\mathbf{Y}}_{TT} & \widetilde{\mathbf{Y}}_{TI} & \widetilde{\mathbf{Y}}_{TR} \\ \widetilde{\mathbf{Y}}_{IT} & \widetilde{\mathbf{Y}}_{II} & \widetilde{\mathbf{Y}}_{IR} \\ \widetilde{\mathbf{Y}}_{RT} & \widetilde{\mathbf{Y}}_{RI} & \widetilde{\mathbf{Y}}_{RR} \end{bmatrix}, \tag{8.37}$$

we have

$$\mathbf{i}_T = \widetilde{\mathbf{Y}}_{TT}\mathbf{i}_{s,T}, \; \mathbf{i}_R = \widetilde{\mathbf{Y}}_{RT}\mathbf{i}_{s,T}. \tag{8.38}$$

As a consequence of (8.8), (8.10), and (8.38), we obtain

$$\mathbf{v}_R = -\mathbf{Y}_R^{-1}\mathbf{i}_R = -\mathbf{Y}_R^{-1}\widetilde{\mathbf{Y}}_{RT}\mathbf{i}_{s,T} = \mathbf{Y}_R^{-1}\widetilde{\mathbf{Y}}_{RT}\left(\widetilde{\mathbf{Y}}_{TT} - \mathbf{I}\right)^{-1}\mathbf{Y}_T\mathbf{v}_T, \tag{8.39}$$

yielding

$$\mathbf{H} = \mathbf{Y}_R^{-1}\widetilde{\mathbf{Y}}_{RT}\left(\widetilde{\mathbf{Y}}_{TT} - \mathbf{I}\right)^{-1}\mathbf{Y}_T, \tag{8.40}$$

which is the general channel model with the Y-parameters.

With the S-parameters, the channel $\mathbf{H}$ is derived by solving the system of (8.11), (8.15), (8.17), and (8.19). According to Section 8.2.1, and using $\mathbf{b} = \mathbf{S}\mathbf{a}$, this system gives

$$\mathbf{b} = \mathbf{S}\left(\mathbf{I} - \mathbf{\Gamma}\mathbf{S}\right)^{-1}\mathbf{b}_s, \tag{8.41}$$

where we define $\mathbf{b}_s \in \mathbb{C}^{N\times 1}$ and $\mathbf{\Gamma} \in \mathbb{C}^{N\times N}$ as

$$\mathbf{b}_s = \begin{bmatrix} \mathbf{b}_{s,T} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \ \mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_T & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Theta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Gamma}_R \end{bmatrix}. \tag{8.42}$$

Thus, by introducing $\widetilde{\mathbf{S}} = \mathbf{S}\left(\mathbf{I} - \mathbf{\Gamma}\mathbf{S}\right)^{-1}$, partitioned as

$$\widetilde{\mathbf{S}} = \begin{bmatrix} \widetilde{\mathbf{S}}_{TT} & \widetilde{\mathbf{S}}_{TI} & \widetilde{\mathbf{S}}_{TR} \\ \widetilde{\mathbf{S}}_{IT} & \widetilde{\mathbf{S}}_{II} & \widetilde{\mathbf{S}}_{IR} \\ \widetilde{\mathbf{S}}_{RT} & \widetilde{\mathbf{S}}_{RI} & \widetilde{\mathbf{S}}_{RR} \end{bmatrix}, \tag{8.43}$$

we have

$$\mathbf{b}_T = \widetilde{\mathbf{S}}_{TT}\mathbf{b}_{s,T}, \ \mathbf{b}_R = \widetilde{\mathbf{S}}_{RT}\mathbf{b}_{s,T}. \tag{8.44}$$

As a consequence of (8.13), (8.15), (8.19), and (8.44), we obtain

$$\mathbf{v}_R = \mathbf{a}_R + \mathbf{b}_R = \left(\mathbf{\Gamma}_R + \mathbf{I}\right)\widetilde{\mathbf{S}}_{RT}\mathbf{b}_{s,T} = \left(\mathbf{\Gamma}_R + \mathbf{I}\right)\widetilde{\mathbf{S}}_{RT}\left(\mathbf{I} + \mathbf{\Gamma}_T\widetilde{\mathbf{S}}_{TT} + \widetilde{\mathbf{S}}_{TT}\right)^{-1}\mathbf{v}_T, \tag{8.45}$$

yielding

$$\mathbf{H} = \left(\mathbf{\Gamma}_R + \mathbf{I}\right)\widetilde{\mathbf{S}}_{RT}\left(\mathbf{I} + \mathbf{\Gamma}_T\widetilde{\mathbf{S}}_{TT} + \widetilde{\mathbf{S}}_{TT}\right)^{-1}, \tag{8.46}$$

which is the general channel model based on the S-parameters, in agreement with [16].

8.2.3 Equivalence Between General Models

We have derived three general channel models in (8.34), (8.40), and (8.46), based on Z-, Y-, and S-parameters, respectively. We now confirm, for the first time, that these channel models are equivalent representations of an RIS-aided system.

For this purpose, it is necessary to relate the vectors $\mathbf{i}_{s,T}$ and $\mathbf{b}_{s,T}$ to the vector $\mathbf{v}_{s,T}$ by observing that (8.3), (8.8), and (8.15) are three equivalent descriptions of the electrical properties at the transmitter. By equating (8.3) and (8.8), and recalling that $\mathbf{Y}_T = \mathbf{Z}_T^{-1}$, we obtain

$$\mathbf{i}_{s,T} = \mathbf{Y}_T\mathbf{v}_{s,T}, \tag{8.47}$$

giving the relationship between $\mathbf{i}_{s,T}$ and $\mathbf{v}_{s,T}$. In addition, by equating (8.3) and (8.15), and using (8.13) and (8.16), we obtain

$$\mathbf{b}_{s,T} = \frac{\mathbf{I} - \mathbf{\Gamma}_T}{2} \mathbf{v}_{s,T}, \tag{8.48}$$

giving the relationship between $\mathbf{b}_{s,T}$ and $\mathbf{v}_{s,T}$.

We first show that the general channel model based on the Z-parameters (8.34) is equivalent to the general channel model based on the Y-parameters (8.40). To this end, we note that (8.34) is a direct consequence of (8.29) and that (8.40) is a direct consequence of (8.35). Thus, it is convenient to just show that (8.29) and (8.35) are equivalent. By using $\overline{\mathbf{Z}} = \overline{\mathbf{Y}}^{-1}$ and $\mathbf{Z} = \mathbf{Y}^{-1}$, we can rewrite (8.29) as

$$\mathbf{v} = \left(\mathbf{I} + \overline{\mathbf{Y}}^{-1}\mathbf{Y}\right)^{-1} \mathbf{v}_s. \tag{8.49}$$

Then, recalling that $\mathbf{i} = \mathbf{Y}\mathbf{v}$ and observing that (8.47) gives $\mathbf{v}_s = \overline{\mathbf{Y}}^{-1}\mathbf{i}_s$, from (8.49) we obtain

$$\mathbf{i} = \mathbf{Y}\left(\mathbf{I} + \overline{\mathbf{Y}}^{-1}\mathbf{Y}\right)^{-1} \overline{\mathbf{Y}}^{-1}\mathbf{i}_s = \left(\mathbf{I} + \overline{\mathbf{Y}}\mathbf{Y}^{-1}\right)^{-1} \mathbf{i}_s, \tag{8.50}$$

giving (8.35). Thus, we have that the general channel models (8.34) and (8.40) are equivalent since they are a direct consequence of equivalent equations.

We now show that the general channel model based on the Z-parameters (8.34) is also equivalent to the general channel model based on the S-parameters (8.46). Since (8.34) is a direct consequence of (8.29) and (8.46) is a direct consequence of (8.41), it is convenient to just show that (8.29) and (8.41) are equivalent. By using $\overline{\mathbf{Z}} = Z_0 (\mathbf{I} + \mathbf{\Gamma}) (\mathbf{I} - \mathbf{\Gamma})^{-1}$ and $\mathbf{Z} = Z_0 (\mathbf{I} + \mathbf{S}) (\mathbf{I} - \mathbf{S})^{-1}$, we can rewrite (8.29) as

$$\mathbf{v} = \left(\mathbf{I} + (\mathbf{I} + \mathbf{\Gamma}) (\mathbf{I} - \mathbf{\Gamma})^{-1} (\mathbf{I} - \mathbf{S}) (\mathbf{I} + \mathbf{S})^{-1}\right)^{-1} \mathbf{v}_s \tag{8.51}$$

$$= (\mathbf{I} + \mathbf{S}) (\mathbf{I} - \mathbf{\Gamma}\mathbf{S})^{-1} (\mathbf{I} - \mathbf{\Gamma}) \mathbf{v}_s/2. \tag{8.52}$$

Then, recalling that $\mathbf{b} = (\mathbf{S}^{-1} + \mathbf{I})^{-1}\mathbf{v}$ and observing that (8.48) gives $\mathbf{v}_s = 2(\mathbf{I} - \mathbf{\Gamma})^{-1}\mathbf{b}_s$, from (8.52) we obtain

$$\mathbf{b} = \left(\mathbf{S}^{-1} + \mathbf{I}\right)^{-1} (\mathbf{I} + \mathbf{S}) (\mathbf{I} - \mathbf{\Gamma}\mathbf{S})^{-1} \mathbf{b}_s = \mathbf{S} (\mathbf{I} - \mathbf{\Gamma}\mathbf{S})^{-1} \mathbf{b}_s, \tag{8.53}$$

giving (8.41). Thus, the general channel models (8.34) and (8.46) are equivalent since they are a direct consequence of equivalent equations, in agreement with microwave theory [74, Chapter 4]. By the transitive property of the equivalence relation, we also have that the general channel models based on the Y- and S-parameters are equivalent.

8.3 RIS-Aided Communication Model with Unilateral Approximation

We have derived three general channel models based on the Z-, Y-, and S-parameters. They account for imperfect matching and mutual coupling at the transmitter, RIS, and receiver as they have been obtained without any approximation or assumption. However, it is difficult to obtain insights into the role of the RIS in the communication model due to the presence of matrix inversion operations. To gain a better understanding of the communication models, in this section, we approximate the general models by assuming sufficiently large transmission distances.

8.3.1 Universal Framework

We begin by simplifying the universal framework introduced in Section 8.2.1. Specifically, we consider the matrix $\mathbf{A}$ to be block lower triangular given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{0} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}. \tag{8.54}$$

Remarkably, system (8.22) with $\mathbf{A}$ given by (8.54) provides a universal framework that can be used to describe the equations based on Z-, Y-, and S-parameters with the unilateral approximation [94], as it will be clarified in Section 8.3.2. With the simplification in (8.54), the solution to system (8.22) given by (8.28) simplifies as

$$\mathbf{x}_1 = \left(\mathbf{I} - \mathbf{A}_1\mathbf{A}_{11}\right)^{-1}\mathbf{c}_1, \tag{8.55}$$

$$\mathbf{x}_2 = \left(\mathbf{I} - \mathbf{A}_2\mathbf{A}_{22}\right)^{-1}\mathbf{A}_2\mathbf{A}_{21}\left(\mathbf{I} - \mathbf{A}_1\mathbf{A}_{11}\right)^{-1}\mathbf{c}_1, \tag{8.56}$$

$$\mathbf{x}_3 = \left(\mathbf{A}_3^{-1} - \mathbf{A}_{33}\right)^{-1}\left(\mathbf{A}_{31} + \mathbf{A}_{32}\left(\mathbf{A}_2^{-1} - \mathbf{A}_{22}\right)^{-1}\mathbf{A}_{21}\right)\left(\mathbf{I} - \mathbf{A}_1\mathbf{A}_{11}\right)^{-1}\mathbf{c}_1, \tag{8.57}$$

allowing to explicitly express $\mathbf{x}$ in terms of $\mathbf{A}$, $\mathbf{c}_1$, $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$. In the following, we use this universal framework to derive the channel model based on Z-, Y-, and S-parameters.

8.3.2 Z-, Y-, and S-Parameters Analysis

To simplify (8.34), we observe that with a large transmission distance between a transmitter and a receiver, the electrical properties at the transmitter are approximately independent of the electrical properties at the receiver. The minimum transmission distance at which this phenomenon occurs is a function of the number of antennas at the transmitter and receiver, their radiation pattern, and the antenna spacing, as discussed in [94]. Thus, assuming that the transmission distance is large enough, we can consider the so-called unilateral approximation and set $\mathbf{Z}_{TI} = \mathbf{0}$, $\mathbf{Z}_{TR} = \mathbf{0}$, and $\mathbf{Z}_{IR} = \mathbf{0}$ [94]. Note that with this approximation we do not assume the wireless channel to be non-reciprocal. Conversely, we set to zero the upper block triangular part of $\mathbf{Z}$ as it does not impact the expression of $\mathbf{H}$. Thus, we can express $\mathbf{v}_T$ and $\mathbf{v}_R$ as

$$\mathbf{v}_T = \left(\mathbf{I} + \mathbf{Z}_T\mathbf{Z}_{TT}^{-1}\right)^{-1}\mathbf{v}_{s,T}, \tag{8.58}$$

$$\mathbf{v}_R = \mathbf{Z}_R\left(\mathbf{Z}_R + \mathbf{Z}_{RR}\right)^{-1}\left(\mathbf{Z}_{RT} - \mathbf{Z}_{RI}\left(\mathbf{Z}_I + \mathbf{Z}_{II}\right)^{-1}\mathbf{Z}_{IT}\right)\left(\mathbf{Z}_T + \mathbf{Z}_{TT}\right)^{-1}\mathbf{v}_{s,T}, \tag{8.59}$$

according to Section 8.3.1. Consequently, the channel model based on the Z-parameters is given by

$$\mathbf{H} = \mathbf{Z}_R \left(\mathbf{Z}_R + \mathbf{Z}_{RR}\right)^{-1} \left(\mathbf{Z}_{RT} - \mathbf{Z}_{RI} \left(\mathbf{Z}_I + \mathbf{Z}_{II}\right)^{-1} \mathbf{Z}_{IT}\right) \mathbf{Z}_{TT}^{-1}, \tag{8.60}$$

explicitly clarifying the impact of $\mathbf{Z}$, $\mathbf{Z}_T$, $\mathbf{Z}_I$, and $\mathbf{Z}_R$, in agreement with [92, Corollary 1].

As discussed for the Z-parameters, we can also simplify (8.40) by considering the unilateral approximation. Setting $\mathbf{Z}_{TI} = \mathbf{0}$, $\mathbf{Z}_{TR} = \mathbf{0}$, and $\mathbf{Z}_{IR} = \mathbf{0}$, and recalling that the entire matrix $\mathbf{Y}$ is related to $\mathbf{Z}$ through $\mathbf{Y} = \mathbf{Z}^{-1}$, we obtain $\mathbf{Y}_{TI} = \mathbf{0}$, $\mathbf{Y}_{TR} = \mathbf{0}$, and $\mathbf{Y}_{IR} = \mathbf{0}$, allowing us to express $\mathbf{i}_T$ and $\mathbf{i}_R$ as

$$\mathbf{i}_T = \left(\mathbf{I} + \mathbf{Y}_T \mathbf{Y}_{TT}^{-1}\right)^{-1} \mathbf{i}_{s,T}, \tag{8.61}$$

$$\mathbf{i}_R = \mathbf{Y}_R \left(\mathbf{Y}_R + \mathbf{Y}_{RR}\right)^{-1} \left(\mathbf{Y}_{RT} - \mathbf{Y}_{RI} \left(\mathbf{Y}_I + \mathbf{Y}_{II}\right)^{-1} \mathbf{Y}_{IT}\right) \left(\mathbf{Y}_T + \mathbf{Y}_{TT}\right)^{-1} \mathbf{i}_{s,T}, \tag{8.62}$$

according to Section 8.3.1. Thus, the channel model based on the Y-parameters with the unilateral approximation is given by

$$\mathbf{H} = \left(\mathbf{Y}_R + \mathbf{Y}_{RR}\right)^{-1} \left(-\mathbf{Y}_{RT} + \mathbf{Y}_{RI} \left(\mathbf{Y}_I + \mathbf{Y}_{II}\right)^{-1} \mathbf{Y}_{IT}\right), \tag{8.63}$$

explicitly defining the effect of $\mathbf{Y}$, $\mathbf{Y}_T$, $\mathbf{Y}_I$, and $\mathbf{Y}_R$.

As done for the Z- and Y-parameters, we simplify (8.46) by considering the unilateral approximation to better understand the role of the RIS. Setting $\mathbf{Z}_{TI} = \mathbf{0}$, $\mathbf{Z}_{TR} = \mathbf{0}$, and $\mathbf{Z}_{IR} = \mathbf{0}$, and recalling (8.14), we obtain $\mathbf{S}_{TI} = \mathbf{0}$, $\mathbf{S}_{TR} = \mathbf{0}$, and $\mathbf{S}_{IR} = \mathbf{0}$, which allow us to express $\mathbf{b}_T$ and $\mathbf{b}_R$ as

$$\mathbf{b}_T = \mathbf{S}_{TT} \left(\mathbf{I} - \mathbf{\Gamma}_T \mathbf{S}_{TT}\right)^{-1} \mathbf{b}_{s,T}, \tag{8.64}$$

$$\mathbf{b}_R = \left(\mathbf{I} - \mathbf{S}_{RR} \mathbf{\Gamma}_R\right)^{-1} \left(\mathbf{S}_{RT} + \mathbf{S}_{RI} \left(\mathbf{I} - \mathbf{\Theta} \mathbf{S}_{II}\right)^{-1} \mathbf{\Theta} \mathbf{S}_{IT}\right) \left(\mathbf{I} - \mathbf{\Gamma}_T \mathbf{S}_{TT}\right)^{-1} \mathbf{b}_{s,T}, \tag{8.65}$$

according to Section 8.3.1. Thus, the channel model based on the S-parameters with the unilateral approximation is given by

$$\mathbf{H} = \left(\mathbf{\Gamma}_R + \mathbf{I}\right) \left(\mathbf{I} - \mathbf{S}_{RR} \mathbf{\Gamma}_R\right)^{-1} \left(\mathbf{S}_{RT} + \mathbf{S}_{RI} \left(\mathbf{I} - \mathbf{\Theta} \mathbf{S}_{II}\right)^{-1} \mathbf{\Theta} \mathbf{S}_{IT}\right) \left(\mathbf{I} + \mathbf{S}_{TT}\right)^{-1}, \tag{8.66}$$

clearly highlighting the role of $\mathbf{S}$, $\mathbf{\Gamma}_T$, $\mathbf{\Theta}$, and $\mathbf{\Gamma}_R$.

Remarkably, the three models (8.60), (8.63), and (8.66) effectively show the impact of $\mathbf{Z}_I$, $\mathbf{Y}_I$, and $\mathbf{\Theta}$ on the communication, while still capturing the impact of imperfect matching and mutual coupling at the transmitter, RIS, and receiver.

8.3.3 Mappings Between Parameters

Under the unilateral approximation, we can derive simplified mappings that allow us to express the Y- and S-parameters as functions of Z-parameters.

Table 8.2: Channel models based on Z-, Y-, and S-parameters with corresponding assumptions and approximations.

	Z-parameters	Y-parameters	S-parameters
Channel model	$\widetilde{\mathbf{Z}}_{RT}\widetilde{\mathbf{Z}}_{TT}^{-1}$	$\mathbf{Y}_R^{-1}\widetilde{\mathbf{Y}}_{RT}\left(\widetilde{\mathbf{Y}}_{TT}-\mathbf{I}\right)^{-1}\mathbf{Y}_T$	$(\mathbf{\Gamma}_R+\mathbf{I})\,\widetilde{\mathbf{S}}_{RT}\left(\mathbf{I}+\mathbf{\Gamma}_T\widetilde{\mathbf{S}}_{TT}+\widetilde{\mathbf{S}}_{TT}\right)^{-1}$
A1: Sufficiently large transmission distances*	$\mathbf{Z}_{TI}=\mathbf{0}$, $\mathbf{Z}_{TR}=\mathbf{0}$, $\mathbf{Z}_{IR}=\mathbf{0}$	$\mathbf{Y}_{TI}=\mathbf{0}$, $\mathbf{Y}_{TR}=\mathbf{0}$, $\mathbf{Y}_{IR}=\mathbf{0}$	$\mathbf{S}_{TI}=\mathbf{0}$, $\mathbf{S}_{TR}=\mathbf{0}$, $\mathbf{S}_{IR}=\mathbf{0}$
Channel model with A1	$\mathbf{Z}_R\left(\mathbf{Z}_R+\mathbf{Z}_{RR}\right)^{-1}\left(\mathbf{Z}_{RT}-\mathbf{Z}_{RI}\left(\mathbf{Z}_I+\mathbf{Z}_{II}\right)^{-1}\mathbf{Z}_{IT}\right)\mathbf{Z}_{TT}^{-1}$	$\left(\mathbf{Y}_R+\mathbf{Y}_{RR}\right)^{-1}\left(-\mathbf{Y}_{RT}+\mathbf{Y}_{RI}\left(\mathbf{Y}_I+\mathbf{Y}_{II}\right)^{-1}\mathbf{Y}_{IT}\right)$	$(\mathbf{\Gamma}_R+\mathbf{I})\left(\mathbf{I}-\mathbf{S}_{RR}\mathbf{\Gamma}_R\right)^{-1}\left(\mathbf{S}_{RT}+\mathbf{S}_{RI}\left(\mathbf{I}-\mathbf{\Theta}\mathbf{S}_{II}\right)^{-1}\mathbf{\Theta}\mathbf{S}_{IT}\right)\left(\mathbf{I}+\mathbf{S}_{TT}\right)^{-1}$
A2: Perfect matching and no mutual coupling at TX and RX	$\mathbf{Z}_{TT}=Z_0\mathbf{I}$, $\mathbf{Z}_{RR}=Z_0\mathbf{I}$	$\mathbf{Y}_{TT}=Y_0\mathbf{I}$, $\mathbf{Y}_{RR}=Y_0\mathbf{I}$	$\mathbf{S}_{TT}=\mathbf{0}$, $\mathbf{S}_{RR}=\mathbf{0}$
A3: Impedances at TX and RX equal to Z_0	$\mathbf{Z}_T=Z_0\mathbf{I}$, $\mathbf{Z}_R=Z_0\mathbf{I}$	$\mathbf{Y}_T=Y_0\mathbf{I}$, $\mathbf{Y}_R=Y_0\mathbf{I}$	$\mathbf{\Gamma}_T=\mathbf{0}$, $\mathbf{\Gamma}_R=\mathbf{0}$
Channel model with A1, A2, and A3	$\frac{1}{2Z_0}\left(\mathbf{Z}_{RT}-\mathbf{Z}_{RI}\left(\mathbf{Z}_I+\mathbf{Z}_{II}\right)^{-1}\mathbf{Z}_{IT}\right)$	$\frac{1}{2Y_0}\left(-\mathbf{Y}_{RT}+\mathbf{Y}_{RI}\left(\mathbf{Y}_I+\mathbf{Y}_{II}\right)^{-1}\mathbf{Y}_{IT}\right)$	$\mathbf{S}_{RT}+\mathbf{S}_{RI}\left(\mathbf{I}-\mathbf{\Theta}\mathbf{S}_{II}\right)^{-1}\mathbf{\Theta}\mathbf{S}_{IT}$
A4: Perfect matching and no mutual coupling at RIS	$\mathbf{Z}_{II}=Z_0\mathbf{I}$	$\mathbf{Y}_{II}=Y_0\mathbf{I}$	$\mathbf{S}_{II}=\mathbf{0}$
Channel model with A1, A2, A3, and A4	$\frac{1}{2Z_0}\left(\mathbf{Z}_{RT}-\mathbf{Z}_{RI}\left(\mathbf{Z}_I+Z_0\mathbf{I}\right)^{-1}\mathbf{Z}_{IT}\right)$	$\frac{1}{2Y_0}\left(-\mathbf{Y}_{RT}+\mathbf{Y}_{RI}\left(\mathbf{Y}_I+Y_0\mathbf{I}\right)^{-1}\mathbf{Y}_{IT}\right)$	$\mathbf{S}_{RT}+\mathbf{S}_{RI}\mathbf{\Theta}\mathbf{S}_{IT}$

* Distances larger than the critical distance provided in [94].

