Phase Shifts Optimization for IRS-Assisted SIMO/MISO Channels

by

Milad Dabiri

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirement for degree of

Doctor of Philosophy in Electrical and Computer Engineering

School of Electrical Engineering and Computer Science

Faculty of Engineering

University of Ottawa

ABSTRACT

Although 5G has significantly improved the capabilities of wireless networks compared to previous generations, the ever-growing number of users and their demands surpass the limits of 5G. As a result, attentions have drawn to new technologies for the next generation, one of which is the smart radio environment. The Intelligent Reflecting Surface (IRS) is a major enabler of this concept by providing some control over the previously untouched radio environment. In this thesis, we consider IRS-assisted SIMO/MISO channels where channel capacity involves optimization over IRS phase shifts, presenting a non-convex problem for which no closed-form solutions are known. We have obtained several closed-form bounds for the general case, which are tight in some special cases and thus provide a globally optimal solution to the original problem. We discuss some practically important cases and provide closed-form global optimal solutions, such as in mmWave/THz channels, single-reflector IRS, mMIMO settings, and multi-IRS channels. We propose a computationally efficient iterative algorithm for the general case based on a closed-form globally optimal solution for the single-reflector case. Its convergence to a local optimum is rigorously proved, and several cases are identified where its convergence point is also globally optimal. Numerical experiments show that the algorithm converges quickly in practice, and its convergence point is close to the global optimum. Further, we define a novel concept of phase dispersion and propose an analytical approach based on a concave lower bound. We show that the lower bound maximization problem is convex and can be solved in closed form. It is further shown that the maximized lower bound is tight (coincides with the global optimum of the original problem) in many cases and provides an accurate approximation in others.

In memory of my father, Towfigh

To

my beloved mother, Nahid

my lovely sister, Maryam.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my sincere gratitude to my supervisor, Professor Sergey Loyka, for his unrelenting support, invaluable guidance, knowledge, and constructive suggestions throughout my studies. Working and studying under his guidance was a great privilege and honor.

Heartfelt gratitude and love go to my mother, Nahid, for providing me with abundant support and continuous encouragement throughout my years of study, and to my late father, Towfigh, for his constant motivation to pursue my Ph.D., even though he did not live to see this achievement. Also, I would like to thank my sister, Maryam, for her unwavering support and inspiration.

Contents

Abbreviations

5G	Fifth Generation
6G	Sixth Generation
AOA	Alternating Optimization Algorithm
ANN	Artificial Neural Network
AP	Access Point
AF	Amplify and Forward
AO	Alternate Optimization
BCD	Block Coordinate Descent
BS	Base Station
BnB	Branch and Bound
CIR	Channel Impulse Response
CLT	Central Limit Theorem
CSI	Channel State Information
D2D	Device-to-Device
DF	Decode and Forward
EB	ExaByte
EE	Energy Efficiency
FPGA	Field Programmable Gate Array
FD	Full Duplex
FPI	Fixed Point Iteration
HCT	Holographic Communication Technology
i.i.d	Independent Identically Distributed
IRS	Intelligent Reflecting Surface
IoE	Internet of Everything

ILP	Integer Linear Programming
KKT	Karush–Kuhn–Tucker
KPI	Key Performance Indicator
LISA	Large Intelligent Surface/Antenna
LoS	Line of Sight
MF	Matched Filter
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MMSE	Minimum Mean Square Error
MO	Manifold Optimization
MRC	Maximum Ratio Combining
OFDM	Orthogonal Frequency Division Multiplexing
PB	PetaByte
QCQP	Quadratically Constrained Quadratic Program
QoS	Quality of Service
RF	Radio Frequency
RIS	Reconfigurable Intelligent Surface
SCA	Successive Convex Approximation
SDP	Semidefinite Program
SDR	Semidefinite Relaxation
SE	Spectral Efficiency
SIMO	Single-Input Multiple-Output
SINR	Signal to Interference and Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
STAR-RIS	Simultaneously Transmit and Reflect-RIS
THz	Terahertz
VR	Virtual Reality

ZF	Zero Forcing

List of Symbols

$\mathbf{A}$	Bold capitals ($\mathbf{A}$) denote matrices
$\mathbf{a}$	Bold lower case letters ($\mathbf{a}$) denote vectors
ϕ_l	The induced phase shift at l-th reflector, p. 18
L	The number of IRS reflectors, p. 18
h_0	The gain of direct Tx-Rx link, p. 18
h_l	The gain of Tx-IRS-Rx link via l-th reflector, p. 18
P_{rx}	The receive power, p. 20
P_{tr}	The transmit power, p. 21
N	The number of antennas at receiver, p. 29
x	The scalar transmitted signal, p. 29
h_l^{Tx}	The gain of link from transmitter to l-th reflector at IRS, p. 29
$h_{l,n}^{Rx}$	The gain of link from l-th reflector at IRS to n-th antenna at receiver, p. 29
$\mathbf{y}$	The receiver signal vector, p. 29
$\mathbf{z}$	Zero-mean white Gaussian circularly-symmetric noise vector at the receiver with variance σ_0^2, p. 29
$\mathbf{h}_0$	The direct Tx-Rx channel vector, p. 29
$\mathbf{h}_l$	The reflected Tx-IRS-Rx channel vector via l-th reflector, p. 29
γ_0	The Rx SNR in the unit-gain channel, p. 31
$\mathbf{h}_{eq}$	The equivalent IRS-assisted channel vector, p. 31
$\boldsymbol{\phi}$	The IRS phase shifts vector, p. 31