From Z- to Y-Parameters

To express the matrices $\mathbf{Y}_{RT}$, $\mathbf{Y}_{RI}$, and $\mathbf{Y}_{IT}$ as functions of $\mathbf{Z}$, we consider the relationship $\mathbf{Y}=\mathbf{Z}^{-1}$ where $\mathbf{Z}$ has been simplified through the unilateral approximation, i.e., $\mathbf{Z}_{TI}=\mathbf{0}$, $\mathbf{Z}_{TR}=\mathbf{0}$, and $\mathbf{Z}_{IR}=\mathbf{0}$. Thus, by computing $\mathbf{Y}$, we have

$$\mathbf{Y}_{RI}=-\mathbf{Z}_{RR}^{-1}\mathbf{Z}_{RI}\mathbf{Z}_{II}^{-1},\ \mathbf{Y}_{IT}=-\mathbf{Z}_{II}^{-1}\mathbf{Z}_{IT}\mathbf{Z}_{TT}^{-1}, \tag{8.67}$$

$$\mathbf{Y}_{RT}=\mathbf{Z}_{RR}^{-1}\left(-\mathbf{Z}_{RT}+\mathbf{Z}_{RI}\mathbf{Z}_{II}^{-1}\mathbf{Z}_{IT}\right)\mathbf{Z}_{TT}^{-1}. \tag{8.68}$$

From Z- to S-Parameters

To express the matrices $\mathbf{S}_{RT}$, $\mathbf{S}_{RI}$, and $\mathbf{S}_{IT}$ as functions of $\mathbf{Z}$, we consider the relationship (8.14) where $\mathbf{Z}_{TI}=\mathbf{0}$, $\mathbf{Z}_{TR}=\mathbf{0}$, and $\mathbf{Z}_{IR}=\mathbf{0}$. Thus, it can be proved that

$$\mathbf{S}_{RI}=2Z_0\left(\mathbf{Z}_{RR}+Z_0\mathbf{I}\right)^{-1}\mathbf{Z}_{RI}\left(\mathbf{Z}_{II}+Z_0\mathbf{I}\right)^{-1}, \tag{8.69}$$

$$\mathbf{S}_{IT}=2Z_0\left(\mathbf{Z}_{II}+Z_0\mathbf{I}\right)^{-1}\mathbf{Z}_{IT}\left(\mathbf{Z}_{TT}+Z_0\mathbf{I}\right)^{-1}, \tag{8.70}$$

$$\mathbf{S}_{RT}=2Z_0\left(\mathbf{Z}_{RR}+Z_0\mathbf{I}\right)^{-1}\left(\mathbf{Z}_{RT}-\mathbf{Z}_{RI}\left(\mathbf{Z}_{II}+Z_0\mathbf{I}\right)^{-1}\mathbf{Z}_{IT}\right)\left(\mathbf{Z}_{TT}+Z_0\mathbf{I}\right)^{-1}. \tag{8.71}$$

Interestingly, these mappings further clarify the equivalence of the three analyses and the meaning of the different terms. In the following, we show that these mappings play a fundamental role in relating the EM-compliant models derived in this chapter with the channel model widely used in related literature.

8.4 RIS-Aided Communication Model with No Mutual Coupling

We have derived three general channel models based on the Z-, Y-, and S-parameters, and we have simplified them by considering the unilateral approximation. In this section, we further simplify the obtained models to gain engineering insights into the role of the RIS in the communication system.

8.4.1 Z-, Y-, and S-Parameters Analysis

To further simplify (8.60), we consider two assumptions in addition to the unilateral approximation. First, we assume that the antennas at the transmitter and receiver are perfectly matched and have no mutual coupling, yielding $\mathbf{Z}_{TT} = Z_0\mathbf{I}$ and $\mathbf{Z}_{RR} = Z_0\mathbf{I}$. Second, we assume that the impedances at the transmitter Z_{T,n_T} and the receiver Z_{R,n_R} are all Z_0, i.e., $\mathbf{Z}_T = Z_0\mathbf{I}$ and $\mathbf{Z}_R = Z_0\mathbf{I}$. With these assumptions, the channel model (8.60) simplifies to

$$\mathbf{H} = \frac{1}{2Z_0}\left(\mathbf{Z}_{RT} - \mathbf{Z}_{RI}\left(\mathbf{Z}_I + \mathbf{Z}_{II}\right)^{-1}\mathbf{Z}_{IT}\right). \tag{8.72}$$

In addition, assuming perfect matching and no mutual coupling at the RIS, i.e., $\mathbf{Z}_{II} = Z_0\mathbf{I}$, we obtain

$$\mathbf{H} = \frac{1}{2Z_0}\left(\mathbf{Z}_{RT} - \mathbf{Z}_{RI}\left(\mathbf{Z}_I + Z_0\mathbf{I}\right)^{-1}\mathbf{Z}_{IT}\right), \tag{8.73}$$

giving the simplified channel model based on the Z-parameters with perfect matching and no mutual coupling. By further assuming the RIS elements to be canonical minimum scattering antennas [96], the channel without the RIS is equivalent to the channel when the RIS elements are open-circuited, i.e., $\mathbf{Z}_I = \infty\mathbf{I}$, given by $\mathbf{H} = \mathbf{Z}_{RT}/(2Z_0)$ following (8.73).

To further simplify (8.63), we consider the two additional assumptions as discussed for the Z-parameters. First, assuming perfect matching and no mutual coupling at the transmitter and receiver, and noting that, with the unilateral approximation, the relationship $\mathbf{Y} = \mathbf{Z}^{-1}$ implies $\mathbf{Y}_{ii} = \mathbf{Z}_{ii}^{-1}$, for $i \in \{T, I, R\}$, we obtain $\mathbf{Y}_{TT} = Y_0\mathbf{I}$ and $\mathbf{Y}_{RR} = Y_0\mathbf{I}$. Second, considering all the impedances at the transmitter and receiver to be Z_0, and recalling that $\mathbf{Y}_T = \mathbf{Z}_T^{-1}$ and $\mathbf{Y}_R = \mathbf{Z}_R^{-1}$, we have $\mathbf{Y}_T = Y_0\mathbf{I}$ and $\mathbf{Y}_R = Y_0\mathbf{I}$. With these assumptions, the channel model (8.63) simplifies to

$$\mathbf{H} = \frac{1}{2Y_0}\left(-\mathbf{Y}_{RT} + \mathbf{Y}_{RI}\left(\mathbf{Y}_I + \mathbf{Y}_{II}\right)^{-1}\mathbf{Y}_{IT}\right), \tag{8.74}$$

In addition, assuming perfect matching and no mutual coupling at the RIS, i.e., $\mathbf{Y}_{II} = Y_0\mathbf{I}$, we obtain

$$\mathbf{H} = \frac{1}{2Y_0}\left(-\mathbf{Y}_{RT} + \mathbf{Y}_{RI}\left(\mathbf{Y}_I + Y_0\mathbf{I}\right)^{-1}\mathbf{Y}_{IT}\right), \tag{8.75}$$

giving the simplified channel model based on the Y-parameters with perfect matching and no mutual coupling.

As done for Z- and Y-parameters, we now further simplify (8.66). First, assuming perfect matching and no mutual coupling at the transmitter and receiver, and noting that, with the unilateral approximation, (8.14) implies $\mathbf{S}_{ii} = \left(\mathbf{Z}_{ii} + Z_0\mathbf{I}\right)^{-1}\left(\mathbf{Z}_{ii} - Z_0\mathbf{I}\right)$, for $i \in \{T, I, R\}$, we obtain $\mathbf{S}_{TT} = \mathbf{0}$ and $\mathbf{S}_{RR} = \mathbf{0}$. Second, considering all the impedances at the transmitter and receiver to be Z_0, and recalling (8.16) and (8.20), we have $\mathbf{\Gamma}_T = \mathbf{0}$ and $\mathbf{\Gamma}_R = \mathbf{0}$. With these two assumptions, (8.66) simplifies to

$$\mathbf{H} = \mathbf{S}_{RT} + \mathbf{S}_{RI}\left(\mathbf{I} - \mathbf{\Theta}\mathbf{S}_{II}\right)^{-1}\mathbf{\Theta}\mathbf{S}_{IT}. \tag{8.76}$$

In addition, assuming perfect matching and no mutual coupling at the RIS, i.e., $\mathbf{S}_{II} = \mathbf{0}$, we obtain

$$\mathbf{H} = \mathbf{S}_{RT} + \mathbf{S}_{RI}\mathbf{\Theta}\mathbf{S}_{IT}, \tag{8.77}$$

which is the simplified channel model based on S-parameters with perfect matching and no mutual coupling, commonly used in communications in agreement with [16]. We summarize the main results of the three analyses based on Z-, Y-, and S-parameters in Tab. 8.2.

8.4.2 Mappings Between Parameters

Under the assumptions of perfect matching and no mutual coupling, we can derive more simplified mappings that allow us to express the Y- and S-parameters as functions of Z-parameters. By using such mappings, it is possible to directly show that the simplified channel models (8.73), (8.75), and (8.77) are equivalent.

From Z- to Y-Parameters

Considering the unilateral approximation and perfect matching and no mutual coupling at the transmitter and receiver, the matrices $\mathbf{Y}_{RT}$, $\mathbf{Y}_{RI}$, and $\mathbf{Y}_{IT}$ can be expressed as functions of $\mathbf{Z}$ by setting $\mathbf{Z}_{TT} = Z_0\mathbf{I}$ and $\mathbf{Z}_{RR} = Z_0\mathbf{I}$ in (8.67) and (8.68), yielding

$$\mathbf{Y}_{RI} = -\frac{\mathbf{Z}_{RI}\mathbf{Z}_{II}^{-1}}{Z_0},\ \mathbf{Y}_{IT} = -\frac{\mathbf{Z}_{II}^{-1}\mathbf{Z}_{IT}}{Z_0},\ \mathbf{Y}_{RT} = \frac{1}{Z_0^2}\left(-\mathbf{Z}_{RT} + \mathbf{Z}_{RI}\mathbf{Z}_{II}^{-1}\mathbf{Z}_{IT}\right). \tag{8.78}$$

In addition, with perfect matching and no mutual coupling at the RIS, i.e., $\mathbf{Z}_{II} = Z_0\mathbf{I}$, (8.78) boils down to

$$\mathbf{Y}_{RI} = -\frac{\mathbf{Z}_{RI}}{Z_0^2},\ \mathbf{Y}_{IT} = -\frac{\mathbf{Z}_{IT}}{Z_0^2},\ \mathbf{Y}_{RT} = \frac{1}{Z_0^2}\left(-\mathbf{Z}_{RT} + \frac{\mathbf{Z}_{RI}\mathbf{Z}_{IT}}{Z_0}\right). \tag{8.79}$$

Interestingly, by substituting $\mathbf{Y}_I = \mathbf{Z}_I^{-1}$ and (8.79) into (8.75), we directly obtain (8.73), confirming that the two analyses based on Z- and Y-parameters lead to the same result.

From Z- to S-Parameters

To express the matrices $\mathbf{S}_{RT}$, $\mathbf{S}_{RI}$, and $\mathbf{S}_{IT}$ as functions of $\mathbf{Z}$ with the unilateral approximation and perfect matching and no mutual coupling at the transmitter and receiver, we can set $\mathbf{Z}_{TT} = Z_0\mathbf{I}$ and $\mathbf{Z}_{RR} = Z_0\mathbf{I}$ in (8.69), (8.70), and (8.71), yielding

$$\mathbf{S}_{RI} = \mathbf{Z}_{RI}\left(\mathbf{Z}_{II} + Z_0\mathbf{I}\right)^{-1},\ \mathbf{S}_{IT} = \left(\mathbf{Z}_{II} + Z_0\mathbf{I}\right)^{-1}\mathbf{Z}_{IT}, \tag{8.80}$$

$$\mathbf{S}_{RT} = \frac{1}{2Z_0}\left(\mathbf{Z}_{RT} - \mathbf{Z}_{RI}\left(\mathbf{Z}_{II} + Z_0\mathbf{I}\right)^{-1}\mathbf{Z}_{IT}\right). \tag{8.81}$$

In addition, with perfect matching and no mutual coupling at the RIS, i.e., $\mathbf{Z}_{II} = Z_0\mathbf{I}$, (8.80) and (8.81) boil down to

$$\mathbf{S}_{RI} = \frac{\mathbf{Z}_{RI}}{2Z_0},\ \mathbf{S}_{IT} = \frac{\mathbf{Z}_{IT}}{2Z_0},\ \mathbf{S}_{RT} = \frac{1}{2Z_0}\left(\mathbf{Z}_{RT} - \frac{\mathbf{Z}_{RI}\mathbf{Z}_{IT}}{2Z_0}\right). \tag{8.82}$$

Interestingly, by substituting (8.18) and (8.82) into (8.77), we directly obtain (8.73), confirming that the two analyses based on the Z- and S-parameters lead to the same conclusion.

8.4.3 Relationship with the Widely Used Model

We have derived three equivalent simplified channel models under the assumptions of perfect matching and no mutual coupling. In this section, we relate these models with the RIS-aided channel model widely used in related literature.

In RIS literature [13], the notation

$$\mathbf{H}_{RT} = \mathbf{S}_{RT},\ \mathbf{H}_{RI} = \mathbf{S}_{RI},\ \mathbf{H}_{IT} = \mathbf{S}_{IT}, \tag{8.83}$$

is commonly used to denote the channel matrices from the transmitter to the receiver, from the RIS to the receiver, and from the transmitter to the RIS, respectively. By substituting (8.83) into the simplified channel based on the S-parameters (8.77), we obtain

$$\mathbf{H} = \mathbf{H}_{RT} + \mathbf{H}_{RI}\boldsymbol{\Theta}\mathbf{H}_{IT}, \tag{8.84}$$

which is the RIS-aided communication model widely used in related literature. Thus, we can conclude that the widely used model (8.84) is derived by considering the unilateral approximation and under the assumptions of perfect matching and no mutual coupling, in agreement with [16].

An interesting aspect of the channel model (8.84) is the dependence of $\mathbf{H}_{RT}$ on $\mathbf{H}_{RI}$ and $\mathbf{H}_{IT}$ as a result of (8.82) and (8.83). However, this aspect is commonly neglected in related literature, where it is implicitly considered

$$\mathbf{S}_{RT} \approx \frac{\mathbf{Z}_{RT}}{2Z_0}, \tag{8.85}$$

as an approximation of (8.82). Remarkably, (8.85) can be obtained by considering the relationship

$$\mathbf{Z} = 2Z_0\left(\mathbf{I} - \mathbf{S}\right)^{-1} - Z_0\mathbf{I}, \tag{8.86}$$

which is derived from (8.14), by applying the first-order Neumann series approximation $(\mathbf{I} - \mathbf{S})^{-1} \approx (\mathbf{I} + \mathbf{S})$. With this approximation, from (8.86) we obtain $\mathbf{Z} \approx Z_0\left(\mathbf{I} + 2\mathbf{S}\right)$, directly giving (8.85). Note that the first-order Neumann series approximation holds as long as $\|\mathbf{S}\|_F$ is small, which is typically valid in wireless communication scenarios given the high losses suffered from the propagating signal.

According to the approximation in (8.85), when the direct channel between transmitter and receiver is completely obstructed, i.e., $\mathbf{Z}_{RT} = \mathbf{0}$, we have $\mathbf{H}_{RT} = \mathbf{0}$, as widely accepted in related literature. However, according to (8.82), a completely obstructed direct channel, i.e., $\mathbf{Z}_{RT} = \mathbf{0}$, gives $\mathbf{H}_{RT} = -\mathbf{H}_{RI}\mathbf{H}_{IT}$, which is in general non zero, as correctly noticed in [93]. The physical meaning of having $\mathbf{H}_{RT} \neq \mathbf{0}$ in the case of a completely obstructed direct channel, i.e., $\mathbf{Z}_{RT} = \mathbf{0}$, lies in the structural scattering of the RIS.

Since the approximation in (8.85) causes an inconsistency, especially when $\mathbf{Z}_{RT} = \mathbf{0}$, it is worth quantifying the difference between (8.84) and the channel model resulting from (8.85), widely used in the literature. To this end, we assume a completely obstructed direct channel, i.e., $\mathbf{Z}_{RT} = \mathbf{0}$, and consider a SISO system, i.e., $N_R = 1$ and $N_T = 1$, where (8.84) boils down to

$$h = -\mathbf{h}_{RI}\mathbf{h}_{IT} + \mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{h}_{IT}, \tag{8.87}$$

where the term $-\mathbf{h}_{RI}\mathbf{h}_{IT}$ represents the structural scattering of the RIS, and, considering the approximation in (8.85), to

$$h \approx \mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{h}_{IT} \triangleq h'. \tag{8.88}$$

In the following, we derive the scaling laws of the received signal power achieved under the channel models in (8.87) and (8.88) when a conventional RIS architecture is considered.

As for the model in (8.87), the received signal power is given by $P_R = |-\mathbf{h}_{RI}\mathbf{h}_{IT} + \mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{h}_{IT}|^2$, where a unit transmit power is assumed with no loss of generality. Thus, the maximum received signal power is given by

$$P_R = \left(|\mathbf{h}_{RI}\mathbf{h}_{IT}| + \sum_{n_I=1}^{N_I}\left|[\mathbf{h}_{RI}]_{n_I}[\mathbf{h}_{IT}]_{n_I}\right|\right)^2, \tag{8.89}$$

achievable by co-phasing the terms $-\mathbf{h}_{RI}\mathbf{h}_{IT}$ and $\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{h}_{IT}$ through an appropriate phase shift introduced by $\boldsymbol{\Theta}$.

As for the widely used model h', the received signal power is equal to $P_R' = |\mathbf{h}_{RI}\boldsymbol{\Theta}\mathbf{h}_{IT}|^2$. Thus, it is possible to optimize the RIS to achieve a received signal power of

$$P_R' = \left(\sum_{n_I=1}^{N_I}\left|[\mathbf{h}_{RI}]_{n_I}[\mathbf{h}_{IT}]_{n_I}\right|\right)^2. \tag{8.90}$$

We define the relative difference between the average received signal power $\mathrm{E}[P_R]$ and $\mathrm{E}[P_R']$ as

$$\delta = \frac{\mathrm{E}[P_R] - \mathrm{E}[P_R']}{\mathrm{E}[P_R]}. \tag{8.91}$$

To gain numerical insights into (8.91), we consider LoS channels given by $\mathbf{h}_{RI} = [e^{\phi_1}, \ldots, e^{\phi_{N_I}}]$ and $\mathbf{h}_{IT} = [e^{\psi_1}, \ldots, e^{\psi_{N_I}}]^T$. With these channels, in the case of the model h, we have

$$\mathrm{E}[P_R] = \mathrm{E}\left[|\mathbf{h}_{RI}\mathbf{h}_{IT}|^2\right] + N_I^2 + 2N_I\mathrm{E}\left[|\mathbf{h}_{RI}\mathbf{h}_{IT}|\right]. \tag{8.92}$$

Approximating the term $\mathbf{h}_{RI}\mathbf{h}_{IT}$ as a sum of N_I independent and identically distributed (i.i.d.) random variables with mean 0 and variance 1, it holds $\mathbf{h}_{RI}\mathbf{h}_{IT} \sim \mathcal{CN}(0, N_I)$ by the Central Limit Theorem. Thus, by using $\mathrm{E}[|\mathbf{h}_{RI}\mathbf{h}_{IT}|] = \sqrt{\frac{\pi}{4}N_I}$ and $\mathrm{E}[|\mathbf{h}_{RI}\mathbf{h}_{IT}|^2] = N_I$, we have

$$\mathrm{E}[P_R] = N_I^2 + \sqrt{\pi N_I}N_I + N_I. \tag{8.93}$$

Besides, in the case of the widely used model h', we have

$$\mathrm{E}\left[P_R'\right] = P_R' = N_I^2. \tag{8.94}$$

Considering (8.93) and (8.94), the relative difference between $\mathrm{E}[P_R]$ and $\mathrm{E}[P_R']$ under LoS

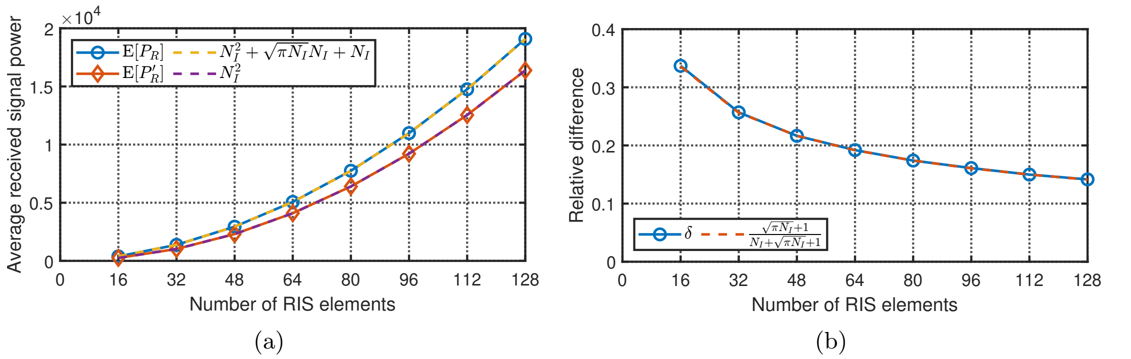


Figure 8.2: (a) Average received signal power obtained with the model h and the widely used model h' and (b) relative difference between them.

channels is equal to

$$\delta = \frac{\sqrt{\pi N_I} + 1}{N_I + \sqrt{\pi N_I} + 1}, \tag{8.95}$$

giving $\delta \to 0$ for $N_I \to \infty$.

In Fig. 8.2(a), we report the average received signal power achieved with h and the approximated model h'. We observe that the received signal power for the channel h is higher because of the additional term $-\mathbf{h}_{RI}\mathbf{h}_{IT}$. In addition, in Fig. 8.2(b), we report the relative difference between the average received signal power achieved with h and with the approximated model h'. We observe that, despite the difference vanishes when $N_I \to \infty$, it is non-negligible for a practical number of RIS elements since it vanishes slowly with N_I.

8.5 RIS Architectures and Optimization

We have shown that Z-, Y-, and S-parameters can be equivalently used to characterize the RIS-aided communication model. Similarly, they can be equivalently used to characterize the N_I-port reconfigurable impedance network implementing the RIS. In this section, we discuss how the matrices $\mathbf{Z}_I$, $\mathbf{Y}_I$, and $\boldsymbol{\Theta}$ can be used to characterize different RIS architectures, including conventional RIS and BD-RIS. In addition, we discuss the advantages of Z-, Y-, and S-parameters in optimizing the RIS.

8.5.1 RIS Architectures

In conventional RIS, also known as single-connected, each RIS port is solely connected to ground through a tunable impedance [16]. Hence, the resulting impedance, admittance, and scattering matrices are diagonal, as given in Tab. 8.3, where Z_{n_I} is the tunable impedance connecting the n_Ith RIS port to ground, $Y_{n_I} = Z_{n_I}^{-1}$, and Θ_{n_I} is the reflection coefficient corresponding to Z_{n_I}, given by $\Theta_{n_I} = (Z_{n_I} - Z_0)/(Z_{n_I} + Z_0)$, according to (8.18), for $n_I = 1, \ldots, N_I$. Furthermore, in the case of a lossless single-connected RIS, Z_{n_I} and Y_{n_I} are purely imaginary, and $|\Theta_{n_I}| = 1$, for $n_I = 1, \ldots, N_I$, as shown in Tab. 8.3. This conventional circuit topology is the simplest RIS circuit topology, offering limited flexibility [16].

Table 8.3: Constraints of lossless and reciprocal RIS architectures based on Z-, Y-, and S-parameters.

	Z-parameters	Y-parameters	S-parameters
Single-connected	$\mathbf{Z}_I = \mathrm{diag}\,(Z_1, \ldots, Z_{N_I})$, $Z_{n_I} = jX_{n_I}$, $X_{n_I} \in \mathbb{R}$, $\forall n_I$	$\mathbf{Y}_I = \mathrm{diag}\,(Y_1, \ldots, Y_{N_I})$, $Y_{n_I} = jB_{n_I}$, $B_{n_I} \in \mathbb{R}$, $\forall n_I$	$\mathbf{\Theta} = \mathrm{diag}\,(\Theta_1, \ldots, \Theta_{N_I})$, $\Theta_{n_I} = e^{j\theta_{n_I}}$, $\theta_{n_I} \in [0, 2\pi)$, $\forall n_I$
Fully-connected	$\mathbf{Z}_I = j\mathbf{X}_I$, $\mathbf{X}_I = \mathbf{X}_I^T$, $\mathbf{X}_I \in \mathbb{R}^{N_I \times N_I}$	$\mathbf{Y}_I = j\mathbf{B}_I$, $\mathbf{B}_I = \mathbf{B}_I^T$, $\mathbf{B}_I \in \mathbb{R}^{N_I \times N_I}$	$\mathbf{\Theta}^H\mathbf{\Theta} = \mathbf{I}$, $\mathbf{\Theta} = \mathbf{\Theta}^T$
Group-connected	$\mathbf{Z}_I = \mathrm{diag}\,(\mathbf{Z}_{I,1}, \ldots, \mathbf{Z}_{I,G})$, $\mathbf{Z}_{I,g} = j\mathbf{X}_{I,g}$, $\mathbf{X}_{I,g} = \mathbf{X}_{I,g}^T$, $\mathbf{X}_{I,g} \in \mathbb{R}^{N_G \times N_G}$, $\forall g$	$\mathbf{Y}_I = \mathrm{diag}\,(\mathbf{Y}_{I,1}, \ldots, \mathbf{Y}_{I,G})$, $\mathbf{Y}_{I,g} = j\mathbf{B}_{I,g}$, $\mathbf{B}_{I,g} = \mathbf{B}_{I,g}^T$, $\mathbf{B}_{I,g} \in \mathbb{R}^{N_G \times N_G}$, $\forall g$	$\mathbf{\Theta} = \mathrm{diag}\,(\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_G)$, $\mathbf{\Theta}_g^H\mathbf{\Theta}_g = \mathbf{I}$, $\mathbf{\Theta}_g = \mathbf{\Theta}_g^T$, $\forall g$
Tree-connected⋆	–	$\mathbf{Y}_I = j\mathbf{B}_I$, $\mathbf{B}_I = \mathbf{B}_I^T$, $\mathbf{B}_I \in \mathbb{R}^{N_I \times N_I}$, $[\mathbf{B}_I]_{i,j} = 0$ if $\vert i-j\vert > 1$	–
Forest-connected⋆	–	$\mathbf{Y}_I = \mathrm{diag}\,(\mathbf{Y}_{I,1}, \ldots, \mathbf{Y}_{I,G})$, $\mathbf{Y}_{I,g} = j\mathbf{B}_{I,g}$, $\mathbf{B}_{I,g} = \mathbf{B}_{I,g}^T$, $\mathbf{B}_{I,g} \in \mathbb{R}^{N_G \times N_G}$, $[\mathbf{B}_{I,g}]_{i,j} = 0$ if $\vert i-j\vert > 1$, $\forall g$	–

⋆ For these RIS families, the tridiagonal RIS architecture is considered [80], and the constraints cannot be expressed in terms of the Z- and S-parameters.

To overcome the limited flexibility of conventional RIS, the ports in BD-RIS can also be connected to each other through additional tunable impedance (or admittance) components. By denoting the tunable admittance connecting the n_Ith port to the m_Ith port as Y_{n_I,m_I}, the (n_I, m_I)th entry of the admittance matrix is given by

$$[\mathbf{Y}_I]_{n_I,m_I} = \begin{cases} -Y_{n_I,m_I} & n_I \neq m_I \\ Y_{n_I} + \sum_{k \neq n_I} Y_{n_I,k} & n_I = m_I \end{cases}, \tag{8.96}$$

as discussed in [16]. Thus, the admittance matrix of BD-RIS is not limited to being diagonal. The same holds for the impedance and scattering matrices of BD-RIS, as it can be easily noticed by applying $\mathbf{Z}_I = \mathbf{Y}_I^{-1}$ and (8.18). Depending on the topology of the interconnections between the RIS ports, multiple RIS architectures have been proposed. In Tab. 8.3, we report the constraint of fully-/group-connected RIS [16], and tree-/forest-connected RIS with the tridiagonal architecture [80] based on the three parameters. In group- and forest-connected RIS, the elements are grouped into G groups, each with size $N_G = N/G$, such that $\mathbf{Z}_{I,g} \in \mathbb{C}^{N_G \times N_G}$, $\mathbf{Y}_{I,g} \in \mathbb{C}^{N_G \times N_G}$, and $\mathbf{\Theta}_g \in \mathbb{C}^{N_G \times N_G}$ are the impedance, admittance, and scattering matrix of the gth group, for $g = 1, \ldots, G$, respectively [16, 80]. The N_I-port network of the RIS is often assumed to be reciprocal (not including non-reciprocal media) and lossless (to maximize the scattered power). If the RIS is reciprocal, the impedance, admittance, and scattering matrices are symmetric. Furthermore, if the RIS is lossless, the impedance and admittance matrices are purely imaginary while the scattering matrix is unitary, as shown in Tab. 8.3.

8.5.2 RIS Optimization Based on the Z-, Y-, and S-Parameters

Since Z-, Y-, and S-parameters can be used interchangeably to characterize the RIS architecture, we now highlight the advantages of using the different parameters to optimize the RIS. To this end, we consider three case studies derived from the received signal power maximization problem in an RIS-aided MIMO system. In this communication system, the transmitted signal is expressed as $\mathbf{v}_T = \mathbf{w}s$, where $\mathbf{w} \in \mathbb{C}^{N_T \times 1}$ is the normalized precoder subject to $\|\mathbf{w}\| = 1$, and $s \in \mathbb{C}$ is the transmitted symbol with average power $P_T = \mathrm{E}[|s|^2]$. Besides, the signal used for detection is given by $z = \mathbf{g}\mathbf{v}_R \in \mathbb{C}$, where $\mathbf{g} \in \mathbb{C}^{1 \times N_R}$ is the normalized combiner subject to the constraint $\|\mathbf{g}\| = 1$, and $\mathbf{v}_R = \mathbf{H}\mathbf{v}_T$ is the received signal. Thus, the received signal power is given by $P_R = P_T|\mathbf{g}\mathbf{H}\mathbf{w}|^2$.

Z-parameters proved to be the effective representation when optimizing the RIS with mutual coupling [97, 98, 99, 100]. Specifically, RIS characterized through the Z-parameters have been optimized by applying the Neumann series approximation [97, 98], by iteratively optimizing the tunable loads in closed-form [99], and trough gradient ascent [100]. Furthermore, it has been shown that Z-parameters facilitate the optimization of BD-RIS in the presence of mutual coupling at the RIS [55].

As a case study, we consider the received signal power maximization problem in an RIS-aided system in the presence of mutual coupling, with the RIS being group-connected, including single- and fully-connected as two special cases [16]. Considering lossless and reciprocal RIS, i.e., with $\mathbf{Z}_I = j\mathbf{X}_I$ and $\mathbf{X}_I = \mathbf{X}_I^T$, where $\mathbf{X}_I$ is the RIS reactance matrix, and $\mathbf{H}$ given as in (8.72), our problem writes as

$$\max_{\mathbf{w},\mathbf{g},\mathbf{X}_I} \frac{P_T}{4Z_0^2} \left|\mathbf{g}\left(\mathbf{Z}_{RT} - \mathbf{Z}_{RI}\left(j\mathbf{X}_I + \mathbf{Z}_{II}\right)^{-1}\mathbf{Z}_{IT}\right)\mathbf{w}\right|^2 \tag{8.97}$$

$$\text{s.t.} \quad \mathbf{X}_I = \operatorname{diag}\left(\mathbf{X}_{I,1}, \ldots, \mathbf{X}_{I,G}\right), \tag{8.98}$$

$$\mathbf{X}_{I,g} = \mathbf{X}_{I,g}^T, \ \forall g, \tag{8.99}$$

$$\|\mathbf{w}\| = 1, \ \|\mathbf{g}\| = 1, \tag{8.100}$$

which is solved by jointly optimizing $\mathbf{w}$, $\mathbf{g}$, and $\mathbf{X}_I$. To solve (8.97)-(8.100), we initialize $\mathbf{X}_I$ to a feasible value and alternate between the following two steps. First, with $\mathbf{X}_I$ fixed, $\mathbf{w}$ and $\mathbf{g}$ are updated as the dominant right and left singular vectors of $\mathbf{Z}_{RT} - \mathbf{Z}_{RI}(j\mathbf{X}_I + \mathbf{Z}_{II})^{-1}\mathbf{Z}_{IT}$, respectively, which is globally optimal. Second, with $\mathbf{w}$ and $\mathbf{g}$ fixed, $\mathbf{X}_I$ is updated by solving

$$\max_{\mathbf{X}_I} \left|z_{RT}^{\text{eff}} - \mathbf{z}_{RI}^{\text{eff}}\left(j\mathbf{X}_I + \mathbf{Z}_{II}\right)^{-1}\mathbf{z}_{IT}^{\text{eff}}\right|^2 \text{ s.t. (8.98), (8.99),} \tag{8.101}$$

where $z_{RT}^{\text{eff}} = \mathbf{g}\mathbf{Z}_{RT}\mathbf{w}$, $\mathbf{z}_{RI}^{\text{eff}} = \mathbf{g}\mathbf{Z}_{RI}$, and $\mathbf{z}_{IT}^{\text{eff}} = \mathbf{Z}_{IT}\mathbf{w}$, as proposed in [55]. These two steps are alternatively repeated until convergence of the objective (8.97).

RIS optimization based on Y-parameters is necessary in the case of BD-RIS whose admittance matrix $\mathbf{Y}_I$ is sparse, i.e., there is only a limited number of tunable admittance components interconnecting the RIS elements, such as in tree- and forest-connected RISs [80]. In this case, Y-parameters are the effective representation since the entries of $\mathbf{Y}_I$ are directly linked to the tunable admittance components in the BD-RIS circuit topology, as given by (8.96).

As a case study, we consider the received signal power maximization problem in an RIS-aided system, with the RIS being forest-connected, including single- and tree-connected as two special cases [80]. Considering lossless and reciprocal RIS, i.e., with $\mathbf{Y}_I = j\mathbf{B}_I$ and $\mathbf{B}_I = \mathbf{B}_I^T$, where $\mathbf{B}_I$ is the RIS susceptance matrix, and $\mathbf{H}$ given as in (8.75), assuming no mutual coupling, our problem writes as

$$\max_{\mathbf{w},\mathbf{g},\mathbf{B}_I} \frac{P_T}{4Y_0^2}\left|\mathbf{g}\left(-\mathbf{Y}_{RT}+\mathbf{Y}_{RI}\left(j\mathbf{B}_I+Y_0\mathbf{I}\right)^{-1}\mathbf{Y}_{IT}\right)\mathbf{w}\right|^2 \tag{8.102}$$

$$\text{s.t.}\quad \mathbf{B}_I=\text{diag}\left(\mathbf{B}_{I,1},\ldots,\mathbf{B}_{I,G}\right), \tag{8.103}$$

$$\mathbf{B}_{I,g}=\mathbf{B}_{I,g}^T,\ [\mathbf{B}_{I,g}]_{i,j}=0 \text{ if } |i-j|>1, \forall g, \tag{8.104}$$

$$\|\mathbf{w}\|=1,\ \|\mathbf{g}\|=1, \tag{8.105}$$

where (8.104) indicates that the gth block of the susceptance matrix $\mathbf{B}_g$ has a tridiagonal architecture, as introduced in [80]. To jointly optimize $\mathbf{w}$, $\mathbf{g}$, and $\mathbf{B}_I$, we initialize $\mathbf{B}_I$ to a feasible value and alternate between the following two steps. First, with $\mathbf{B}_I$ fixed, $\mathbf{w}$ and $\mathbf{g}$ are updated as the dominant right and left singular vectors of $-\mathbf{Y}_{RT}+\mathbf{Y}_{RI}(j\mathbf{B}_I+Y_0\mathbf{I})^{-1}\mathbf{Y}_{IT}$, respectively, which is a global optimal solution. Second, with $\mathbf{w}$ and $\mathbf{g}$ fixed, $\mathbf{B}_I$ is updated by solving

$$\max_{\mathbf{B}_I}\left|-y_{RT}^{\text{eff}}+\mathbf{y}_{RI}^{\text{eff}}\left(j\mathbf{B}_I+Y_0\mathbf{I}\right)^{-1}\mathbf{y}_{IT}^{\text{eff}}\right|^2 \text{ s.t. (8.103), (8.104)}, \tag{8.106}$$

where $y_{RT}^{\text{eff}}=\mathbf{g}\mathbf{Y}_{RT}\mathbf{w}$, $\mathbf{y}_{RI}^{\text{eff}}=\mathbf{g}\mathbf{Y}_{RI}$, and $\mathbf{y}_{IT}^{\text{eff}}=\mathbf{Y}_{IT}\mathbf{w}$. Remarkably, (8.106) has been solved in [80], where a globally optimal solution for each block $\mathbf{B}_{I,g}$ has been proposed. These two steps are alternatively repeated until convergence of (8.102).

S-parameters have been commonly used to optimize conventional RIS in the vast majority of related works [13]. S-parameters own their popularity to the widely used channel model (8.77), in which $\mathbf{\Theta}=\text{diag}\left(e^{j\theta_1},\ldots,e^{j\theta_{N_I}}\right)$ in the case of a lossless single-connected RIS. For single-connected RIS, S-parameters allow to directly optimize the phase shifts $\theta_{n_I}\in[0,2\pi)$ through several optimization techniques, such as semidefinite relaxation (SDR) [17]. For BD-RIS, the S-parameters enable the use of manifold optimization by exploiting the unitary constraint on $\mathbf{\Theta}$ [65]-[67]. In addition, the S-parameters allow for efficient optimization of BD-RIS by using tailored decompositions of $\mathbf{\Theta}$ [75], through symmetric unitary projection [50], and by solving the orthogonal Procrustes problem [68].

As a case study, we consider the received signal power maximization problem in an RIS-aided system, with the RIS being group-connected. Considering lossless and reciprocal RIS, i.e., with $\mathbf{\Theta}^H\mathbf{\Theta}=\mathbf{I}$ and $\mathbf{\Theta}=\mathbf{\Theta}^T$, and $\mathbf{H}$ given as in (8.77), assuming no mutual coupling, our problem writes as

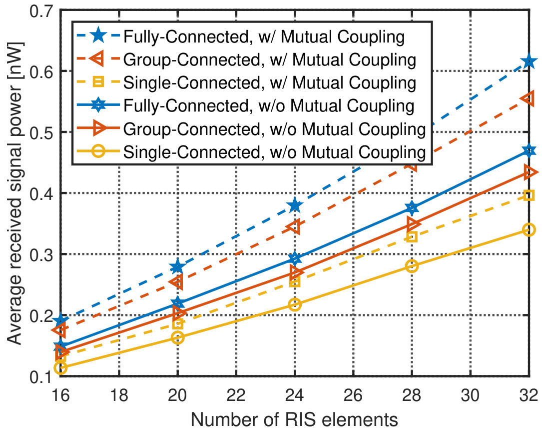
$$\max_{\mathbf{w},\mathbf{g},\mathbf{\Theta}} P_T\left|\mathbf{g}\left(\mathbf{S}_{RT}+\mathbf{S}_{RI}\mathbf{\Theta}\mathbf{S}_{IT}\right)\mathbf{w}\right|^2 \tag{8.107}$$

$$\text{s.t.}\quad \mathbf{\Theta}=\text{diag}\left(\mathbf{\Theta}_1,\ldots,\mathbf{\Theta}_G\right), \tag{8.108}$$

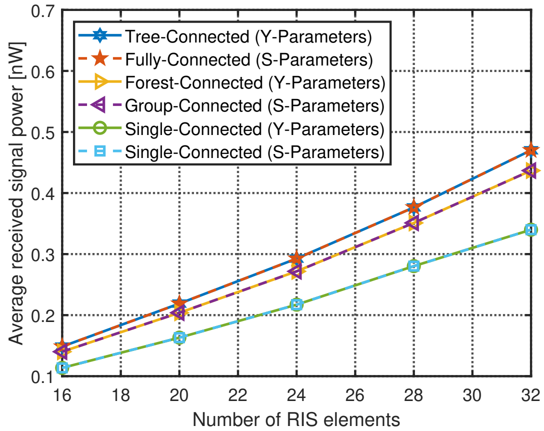
$$\mathbf{\Theta}_g=\mathbf{\Theta}_g^T,\ \mathbf{\Theta}_g^H\mathbf{\Theta}_g=\mathbf{I},\ \forall g, \tag{8.109}$$

$$\|\mathbf{w}\|=1,\ \|\mathbf{g}\|=1, \tag{8.110}$$

which is solved by jointly optimizing $\mathbf{w}$, $\mathbf{g}$, and $\mathbf{\Theta}$. To this end, we initialize $\mathbf{\Theta}$ to a feasible value and alternate between the following two steps. First, with $\mathbf{\Theta}$ fixed, $\mathbf{w}$ and $\mathbf{g}$ are updated as the dominant right and left singular vectors of $\mathbf{S}_{RT}+\mathbf{S}_{RI}\mathbf{\Theta}\mathbf{S}_{IT}$, respectively,



(a) Z-parameters, w/ and w/o mutual coupling.



(b) Y- and S-parameters, w/o mutual coupling.