$\gamma(\boldsymbol{\phi})$	The IRS-assisted Rx SNR for give $\boldsymbol{\phi}$, p. 31
$g(\boldsymbol{\phi})$	The IRS gain for given $\boldsymbol{\phi}$, p. 31
$R(\boldsymbol{\phi})$	The achievable rate per unit bandwidth for IRS-assisted channel as a function of $\boldsymbol{\phi}$, p. 32
C_{IRS}	The IRS-assisted channel capacity, p. 32
$\mathbf{H}$	The channel matrix, p. 35
$\sigma_l(\mathbf{H})$	The l-th largest singular value of matrix $\mathbf{H}$, p. 35
g_{lb}	The lower bound for $g(\boldsymbol{\phi})$, p. 35
g_{ub}	The upper bound for $g(\boldsymbol{\phi})$, p. 35
g_{UB}	The upper bound for $g(\boldsymbol{\phi})$, p. 39
$\mathbf{a}(\theta)$	The local array response vector for a ULA with the angle of arrival/departure of θ, p. 50
θ_{Rx}	The angle of arrival at the Rx, p. 50
θ_{IRS}^{I}	The incidence angle at IRS, p. 50
d	The element spacing in a ULA, p. 50
λ	The wavelength, p. 50
$\mathbf{H}_m$	The channel matrix for m-th IRS in a multi-IRS setup, p. 57
L_m	The number of reflectors for m-th IRS in a multi-IRS setup, p. 58
$\mathcal{I}_m$	The index set of columns in $\mathbf{H}_m$, p. 58
$\boldsymbol{\phi}_0$	The initial phase shift vector for IRS, p. 67
$g^{(i)}$	The IRS gain after i-th outer loop iteration in Algorithm 1, p. 67
$g^{(i,l)}$	The IRS gain after l-th inner loop of i-th outer loop iteration in Algorithm 1, p. 67
Δg	The threshold set for stopping criteria in Algorithm 1, p. 67
$\boldsymbol{\phi}^{(i,l)}$	The phase shift vector after step l-th of inner loop, at step $i-th$ of the outer loop in Algorithm 1, p. 69

$m_{\max}$	The maximum number of initial points in Algorithm 2, p. 73
g_{out}	The output IRS gain from Algorithm 1, p. 73
$\boldsymbol{\phi}_c$	The phase shift vector at convergence point, p. 74
$g^{(c)}$	The IRS gain at convergence point , p. 74
$g_a(\phi)$	The approximated IRS gain for given $\boldsymbol{\phi}$, p. 83
$\Delta\theta_{nl}$	The phase difference of the IRS-reflected and LoS links for n-th antenna and l-th reflector, p. 92
Θ	The largest phase difference between the reflected and LoS links, across all antennas and reflectors, p. 92

Chapter 1

Introduction

1.1 Motivation

In recent decades, mobile technologies have undergone dramatic revolutions, with almost one significant change every 10 years. Each generation of mobile technology has significantly improved performance (spectral/energy efficiency, latency, etc.) [1]. These rapid transitions are a response to the unceasing increase in the number of mobile devices and their corresponding rate demands. For instance, the total number of mobile subscribers increased from 0.3 million in 1985 to 5.7 billion by 2023, representing a 19,000-fold increase [2,3]. Meanwhile, the demand for content continues to grow at an extreme rate, and trends indicate that video traffic accounts for the highest share of traffic across all device types, followed by social networking [4]. There seems to be no saturation to this growth as mobile data traffic has grown more than 10 times in the past 6 years, surpassing forecasts [4].

For the upcoming years, mobile data traffic is forecast to experience an exponential increase, from 100 PB [1] per month in 2008 to 562.7 EB [2] per month in 2029 [4,5]. Consequently, efforts and studies in academia and industry are intensively focused on designing new wireless network technologies to satisfy this growth. These efforts have led to the development of the Fifth Generation (5G) technology, the latest commercialized wireless technology. Not only does it significantly improve data rates, but it also enables substantial enhancements in terms of reliability, latency, power

[1] PetaByte, 10^{15} bytes.
[2] ExaByte, 10^{18} bytes.

	5G	**6G**
Peak data rate	$\geq$ 1 Gbps	$\geq$ 1 Tbps
Experienced data rate	Downlink: $\geq$ 100 Mbps Uplink: $\geq$ 50 Mbps	Downlink: $\geq$ 1 Gbps Uplink: $\geq$ 500 Mbps
Maximum bandwidth	1 GHz	100 GHz
Spectrum efficiency	3$\times$ that of 4G	$5-10\times$ that of 5G
Network energy efficiency	$10-100\times$ that of 4G	$10-100\times$ that of 5G
Connection density	$10^6/\text{km}^2$	$10^7/\text{km}^2$
Reliability	99.999%	99.99999%
Mobility	500 km/h	1000 km/h
Latency	1 ms	10 - 100 μs

Table 1.1: Comparision of the network KPIs for 5G and 6G [9–11].

efficiency, and massive connectivity [6].

The 5G cellular system was marketed as the key enabler for services such as the Internet of Everything (IoE), Virtual Reality (VR), autonomous vehicles, and distributed automation [7]. However, the requirements of these applications - ultra-high data rates, extremely low latency, high reliability, and real-time access to powerful computing resources - often surpass the capabilities currently offered by 5G. This has directed research and studies towards the sixth generation (6G), which aims to satisfy more demanding requirements than 5G. Table 1.1 summarizes the Key Performance Indicators (KPIs) for 5G and 6G. The key enabling technologies for 6G include Terahertz communication, Intelligent Reflecting Surface (IRS), Blockchain, and Holographic Communication Technology (HCT) [8].

IRS is an enabling technology for 6G and has become an active area of research for enhancing the performance of the next generation wireless networks [12–14]. The significant increase in the number of studies on 'intelligent reflecting surface' technology is illustrated in Fig. 1.1, highlighting the active research in this area. This trend is reflected in the number of overview/survey/tutorial studies conducted from different perspectives on IRS, such as in [15–23]. Additionally, several journals have special issues on this topic, e.g., [24].