Figure 8.3: Average received signal power maximized using Z-, Y-, and S-parameters.

which is globally optimal. Second, with $\mathbf{w}$ and $\mathbf{g}$ fixed, $\boldsymbol{\Theta}$ is updated by solving

$$\max_{\boldsymbol{\Theta}} \left| s_{RT}^{\text{eff}} + \mathbf{s}_{RI}^{\text{eff}} \boldsymbol{\Theta} \mathbf{s}_{IT}^{\text{eff}} \right|^2 \text{ s.t. } (8.108),\ (8.109), \tag{8.111}$$

where $s_{RT}^{\text{eff}} = \mathbf{g}\mathbf{S}_{RT}\mathbf{w}$, $\mathbf{s}_{RI}^{\text{eff}} = \mathbf{g}\mathbf{S}_{RI}$, and $\mathbf{s}_{IT}^{\text{eff}} = \mathbf{S}_{IT}\mathbf{w}$. Remarkably, (8.111) has been solved for group-connected RIS through a closed-form global optimally solution in [75]. These two steps are alternatively repeated until convergence of (8.107).

8.6 Performance Evaluation

In this section, we evaluate the performance obtained by solving the three optimization problems presented in Section 8.5, using Z-, Y-, and S-parameters. The transmitter, RIS, and receiver are located at $(0,0)$, $(50,2)$, and $(52,0)$ in meters (m), respectively. We set $N_T = 2$ and $N_R = 2$, and assume that the direct channel between the transmitter and receiver is completely obstructed, i.e., $\mathbf{Z}_{RT} = \mathbf{0}$. For the large-scale path loss of the channels from the RIS to the receiver and from the transmitter to the RIS, we use the distance-dependent path loss model $L_{ij}(d_{ij}) = L_0 d_{ij}^{-\alpha_{ij}}$, where L_0 is the reference path loss at distance 1 m, d_{ij} is the distance, and α_{ij} is the path loss exponent, for $ij \in \{RI, IT\}$. We set $L_0 = -30$ dB, $\alpha_{RI} = 2.8$, $\alpha_{IT} = 2$, and $P_T = 10$ mW. For the small-scale fading, we model the channels as i.i.d. Rayleigh, i.e., $\mathbf{S}_{RI} \sim \mathcal{CN}(\mathbf{0}, L_{RI}\mathbf{I})$ and $\mathbf{S}_{IT} \sim \mathcal{CN}(\mathbf{0}, L_{IT}\mathbf{I})$. Given $\mathbf{Z}_{RT}$, $\mathbf{S}_{RI}$, and $\mathbf{S}_{IT}$, we obtain the rest of the off-diagonal blocks of $\mathbf{Z}$, $\mathbf{Y}$, and $\mathbf{S}$ through (8.79) and (8.82).

In Fig. 8.3, we report the average received signal power obtained by optimizing the RIS through Z-, Y-, and S-parameters. In the case of group- and forest-connected RISs, the group size is $N_G = 4$. The schemes "w/o Mutual Coupling" are obtained by setting $\mathbf{Z}_{II} = Z_0\mathbf{I}$. Besides, the schemes "w/ Mutual Coupling" in Fig. 8.3(a) are obtained by modeling $\mathbf{Z}_{II}$ as follows: its diagonal entries are set to Z_0, assuming perfectly matched RIS elements, while its off-diagonal entries are modeled as in [55], considering the RIS antennas being dipoles with length $\ell = \lambda/4$ and inter-element distance $d = \lambda/4$, where $\lambda = c/f$ is the wavelength with frequency $f = 28$ GHz. We make the following observations. *First*, BD-

RIS outperforms single-connected RIS in the presence of mutual coupling because of its higher flexibility, in agreement with [55], and also with no mutual coupling since Rayleigh small-scale fading is considered [16, 80]. *Second*, the presence of mutual coupling between the RIS elements increases the performance, in accordance with [100]. We observe this gain for all the considered RIS architectures by assuming that the RIS mutual coupling matrix $\mathbf{Z}_{II}$ is perfectly known during the optimization process. *Third*, in the absence of mutual coupling, single-/group-/fully-connected RIS architectures optimized with the Z-parameters achieve the same performance as the same architectures optimized with the S-parameters, providing additional evidence of the equivalence of the two analyses. In our case study, optimizing based on the S-parameters is preferred since closed-form solutions are available, leading to a low-complexity optimization process [75]. *Fourth*, in the absence of mutual coupling, single-/group-/fully-connected RISs optimized with the S-parameters achieve the same performance as single-/forest-/tree-connected RISs optimized with the Y-parameters, respectively, in agreement with [80]. In our case study, forest-/tree-connected RISs are preferred over group-/fully-connected RISs since they are characterized by a reduced circuit complexity [80].

8.7 Conclusion

We introduced a universal framework to perform multiport network analysis of an RIS-aided communication system. The proposed framework is used to analyze the RIS-aided system based on the Z-, Y-, and S-parameters. Based on these three independent analyses, three equivalent channel models are derived accounting for the effects of impedance mismatching and mutual coupling at the transmitter, RIS, and receiver. Subsequently, to gain insights into the role of the RIS in the communication model, these models are simplified by assuming large transmission distances, perfect matching, and no mutual coupling. The equivalence between the obtained simplified models is shown by providing the mappings between the different parameters of the three representations. The derived simplified channel model is consistent with the channel model widely used in related literature. However, we show that an additional approximation is commonly considered in related literature, whose impact in terms of received signal power vanishes as the number of RIS elements increases. To analyze the impact of this approximation, i.e., the omission of the RIS structural scattering, in practical RIS-aided scenarios, including when the RIS is used in hybrid reflective and transmissive mode, constitutes an interesting research direction.

Since Z-, Y-, and S-parameters are equivalent representations, we discuss the advantages of each of them in the characterization of the RIS architecture and its optimization. To this end, we present three case studies and show that it is convenient to solve each of them by using a different multiport network model representation. Numerical results further support the equivalence of the three analyses.

Chapter 9

BD-RIS With Long and Lossy Interconnections

Although BD-RIS working in hybrid mode and multi-sector BD-RIS have shown great coverage improvements [65, 66], an RIS effectively enhances the link between a transmitter and a receiver only when placed close to one of them [13]. For this reason, an RIS can be particularly beneficial only to users in its proximity, yielding reduced coverage capabilities. This reduced coverage limitation arises from the fact that RIS has been commonly regarded as a localized array of scattering elements, here denoted as *localized RIS*. In this chapter, we overcome this limitation by investigating RIS made of a distributed array of elements, in which the inter-element distance can be much longer than the wavelength, referred to as *distributed RIS*. Our focus will extend specifically to distributed BD-RIS, as they are expected to achieve particularly enhanced performance due to the presence of tunable impedance components interconnecting the RIS elements. These interconnections have the potential to effectively guide the EM signal toward the receiver by enabling its propagation within the RIS. Distributed BD-RIS could be implemented by integrating the RIS elements into long cables, forming long linear arrays. Specifically, distributed BD-RIS could drive the EM signal from a fixed base station (BS) to the mobile user equipments (UEs) within a coverage area proportional to their array length in highly obstructed indoor environments, e.g., smart factories, as well as in urban outdoor settings.

The investigation of distributed BD-RIS requires modeling the interconnections in the BD-RIS architecture by accounting for their losses, which is a challenge overlooked in previous literature on BD-RIS. In previous works, BD-RIS has been modeled through a lumped-element circuit model, valid when the physical dimensions of the RIS reconfigurable impedance network are much smaller than the wavelength. According to the lumped-element model, voltages and currents are constant along the interconnections, which simplifies the BD-RIS modeling and optimization. However, the length of the interconnections between the RIS elements in BD-RIS can be a considerable fraction of the wavelength in localized BD-RIS, or many wavelengths in distributed BD-RIS. For example, even when the inter-element distance is half-wavelength, the interconnections are at least half-wavelength long. In the case the interconnections are not much smaller than the wavelength, two critical effects take place. First, voltages and currents vary in phase along the interconnections. Second, voltages and currents may vary in magnitude along the interconnections due to losses. Thus, it is crucial to develop a BD-RIS model accounting for these two effects, as they may lead to undesired changes of phase and dissipation of power within the BD-RIS circuit. To this end, in this chapter, we model and optimize localized and distributed BD-RIS by characterizing its interconnections through transmission line theory. Our contributions are summarized as follows.

First, we propose the concept of distributed RIS, in which the antenna elements are not localized in a specific site but distributed over a wide region. We analyze their scaling laws and quantify the gain of distributed RIS over localized RIS, and of distributed BD-RIS over distributed conventional RIS. The derived scaling laws show that lossless distributed BD-RIS can offer substantial gains over distributed conventional RIS and localized BD-RIS, as high as several orders of magnitude. These gains are enabled by the interconnections within BD-RIS, which allow the guided propagation of the EM signal within the BD-RIS.

Second, we model BD-RIS (localized and distributed) with lossy interconnections by using transmission line theory. Specifically, we model the interconnections present in the BD-RIS circuit topology as tunable impedance components in series with lossy transmission lines. Thus, we derive the expression of the BD-RIS admittance matrix as a function of the tunable impedance components and the transmission line parameters.

Third, since it is difficult to derive engineering insights from the derived BD-RIS model, we obtain three simplified models. First, by making specific assumptions on the interconnection lengths, we show how losses impact the BD-RIS admittance matrix elements. Second, by assuming lossless interconnections, we illustrate how the presence of long interconnections affects the lossless BD-RIS model. Third, by assuming lossless interconnections with specific lengths, it is shown how the derived model boils down to the BD-RIS model widely considered in previous literature.

Fourth, we optimize lossy BD-RIS based on the proposed models and assess its performance, compared with conventional RIS. Numerical results show that the performance of localized BD-RIS is only slightly impacted by losses given the short interconnection lengths. Besides, the performance of distributed RIS, even with losses, can achieve orders of magnitude of gains over conventional RIS.

Organization: In Section 9.1, we introduce the localized and distributed BD-RIS-aided system model. In Section 9.2, we derive and compare the scaling laws of localized and distributed BD-RIS. In Section 9.3, we model BD-RIS with lossy interconnections through transmission line theory. In Section 9.4, we derive three simplified models for lossy BD-RIS. In Section 9.5, we optimize BD-RIS characterized with the proposed models. In Section 9.6, we present the numerical results. Finally, Section 9.7 concludes this chapter.

The content of this chapter has appeared in [101].

9.1 Localized and Distributed BD-RIS-Aided System Model

Consider a MIMO system between an N_T-antenna transmitter and an N_R-antenna receiver aided by an N-element BD-RIS. The N elements of the BD-RIS are connected to a N-port reconfigurable impedance network, with scattering matrix $\mathbf{\Theta} \in \mathbb{C}^{N \times N}$. Given the matrix $\mathbf{\Theta}$, the wireless channel $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ writes as $\mathbf{H} = \mathbf{H}_{RT} + \mathbf{H}_R \mathbf{\Theta} \mathbf{H}_T$, where $\mathbf{H}_{RT} \in \mathbb{C}^{N_R \times N_T}$, $\mathbf{H}_R \in \mathbb{C}^{N_R \times N}$, and $\mathbf{H}_T \in \mathbb{C}^{N \times N_T}$ refer to the channels from transmitter to receiver, RIS to receiver, and transmitter to RIS, respectively [16].

In related literature, an RIS has been regarded as an antenna array with an inter-element distance comparable with the wavelength. We refer to this type of RIS as localized RIS, as its elements are all localized in a specific site. With a localized RIS, the channels $\mathbf{H}_R$ and $\mathbf{H}_T$ have been commonly modeled as $\mathbf{H}_R^{\text{Loc}} = \sqrt{\rho_R}\widetilde{\mathbf{H}}_R$ and $\mathbf{H}_T^{\text{Loc}} = \sqrt{\rho_T}\widetilde{\mathbf{H}}_T$, where ρ_R and ρ_T are the path-gains while $\widetilde{\mathbf{H}}_R \in \mathbb{C}^{N_R \times N}$ and $\widetilde{\mathbf{H}}_T \in \mathbb{C}^{N \times N_T}$ are the small-scale fading effects. Opposed to localized RISs, we investigate in this chapter the performance of distributed RISs, i.e., RISs whose elements are distributed over a wide region, with an inter-element distance that can be much longer than the wavelength. In the presence of a distributed RIS, the channel $\mathbf{H}_R$ is expressed as $\mathbf{H}_R^{\text{Dis}} = \widetilde{\mathbf{H}}_R \mathbf{R}_R^{1/2}$, where $\mathbf{R}_R = \text{diag}(\boldsymbol{\rho}_R)$, with $\boldsymbol{\rho}_R \in \mathbb{R}^{N \times 1}$ introduced such that $[\boldsymbol{\rho}_R]_n$ is the path-gain from the nth RIS element to the receiver. Similarly, we have $\mathbf{H}_T^{\text{Dis}} = \mathbf{R}_T^{1/2}\widetilde{\mathbf{H}}_T$, where $\mathbf{R}_T = \text{diag}(\boldsymbol{\rho}_T)$, with $\boldsymbol{\rho}_T \in \mathbb{R}^{N \times 1}$ introduced such that $[\boldsymbol{\rho}_T]_n$ is the path-gain from the transmitter to the nth RIS element.

In this BD-RIS-aided system, we denote the transmitted signal as $\mathbf{x} = \mathbf{w}s \in \mathbb{C}^{N_T \times 1}$, where $\mathbf{w} \in \mathbb{C}^{N_T \times 1}$ is the precoding vector subject to $\|\mathbf{w}\| = 1$, and $s \in \mathbb{C}$ is the transmitted symbol with average power $P_T = \text{E}[|s|^2]$. Denoting the received signal as $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$, we have $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is the AWGN at the receiver with power σ_n^2. By using a combining vector $\mathbf{g} \in \mathbb{C}^{1 \times N_R}$ subject to $\|\mathbf{g}\| = 1$, the signal used for detection $\hat{s} = \mathbf{g}\mathbf{y}$ can be expressed as

$$\hat{s} = \mathbf{g}\left(\mathbf{H}_{RT} + \mathbf{H}_R \mathbf{\Theta} \mathbf{H}_T\right)\mathbf{w}s + \tilde{n}, \tag{9.1}$$

where $\tilde{n} = \mathbf{g}\mathbf{n}$ is the AWGN with power σ_n^2. Thus, when reconfiguring the BD-RIS, $\mathbf{\Theta}$ is optimized jointly with $\mathbf{g}$ and $\mathbf{w}$ to maximize the received signal power, given by

$$P_R = P_T \left|\mathbf{g}\left(\mathbf{H}_{RT} + \mathbf{H}_R \mathbf{\Theta} \mathbf{H}_T\right)\mathbf{w}\right|^2. \tag{9.2}$$

In the following, we analyze and compare the received signal power scaling laws of localized and distributed BD-RIS.

9.2 Scaling Laws

To compare the fundamental limits of localized and distributed BD-RIS, we derive their received signal power scaling laws. For simplicity, we consider a SISO system, i.e., $N_R = 1$ and $N_T = 1$, with obstructed direct channel, i.e., $\mathbf{H}_{RT} = \mathbf{0}$, and lossless BD-RIS, i.e., with unitary $\mathbf{\Theta}$. With no loss of generality, we assume transmit power $P_T = 1$, such that the received signal power writes as $P_R = |\mathbf{h}_R \mathbf{\Theta} \mathbf{h}_T|^2$. We model the path-gain of the channels $\mathbf{h}_R$ and $\mathbf{h}_T$ through the distance-dependent model and their small-scale fading as i.i.d. Rayleigh distributed. Specifically, for localized BD-RIS, we have $\mathbf{h}_i^{\text{Loc}} \sim \mathcal{CN}(\mathbf{0}, \rho_i \mathbf{I})$, with $\rho_i = C_0 d_i^{-a}$, for $i \in \{R, T\}$, where C_0 refers to the path-gain at the reference distance 1 m, a is the path-loss exponent, and d_R (resp. d_T) is the distance between the RIS and the receiver (resp. the transmitter). Similarly, for distributed BD-RIS, we have $\mathbf{h}_i^{\text{Dis}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_i)$, with $\mathbf{R}_i = \text{diag}(\boldsymbol{\rho}_i)$, $[\boldsymbol{\rho}_i]_n = C_0[\mathbf{d}_i]_n^{-a}$, for $i \in \{R, T\}$, where $[\mathbf{d}_R]_n$

(resp. $[\mathbf{d}_T]_n$) is the distance between the nth RIS element and the receiver (resp. the transmitter).

In the following, we study the localized and distributed versions of single- and fully-connected RIS, which are the least and the most complex BD-RIS architectures, respectively. Single-connected RIS is the conventional RIS characterized by a unitary and diagonal scattering matrix, i.e., $\boldsymbol{\Theta} = \mathrm{diag}(e^{j\theta_1}, \ldots, e^{j\theta_N})$, with $\theta_n \in [0, 2\pi)$, for $n = 1, \ldots, N$. Besides, fully-connected RIS is the BD-RIS architecture with the highest flexibility as it is characterized by an arbitrary symmetric and unitary scattering matrix [16].

For localized single-connected RIS, it is possible to optimize the phase shifts θ_n to obtain a received signal power given by

$$P_R^{\text{Loc-SC}} = \left(\sum_{n=1}^{N} \left| \left[\mathbf{h}_R^{\text{Loc}} \right]_n \right| \left| \left[\mathbf{h}_T^{\text{Loc}} \right]_n \right| \right)^2 . \tag{9.3}$$

Furthermore, the expected value of (9.3) can be derived by exploiting the i.i.d. channels assumption, and that $\mathrm{E}[|[\mathbf{h}_i^{\text{Loc}}]_n|^2] = \rho_i$ and $\mathrm{E}[|[\mathbf{h}_i^{\text{Loc}}]_n|] = \frac{\sqrt{\pi}}{2}\sqrt{\rho_i}$, for $i \in \{R, T\}$ and $n = 1, \ldots, N$, because of the moments of the chi distribution with 2 degrees of freedom. Thus, it is possible to show that

$$\mathrm{E}\left[P_R^{\text{Loc-SC}}\right] = \left(N + \frac{\pi^2}{16} N (N-1) \right) C_0^2 d_R^{-a} d_T^{-a}, \tag{9.4}$$

giving the scaling law of localized single-connected RIS [16].

In the case of localized fully-connected RIS, it has been shown that it is always possible to optimize the RIS to achieve a maximum received signal power given by

$$P_R^{\text{Loc-FC}} = \left\| \mathbf{h}_R^{\text{Loc}} \right\|^2 \left\| \mathbf{h}_T^{\text{Loc}} \right\|^2 , \tag{9.5}$$

whose expected value writes as

$$\mathrm{E}\left[P_R^{\text{Loc-FC}}\right] = N^2 C_0^2 d_R^{-a} d_T^{-a}, \tag{9.6}$$

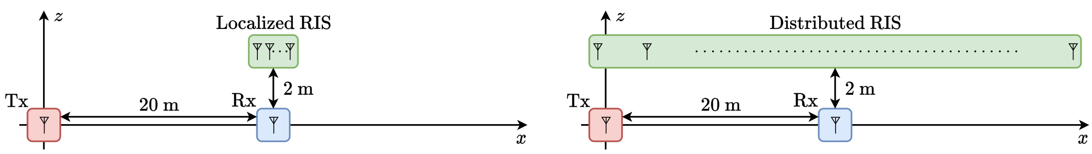
following the i.i.d. channels assumption and the moments of the chi distribution with 2 degrees of freedom [16].

For distributed single-connected RIS, the achievable received signal power writes as

$$P_R^{\text{Dis-SC}} = \left(\sum_{n=1}^{N} \left| \left[\mathbf{h}_R^{\text{Dis}} \right]_n \right| \left| \left[\mathbf{h}_T^{\text{Dis}} \right]_n \right| \right)^2 , \tag{9.7}$$

similarly to (9.3). By exploiting the i.i.d. channels assumption, and that $\mathrm{E}[|[\mathbf{h}_i^{\text{Dis}}]_n|^2] = [\boldsymbol{\rho}_i]_n$ and $\mathrm{E}[|[\mathbf{h}_i^{\text{Dis}}]_n|] = \frac{\sqrt{\pi}}{2}\sqrt{[\boldsymbol{\rho}_i]_n}$, for $i \in \{R, T\}$ and $n = 1, \ldots, N$, it is possible to show that

$$\mathrm{E}\left[P_R^{\text{Dis-SC}}\right] = C_0^2 \left(\sum_{n=1}^{N} \left([\mathbf{d}_R]_n [\mathbf{d}_T]_n \right)^{-a} + \frac{\pi^2}{16} \sum_{n \neq m} \left([\mathbf{d}_R]_n [\mathbf{d}_T]_n [\mathbf{d}_R]_m [\mathbf{d}_T]_m \right)^{-\frac{a}{2}} \right), \tag{9.8}$$



(a) (b)

Figure 9.1: (a) Localized and (b) distributed RIS-aided communication system.

giving the scaling law of distributed single-connected RIS.

Finally, in the case of distributed fully-connected RIS, the achievable received signal power is

$$P_R^{\text{Dis-FC}} = \left\|\mathbf{h}_R^{\text{Dis}}\right\|^2 \left\|\mathbf{h}_T^{\text{Dis}}\right\|^2, \tag{9.9}$$

with expected value

$$\mathrm{E}\left[P_R^{\text{Dis-FC}}\right] = C_0^2 \sum_{n=1}^{N} [\mathbf{d}_R]_n^{-a} \sum_{n=1}^{N} [\mathbf{d}_T]_n^{-a}, \tag{9.10}$$

following again the i.i.d. channels assumption and the moments of the chi distribution with 2 degrees of freedom [1].

In the following, we compare the four scaling laws derived in (9.4), (9.6), (9.8), and (9.10) by analyzing four gains: *i)* the gain of fully- over single-connected localized RIS, *ii)* the gain of fully- over single-connected distributed RIS, *iii)* the gain of distributed over localized single-connected RIS, and *iv)* the gain of distributed over localized fully-connected RIS. First, we carry out a theoretical discussion valid for any values of the distances d_R, d_T, $\mathbf{d}_R$, and $\mathbf{d}_T$. Second, to numerically quantify these gains, we consider a three-dimensional coordinate system where the transmitter and receiver are located at $(0, 0, 0)$ and $(20, 0, 0)$ in meters (m), respectively. In the case of localized RIS, the RIS is located in $(20, 0, 2)$, as shown in Fig. 9.1(a). In the case of distributed RIS, the RIS elements are spread out over a long line. Specifically, the RIS is a ULA with elements uniformly placed between $(0, 0, 2)$ and $(40, 0, 2)$, as shown in Fig. 9.1(b).

9.2.1 Gain of Fully- over Single-Connected Localized RIS

The gain of fully-connected over single-connected localized RIS, which is defined as $\mathcal{G}^{\text{Loc}} = \mathrm{E}[P_R^{\text{Loc-FC}}]/\mathrm{E}[P_R^{\text{Loc-SC}}]$, writes as

$$\mathcal{G}^{\text{Loc}} = \frac{N}{1 + \frac{\pi^2}{16}(N-1)}, \tag{9.11}$$

following (9.4) and (9.6), in agreement with [16]. Remarkably, we observe that $\mathcal{G}^{\text{Loc}}$, which depends only on N, is lower bounded by $\mathcal{G}^{\text{Loc}} \geq 1$ and upper bounded by $\mathcal{G}^{\text{Loc}} < \frac{16}{\pi^2} \approx$

[1]Under LoS small-scale fading, i.e., $[\mathbf{h}_i^{\text{Loc}}]_n = \sqrt{\rho_i} e^{j\phi_{i,n}}$ and $[\mathbf{h}_i^{\text{Dis}}]_n = \sqrt{[\boldsymbol{\rho}_i]_n} e^{j\psi_{i,n}}$, for $i \in \{R, T\}$ and $n = 1, \ldots, N$, it is possible to prove that $\mathrm{E}[P_R^{\text{Loc-FC}}]$ and $\mathrm{E}[P_R^{\text{Dis-FC}}]$ remain as in (9.6) and (9.10), $\mathrm{E}[P_R^{\text{Loc-SC}}] = \mathrm{E}[P_R^{\text{Loc-FC}}]$, and $\mathrm{E}[P_R^{\text{Dis-SC}}] = C_0^2(\sum_n([\mathbf{d}_R]_n[\mathbf{d}_T]_n)^{-a/2})^2$. Given the similarity between the scaling laws under Rayleigh and LoS fading, the following insights on localized and distributed RIS obtained considering Rayleigh fading are also valid under more general Rician fading.

1.62. Thus, localized fully-connected RIS is beneficial over localized single-connected RIS, enabling a gain of at most 62% in a SISO system [16].

9.2.2 Gain of Fully- over Single-Connected Distributed RIS

We define the gain of fully-connected over single-connected distributed RIS as $\mathcal{G}^{\text{Dis}} = \text{E}[P_R^{\text{Dis-FC}}]/\text{E}[P_R^{\text{Dis-SC}}]$. By noticing that (9.8) is upper bounded by

$$\text{E}\left[P_R^{\text{Dis-SC}}\right] < C_0^2 \left(\sum_{n=1}^{N} \left(\left[\mathbf{d}_R\right]_n \left[\mathbf{d}_T\right]_n\right)^{-\frac{a}{2}}\right)^2, \tag{9.12}$$

we obtain that

$$\mathcal{G}^{\text{Dis}} > \frac{\sum_{n=1}^{N} \left[\mathbf{d}_R\right]_n^{-a} \sum_{n=1}^{N} \left[\mathbf{d}_T\right]_n^{-a}}{\left(\sum_{n=1}^{N} \left(\left[\mathbf{d}_R\right]_n \left[\mathbf{d}_T\right]_n\right)^{-\frac{a}{2}}\right)^2} = \left(\frac{M_{-\frac{a}{2}}\left(\mathbf{d}_R \odot \mathbf{d}_T\right)}{M_{-a}\left(\mathbf{d}_R\right) M_{-a}\left(\mathbf{d}_T\right)}\right)^a, \tag{9.13}$$

where $M_p(\boldsymbol{\xi}) = (\sum_{n=1}^{N}[\boldsymbol{\xi}]_n^p/N)^{1/p}$ denotes the generalized mean of the elements of $\boldsymbol{\xi} \in \mathbb{R}_+^{N\times 1}$ with exponent $p \in \mathbb{R}_*$, and $\odot$ denotes the Hadamard product. Thus, we have $\mathcal{G}^{\text{Dis}} \geq 1$ from the Cauchy-Schwarz inequality. Furthermore, since $\min(\boldsymbol{\xi}) < M_{-p}(\boldsymbol{\xi}) < \sqrt[p]{N}\min(\boldsymbol{\xi})$, we have

$$\mathcal{G}^{\text{Dis}} > \left(\frac{\min\left(\mathbf{d}_R \odot \mathbf{d}_T\right)}{\sqrt[a]{N^2}\min\left(\mathbf{d}_R\right)\min\left(\mathbf{d}_T\right)}\right)^a, \tag{9.14}$$

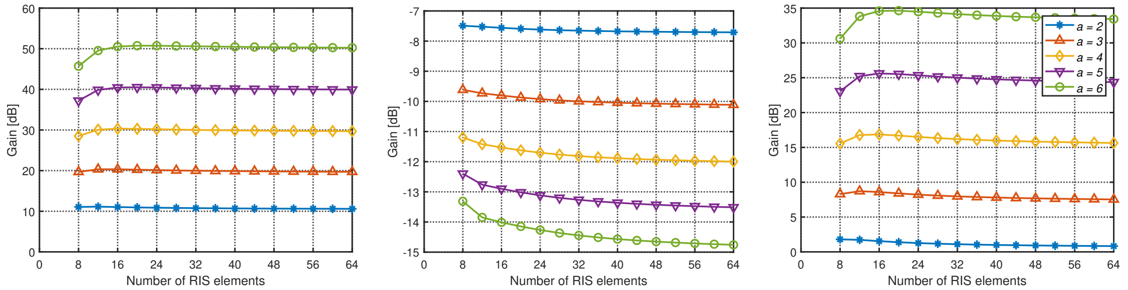
indicating that $\mathcal{G}^{\text{Dis}}$ increases exponentially with a and can be significantly high when both $\min\left(\mathbf{d}_R\right)$ and $\min\left(\mathbf{d}_T\right)$ are reduced, i.e., there is at least an RIS element close to the receiver and another one close to the transmitter.

We numerically quantify $\mathcal{G}^{\text{Dis}}$ in Fig. 9.2(a), where we consider the system represented in Fig. 9.1(b) with different values of path-loss exponent a and number of RIS elements N. From Fig. 9.2(a), we observe that $\mathcal{G}^{\text{Dis}}$ can be as high as several orders of magnitude, differently from $\mathcal{G}^{\text{Loc}}$, which is limited to 1.62. Fig. 9.2(a) confirms that $\mathcal{G}^{\text{Dis}}$ grows exponentially with a, and shows that the gain is only slightly affected by N.

Besides, in Fig. 9.3(a), we report the gain $\mathcal{G}^{\text{Dis}}$ for different locations of the receiver, located in $(x, y, 0)$, where $x \in [-10, 70]$ and $y \in [-40, 40]$, with $a = 4$ and $N = 64$. From Fig. 9.3(a), we observe that $\mathcal{G}^{\text{Dis}}$ is particularly high when the receiver is close to the distributed RIS and at the same time far from the transmitter. As visible from (9.14), $\mathcal{G}^{\text{Dis}}$ assumes high values when both the transmitter and the receiver are close to the distributed RIS. Thus, the high gains in Fig. 9.3(a) can be observed also for different locations of the transmitter, as long as it is located in proximity to an RIS element. Conversely, $\mathcal{G}^{\text{Dis}}$ decreases, approaching 1, when both the transmitter and the receiver are located far from all the RIS elements.

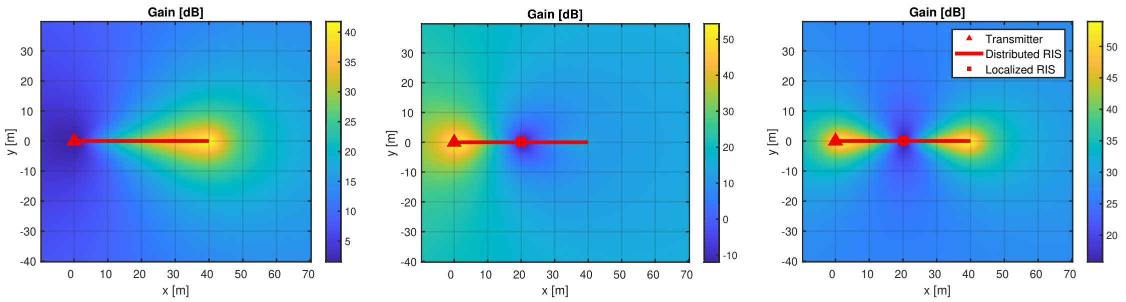
9.2.3 Gain of Distributed over Localized Single-Connected RIS

Consider now the gain of distributed over localized single-connected RIS, defined as $\mathcal{G}^{\text{SC}} = \text{E}[P_R^{\text{Dis-SC}}]/\text{E}[P_R^{\text{Loc-SC}}]$. To gain useful insights into $\mathcal{G}^{\text{SC}}$, we approximate this gain by



(a) $\mathcal{G}^{\text{Dis}}$. (b) $\mathcal{G}^{\text{SC}}$. (c) $\mathcal{G}^{\text{FC}}$.

Figure 9.2: (a) $\mathcal{G}^{\text{Dis}}$, (b) $\mathcal{G}^{\text{SC}}$, and (c) $\mathcal{G}^{\text{FC}}$ for different values of path-loss exponent and number of RIS elements.



(a) $\mathcal{G}^{\text{Dis}}$. (b) $\mathcal{G}^{\text{SC}}$. (c) $\mathcal{G}^{\text{FC}}$.

Figure 9.3: (a) $\mathcal{G}^{\text{Dis}}$, (b) $\mathcal{G}^{\text{SC}}$, and (c) $\mathcal{G}^{\text{FC}}$ for different locations of the receiver.

deriving an upper and a lower bound. First, by noticing that (9.8) is lower bounded by

$$\mathrm{E}\left[P_R^{\text{Dis-SC}}\right] > C_0^2 \frac{\pi^2}{16} \left(\sum_{n=1}^{N} \left([\mathbf{d}_R]_n [\mathbf{d}_T]_n \right)^{-\frac{a}{2}} \right)^2, \tag{9.15}$$

and that (9.4) is upper bounded by

$$\mathrm{E}\left[P_R^{\text{Loc-SC}}\right] < N^2 C_0^2 d_R^{-a} d_T^{-a}, \tag{9.16}$$

we obtain that

$$\mathcal{G}^{\text{SC}} > \frac{\frac{\pi^2}{16} \left(\sum_{n=1}^{N} \left([\mathbf{d}_R]_n [\mathbf{d}_T]_n \right)^{-\frac{a}{2}} \right)^2}{N^2 d_R^{-a} d_T^{-a}} = \frac{\pi^2}{16} \left(\frac{d_R d_T}{M_{-\frac{a}{2}} \left(\mathbf{d}_R \odot \mathbf{d}_T \right)} \right)^a. \tag{9.17}$$

Second, by recalling (9.12) and noticing that (9.4) is lower bounded by

$$\mathrm{E}\left[P_R^{\text{Loc-SC}}\right] > \frac{\pi^2}{16} N^2 C_0^2 d_R^{-a} d_T^{-a}, \tag{9.18}$$

we have

$$\mathcal{G}^{\text{SC}} < \frac{\left(\sum_{n=1}^{N} \left([\mathbf{d}_R]_n [\mathbf{d}_T]_n \right)^{-\frac{a}{2}} \right)^2}{\frac{\pi^2}{16} N^2 d_R^{-a} d_T^{-a}} = \frac{16}{\pi^2} \left(\frac{d_R d_T}{M_{-\frac{a}{2}} \left(\mathbf{d}_R \odot \mathbf{d}_T \right)} \right)^a. \tag{9.19}$$

Thus, considering (9.17) and (9.19), we can approximate $\mathcal{G}^{\text{SC}}$ as

$$\mathcal{G}^{\text{SC}} \approx \left(\frac{d_R d_T}{M_{-\frac{a}{2}}\left(\mathbf{d}_R \odot \mathbf{d}_T\right)}\right)^a \triangleq \tilde{\mathcal{G}}^{\text{SC}}, \tag{9.20}$$

which can be greater or less than one, depending on the positions of the localized and distributed RISs.

Since distributed single-connected RIS is not always beneficial over localized single-connected RIS, we provide intuitive sufficient and necessary conditions for $\tilde{\mathcal{G}}^{\text{SC}} > 1$. First, since $M_{-\frac{a}{2}}(\mathbf{d}_R \odot \mathbf{d}_T) < \sqrt[a]{N^2}\min\left(\mathbf{d}_R \odot \mathbf{d}_T\right)$, we have

$$\tilde{\mathcal{G}}^{\text{SC}} > \left(\frac{d_R d_T}{\sqrt[a]{N^2}\min\left(\mathbf{d}_R \odot \mathbf{d}_T\right)}\right)^a, \tag{9.21}$$

and a sufficient condition for $\tilde{\mathcal{G}}^{\text{SC}} > 1$ is given by $d_R d_T > \sqrt[a]{N^2}\min\left(\mathbf{d}_R \odot \mathbf{d}_T\right)$. Second, since $M_{-\frac{a}{2}}(\mathbf{d}_R \odot \mathbf{d}_T) > \min(\mathbf{d}_R \odot \mathbf{d}_T)$, we have

$$\tilde{\mathcal{G}}^{\text{SC}} < \left(\frac{d_R d_T}{\min\left(\mathbf{d}_R \odot \mathbf{d}_T\right)}\right)^a, \tag{9.22}$$

and a necessary condition for $\tilde{\mathcal{G}}^{\text{SC}} > 1$ is given by $d_R d_T > \min(\mathbf{d}_R \odot \mathbf{d}_T)$.

In Fig. 9.2(b), we numerically evaluate $\mathcal{G}^{\text{SC}}$ considering the systems represented in Fig. 9.1. Interestingly, we always have $\mathcal{G}^{\text{SC}} < 1$, meaning that the distributed single-connected RIS is not beneficial over the localized single-connected RIS in the considered scenario. This is because the receiver is located close to the localized RIS, yielding a particularly small $d_R d_T$, which deteriorates the gain, as given in (9.20).

In addition, in Fig. 9.3(b), we report $\mathcal{G}^{\text{SC}}$ for different locations of the receiver. Remarkably, we observe that $\mathcal{G}^{\text{SC}} < 1$ only when the receiver is in close proximity to the localized RIS, while it can reach high values when the receiver is far from the localized RIS. Thus, distributed single-connected RIS offers better coverage than localized single-connected RIS, improving the received signal power over a much wider area.

9.2.4 Gain of Distributed over Localized Fully-Connected RIS

Finally, we consider the gain of distributed over localized fully-connected RIS, defined as $\mathcal{G}^{\text{FC}} = \text{E}[P_R^{\text{Dis-FC}}]/\text{E}[P_R^{\text{Loc-FC}}]$. From (9.6) and (9.10), we directly obtain

$$\mathcal{G}^{\text{FC}} = \frac{\sum_{n=1}^{N}\left[\mathbf{d}_R\right]_n^{-a}\sum_{n=1}^{N}\left[\mathbf{d}_T\right]_n^{-a}}{N^2 d_R^{-a} d_T^{-a}} = \left(\frac{d_R d_T}{M_{-a}\left(\mathbf{d}_R\right) M_{-a}\left(\mathbf{d}_T\right)}\right)^a, \tag{9.23}$$

indicating that $\mathcal{G}^{\text{FC}} \geq \mathcal{G}^{\text{SC}}$ by the Cauchy-Schwarz inequality. Since

$$\mathcal{G}^{\text{FC}} > \left(\frac{d_R d_T}{\sqrt[a]{N^2}\min\left(\mathbf{d}_R\right)\min\left(\mathbf{d}_T\right)}\right)^a, \tag{9.24}$$

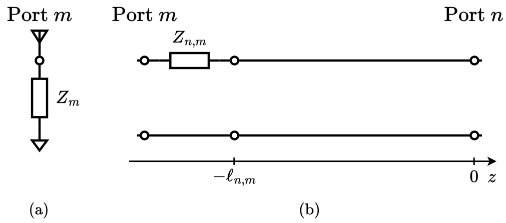


Figure 9.4: (a) Port m connected to ground through Z_m and (b) ports m and n interconnected through $Z_{n,m}$ and a transmission line of length $\ell_{n,m}$.

a sufficient condition for $\mathcal{G}^{\text{FC}} > 1$ is $d_R d_T > \sqrt[a]{N^2}\min(\mathbf{d}_R)\min(\mathbf{d}_T)$, which is more relaxed than the sufficient condition for $\tilde{\mathcal{G}}^{\text{SC}} > 1$. In addition, since

$$\mathcal{G}^{\text{FC}} < \left(\frac{d_R d_T}{\min(\mathbf{d}_R)\min(\mathbf{d}_T)}\right)^a, \tag{9.25}$$

a necessary condition for $\mathcal{G}^{\text{FC}} > 1$ is $d_R d_T > \min(\mathbf{d}_R)\min(\mathbf{d}_T)$, which is more relaxed than the necessary condition for $\tilde{\mathcal{G}}^{\text{SC}} > 1$.

In Fig. 9.2(c), we numerically evaluate $\mathcal{G}^{\text{FC}}$ considering the systems represented in Fig. 9.1. Remarkably, we always observe $\mathcal{G}^{\text{FC}} > 1$, meaning that the distributed fully-connected RIS always outperforms localized fully-connected RIS in the considered scenario. Furthermore, Fig. 9.2(c) shows that $\mathcal{G}^{\text{FC}}$ grows exponentially with a, reaching values as high as several orders of magnitude, and is only minimally impacted by N.

Besides, we report $\mathcal{G}^{\text{FC}}$ for different locations of the receiver in Fig. 9.3(c). Remarkably, we observe that $\mathcal{G}^{\text{FC}} > 15$ dB for all receiver locations, indicating that distributed fully-connected RIS offers better performance than localized fully-connected RIS over the whole considered area. In particular, $\mathcal{G}^{\text{FC}}$ is significantly high when the receiver is far from the localized RIS while being close to the distributed RIS.

9.3 General Lossy BD-RIS Model

We have introduced the concept of distributed RIS, showing that distributed BD-RIS can enable significant gains over distributed single-connected RIS and localized BD-RIS. In this section, we study the impact of the losses within the BD-RIS architecture by developing a model for BD-RIS that accounts for lossy interconnections.

We model a BD-RIS with lossy interconnections as an N-port network with a circuit topology described as follows. First, port m is connected to ground through a tunable impedance Z_m, for $m = 1, \ldots, N$, as represented in Fig. 9.4(a). Second, port m and port n, if interconnected in the BD-RIS architecture, are interconnected through a tunable impedance $Z_{n,m}$ in series with a transmission line of length $\ell_{n,m}$, for $n, m = 1, \ldots, N$, as represented in Fig. 9.4(b).

Given an N-port network implemented with this circuit topology, our goal is to characterize its scattering matrix $\mathbf{\Theta}$. Since it is hard to directly derive $\mathbf{\Theta}$ from the BD-RIS circuit topology, we characterize the BD-RIS admittance matrix $\mathbf{Y} \in \mathbb{C}^{N\times N}$, which is related to $\mathbf{\Theta}$ by (2.3), where Z_0 is the reference impedance, typically set to $Z_0 = 50\ \Omega$ [74,

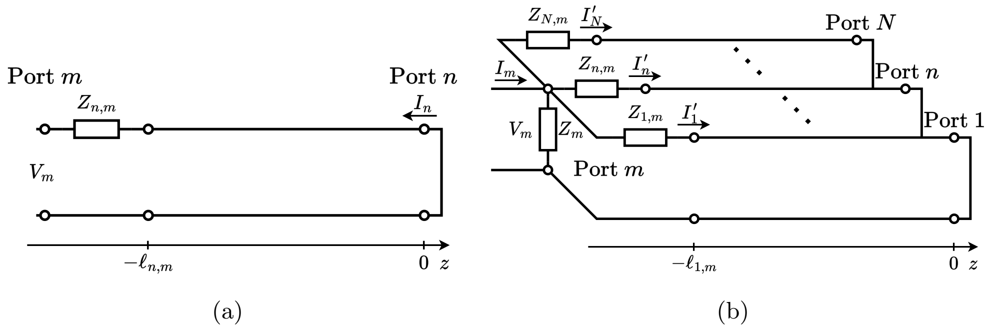


Figure 9.5: Equivalent circuit to be studied to compute (a) the off-diagonal entry $[\mathbf{Y}]_{n,m}$ and (b) the diagonal entry $[\mathbf{Y}]_{m,m}$.

Chapter 4]. According to the definition of $\mathbf{Y}$, each entry $[\mathbf{Y}]_{n,m}$ is obtained by driving port m with a voltage V_m, short-circuiting all other ports, measuring the short-circuit current I_n at port n, and computing the ratio

$$[\mathbf{Y}]_{n,m} = \left.\frac{I_n}{V_m}\right|_{V_k=0,\forall k\neq m}, \tag{9.26}$$

for $n, m = 1, \ldots, N$ [74, Chapter 4]. In the following, we separately model the off-diagonal entries $[\mathbf{Y}]_{n,m}$, with $n \neq m$, and the diagonal entries $[\mathbf{Y}]_{m,m}$.

9.3.1 Modeling the Admittance Matrix: Off-Diagonal Entries

To derive the off-diagonal entries $[\mathbf{Y}]_{n,m}$, with $n \neq m$, given by (9.26), we make two observations that allow us to study a simplified circuit. First, the current flowing on the transmission lines interconnecting port j and port k, with $j, k \neq m$, is zero since these transmission lines are short-circuited at both ends according to (9.26). Consequently, these transmission lines can be removed from the circuit when computing (9.26). Second, the current flowing on the transmission lines interconnecting port m and port k, with $k \neq n$, do not influence the current I_n. Thus, also these transmission lines can be removed from the circuit as they do not impact on (9.26). Following these two observations, we have that each off-diagonal entry $[\mathbf{Y}]_{n,m}$, with $n \neq m$, can be derived by studying an equivalent circuit made of a single transmission line, as represented in Fig. 9.5(a).