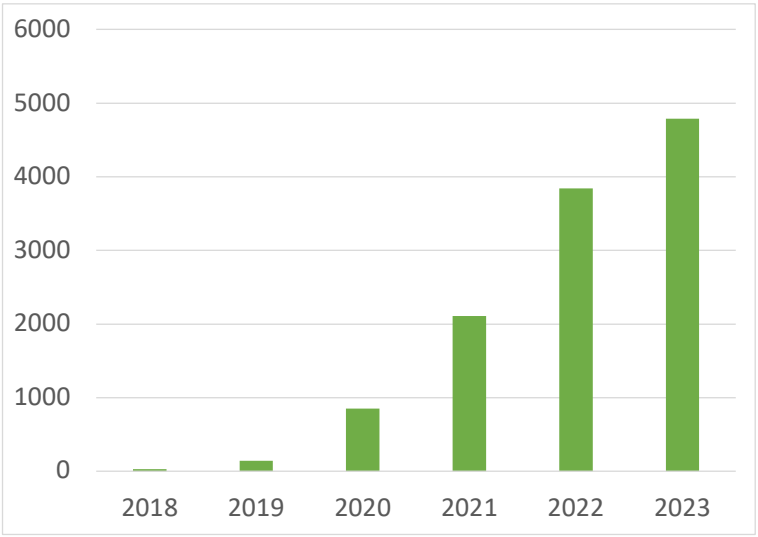


Figure 1.1: The number of publications including the term 'intelligent reflecting surface' or other equivalent terms, as returned by the Google Scholar search engine.

Research on IRS has not been limited to theoretical aspects; a significant number of experimental studies have also been conducted, including the creation of prototypes and performance measurements that are obtained via IRS. Several of these studies are referenced in [25–31]. Additionally, the industry has turned its attention to IRS, and various companies and project groups have launched several large-scale tests and trial projects on this technology [32].

An IRS typically comprises a planar surface incorporating a large number of passive reflectors, each capable of being programmed to induce specific amplitudes and/or phase shifts to the incident signal independently. IRSs can be distributed throughout a wireless network and smartly programmed to configure the propagation environment between the transmitter and receiver in a desired manner, thus achieving specific goals.

Several alternative terms are used in the literature to refer to IRS. These terms include intelligent walls [33], smart reflect-arrays [34], Reconfigurable Intelligent Surface (RIS) [16, 35], Large Intelligent Surface/Antenna (LISA) [36], Simultaneously Transmitting and Reflecting-RIS (STAR-RIS) [37], and RFocus [31]. While there

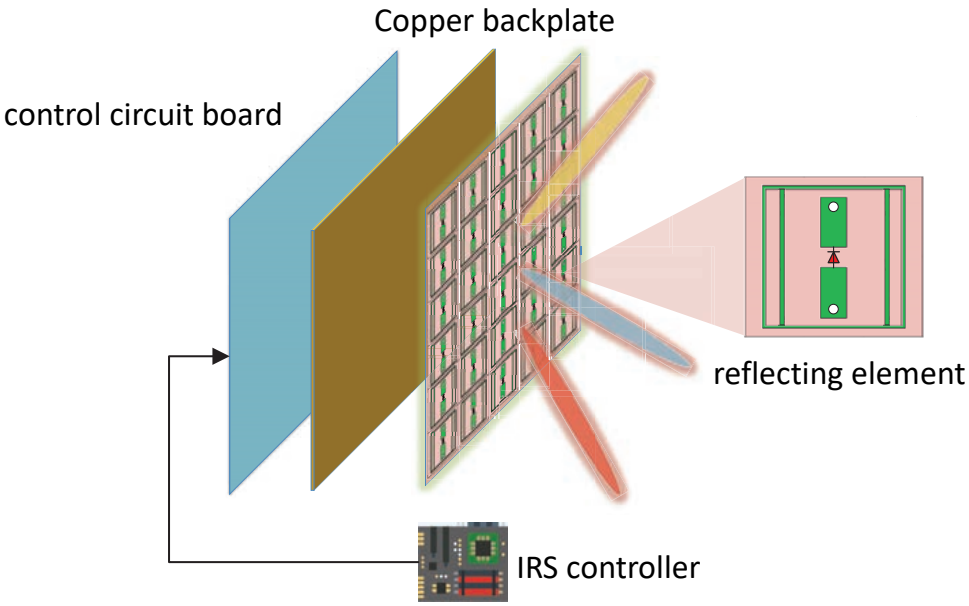


Figure 1.2: Architecture of IRS [12].

are minor differences in terminology among some of them, they all share the same principle of using passive and reconfigurable reflective elements.

Although several hardware implementations of IRS have been suggested in the literature, they all share the same underlying principles. Fig. 1.2 demonstrates a basic example of IRS's architecture, as described in [12]. This IRS implementation consists of three layers and a controller. The outer layer, designed to interact with incident signals, incorporates a large number of patches made of metamaterials printed on a dielectric surface. A copper plate is used in the second layer to minimize the energy loss of the incident signal, while the last layer applies the phase shifts and reflection coefficients, as set by the controller for each reflector individually. The controller, typically a Field-Programmable Gate Array (FPGA), optimizes IRS elements in conjunction with other network elements and uses its own dedicated link to the network.

Overall, integrating IRS into the wireless network allows some control over the propagation environment, which was untouched by previous generations of wireless communication. This additional control naturally creates different optimization

problems for the IRS to solve to achieve specific objectives. For example, studies in [38–41] discuss how IRS assistance enhances energy/spectral efficiencies. A measurement-based study in [25] reports that the receiver's Signal-to-Noise Ratio (SNR) improves by 12.65 dB with IRS assistance. Further, IRS is shown capable of interference nulling under certain conditions, as discussed in [42]. A more detailed discussion of IRS functionalities is provided in Chapter 2.

1.2 Objectives and Scope

There are several gaps in the existing studies on IRS optimization. A common feature of known studies is the absence of analytical solutions due to the difficult analytical structure of the considered problem, even in its simplest (but non-trivial) form, i.e., for a Multiple-Input Single-Output (MISO) or Single-Input Multiple-Output (SIMO) channel. Further, the proposed algorithms converge to local rather than global optimum; it is not clear how far these local optimal points are from the global optimum, or whether the algorithms can converge to a global optimum under certain conditions. Further, no bounds have been established for the global optimality gap of the locally optimal solutions found in previous studies.