To study the circuit in Fig. 9.5(a) and determine $[\mathbf{Y}]_{n,m}$, we begin by solving the telegrapher equations in the sinusoidal steady-state condition [74, Chapter 2], giving the voltage $V(z)$ and current $I(z)$ on the transmission line interconnecting port m and port n as

$$V(z) = V_0^+ e^{-\gamma z} + V_0^- e^{\gamma z},\ I(z) = \frac{V_0^+ e^{-\gamma z} - V_0^- e^{\gamma z}}{Z_0}, \tag{9.27}$$

respectively, where $V_0^+ \in \mathbb{C}$ and $V_0^- \in \mathbb{C}$ are constants, $\gamma \in \mathbb{C}$ is the propagation constant of the transmission line, and $Z_0 \in \mathbb{R}$ is the characteristic impedance of the transmission line[2]. In the following, we characterize the voltage and current on the transmission line by determining the constants V_0^+ and V_0^-.

[2]Note that Z_0 is real for both lossless and the so-called low-loss transmission lines [74, Chapter 2], which are considered in this chapter.

Since the transmission line in Fig. 9.5(a) is terminated in a short circuit, the voltage at $z = 0$ is zero. Thus, by using (9.27), we have $V(0) = V_0^+ + V_0^- = 0$, giving $V_0^- = -V_0^+$. By substituting $V_0^- = -V_0^+$ into (9.27), $V(z)$ and $I(z)$ simplify to

$$V(z) = V_0^+ \left(e^{-\gamma z} - e^{\gamma z}\right), \; I(z) = \frac{V_0^+}{Z_0}\left(e^{-\gamma z} + e^{\gamma z}\right), \tag{9.28}$$

respectively. Furthermore, we can determine V_0^+ by using the input impedance of the transmission line, defined as

$$Z_{\text{in}} = \frac{V\left(-\ell_{n,m}\right)}{I\left(-\ell_{n,m}\right)}, \tag{9.29}$$

and that can be expressed as

$$Z_{\text{in}} = Z_0 \frac{e^{\gamma \ell_{n,m}} - e^{-\gamma \ell_{n,m}}}{e^{\gamma \ell_{n,m}} + e^{-\gamma \ell_{n,m}}}, \tag{9.30}$$

by using (9.28). Specifically, we determine V_0^+ from the voltage at $z = -\ell_{n,m}$, given by

$$V(-\ell_{n,m}) = V_m \frac{Z_{\text{in}}}{Z_{\text{in}} + Z_{n,m}} \tag{9.31}$$

$$= V_0^+ \left(e^{\gamma \ell_{n,m}} - e^{-\gamma \ell_{n,m}}\right), \tag{9.32}$$

where (9.31) is a direct consequence of (9.29), and (9.32) is obtained by expressing $V(-\ell_{n,m})$ through (9.28). By equating (9.31) and (9.32), we obtain the expression of V_0^+ as

$$V_0^+ = V_m \frac{Z_{\text{in}}}{\left(Z_{\text{in}} + Z_{n,m}\right)\left(e^{\gamma \ell_{n,m}} - e^{-\gamma \ell_{n,m}}\right)} \tag{9.33}$$

$$= V_m \frac{Z_0}{Z_{n,m}\left(e^{\gamma \ell_{n,m}} + e^{-\gamma \ell_{n,m}}\right) + Z_0\left(e^{\gamma \ell_{n,m}} - e^{-\gamma \ell_{n,m}}\right)}, \tag{9.34}$$

where (9.34) is derived by using (9.30).

We can now derive the off-diagonal entry $[\mathbf{Y}]_{n,m}$, with $n \neq m$, given by (9.26). Specifically, I_n is obtained through (9.28) as

$$I_n = -I(0) = -2\frac{V_0^+}{Z_0}, \tag{9.35}$$

with V_0^+ given by (9.34). Thus, by substituting (9.35) into (9.26), we finally obtain

$$[\mathbf{Y}]_{n,m} = \frac{-2}{Z_{n,m}\left(e^{\gamma \ell_{n,m}} + e^{-\gamma \ell_{n,m}}\right) + Z_0\left(e^{\gamma \ell_{n,m}} - e^{-\gamma \ell_{n,m}}\right)}. \tag{9.36}$$

Remarkably, (9.36) provides the general expression of the off-diagonal entry $[\mathbf{Y}]_{n,m}$ as a function of the tunable impedance $Z_{n,m}$ and the transmission line parameters $\ell_{n,m}$, γ, and Z_0. Because of the reciprocity of the circuit, $\mathbf{Y}$ is symmetric and we have $[\mathbf{Y}]_{n,m} = [\mathbf{Y}]_{m,n}$. Furthermore, if port m and n are not interconnected in the BD-RIS architecture, $Z_{n,m}$ is an open circuit, i.e., $Z_{n,m} = \infty$, yielding $[\mathbf{Y}]_{n,m} = [\mathbf{Y}]_{m,n} = 0$.

9.3.2 Modeling the Admittance Matrix: Diagonal Entries

To derive the diagonal entries $[\mathbf{Y}]_{m,m}$, we observe that the current flowing on the transmission lines interconnecting port j and port k, with $j, k \neq m$, is zero since these transmission lines are short-circuited at both ends according to (9.26). Thus, these transmission lines do not impact on (9.26) and can be removed from the circuit. Following this observation, we have that each diagonal entry $[\mathbf{Y}]_{m,m}$ can be derived by studying the equivalent circuit represented in Fig. 9.5(b) to find the ratio $[\mathbf{Y}]_{m,m} = I_m/V_m$.

From the circuit in Fig. 9.5(b), we directly notice by Kirchhoff's first law that

$$I_m = \frac{V_m}{Z_m} + \sum_{n \neq m} I'_n, \tag{9.37}$$

where I'_n can be computed through (9.28) as

$$I'_n = I(-\ell_{n,m}) = \frac{V_0^+}{Z_0}\left(e^{\gamma\ell_{n,m}} + e^{-\gamma\ell_{n,m}}\right), \tag{9.38}$$

$\forall n \neq m$, with V_0^+ given by (9.34). Thus, by substituting (9.37) into (9.26), we obtain

$$[\mathbf{Y}]_{m,m} = \frac{1}{Z_m} + \sum_{n \neq m} \frac{I'_n}{V_m}, \tag{9.39}$$

readily giving

$$[\mathbf{Y}]_{m,m} = \frac{1}{Z_m} - \sum_{n \neq m} \frac{e^{\gamma\ell_{n,m}} + e^{-\gamma\ell_{n,m}}}{2}[\mathbf{Y}]_{n,m}, \tag{9.40}$$

by using (9.38), where $[\mathbf{Y}]_{n,m}$ is given by (9.36). Remarkably, the diagonal entry $[\mathbf{Y}]_{m,m}$ is a function of Z_m, $Z_{n,m}$, $\forall n \neq m$, and the transmission line parameters.

9.4 Simplified Lossy BD-RIS Models

We have modeled the RIS admittance matrix $\mathbf{Y}$ by determining its entries in (9.36) and (9.40). In (9.36) and (9.40), the propagation constant of the transmission line $\gamma = \alpha + j\beta$ is in general a complex number, with real part α denoted as the attenuation constant and imaginary part β denoted as the phase constant [74, Chapter 2]. Given the complex nature of γ, it is hard to gain engineering insights into the entries of $\mathbf{Y}$ from (9.36) and (9.40). For this reason, we simplify this general model into three simplified models by making different assumptions on the interconnection lengths $\ell_{n,m}$ and the attenuation constant α.

9.4.1 Model with Specific Interconnection Lengths

Consider the case when $\ell_{n,m}$ is a multiple of $\lambda/2$, with the wavelength λ defined as $\lambda = \frac{2\pi}{\beta}$, i.e., when $\ell_{n,m} = \frac{\pi}{\beta}K_{n,m}$, for some $K_{n,m} \in \mathbb{Z}$. In this case, we have $e^{j\beta\ell_{n,m}} = e^{-j\beta\ell_{n,m}} = (-1)^{K_{n,m}}$. In addition, we assume the tunable impedance components of the RIS to be purely reactive, i.e., $Z_m = jX_m$ with $X_m \in \mathbb{R}$ and $Z_{n,m} = jX_{n,m}$ with $X_{n,m} \in \mathbb{R}$. Thus,

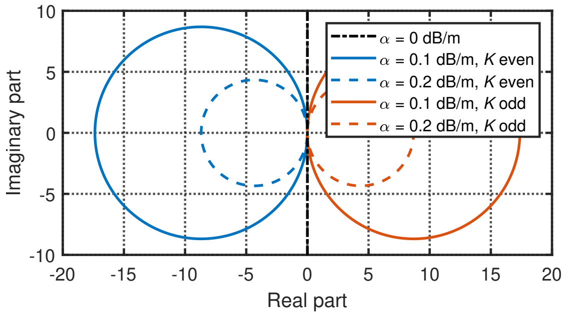


Figure 9.6: Possible values of $[\mathbf{Y}]_{n,m}$ modeled as in (9.42), with $Z_0 = 50\ \Omega$ and $\ell_{n,m} = 0.1$ m.

(9.36) can be simplified into

$$[\mathbf{Y}]_{n,m} = \frac{-2(-1)^{K_{n,m}}}{jX_{n,m}\left(e^{\alpha\ell_{n,m}} + e^{-\alpha\ell_{n,m}}\right) + Z_0\left(e^{\alpha\ell_{n,m}} - e^{-\alpha\ell_{n,m}}\right)} \tag{9.41}$$

$$= \frac{-(-1)^{K_{n,m}}}{jX_{n,m}\cosh\left(\alpha\ell_{n,m}\right) + Z_0\sinh\left(\alpha\ell_{n,m}\right)}. \tag{9.42}$$

Remarkably, in (9.42) the term $jX_{n,m}\cosh\left(\alpha\ell_{n,m}\right)$ is purely imaginary and the term $Z_0\sinh\left(\alpha\ell_{n,m}\right)$ is purely real. Thus, if $[\mathbf{Y}]_{n,m} \neq 0$, $[\mathbf{Y}]_{n,m}$ must have a non-zero real part, which dissipates real power. More interestingly, it can be easily verified that the set of possible values of $[\mathbf{Y}]_{n,m}$ given by (9.42) is a circle in the complex plane with radius

$$r = \frac{1}{2Z_0\sinh\left(\alpha\ell_{n,m}\right)}, \tag{9.43}$$

and centered in $c = (-(-1)^{K_{n,m}}r, 0)$, where any point on the circle can be achieved by properly tuning $X_{n,m}$. Thus, $[\mathbf{Y}]_{n,m}$ has always a negative real part when $K_{n,m}$ is even, i.e., the interconnection length $\ell_{n,m}$ is multiple of the wavelength λ, and has always a positive real part when $K_{n,m}$ is odd.

For better clarity, we report in Fig. 9.6 the possible values of $[\mathbf{Y}]_{n,m}$ considering different parameters, by fixing $Z_0 = 50\ \Omega$ and $\ell_{n,m} = 0.1$ m. From Fig. 9.6, we can make the following two observations. First, when $\alpha\ell_{n,m} = 0$, we have $r = +\infty$ according to (9.43), meaning that $[\mathbf{Y}]_{n,m}$ can assume an arbitrary imaginary part with always zero real part. In this ideal case, the BD-RIS has full flexibility and does not dissipate real power. Second, r decreases as the value of $\alpha\ell_{n,m}$ increases, reducing the flexibility of the BD-RIS since $[\mathbf{Y}]_{n,m}$ has imaginary part constrained within the interval $[-r, +r]$.

In the case $\ell_{n,m}$ is a multiple of $\lambda/2$ for all the interconnections departing from port m, i.e., $\ell_{n,m} = \frac{\pi}{\beta}K_{n,m}$, for some $K_{n,m} \in \mathbb{Z}$, $\forall n \neq m$, (9.40) simplifies as

$$[\mathbf{Y}]_{m,m} = \frac{1}{jX_m} - \sum_{n\neq m}(-1)^{K_{n,m}}\cosh\left(\alpha\ell_{n,m}\right)[\mathbf{Y}]_{n,m} \tag{9.44}$$

$$= \frac{1}{jX_m} + \sum_{n\neq m}\frac{1}{jX_{n,m} + Z_0\tanh\left(\alpha\ell_{n,m}\right)}. \tag{9.45}$$

Remarkably, $[\mathbf{Y}]_{m,m}$ has always a positive real part, $\forall K_{n,m} \in \mathbb{Z}$, given by

$$\Re\{[\mathbf{Y}]_{m,m}\} = \sum_{n\neq m} \frac{Z_0 \tanh(\alpha\ell_{n,m})}{X_{n,m}^2 + Z_0^2 \tanh^2(\alpha\ell_{n,m})}, \tag{9.46}$$

while it can have an arbitrary imaginary part depending on the tunable value X_m.

9.4.2 Model with Lossless Interconnections

Consider the case of lossless interconnections, i.e., $\alpha = 0$, and do not assume any constraints on the lengths $\ell_{n,m}$. In addition, we assume purely reactive tunable impedance components, i.e., $Z_m = jX_m$ and $Z_{n,m} = jX_{n,m}$. In this case, (9.36) can be simplified into

$$[\mathbf{Y}]_{n,m} = \frac{-1}{jX_{n,m}\cos(\beta\ell_{n,m}) + jZ_0 \sin(\beta\ell_{n,m})}, \tag{9.47}$$

which is purely imaginary. Interestingly, $[\mathbf{Y}]_{n,m}$ can be any imaginary number according to the tunable impedance component $X_{n,m}$.

Besides, with $\alpha = 0$, (9.40) boils down to

$$[\mathbf{Y}]_{m,m} = \frac{1}{jX_m} - \sum_{n\neq m} \cos(\beta\ell_{n,m})[\mathbf{Y}]_{n,m} \tag{9.48}$$

$$= \frac{1}{jX_m} + \sum_{n\neq m} \frac{1}{jX_{n,m} + jZ_0 \tan(\beta\ell_{n,m})}, \tag{9.49}$$

which is also purely imaginary. Note that $[\mathbf{Y}]_{m,m}$ can be set independently of the values of $[\mathbf{Y}]_{n,m}$ by properly choosing X_m. Thus, the whole admittance matrix $\mathbf{Y}$ is purely imaginary, indicating that the RIS is lossless and does not dissipate any real power. This is consistent with the fact that the RIS is assumed to be made of lossless tunable components and lossless transmission lines. Remarkably, when the transmission lines of the RIS interconnections are lossless, their only effect is the variation in phase of voltages and currents along them.

9.4.3 Model with Lossless Interconnections of Specific Lengths

Consider the case when $\ell_{n,m}$ is a multiple of $\lambda/2$, i.e., when $\ell_{n,m} = \frac{\pi}{\beta}K_{n,m}$, for some $K_{n,m} \in \mathbb{Z}$, and the interconnections are lossless, i.e., $\alpha = 0$. In this case, we have $e^{\gamma\ell_{n,m}} = e^{-\gamma\ell_{n,m}} = (-1)^{K_{n,m}}$. In addition, we assume purely reactive tunable impedance components, i.e., $Z_m = jX_m$ and $Z_{n,m} = jX_{n,m}$. Thus, (9.36) can be simplified into

$$[\mathbf{Y}]_{n,m} = \frac{-(-1)^{K_{n,m}}}{jX_{n,m}}, \tag{9.50}$$

which can be an arbitrary imaginary number depending on the tunable impedance component $X_{n,m}$.

In the case $\ell_{n,m}$ is a multiple of $\lambda/2$ for all the interconnections departing from port

m, i.e., $\ell_{n,m} = \frac{\pi}{\beta} K_{n,m}$, for some $K_{n,m} \in \mathbb{Z}$, $\forall n \neq m$, (9.40) simplifies as

$$[\mathbf{Y}]_{m,m} = \frac{1}{jX_m} - \sum_{n \neq m} (-1)^{K_{n,m}} [\mathbf{Y}]_{n,m} \tag{9.51}$$

$$= \frac{1}{jX_m} + \sum_{n \neq m} \frac{1}{jX_{n,m}}, \tag{9.52}$$

which can be an arbitrary imaginary number. Remarkably, when all the interconnection lengths $\ell_{n,m}$ are a multiple of λ, i.e., $K_{n,m}$ is even, $\forall n, m$, the transmission lines have no visible effect as they apply a phase shift of a multiple of 2π to the voltages and currents propagating along them. In this case, the simplified model in (9.50) and (9.52) is equivalent to the model with no transmission lines, i.e., with $\ell_{n,m} = 0$, widely considered in previous literature on BD-RIS [16, 80].

9.5 Lossy BD-RIS Optimization

We have modeled the admittance matrix $\mathbf{Y}$ of BD-RIS with lossy interconnections and derived three simplified models. In this section, we optimize BD-RIS characterized by these simplified models to assess the impact of losses on the performance of BD-RIS.

9.5.1 Optimizing Lossless BD-RIS

Consider the lossless BD-RIS model of Section 9.4.2. The received signal power maximization problem in a BD-RIS-aided MIMO system writes as

$$\max_{X_{n,m}, X_m, \mathbf{g}, \mathbf{w}} \quad P_T \left| \mathbf{g} \left(\mathbf{H}_{RT} + \mathbf{H}_R \boldsymbol{\Theta} \mathbf{H}_T \right) \mathbf{w} \right|^2 \tag{9.53}$$

$$\text{s.t.} \quad (2.3),\ (9.47),\ (9.49),\ \|\mathbf{g}\| = 1,\ \|\mathbf{g}\| = 1, \tag{9.54}$$

which is solved by jointly optimizing $X_{n,m}$, X_m, $\mathbf{g}$, $\mathbf{w}$. To this end, we initialize $\mathbf{g}$ and $\mathbf{w}$ to feasible values and alternate between the following two steps.

When $\mathbf{g}$ and $\mathbf{w}$ are fixed, we update $X_{n,m}$ and X_m by solving

$$\max_{X_{n,m}, X_m} \quad P_T \left| \bar{h}_{RT} + \bar{\mathbf{h}}_R \boldsymbol{\Theta} \bar{\mathbf{h}}_T \right|^2 \ \text{ s.t. } \ (2.3),\ (9.47),\ (9.49), \tag{9.55}$$

where $\bar{h}_{RT} = \mathbf{g}\mathbf{H}_{RT}\mathbf{w}$, $\bar{\mathbf{h}}_R = \mathbf{g}\mathbf{H}_R$, and $\bar{\mathbf{h}}_T = \mathbf{H}_T\mathbf{w}$, which is apparently hard given the non-convex constraints. Nevertheless, (9.55) can be simplified by noticing that $\mathbf{Y}$ given by (9.47) and (9.49) can be any purely imaginary symmetric matrix with $[\mathbf{Y}]_{n,m} = 0$ if ports m and n are not interconnected, i.e., $\mathbf{Y} = j\mathbf{B}$, $\mathbf{B} = \mathbf{B}^T$, with $[\mathbf{B}]_{n,m} = 0$ if $Z_{n,m} = \infty$, where $\mathbf{B} \in \mathbb{R}^{N \times N}$ denotes the RIS susceptance matrix. Thus, (9.55) can be rewritten as

$$\max_{\mathbf{B}} \ P_T \left| \bar{h}_{RT} + \bar{\mathbf{h}}_R \left(\mathbf{I} + jZ_0\mathbf{B} \right)^{-1} \left(\mathbf{I} - jZ_0\mathbf{B} \right) \bar{\mathbf{h}}_T \right|^2 \tag{9.56}$$

$$\text{s.t.} \quad \mathbf{B} = \mathbf{B}^T,\ [\mathbf{B}]_{n,m} = 0 \text{ if } Z_{n,m} = \infty, \tag{9.57}$$

which is an unconstrained problem in the upper triangular entries of $\mathbf{B}$ that are not forced to zero. This problem has been solved through the quasi-Newton method in [16] for group-/fully-connected RISs, and through global optimal closed-form solutions in [75] and [80], for group-/fully- and forest-/tree-connected RISs, respectively. Given the optimal susceptance matrix $\mathbf{B}^{\star}$ solution to (9.56)-(9.57), the values of the tunable reactance components $X_{n,m}$ and X_m can be obtained by inverting (9.47) and (9.48) with $\mathbf{Y} = j\mathbf{B}^{\star}$, respectively, i.e.,

$$X_{n,m} = \frac{1}{\cos\left(\beta\ell_{n,m}\right)\left[\mathbf{B}^{\star}\right]_{n,m}} - Z_0 \tan\left(\beta\ell_{n,m}\right), \tag{9.58}$$

$$X_m = \frac{-1}{\left[\mathbf{B}^{\star}\right]_{m,m} + \sum_{n\neq m}\cos\left(\beta\ell_{n,m}\right)\left[\mathbf{B}^{\star}\right]_{n,m}}. \tag{9.59}$$

When $X_{n,m}$ and X_m are fixed, also $\boldsymbol{\Theta}$ is fixed, as given by (2.3), where $\mathbf{Y}$ is determined by (9.47) and (9.49). Thus, $\mathbf{g}$ and $\mathbf{w}$ are updated as the dominant left and right singular vectors of $\mathbf{H}_{RT} + \mathbf{H}_R\boldsymbol{\Theta}\mathbf{H}_T$, respectively, which is globally optimal.

These two steps are alternatively repeated until convergence of the objective (9.53). Notably, this optimization framework is also valid for the model of Section 9.4.3 as it is a special case of the model of Section 9.4.2.