The major focus of this thesis is on phase shift optimization for IRS-assisted SIMO/MISO channels. The receiver's SNR serves as our main performance metric, with IRS gain defined based on this, as discussed later in the thesis. The IRS phase shift optimization problem is then formulated, which has a difficult analytical structure and is not convex. This underlines the importance of analytical study in determining the maximum achievable IRS gain and determining the conditions under which it can be achieved. Additionally, we examine how IRS performance is affected by system parameters, particularly the number of IRS reflectors.

1.3 Thesis Contributions

In this thesis, we discuss the IRS optimization problem for SIMO/MISO channels in several ways. The major contributions of this thesis are summarized as follows:

- IRS gain is our performance metric, in which maximizing IRS gain leads to maximizing receiver achievable rates. We establish upper and lower bounds to the globally optimal IRS gain in the general case, which are tight in some special cases, thus providing a globally optimal solution in those cases. These bounds indicate that the globally optimal IRS gain scales at least as the largest singular value squared of the equivalent (IRS-assisted) channel. These bounds are then used to set up an initial point that can be used for a numerical method, thereby setting an upper bound on the global optimality gap.

- We obtain a closed-form, globally optimal solution for several practically important special cases, including mmWave/THz channels, single-reflector IRS, massive MIMO settings, and multi-IRS channels. It is demonstrated that in mmWave/THz channels, the IRS-assisted globally optimal SNR scales either linearly or quadratically with the number of reflectors.

- We propose the Alternating Optimization Algorithm (AOA), an iterative algorithm based on the closed-form globally optimal solution for the single-reflector IRS. We then use the previously obtained bounds to establish a guaranteed global optimality gap for this method. It is demonstrated that this algorithm outperforms those in [38, 39, 43], and its convergence point is within 1 dB of the global optimum for the considered channels. To the best of our knowledge, this is the first time an algorithm for IRS optimization has been proposed with a guaranteed global optimality gap, which is relatively small in practice.

- We propose a novel analytical approach to tackle the IRS optimization problem, which is based on a concave lower bound. This convex optimization problem can be solved for a global maximum in closed form. It is further shown that the maximized lower bound is tight (coinciding with the global optimum of the original problem) in many cases and is an accurate approximation in some other cases. To this end, a novel concept of the small phase dispersion regime is introduced, and the optimized lower bound is shown to be tight in the small phase dispersion regime, from which a global optimality gap of simplified 1-bit phase sifters is obtained.

1.4 Publications

The results of this thesis have been presented in:

1. M. Dabiri and S. Loyka, "Optimizing IRS-Assisted SIMO/MISO Channels: An Analytical Approach," International Zurich Seminar on Information and Communication (IZS), Zurich, Switzerland, Mar. 2024, accepted.

2. M. Dabiri and S. Loyka, "Globally-Optimal IRS Designs for SIMO/MISO Channels: An Analytical Approach," IEEE Transactions on Wireless Communications, submitted, Nov. 2023.

3. M. Dabiri and S. Loyka, "IRS-Assisted SIMO/MISO Channels in the Small Phase Dispersion Regime," IEEE Transactions on Communications, submitted, Nov. 2023.

4. M. Dabiri and S. Loyka, "On Globally-Optimal IRS Design for SIMO/MISO Channels," 2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Toronto, ON, Canada, pp. 1-6, Sep. 2023.

5. M. Dabiri and S. Loyka, "On The Capacity of IRS-Assisted Gaussian SIMO Channels," 2022 17th Canadian Workshop on Information Theory (CWIT), Ottawa, ON, Canada, pp. 6-10, Jun. 2022.

1.5 Organization of the Thesis

The rest of this thesis is organized as follows:

Chapter 2: In this chapter, we present an overview of the literature on IRS optimization for SIMO/MISO channels, considering different channel conditions and setups. Both theory and measurements show that IRS assistance enhances performance. However, due to the non-convex nature of the underlying optimization problems, no globally optimal solution has been presented in previous studies, and the results obtained are either approximations or, at best, locally optimal solutions. Several studies investigate the rate of performance improvement achievable via IRS as a function of the number of reflectors. Additionally, some studies compare the performance obtained from IRS with other techniques, such as relaying, and discuss the scenarios in which each technique can outperform the others.

Chapter 3: A SIMO uplink channel assisted by an IRS equipped with L reflectors is presented, forming the basis of our further analysis in this thesis. Using matched filtering, which remains optimal for this setup, the SNR on the receiver side is derived as a function of IRS phase shifts. While our model is primarily based on a SIMO channel, the results obtained are also applicable to MISO channels, due to uplink/downlink duality. The equal IRS-assisted channel is then defined, and based on it, the IRS gain as a performance metric. Then, we define the IRS gain optimization problem, which will be investigated using different approaches in subsequent chapters.

Chapter 4: To establish globally optimal IRS phase shifts in some cases and to assess the global optimality gap in others, a number of bounds are established in this chapter. Through analysis and numerical experiments, we demonstrate that these bounds are achievable in certain special cases and reasonably tight in others. Furthermore, based on these bounds, we obtain an initial point that can be used for methods requiring initialization. Using this initial point automatically sets an upper bound on the global optimality gap, which, to the best of our knowledge, is being established for the first time.

Chapter 5: This chapter lists a number of practically important cases in which the non-convex IRS optimization problem can be solved with a globally optimal solution in closed form. These cases include mmWave/THz channels, single-reflector IRS, massive MIMO settings, and multi-IRS channels. For each one of these cases, we obtain the globally optimal phase shifts, supported by rigorous proofs. We further discuss the IRS gain scaling as a function of the number of reflectors and identify the channel conditions that can affect this scaling.

Chapter 6: In this chapter, we propose the AOA method, an iterative method based on the closed-form globally optimal solution obtained for a single-reflector IRS. This method applies to the general problem and, since it uses a closed-form solution for updates in each iteration, is computationally efficient. However, because this closed-form solution is applied iteratively, optimizing one phase shift at a time while keeping the rest fixed, it cannot guarantee a globally optimal convergence point in the general case. To deal with this issue, we propose the multi-start AOA method to reduce its sensitivity to the initial point. Numerical results for various channels and different numbers of reflectors demonstrate that the AOA method can outperform the known algorithms [38, 39, 43].

Chapter 7: A novel analytical approach is presented in this chapter, based on the concave lower bound. The maximized lower bound is shown to be tight in many cases and provides an accurate approximation in others. To this end, we introduce the novel concept of phase dispersion and consider the small phase dispersion regime. Further, it is demonstrated that the optimized lower bound is tight in the small phase dispersion regime, from which a globally optimal gap for simplified IRS with 1-bit phase shifters is obtained.

Chapter 2

Literature review

To put our work into perspective, in this chapter, we provide a brief overview of the existing literature on intelligent reflective surface technology. We review the increasing trend of studies on various aspects of IRS, exploring analytical and experimental studies conducted in both academy and industry sectors. Since the main focus of this study is the phase shift optimization problem for IRS, we discuss the most relevant studies in greater detail and organize them into three categories: analytical results, numerical methods, and measurement results.

2.1 Smart Radio Environment

To accommodate the ever-increasing number of mobile users, along with their demanding data rates and diverse quality of service (QoS) requirements, efforts to improve energy and spectrum efficiency in wireless networks have continued unabated. Previously, performance optimization in wireless networks focused either on the user side or on the network side, i.e., the base station (BS) or access point (AP). On the network side, the growing traffic demand can be accommodated by designing energy-efficient small cells in dense networks or by using multiple antennas to achieve higher spectrum efficiency [44]. Further, power allocation and transmit beamforming design problems can be optimized for adapting to channel variations. At the user side, devices can be equipped with multiple antennas, such as in MIMO and massive MIMO, or different users can collaborate in relaying transmission, as in Device-to-Device (D2D) scenarios [45, 46].

However, these technologies have faced several challenges in their implementation in 5G, and it is questionable whether they will be able to meet the goals set for the next generations of wireless networks, e.g., 6G [32]. For example, massive MIMO requires a large number of active transmit radio-frequency (RF) chains, which significantly increases energy consumption and raises the hardware cost. Shrinking cell size to support higher data rates also requires a greater number of BSs and APs, not only increasing energy consumption and maintenance costs but also leading to additional practical issues such as interference management [32]. Further, mmWave and THz communications require additional active network components to compensate for their more severe propagation loss.

To overcome these challenges, attention has been drawn to the radio environment, which remained an uncontrollable element in the previous framework of wireless networks. Due to the randomness of the radio environment, the signals propagated from the transmitter experience reflection, diffraction, and scattering before reaching the receiver. This results in multiple delayed and attenuated copies of the transmitted signal. This fading effect is one of the major limiting factors in maximizing energy and spectral efficiency performance in wireless networks so far [20]. Accordingly, the vision of a smart radio environment is to gain some control over the radio environment using technologies such as IRS [36, 47].

2.2 Intelligent Reflecting Surface

An innovative IRS concept has recently been introduced in [13, 15] to enable a smart radio environment. IRS is typically described as a two-dimensional surface composed of a large array of passive reflecting elements made of metamaterials. Each reflecting element can be individually controlled in a software-defined manner, allowing it to impose a specific phase shift on the incident signal. By performing a joint phase shift optimization over all reflectors, IRS enables creating desirable beamforming

toward the receiver. In particular, the reflected signals can be configured to have a constructive effect to improve the received signal power, or they can be adjusted to have a destructive effect, thereby mitigating interference among users, depending on the design goals.

The low hardware requirements and passive nature of IRS make it an ideal technology to deploy in network environments, for example, on the walls of buildings. IRS deployment also does not require modifying the hardware on the transmitter or receiver side and can seamlessly integrate with the current design of wireless networks. Further, IRS does not impose a significant energy consumption burden on the network due to its passive reflectors. IRS assistance does not complicate the interference management problem, unlike the use of more active elements in the network.

IRS-assisted wireless networks create additional optimization problems and are expected to revolutionize the current network optimization paradigm, which is expected to be critical to future wireless networks [44, 48]. This potential has attracted significant attention toward IRS, leading to numerous studies and research projects exploring its various aspects. The number of overviews, surveys, and tutorials conducted from different perspectives on IRS reflects this growing interest. Table 2.1 lists some of these studies. In the meantime, multiple experimental studies using prototypes have been conducted to validate the theoretical research and further prove the concept of IRS. Table 2.2 lists some of these experimental studies.

IRS technology has also received substantial interest within the industry, as demonstrated by the large-scale tests and trials, and funded projects launched by various companies and project groups worldwide. Table 2.3 lists some of these funded research projects on IRS, while a more detailed list can be found in [32].

Ref.	type/year	contribution
[15]	overview/2019	An introduction to smart radio environments enabled by IRS, the functionalities that can be achieved, and the fundamental knowledge gaps.
[16]	survey/2020	A comprehensive overview of IRS from theory to design, covering implementation challenges and various use cases.
[17]	tutorial/2021	A discussion on three main challenges in IRS technology: reflections optimization, channel estimation, and deployment from a system designer perspective.
[18]	survey/2020	A comprehensive survey of optimization problems in IRS, along with the approaches employed to solve them.
[19]	tutorial/2022	A discussion on system modelling, providing a tutorial on IRS fundamentals from a signal processing perspective.
[20]	survey/2020	A literature review on performance analysis, considering different metrics for IRS-assisted networks.
[21]	survey/2021	A discussion on different IRS optimization problems, with a focus on machine learning methods.
[22]	survey/2022	A survey on practical design issues, including channel estimation and beamforming design.
[23]	survey/2021	A survey focusing on THz communication via IRS, from theory to implementation.

Table 2.1: List of overview/survey/tutorial papers on IRS.