9.5.2 Optimizing Lossy BD-RIS

Considering the lossy BD-RIS model of Section 9.4.1, the received signal power maximization problem writes as

$$\max_{X_{n,m},X_m,\mathbf{g},\mathbf{w}} \quad P_T\left|\mathbf{g}\left(\mathbf{H}_{RT} + \mathbf{H}_R\boldsymbol{\Theta}\mathbf{H}_T\right)\mathbf{w}\right|^2 \tag{9.60}$$

$$\text{s.t.} \quad (2.3),\ (9.42),\ (9.45),\ \|\mathbf{g}\| = 1,\ \|\mathbf{g}\| = 1, \tag{9.61}$$

which is solved by initializing $\mathbf{g}$ and $\mathbf{w}$ to feasible values and by alternating between the following two steps.

When $\mathbf{g}$ and $\mathbf{w}$ are fixed, $X_{n,m}$ and X_m are updated by solving

$$\max_{X_{n,m},X_m} \quad P_T\left|\bar{h}_{RT} + \bar{\mathbf{h}}_R\boldsymbol{\Theta}\bar{\mathbf{h}}_T\right|^2 \ \text{ s.t. } (2.3),\ (9.42),\ (9.45). \tag{9.62}$$

Interestingly, the constraints of problem (9.62) can be tightened as a consequence of the following proposition.

Proposition 9. *The average real power dissipated by an RIS with admittance matrix* $\mathbf{Y}$ *modeled as in Section 9.4.1 is reduced if any* $|\Re\{[\mathbf{Y}]_{i,j}\}|$ *is reduced, with* $i \neq j$.

Proof. Consider the following two RISs: the first has admittance matrix $\mathbf{Y}$ and dissipates a power P_{av}; the second has admittance matrix $\mathbf{Y}'$ obtained from $\mathbf{Y}$ by reducing $|\Re\{[\mathbf{Y}]_{i,j}\}|$ by $\Delta > 0$ for some i, j, i.e., $|\Re\{[\mathbf{Y}']_{i,j}\}| = |\Re\{[\mathbf{Y}]_{i,j}\}| - \Delta$, and dissipates a power P'_{av}. According to (9.42) and (9.44), introducing $\mathbf{G} = \Re\{\mathbf{Y}\}$ and $\mathbf{G}' = \Re\{\mathbf{Y}'\}$, we have $[\mathbf{G}']_{n,m} = [\mathbf{G}]_{n,m}$ for $m, n \notin \{i, j\}$, and

$$\left[\mathbf{G}'\right]_{i,j} = [\mathbf{G}]_{i,j} + (-1)^{K_{i,j}}\Delta,\ \left[\mathbf{G}'\right]_{j,i} = [\mathbf{G}]_{j,i} + (-1)^{K_{i,j}}\Delta, \tag{9.63}$$

$$\left[\mathbf{G}'\right]_{k,k} = [\mathbf{G}]_{k,k} - \cosh\left(\alpha\ell_{i,j}\right)\Delta, \tag{9.64}$$

for $k \in \{i,j\}$. Given these two RISs, our goal is to prove that $P'_{\text{av}} \leq P_{\text{av}}$. To this end, we first recall that average real power dissipated by an N-port network is defined as $P_{\text{av}} = \Re\{\mathbf{v}^T\mathbf{i}^*\}/2$, where $\mathbf{v}, \mathbf{i} \in \mathbb{C}^{N\times 1}$ are the voltage and current vectors at the N ports, respectively [74, Chapter 4]. Thus, if the admittance matrix $\mathbf{Y}$ of this N-port network is symmetric, it is possible to show that

$$P_{\text{av}} = \frac{1}{2}\mathbf{v}^T\mathbf{G}\mathbf{v}^* = \frac{1}{2}\sum_{m,n} v_n v_m^* \left[\mathbf{G}\right]_{n,m}, \tag{9.65}$$

where $\mathbf{v} = [v_1, \ldots, v_N]^T$, giving the power dissipated by the first RIS. Similarly, the power dissipated by the second RIS is

$$\begin{aligned} P'_{\text{av}} &= \frac{1}{2}\sum_{m,n} v_n v_m^* \left[\mathbf{G}'\right]_{n,m} && (9.66)\\ &= \frac{1}{2}\sum_{m,n\notin\{i,j\}} v_n v_m^* \left[\mathbf{G}\right]_{n,m} + \frac{1}{2}\sum_{m,n\in\{i,j\}} v_n v_m^* \left[\mathbf{G}'\right]_{n,m}, && (9.67) \end{aligned}$$

since $[\mathbf{G}']_{n,m} = [\mathbf{G}]_{n,m}$ for $m,n \notin \{i,j\}$. By substituting (9.63) and (9.64) into (9.67), we obtain

$$\begin{aligned} P'_{\text{av}} = P_{\text{av}} - \frac{1}{2}\big(&|v_i|^2 \cosh(\alpha\ell_{i,j})\,\Delta + \\ &|v_j|^2 \cosh(\alpha\ell_{i,j})\,\Delta - (-1)^{K_{i,j}}\left(v_i v_j^* + v_i^* v_j\right)\Delta\big). \end{aligned} \tag{9.68}$$

Finally, since $\cosh(\alpha\ell_{i,j}) \geq 1$, we have

$$P'_{\text{av}} \leq P_{\text{av}} - \frac{1}{2}|v_i - (-1)^{K_{i,j}} v_j|^2\Delta, \tag{9.69}$$

indicating that $P'_{\text{av}} \leq P_{\text{av}}$. □

From Proposition 9, we deduce that it is convenient to select the off-diagonal entries $[\mathbf{Y}]_{n,m}$ with a small real part in magnitude to reduce the power dissipated by the RIS. Interestingly, Fig. 9.6 shows that we can select $[\mathbf{Y}]_{n,m}$ with any imaginary part in the interval $[-r,+r]$, with the real part constrained by $|\Re\{[\mathbf{Y}]_{n,m}\}| \leq r$. This constraint means that we can select $[\mathbf{Y}]_{n,m}$ in the right semicircle of its possible values if $K_{n,m}$ is even, and in the left semicircle if $K_{n,m}$ is odd, to reduce the dissipated power.

By adding the constraint $|\Re\{[\mathbf{Y}]_{n,m}\}| \leq r$ for $n \neq m$ to (9.62), the resulting problem can be reformulated by expressing $\mathbf{Y}$ as a function of an auxiliary variable $\mathbf{A} \in \mathbb{R}^{N\times N}$ as

$$[\mathbf{Y}]_{n,m} = (-1)^{K_{n,m}}\left(\frac{r}{\sqrt{\frac{[\mathbf{A}]_{n,m}^2}{r^2}+1}} - r\right) + j\frac{[\mathbf{A}]_{n,m}}{\sqrt{\frac{[\mathbf{A}]_{n,m}^2}{r^2}+1}}, \tag{9.70}$$

$$[\mathbf{Y}]_{m,m} = \sum_{n\neq m} \cosh(\alpha\ell_{n,m})\,|\Re\{[\mathbf{Y}]_{n,m}\}| + j[\mathbf{A}]_{m,m}, \tag{9.71}$$

where $\mathbf{A}$ has been introduced such that $\Im\{[\mathbf{Y}]_{n,m}\}$ varies in the interval $[-r,+r]$ as $[\mathbf{A}]_{n,m}$ varies in $[-\infty,+\infty]$, with $n \neq m$, and $\Im\{[\mathbf{Y}]_{m,m}\} = [\mathbf{A}]_{m,m}$. Furthermore, in the case of a lossless RIS, i.e., $r = +\infty$, (9.70) and (9.71) boil down to $[\mathbf{Y}]_{n,m} = j[\mathbf{A}]_{n,m}$ and $[\mathbf{Y}]_{m,m} = j[\mathbf{A}]_{m,m}$, respectively. Thus, the resulting problem writes as

$$\max_{\mathbf{A}} P_T \left| \bar{h}_{RT} + \bar{\mathbf{h}}_R \left(\mathbf{I} + Z_0\mathbf{Y}\right)^{-1} \left(\mathbf{I} - Z_0\mathbf{Y}\right) \bar{\mathbf{h}}_T \right|^2 \tag{9.72}$$

$$\text{s.t.} \;\; (9.70),\; (9.71),\; \mathbf{A} = \mathbf{A}^T,\; [\mathbf{A}]_{n,m} = 0 \text{ if } Z_{n,m} = \infty, \tag{9.73}$$

which is an unconstrained problem in the upper triangular entries of $\mathbf{A}$ that are not forced to zero. Remarkably, (9.72)-(9.73) is a generalization of (9.56)-(9.57) since in the case of a lossless RIS, constraints (9.70) and (9.71) become $[\mathbf{Y}]_{n,m} = j[\mathbf{A}]_{n,m}$ and $[\mathbf{Y}]_{m,m} = j[\mathbf{A}]_{m,m}$, respectively, reducing (9.72)-(9.73) to (9.56)-(9.57). Since (9.72)-(9.73) is unconstrained, it can be solved with the quasi-Newton method to find a stationary point. Given the optimal $\mathbf{A}^\star$ solution to (9.72)-(9.73), the admittance matrix $\mathbf{Y}$ is first obtained throguh (9.70) and (9.71). Then, the values of the tunable reactance components $X_{n,m}$ and X_m are obtained from $\mathbf{Y}$ by inverting (9.42) and (9.44), respectively.

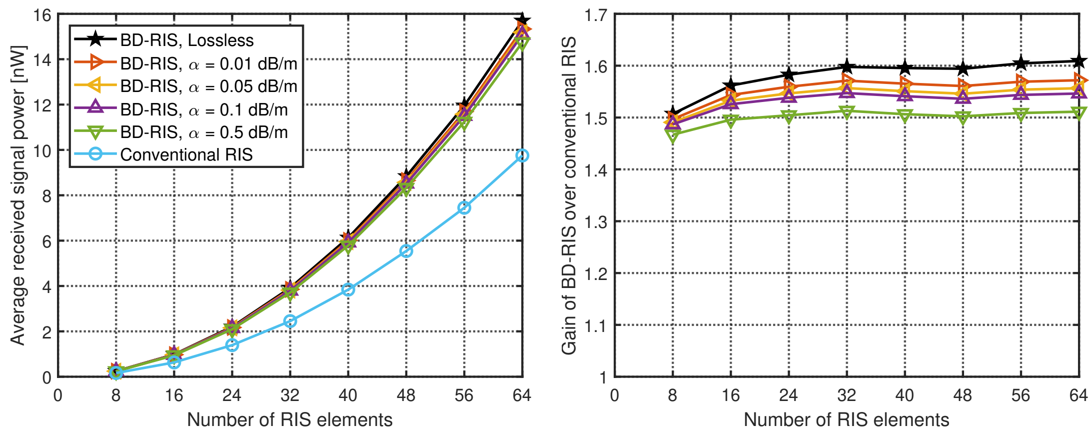
When $X_{n,m}$ and X_m are fixed, also $\boldsymbol{\Theta}$ is fixed, as given by (2.3), where $\mathbf{Y}$ is determined by (9.42) and (9.45). Thus, $\mathbf{g}$ and $\mathbf{w}$ are updated as the dominant left and right singular vectors of $\mathbf{H}_{RT} + \mathbf{H}_R\boldsymbol{\Theta}\mathbf{H}_T$, respectively, which is globally optimal.

These two steps are repeated until convergence of (9.60).

9.6 Numerical Results

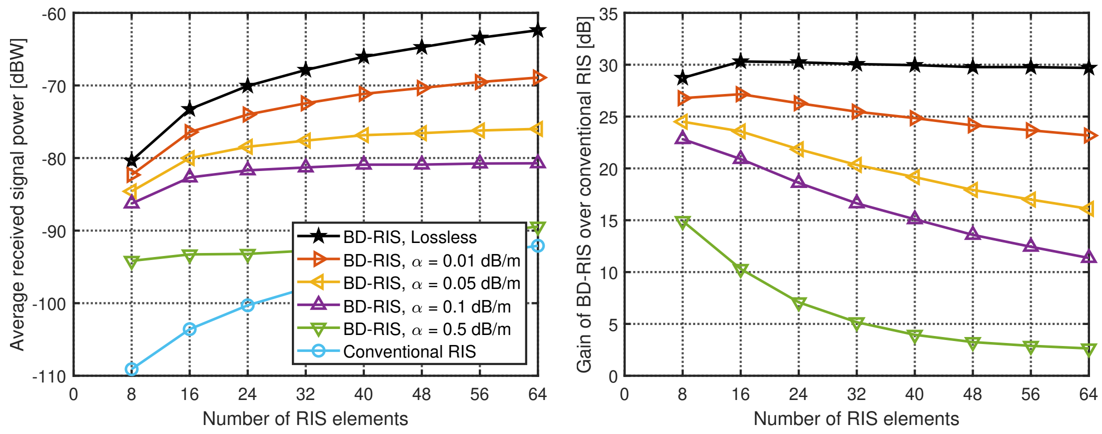
To evaluate the performance of localized and distributed BD-RIS with lossy interconnections, we consider the system represented in Fig. 9.1, where the transmitter and receiver are located at $(0, 0, 0)$ and $(20, 0, 0)$, respectively. In the case of localized RIS, the RIS is a ULA centered in $(20, 0, 2)$ and with an inter-element distance 0.05 m, as shown in Fig. 9.1(a). We model the path-gain of the channels $\mathbf{H}_R^{\text{Loc}}$ and $\mathbf{H}_T^{\text{Loc}}$ through the distance-dependent model and their small-scale fading as i.i.d. Rayleigh distributed, i.e., $\mathbf{H}_i^{\text{Loc}} = \sqrt{\rho_i}\widetilde{\mathbf{H}}_i$, where $\rho_i = C_0 d_i^{-a}$ and $\widetilde{\mathbf{H}}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, for $i \in \{R, T\}$. In the case of distributed RIS, the RIS is a ULA with elements uniformly placed between $(0, 0, 2)$ and $(40, 0, 2)$, as shown in Fig. 9.1(b). Accordingly, the channels are given by $\mathbf{H}_R^{\text{Dis}} = \widetilde{\mathbf{H}}_R\mathbf{R}_R^{1/2}$ and $\mathbf{H}_T^{\text{Dis}} = \mathbf{R}_T^{1/2}\widetilde{\mathbf{H}}_T$, where $\mathbf{R}_i = \text{diag}(\boldsymbol{\rho}_i)$, $[\boldsymbol{\rho}_i]_n = C_0[\mathbf{d}_i]_n^{-a}$, and $\widetilde{\mathbf{H}}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, for $i \in \{R, T\}$. The direct channel is assumed to be fully obstructed, i.e., $\mathbf{H}_{RT} = \mathbf{0}$, and we set $C_0 = -30$ dB, $a = 4$, and $P_T = 10$ W. In the following, we report the performance of tridiagonal BD-RIS, a tree-connected BD-RIS that is proven to be the least complex BD-RIS architecture able to achieve the performance upper bound in single-user systems [80], and compare it with conventional RIS. Recalling that tridiagonal BD-RIS presents interconnections only between adjacent RIS elements [80], its interconnection lengths are all equal to the inter-element distance.

In Fig. 9.7, we report the received signal power obtained with localized RIS versus the number of RIS elements, with $N_R = 1$ and $N_T = 1$ for simplicity. We can make the following observations. *First*, BD-RIS achieves a higher received signal power than conventional RIS thanks to the additional flexibility, both with lossless and lossy interconnections. *Second*, lossless BD-RIS achieves the highest received signal power since no



(a) (b)

Figure 9.7: (a) Received signal power and (b) gain of BD-RIS over conventional RIS versus the number of RIS elements, for localized RIS.

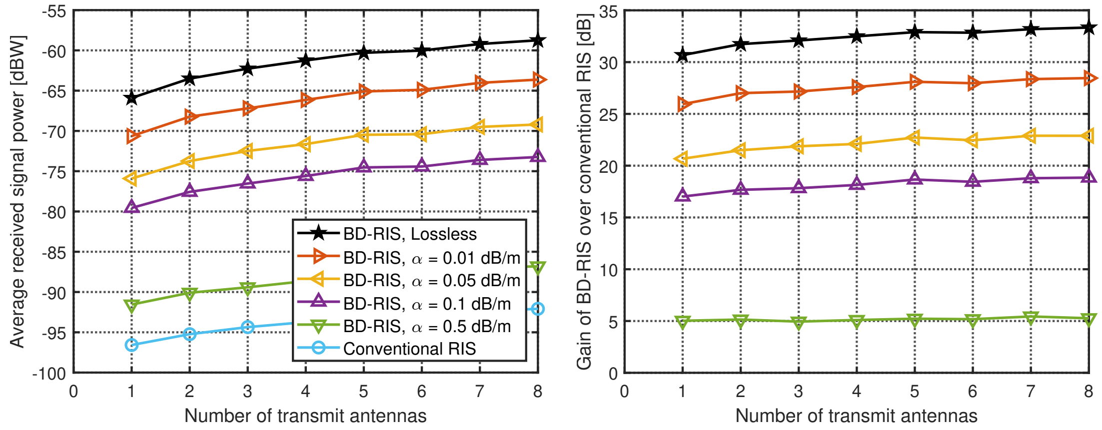


(a) (b)

Figure 9.8: (a) Received signal power and (b) gain of BD-RIS over conventional RIS versus the number of RIS elements, for distributed RIS.

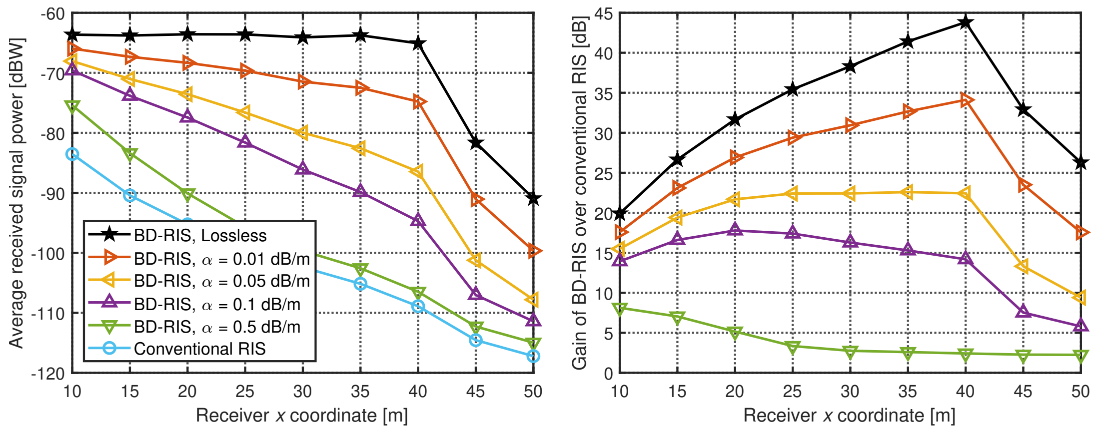
power is dissipated by the BD-RIS circuit. Specifically, Fig. 9.7(b) shows that lossless BD-RIS offers a gain over conventional RIS of up to 1.62, as discussed in Section 9.2.1. *Third*, the received signal power achieved by BD-RIS decreases as the attenuation constant α increases. However, the performance of localized BD-RIS is only slightly impacted by the losses given the short inter-element distance.

In Fig. 9.8, we report the received signal power obtained with distributed RIS versus the number of RIS elements, with $N_R = 1$ and $N_T = 1$. We can make the following observations. *First*, BD-RIS always achieves a higher received signal power than conventional RIS. *Second*, lossless BD-RIS achieves the highest received signal power, offering gains of up to 30 dB over conventional RIS. These massive gains are because the EM signal can propagate without losses between the BD-RIS elements, which are spread over a long line. *Third*, the received signal power of distributed BD-RIS visibly decreases as the attenuation constant α increases, given the significantly long inter-element distance. Nevertheless, lossy BD-RIS can still achieve more than an order of magnitude of gain over conventional RIS.



(a) (b)

Figure 9.9: (a) Received signal power and (b) gain of BD-RIS over conventional RIS versus the number of transmit antennas, for distributed RIS.



(a) (b)

Figure 9.10: (a) Received signal power and (b) gain of BD-RIS over conventional RIS versus the x coordinate of the receiver, for distributed RIS.

To investigate the impact of N_R and N_T on the system performance, we report in Fig. 9.9 the received signal power of distributed RIS versus the number of transmit antennas, with $N = 32$ and $N_R = 2$. We observe that the received signal power of BD-RIS and conventional RIS increases with N_T, given the diversity gain offered by the multiple transmit antennas. Besides, the gain of BD-RIS over conventional RIS is approximately stable for any value of N_T.

In Fig. 9.10, we report the received signal power obtained with distributed RIS for different locations of the receiver. Specifically, the receiver is located in $(x, 0, 0)$, where $x \in [10, 50]$, while we fix $N = 32$, $N_R = 2$, and $N_T = 2$. We observe that the received signal power decreases as x increases since x also represents the distance between the transmitter and receiver. The received signal power decreases significantly after $x = 40$ m, as this is the location of the last distributed RIS element. Besides, Fig. 9.10(b) shows that the gain of lossless BD-RIS over conventional RIS increases with x reaching values as high as several orders of magnitude until $x = 40$ m.

9.7 Conclusion

We propose the concept of distributed RIS, whose elements are distributed over a wide region, in opposition to localized RIS, as commonly considered in previous literature. We derive its scaling laws and analyze the gain of distributed RIS over localized RIS, and of distributed BD-RIS over distributed conventional RIS. The derived gains show that lossless distributed BD-RIS can offer gains as high as several orders of magnitude over distributed conventional RIS and localized BD-RIS. In particular, these substantial gains are enabled by the tunable impedance components interconnecting the BD-RIS elements, which allow the EM signal to effectively reach the receiver by propagating within the BD-RIS circuit.

To assess the practical performance of distributed BD-RIS, we model BD-RIS with lossy interconnections by using transmission line theory. More precisely, we model the BD-RIS interconnections as tunable impedance components in series with lossy transmission lines and derive the expression of the BD-RIS admittance matrix. Since it is hard to gain engineering insights from the obtained model, we also derive three simplified models through different assumptions. Finally, we optimize lossy BD-RIS based on the proposed models and evaluate its performance. Numerical results show that the performance of localized BD-RIS is only slightly impacted by losses because of the short interconnection lengths. Furthermore, distributed BD-RIS can achieve orders of magnitude of gains over conventional RIS, even with low losses. The validation of the proposed modeling of distributed RIS through full-wave simulators constitutes an intriguing future research avenue.

Chapter 10

Conclusion and Future Works

In this thesis, we investigate BD-RIS, a technology recently proposed as a generalization of conventional RIS. We analyze the fundamental performance limits of BD-RIS with ideal hardware and communication models and explore how hardware non-idealities impact the performance limits. Specifically, our three main contributions are as follows. First, we optimize the scattering matrix of group- and fully-connected RISs through a closed-form global optimal solution and optimize the grouping strategy in group-connected RIS based on the channel statistics. Second, we propose two novel BD-RIS architectures, namely forest- and tree-connected RISs, to approach the best trade-off between performance and circuit complexity, and characterize the Pareto frontier of such a trade-off. Third, we model and optimize BD-RIS architectures in the presence of hardware impairments, namely discrete-value scattering matrix, presence of EM mutual coupling, and lossy interconnections. In the following, we delineate six future research directions for BD-RIS.

Further BD-RIS Architectures

A crucial challenge is to identify the BD-RIS architectures offering the optimal trade-off between achieved performance and circuit complexity. In this thesis, we have presented the forest- and tree-connected BD-RIS architectures and showed how to achieve the Pareto frontier of the performance-complexity trade-off in single-user systems [80, 85]. However, the fundamental limits of such a trade-off are still unknown for more complex wireless systems, such as multi-user, which constitute an interesting future research direction.

New Applications for BD-RIS

Beyond improving the performance, BD-RIS enables full-space coverage through the hybrid and multi-sector modes [65, 66]. Interestingly, this application is enabled by the interconnections present in BD-RIS and is not feasible through conventional RIS. Thus, an open research direction is to investigate whether BD-RIS can unlock completely new opportunities that are not possible with conventional RIS.

Interplay Between BD-RIS and Other Technologies

Recent works explored the interplay of BD-RIS with other technologies, namely RSMA [56, 57, 58], massive MIMO [59], UAV [60], MEC [61], and SWIPT [62], showing promising advantages. However, additional research is needed to show the benefits of BD-RIS integrated with novel technologies and in other wireless scenarios such as WPT, RF sensing, and integrated sensing, communication, and WPT. In particular, a new direction has been recently opened by implementing stacked intelligent metasurface (SIM), i.e., an emerging technology enabling wave domain beamforming through multiple stacked RISs, through BD-RIS [102]. Given the promising benefits of BD-RIS in realizing SIM [102], the interplay between BD-RIS and SIM represents a brand new avenue for research.

Codebook-Based BD-RIS Design and Optimization

In recent literature on BD-RIS, it is commonly assumed that the BD-RIS scattering matrix can be optimized with arbitrary values based on perfect channel knowledge. However, in practice, the BD-RIS scattering matrix cannot be reconfigured with arbitrary precision and channel estimation protocols may cause high communication overhead [54, 87]. To address these two problems, a codebook-based optimization for BD-RIS could be investigated. Specifically, a low-cardinality codebook of candidate scattering matrices could be built in an offline stage according to the channel statistics. Based on this codebook, at every coherence time interval, the BD-RIS switches between all the candidate scattering matrices while pilot signals are sent from the transmitter to the receiver. After comparing the received pilots, the receiver feeds back the index of the best BD-RIS scattering matrix, allowing BD-RIS optimization based on discrete values with no explicit channel estimation.

Physically Consistent Models of BD-RIS

In most of the related literature, BD-RIS has been modeled by assuming ideal hardware and RF design. Despite BD-RIS with hardware impairments has been studied in this thesis, more work has to be done to develop physically consistent models of BD-RIS. Specifically, we identify three relevant research avenues. First, models for BD-RIS with lossy tunable impedance components should be studied, and the corresponding achievable performance should be assessed. Second, in this thesis and in the related literature, the studied BD-RIS-aided communication systems are implicitly assumed to be narrowband, such that the RIS scattering matrix is constant in frequency. However, in wideband communication systems, e.g., employing OFDM, the properties of the RIS tuning circuit made of positive-intrinsic negative (PIN) diodes or voltage-controlled varactor change with the frequency. Thus, further work is needed to model these frequency-dependent variations and to optimize BD-RIS in wideband systems. Third, further effort is needed to bridge the model of RIS-aided systems based on multiport network theory developed in Chapter 8 with other models based on different formalisms, such as the method of moments [103].

BD-RIS Prototyping

As BD-RIS is a very recent concept [16], the deployment of BD-RIS prototypes is still in progress. The main difference over conventional RIS prototypes lies in the reconfigurable impedance network connected to the RIS antenna array, which includes additional tunable impedance components interconnecting the RIS elements. Since such components can be realized through passive and low-cost devices, such as PIN diodes or voltage-controlled varactor, the cost of BD-RIS is expected to be similar to that of conventional RIS.

Bibliography

[1] B. Clerckx and C. Oestges, *MIMO wireless networks: channels, techniques and standards for multi-antenna, multi-user and multi-cell systems.* Academic Press, 2013.

[2] A. Paulraj and C. Papadias, "Space-time processing for wireless communications," *IEEE Signal Process. Mag.*, vol. 14, no. 6, pp. 49–83, 1997.

[3] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[4] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–20, 2019.

[5] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.

[6] C. Holloway, M. Mohamed, E. Kuester, and A. Dienstfrey, "Reflection and transmission properties of a metafilm: with an application to a controllable surface composed of resonant particles," *IEEE Trans. Electromagn. Compat.*, vol. 47, no. 4, pp. 853–865, 2005.

[7] L. Subrt and P. Pechac, "Intelligent walls as autonomous parts of smart indoor environments," *IET communications*, vol. 6, no. 8, pp. 1004–1010, 2012.

[8] N. Kaina, M. Dupré, G. Lerosey, and M. Fink, "Shaping complex microwave fields in reverberating media with binary tunable metasurfaces," *Scientific reports*, vol. 4, no. 1, p. 6693, 2014.

[9] N. Kaina, M. Dupré, M. Fink, and G. Lerosey, "Hybridized resonances to design tunable binary phase metasurface unit cells," *Optics express*, vol. 22, no. 16, pp. 18 881–18 888, 2014.

[10] M. Dupré, P. Del Hougne, M. Fink, F. Lemoult, and G. Lerosey, "Wave-field shaping in cavities: Waves trapped in a box with controllable boundaries," *Physical review letters*, vol. 115, no. 1, p. 017701, 2015.

[11] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[12] 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.

[13] 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.

[14] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 3, pp. 1546–1577, 2021.

[15] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov, and S. Shamai, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, 2020.

[16] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1229–1243, 2022.

[17] 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.

[18] 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.

[19] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, 2020.

[20] Y. Liu, J. Zhao, M. Li, and Q. Wu, "Intelligent reflecting surface aided MISO uplink communication network: Feasibility and power minimization for perfect and imperfect CSI," *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 1975–1989, 2021.

[21] 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.

[22] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522–4535, 2020.

[23] H. Li, W. Cai, Y. Liu, M. Li, Q. Liu, and Q. Wu, "Intelligent reflecting surface enhanced wideband MIMO-OFDM communications: From practical model to reflection optimization," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4807–4820, 2021.

[24] Y. Xiu, J. Zhao, W. Sun, M. Di Renzo, G. Gui, Z. Zhang, and N. Wei, "Reconfigurable intelligent surfaces aided mmwave NOMA: Joint power allocation, phase shifts, and hybrid beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8393–8409, 2021.

[25] A. Bansal, K. Singh, B. Clerckx, C.-P. Li, and M.-S. Alouini, "Rate-splitting multiple access for intelligent reflecting surface aided multi-user communications," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 9217–9229, 2021.

[26] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 3187–3196, 2020.

[27] S. Guo, S. Lv, H. Zhang, J. Ye, and P. Zhang, "Reflecting modulation," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2548–2561, 2020.

[28] Z. Feng, B. Clerckx, and Y. Zhao, "Waveform and beamforming design for intelligent reflecting surface aided wireless power transfer: Single-user and multi-user solutions," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5346–5361, 2022.

[29] Y. Zhao, B. Clerckx, and Z. Feng, "IRS-aided SWIPT: Joint waveform, active and passive beamforming design under nonlinear harvester model," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1345–1359, 2022.

[30] J. Hu, H. Zhang, B. Di, L. Li, K. Bian, L. Song, Y. Li, Z. Han, and H. V. Poor, "Reconfigurable intelligent surface based RF sensing: Design, optimization, and implementation," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2700–2716, 2020.

[31] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Tran. Signal Process.*, vol. 69, pp. 5386–5402, 2021.

[32] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, 2020.

[33] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, 2021.

[34] Z. Peng, Z. Chen, C. Pan, G. Zhou, and H. Ren, "Robust transmission design for RIS-aided communications with both transceiver hardware impairments and imperfect CSI," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 528–532, 2022.

[35] Y. Chen, Y. Wang, and L. Jiao, "Robust transmission for reconfigurable intelligent surface aided millimeter wave vehicular communications with statistical CSI," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 928–944, 2022.

[36] B. Zheng, C. You, and R. Zhang, "Double-IRS assisted multi-user MIMO: Cooperative passive beamforming design," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4513–4526, 2021.

[37] W. Mei and R. Zhang, "Intelligent reflecting surface for multi-path beam routing with active/passive beam splitting and combining," *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1165–1169, 2022.

[38] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, 2020.

[39] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, 2020.

[40] C. You, B. Zheng, and R. Zhang, "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2604–2620, 2020.

[41] H. Guo and V. K. N. Lau, "Uplink cascaded channel estimation for intelligent reflecting surface assisted multiuser MISO systems," *IEEE Trans. Signal Process.*, vol. 70, pp. 3964–3977, 2022.

[42] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel estimation with reconfigurable intelligent surfaces – a general framework," *Proc. IEEE*, vol. 110, no. 9, pp. 1312–1338, 2022.

[43] R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson, "Active reconfigurable intelligent surface-aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4962–4975, 2021.

[44] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1707–1725, 2023.

[45] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. Di Renzo, C.-B. Chae, and L. Hanzo, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45 913–45 923, 2020.

[46] S. Zhao, R. Langwieser, and C. F. Mecklenbraeuker, "Reconfigurable digital metasurface for 3-bit phase encoding," in *WSA 2021; 25th International ITG Workshop on Smart Antennas*, 2021.

[47] J. Rao, Y. Zhang, S. Tang, Z. Li, S. Shen, C.-Y. Chiu, and R. Murch, "A novel reconfigurable intelligent surface for wide-angle passive beamforming," *IEEE Trans. Microw. Theory Tech.*, vol. 70, no. 12, pp. 5427–5439, 2022.

[48] H. Li, S. Shen, M. Nerini, and B. Clerckx, "Reconfigurable intelligent surfaces 2.0: Beyond diagonal phase shift matrices," *IEEE Commun. Mag.*, vol. 62, no. 3, pp. 102–108, 2024.

[49] I. Santamaria, M. Soleymani, E. Jorswieck, and J. Gutiérrez, "SNR maximization in beyond diagonal RIS-assisted single and multiple antenna links," *IEEE Signal Process. Lett.*, vol. 30, pp. 923–926, 2023.

[50] T. Fang and Y. Mao, "A low-complexity beamforming design for beyond-diagonal RIS aided multi-user networks," *IEEE Commun. Lett.*, vol. 28, no. 1, pp. 203–207, 2024.

[51] M.-A. Kim, S.-G. Yoo, H.-D. Kim, K.-H. Shin, and H.-K. Song, "Scattering matrix design of reconfigurable intelligent surface based on group connected impedance network in MU-MIMO system," in *2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2023, pp. 642–645.

[52] Y. Zhou, Y. Liu, H. Li, Q. Wu, S. Shen, and B. Clerckx, "Optimizing power consumption, energy efficiency and sum-rate using beyond diagonal RIS – a unified approach," *IEEE Trans. Wireless Commun.*, 2023.

[53] G. Bartoli, A. Abrardo, N. Decarli, D. Dardari, and M. Di Renzo, "Spatial multiplexing in near field MIMO channels with reconfigurable intelligent surfaces," *IET Signal Processing*, vol. 17, no. 3, p. e12195, 2023.

[54] H. Li, Y. Zhang, and B. Clerckx, "Channel estimation for beyond diagonal reconfigurable intelligent surfaces with group-connected architectures," in *2023 IEEE 9th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2023, pp. 21–25.

[55] H. Li, S. Shen, M. Nerini, M. Di Renzo, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces with mutual coupling: Modeling and optimization," *IEEE Commun. Lett.*, 2024.

[56] T. Fang, Y. Mao, S. Shen, Z. Zhu, and B. Clerckx, "Fully connected reconfigurable intelligent surface aided rate-splitting multiple access for multi-user multi-antenna transmission," in *2022 IEEE ICC Workshops*, 2022, pp. 675–680.

[57] M. Soleymani, I. Santamaria, E. Jorswieck, and B. Clerckx, "Optimization of rate-splitting multiple access in beyond diagonal RIS-assisted URLLC systems," *IEEE Trans. Wireless Commun.*, 2023.

[58] M.-A. Kim, S.-G. Yoo, H.-D. Kim, K.-H. Shin, Y.-H. You, and H.-K. Song, "Group-connected impedance network of RIS-assisted rate-splitting multiple access in MU-MIMO wireless communication systems," *Sensors*, vol. 23, no. 8, p. 3934, 2023.

[59] A. Mishra, Y. Mao, C. D'Andrea, S. Buzzi, and B. Clerckx, "Transmitter side beyond-diagonal reconfigurable intelligent surface for massive MIMO networks," *IEEE Wireless Commun. Lett.*, vol. 13, no. 2, pp. 352–356, 2024.

[60] A. M. Huroon, Y.-C. Huang, and L.-C. Wang, "Optimized transmission strategy for UAV-RIS 2.0 assisted communications using rate splitting multiple access," in *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*, 2023.

[61] A. Mahmood, T. X. Vu, W. U. Khan, S. Chatzinotas, and B. Ottersten, "Joint computation and communication resource optimization for beyond diagonal UAV-IRS empowered MEC networks," *arXiv preprint arXiv:2311.07199*, 2023.

[62] T. D. Hua, M. Mohammadi, H. Q. Ngo, and M. Matthaiou, "Cell-free massive MIMO SWIPT with beyond diagonal reconfigurable intelligent surfaces," *arXiv preprint arXiv:2402.00646*, 2024.

[63] J. Xu, Y. Liu, X. Mu, and O. A. Dobre, "STAR-RISs: Simultaneous transmitting and reflecting reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3134–3138, 2021.

[64] H. Zhang, S. Zeng, B. Di, Y. Tan, M. Di Renzo, M. Debbah, Z. Han, H. V. Poor, and L. Song, "Intelligent omni-surfaces for full-dimensional wireless communications: Principles, technology, and implementation," *IEEE Commun. Mag.*, vol. 60, no. 2, pp. 39–45, 2022.

[65] H. Li, S. Shen, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces: From transmitting and reflecting modes to single-, group-, and fully-connected architectures," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2311–2324, 2023.

[66] ——, "Beyond diagonal reconfigurable intelligent surfaces: A multi-sector mode enabling highly directional full-space wireless coverage," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2446–2460, 2023.

[67] ——, "Synergizing beyond diagonal reconfigurable intelligent surface and rate-splitting multiple access," *IEEE Trans. Wireless Commun.*, 2024.

[68] B. Wang, H. Li, Z. Cheng, S. Shen, and B. Clerckx, "A dual-function radar-communication system empowered by beyond diagonal reconfigurable intelligent surface," *arXiv preprint arXiv:2301.03286*, 2023.

[69] H. Li, S. Shen, and B. Clerckx, "A dynamic grouping strategy for beyond diagonal reconfigurable intelligent surfaces with hybrid transmitting and reflecting mode," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16 748–16 753, 2023.

[70] Q. Li, M. El-Hajjar, I. Hemadeh, A. Shojaeifard, A. A. M. Mourad, B. Clerckx, and L. Hanzo, "Reconfigurable intelligent surfaces relying on non-diagonal phase shift matrices," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6367–6383, 2022.

[71] Q. Li, M. El-Hajjar, I. Hemadeh, A. Shojaeifard, and L. Hanzo, "Coordinated reconfigurable intelligent surfaces: Non-diagonal group-connected design," *IEEE Trans. Veh. Technol.*, 2024.

[72] H. Wang, Z. Han, and L. Swindlehurst, "Channel reciprocity attacks using intelligent surfaces with non-diagonal phase shifts," *arXiv preprint arXiv:2309.11665*, 2023.

[73] M. Nerini, H. Li, S. Shen, and B. Clerckx, *Beyond diagonal RIS.* Springer, 2024 (to appear as a book chapter).

[74] D. M. Pozar, *Microwave engineering.* John wiley & sons, 2011.

[75] M. Nerini, S. Shen, and B. Clerckx, "Closed-form global optimization of beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1037–1051, 2024.

[76] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 586–590, 2020.

[77] M. Fiedler, "Bounds for the determinant of the sum of hermitian matrices," *Proceedings of the American Mathematical Society*, vol. 30, no. 1, pp. 27–31, 1971.

[78] R. A. Horn and C. R. Johnson, *Matrix analysis.* Cambridge university press, 2012.

[79] M. Nerini, S. Shen, and B. Clerckx, "Static grouping strategy design for beyond diagonal reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, 2024.

[80] M. Nerini, S. Shen, H. Li, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces utilizing graph theory: Modeling, architecture design, and optimization," *IEEE Trans. Wireless Commun.*, 2024.

[81] S. Shen and R. D. Murch, "Impedance matching for compact multiple antenna systems in random RF fields," *IEEE Trans. Antennas Propag.*, vol. 64, no. 2, pp. 820–825, 2016.

[82] J. A. Bondy and U. S. R. Murty, *Graph theory with applications.* Macmillan London, 1976, vol. 290.

[83] D. O'Leary and G. Stewart, "Computing the eigenvalues and eigenvectors of symmetric arrowhead matrices," *J. Comput. Phys.*, vol. 90, no. 2, pp. 497–505, 1990.

[84] G. Strang, *Introduction to linear algebra.* Wellesley-Cambridge Press Wellesley, MA, 1993, vol. 3.

[85] M. Nerini and B. Clerckx, "Pareto frontier for the performance-complexity trade-off in beyond diagonal reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2842–2846, 2023.

[86] L. V. Ahlfors, *Complex analysis: An introduction to the theory of analytic functions of one complex variable.* McGraw-Hill, 1953.

[87] M. Nerini, S. Shen, and B. Clerckx, "Discrete-value group and fully connected architectures for beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16 354–16 368, 2023.

[88] C. Audet and J. E. Dennis Jr, "Analysis of generalized pattern searches," *SIAM Journal on optimization*, vol. 13, no. 3, pp. 889–903, 2002.

[89] C. M. Bishop, *Pattern Recognition and Machine Learning.* Springer Verlag, Aug. 2006.

[90] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, 2014.

[91] M. Di Renzo, F. H. Danufane, and S. Tretyakov, "Communication models for reconfigurable intelligent surfaces: From surface electromagnetics to wireless networks optimization," *Proc. IEEE*, vol. 110, no. 9, pp. 1164–1209, 2022.

[92] G. Gradoni and M. Di Renzo, "End-to-end mutual coupling aware communication model for reconfigurable intelligent surfaces: An electromagnetic-compliant approach based on mutual impedances," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 938–942, 2021.

[93] J. A. Nossek, D. Semmler, M. Joham, and W. Utschick, "Physically consistent modelling of wireless links with reconfigurable intelligent surfaces using multiport network analysis," *arXiv preprint arXiv:2308.12223*, 2023.

[94] M. T. Ivrlač and J. A. Nossek, "Toward a circuit theory of communication," *IEEE Trans. Circuits Syst. I: Regul. Pap.*, vol. 57, no. 7, pp. 1663–1683, 2010.

[95] M. Nerini, S. Shen, H. Li, M. Di Renzo, and B. Clerckx, "A universal framework for multiport network analysis of reconfigurable intelligent surfaces," *arXiv preprint arXiv:2311.10561*, 2023.

[96] W. Kahn and H. Kurss, "Minimum-scattering antennas," *IEEE Trans. Antennas Propag.*, vol. 13, no. 5, pp. 671–675, 1965.

[97] X. Qian and M. Di Renzo, "Mutual coupling and unit cell aware optimization for reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1183–1187, 2021.

[98] P. Mursia, S. Phang, V. Sciancalepore, G. Gradoni, and M. Di Renzo, "SARIS: Scattering aware reconfigurable intelligent surface model and optimization for complex propagation channels," *IEEE Wireless Commun. Lett.*, vol. 12, no. 11, pp. 1921–1925, 2023.

[99] H. E. Hassani, X. Qian, S. Jeong, N. S. Perović, M. Di Renzo, P. Mursia, V. Sciancalepore, and X. Costa-Pérez, "Optimization of RIS-aided MIMO – a mutually coupled loaded wire dipole model," *IEEE Wireless Commun. Lett.*, vol. 13, no. 3, pp. 726–730, 2024.

[100] M. Akrout, F. Bellili, A. Mezghani, and J. A. Nossek, "Physically consistent models for intelligent reflective surface-assisted communications under mutual coupling and element size constraint," *arXiv preprint arXiv:2302.11130*, 2023.

[101] M. Nerini, G. Ghiaasi, and B. Clerckx, "Localized and distributed beyond diagonal reconfigurable intelligent surfaces with lossy interconnections: Modeling and optimization," *arXiv preprint arXiv:2402.05881*, 2024.

[102] M. Nerini and B. Clerckx, "Physically consistent modeling of stacked intelligent metasurfaces implemented with beyond diagonal RIS," *IEEE Commun. Lett.*, 2024.

[103] K. Konno, S. Terranova, Q. Chen, and G. Gradoni, "Generalised impedance model of wireless links assisted by reconfigurable intelligent surfaces," *arXiv preprint arXiv:2306.03761*, 2023.