Ref.	setup	validation
[25]	SISO indoor environment with a 16×16 1-bit IRS at 5.8 GHz	BER improvement and SNR enhancement of 12.65 dB.
[26]	SISO indoor environment with a 16×16 2-bit IRS at 2.3 GHz and 28.5 GHz	Antenna gains of 21.7 dBi and 19.1 dBi at 2.3 GHz and 28.5 GHz, respectively.
[27]	SISO indoor environment with a 20×20 1-bit IRS at 5.4 GHz	Analysis of SNR improvement using IRS in near-field and far-field mobile scenarios.
[28]	SISO indoor environment with a 2×2 active 2-bit IRS at 2.4 GHz	8.5 dB SNR gain for active IRS compared to passive settings.
[29]	SISO indoor environment with 3 prototypes, 8×32-100×102 at 4.25 GHz and 10.5 GHz	Measurements validate the analytical results for different near field and far field scenarios.
[30]	SISO indoor environment with a 640 1-bit elements IOS-IRS at 3.6 GHz	Analyzing the pattern configuration and the half-power beam width.
[31]	SISO indoor environment with 3200 1-bit elements IRS at 2.4 GHz	Prototype improves the channel capacity by 2 times.

Table 2.2: Experimental studies on IRS.

title	objective	duration	budget
VisoSurf	Joint hardware and software design for IRS, featuring two experimental prototypes.	Jan 2017 Dec 2020	€5,748,000
ARIADNE	Sets the theoretical and algorithmic bases for IRS-empowered wireless networks, with an emphasis on AI and high-frequency transmission.	May 2021 Apr 2023	€184,707.84
META WIRELESS	Studying the manipulation of wireless propagation environments via IRS.	Dec 2020 Dec 2024	€3,995,128.44
RISE-6G	Studies and standardization practices for transforming the technically advanced vision of IRS into industrial exploitation.	Jan 2021 Dec 2023	€6,499,613.75

Table 2.3: A number of funded research projects on IRS from [32].

2.3 IRS Advantages

Previous studies on IRS have discussed the various functionalities that IRS can offer under specific scenarios; this implies that different optimization objectives can be defined for IRS-assisted communication. A number of IRS functionalities are listed as follows:

- **Coverage enhancement**: deploying IRS allows to overcome physical obstacles between the transmitter and the receiver, providing better coverage for blind spots and virtual Line-of-Sight (LoS) [49].

- **Higher spectral/energy efficiency**: IRS-assistance improves spectral and energy efficiency in specific scenarios, as demonstrated in [50,51]. For instance, this can be achieved by enhancing multiplexing gain by creating multi-path transmission environments or by constructively manipulating reflecting links toward the receiver.

- **Channel improvement**: Using IRS, we gain some control to manipulate the effective channel to our advantage. This may include converting channels from fast-fading to slow-fading [52], improving channel rank in MIMO networks [53], and achieving channel hardening[1] [55].

[1]The phenomena in which small-scale fading and frequency dependence disappear for a large number of antennas [54, pp. 11-16].

- **Interference management**: IRS is capable of effectively cancelling inter/intra cell interference, and thus, it can serve the purpose of interference nulling under specific conditions, such as those described in [42].

2.4 Key Challenges

Although IRS-assisted wireless communication offers significant potential and functionalities, it also presents significant challenges in theory and practice. Often discussed in the literature, such as in [17, 20, 22], these issues can be divided into the following three categories:

- **Phase shift optimization**: phase shift optimization for the passive reflectors of the IRS creates challenging optimization problems. Based on the objective function (e.g., achievable rate maximization, interference minimization), different scenarios, such as single/multi IRS or single/multi-antenna BS/user, lead to different optimization problems.

- **Channel estimation**: to optimize IRS reflectors, access to Channel State Information (CSI) is valuable. However, this presents a challenging issue for two reasons. First, the passive nature of IRS reflectors means they are unable to transmit or receive pilot signals. Second, IRS typically incorporates a large number of reflectors, each with different channel coefficients, which makes acquiring this amount of information difficult.

- **Hardware constraints**: in numerous studies, IRS is often treated as ideal for the sake of analytical simplicity. Nevertheless, realistic implementations are subject to hardware constraints and impairments that impact the system's performance.

2.5 Phase Shift Optimization

This thesis concentrates on the IRS phase shift optimization problem, aiming to maximize channel capacity/receiver SNR for SIMO/MISO channels[2]. A globally optimal solution to this problem identifies the fundamental performance limits that can be achieved by IRS assistance. Although theoretical limitations (e.g., simplifying assumptions), along with practical constraints (e.g., non-ideal hardware), can have adverse effects on both theoretical analysis and the practical performance of the IRS, a globally optimal solution will serve as an ultimate upper bound and can be used as a benchmark to evaluate the performance of known algorithms in terms of their global optimality gap, which remains unknown at this time.

Using IRS gives system designers some control to configure the channel in accordance with their design objectives. The effective channel becomes dependent on IRS phase shift selection. In practice, the phase and amplitude for each IRS reflector are not independently controllable, but coupled [57, 58]. To characterize the ultimate IRS performance, we follow the previous studies [30, 38, 39, 43, 47], and consider the ideal case where we can optimize the phases and amplitudes independently with arbitrary precision. Also, each reflector is assumed to have a constant phase shift over the entire bandwidth. Our focus will be on the phase shift optimization problem. Although previous studies [38, 39, 43, 59, 60] have investigated this ideal model, a globally optimal solution to the IRS phase shift optimization problem for general SIMO/MISO case is not yet known either analytically or numerically. The strategy for tackling the SNR maximization problem varies depending on channel specifications (e.g., the number of antennas, the number of users, and continuous/discrete phase shifts). This thesis investigates IRS-assisted SIMO channels for single-user applications with continuous phase shifts, one of the cases for which a globally optimal

[2]To simplify the notation, the IRS phase shift optimization to maximize receive SNR will henceforth be referred to as the SNR maximization problem, which has the same solution as channel gain maximization and transmit power minimization [56].

solution has not yet been identified. This SIMO channel setting can also be extended to a MISO channel via the uplink-downlink duality [56]. While extensive literature on IRS optimization under various setups exists, Table 2.4 lists studies that are more closely related to what we have studied. A more detailed description of the relevant studies falls into the following categories: analytical results, numerical methods, and measurement results.

2.5.1 Analytical Results

Analytical solutions for SNR maximization in SIMO/MISO channels are not feasible due to the complexity of the underlying optimization problem. To overcome this challenge, earlier studies [38, 39, 43, 57, 59–61] have proposed various practically valuable numerical methods. However, they suffer from the fundamental weakness that their convergence point is locally optimal at best and can be far from a globally optimal one. This makes it challenging to evaluate the fundamental performance limits from a communication/information-theoretic perspective.

Unlike SIMO/MISO channels, where no analytical results are available, the SNR maximization for SISO channels is straightforward using the globally optimal solution presented in [30, 47, 62–67]. In this setup, the phase shift for each reflector is adjusted to compensate for the phase difference between the reflected and direct links, which we refer to as "co-phasing." The global optimality of this method is based on triangle inequality:

$$|h_0 + \sum_{l=1}^{L} e^{j\phi_l} h_l| \leq |h_0| + \sum_{l=1}^{L} |h_l|, \tag{2.1}$$

where ϕ_l is the induced phase shift at l-th IRS reflector, L is the number of IRS reflectors, h_0 and h_l denote direct Tx-Rx link and Tx-IRS-Rx link via l-th reflector, respectively. Note that the upper bound in (2.1) is independent of ϕ_l and the

Ref.	Optimization	System model	Phase shifts	CSI	Optimality	Method
[30]	Max. SNR	SU-SISO	Discrete	Perfect	Local	Analytical
[47]	Max. SNR	SU-SISO	Continuous	Perfect	Global	Analytical
[63]	Max. SINR	MU-SISO	Continuous	Perfect	Local	Analytical
[64]	Max. SINR	MU-SISO	Continuous	Perfect	Local	Analytical
[65]	Max. SNR	SU-SISO	Continuous	Perfect	Global	Analytical
[68]	Max. SINR	SU-SISO	Discrete	Estimate	Local	Semidefinite relaxation (SDR)
[69]	Max. SNR	SU-SISO	Discrete	Perfect	Local	Approximation algorithm, alternating optimization (AO), SDR
[38, 39]	Max. SNR Min. Tx power	SU/MU-MISO	Continuous	Perfect	Local	AO, SDR, distributed algorithm
[43]	Max. SNR	SU-MISO	Continuous	Perfect	Local	Fixed-point iteration (FPI), manifold optimization (MO)
[70]	Max. ergodic capacity	SU-MISO	Continuous	Statistical	Local	AO, analytical
[60]	Max. Min. SINR	MU-MISO	Continuous	Perfect	Local	AO, SDR, successive convex approximation (SCA)
[71]	Max. SNR/SINR	MU/SU-MISO	Discrete	Perfect	Local	AO, zero-forcing (ZF) method, minimum mean square error (MMSE) method
[57]	Min. Tx power	SU/MU-MISO	Continuous	Perfect	Local	AO, two-stage and penalty-based algorithms
[59]	Max. sum rate	MU-MISO	Continuous	Perfect/im-perfect	Local	AO, MO, block coordinate descent (BCD), SCA

Table 2.4: The studies of SNR/capacity maximization in IRS-assisted channels.

inequality holds with equality if phase shifts are set as follows:

$$\phi_l^* = \arg(h_0) - \arg(h_l), \quad l = 1...L, \tag{2.2}$$

where for a complex number z, $\arg(z)$ returns its argument (phase). The results in (2.2) hold for any SISO channel, e.g., deterministic or fading channels. In [47], a Rayleigh fading IRS-assisted SISO channel is considered; it is shown that by optimizing IRS due to (2.2), the average receive SNR is proportional to L^2 according to the Central Limit Theorem (CLT). The study in [39] also considers a Rayleigh fading channel; however, it is shown that $P_{rx} \sim L^2$ only applies to optimal phase shifts as in (2.2) where P_{rx} denotes the received power [3]. However, for random or zero phase shifts ($\phi_l = 0$), $P_{rx} \sim L$. This result, referred to as the "power scaling law" in [39], claims that the "squared gain" occurs because the IRS benefits not only from the transmit beamforming gain of the order of L in the IRS-Rx link but also from the inherent aperture gain of the order of L by collecting more signal power in the Tx-IRS link. As a result, [39] compares the power scaling for IRS with the power scaling for a Full-Duplex (FD) Amplify-and-Forward (AF) relay equipped with L transmit and L receive antennas. This comparison argues that when L is asymptotically large, the received SNR with the FD AF relay increases linearly with L, whereas the IRS surpasses this by having a quadratic growth with L. The obtained power scaling for IRS, however, does not hold under the asymptotic regime where $L \to \infty$, as it violates the law of conservation of energy. Since it is not possible to receive more power than was transmitted, the power scaling law must saturate.

The two-ray model [73, pp. 34-37] is used in [47] to derive the power scaling law and to demonstrate that $P_{rx} \sim L^2$. However, this result is not physically achievable due to the limited surface area of each reflector and its inability to cover the entire

[3]For functions $f(x)$ and $g(x)$ of a real variable x, if $f(x)/g(x)$ tends to some non-zero constant as $x \to \infty$, we write $f(x) \sim g(x)$ [72, pp. 4-8].

first Fresnel zone. To clarify the power scaling law for IRS, Björnson and Sanguinetti in [40] pose the following question:

"How can an IRS with L passive reflecting elements be more power-efficient than massive MIMO (mMIMO) with L active antennas?"

This question is answered in [40] by comparing the SNR expressions achieved by mMIMO with that of an IRS-assisted transmission. It is demonstrated that the receive SNR achieved by the IRS grows at a faster rate than that of mMIMO, with growth rates of L^2 and L, respectively. Yet, for any given L, IRS can never achieve a higher SNR than a mMIMO setup with a matching array size and transmit power. In far-field transmissions, however, an IRS with more reflectors (compared to the number of antennas in a mMIMO receiver) will provide higher spectral efficiency (SE).

Also, the IRS requires a considerable number of reflectors (L) to obtain SNR comparable to mMIMO. The upper bound for the IRS power scaling is discussed in [40, 41]. To obtain the asymptotic behaviour as $L \to \infty$, the studies in [38, 39, 47] assume that Tx remains in the far-field of the IRS. This assumption is not physically possible since as L grows asymptotically, the far-field approximation eventually breaks down. In [41], the upper bound for the received power is obtained as follows:

$$P_{rx} \to \frac{P_{tr}}{9}, \text{ as } L \to \infty, \tag{2.3}$$

where P_{tr} represents the transmit power.

A comparison between IRS and Decode-and-Forward (DF) relaying is performed in [41, 74], with both studies considering the classic repetition-coded DF relaying protocol. It is demonstrated that the choice between an IRS and relay depends on the number of reflectors (L) and required receive SNR. The IRS achieves higher energy efficiency (EE) than DF relaying if high transmission rates are required. Also, using an IRS only provides results comparable to DF relaying for large enough values

of L.

The IRS-assisted Rayleigh fading channel in [49,51] is further analyzed in [55,64]. In [64], it is shown that the ratio of the mean value to the standard deviation of the receive SNR increases proportionally with $\sqrt{L}$. So, the SNR variations average out (in relative terms) as the number of IRS reflectors increases and channel hardening occurs. This result is further investigated in [55] for planner array IRS setups and compared to the channel hardening in mMIMO channels. It is shown that IRS-assisted channels are subject to channel hardening under specific circumstances: IRS phase shifts should be optimal and not random. Further, once a stronger Tx-Rx link exists, more reflectors are required to achieve channel hardening.

The interference-nulling capability of IRS in a K-user single-antenna Line-of-Sight (LoS) interference channel has been studied in [42] and it is shown that, if no direct Tx-Rx links are present, large-enough IRS (with $\approx 2K^2$ reflectors) is able to cancel interference completely.

2.5.2 Numerical Methods

Several numerical methods have been proposed for SNR maximization in different setups of IRS-aided SIMO/MISO channels. In this section, we discuss in more detail those that are more relevant to our study.

The joint transmit beamforming vector and IRS phase shifts optimization problem for an IRS-assisted MISO channel is studied in [38, 39], where two methods are proposed: semidefinite relaxation (SDR) and distributed methods. Using the Alternating Optimization (AO) method, the joint optimization problem is first decoupled into two sub-problems: optimizing the transmit beamforming vector for fixed IRS phase shifts and IRS phase shifts optimization for a fixed beamforming vector. These two sub-problems are then solved iteratively until convergence. The transmit beamforming vector optimization for fixed IRS phase shifts is solved using Maximum Ratio

Combining (MRC), which results in a closed-form optimal solution [56]. The phase shift optimization sub-problem for a fixed transmit beamforming vector, which is a non-convex optimization problem, is then reformulated as a Quadratically Constrained Quadratic Program (QCQP). This optimization problem is subject to the following constraints:

$$\mathbf{W} \geq \mathbf{0} \quad \text{and} \quad \text{rank}(\mathbf{W}) = 1, \tag{2.4}$$

where $\mathbf{w} = [1, e^{j\phi_1}, ..., e^{j\phi_L}]^T$ and $\mathbf{W} = \mathbf{w}\mathbf{w}^H$. Since the rank-one constraint is non-convex, it is then relaxed (i.e., SDR method). The relaxed problem is a standard convex Semidefinite Program (SDP), which is then solved using CVX [75, 76]. The obtained solution for the relaxed problem, $\mathbf{W}$, can be of a higher rank. It is then approximated with a rank-one matrix using the Gaussian randomization method. However, it is not known whether the phase shifts obtained for the relaxed problem are also optimal for the original problem, either locally or globally. Further, IRS is typically considered to be equipped with many reflectors; since the computational complexity of the SDR method scales polynomial with the number of reflectors ($L^{6.5}$), this poses a practical challenge, as discussed in [77–79]. Another method proposed in [38] is the distributed method, which uses the AO method and then solves the phase shift optimization problem based on the triangle inequality. However, similar to the SDR method, it is not known if these methods can guarantee optimality, either locally or globally.

The same joint optimization problem as in [38,39] is considered in [43]. Likewise, the AO method is applied, and the transmit beamforming optimization sub-problem is solved with an optimal closed-form solution using MRC. Two novel approaches are proposed for the phase shift optimization sub-problem: the Fixed-Point Iteration (FPI) and Manifold Optimization (MO) methods. The FPI method uses an auxiliary variable to rewrite the objective function as a quadratic programming problem. The

phase shift optimization problem is then solved iteratively, where, within each step, one phase shift vector is kept fixed from the previous iteration, and the other is updated to maximize the objective function.

In the MO method, the phase shift vector is iteratively updated until convergence, and this method requires initialization. During each iteration, the update vector is determined by the Riemannian gradient and the selection of a step size. The provided numerical results demonstrate that both methods outperform the SDR method, and their computational complexity is lower in comparison. However, it is not clear if the FPI and MO methods outperform the SDR method only for the channels used in the provided simulation results or if it holds in the general case. Based on the provided comparison between these two methods, MO achieves a slightly higher SNR at the expense of higher computational complexity. Yet, neither method can guarantee local or global optimality, even in special cases. Further, due to the absence of high-quality bounds, it is not possible to determine the global optimality gap.

The phase shift optimization problem for IRS-assisted MISO channels is also examined in [61], where Gradient Projection (GP) and Cross-Entropy (CE) methods are proposed. In the GP method, the objective function is reformulated as a quadratic programming problem, and the phase shifts are updated iteratively using the gradient descent method. It is claimed that the GP method is more computationally efficient compared to the MO method. The CE method, which is based on reinforcement learning techniques, is also proposed for cases with low-resolution discrete phase shift values (1-2 bits). The numerical results show that the GP method performs close to the MO method but with reduced complexity.

An IRS-assisted MISO channel in a multi-user setup is discussed in [57]. The optimization problem is defined to minimize the total transmit power at the BS by jointly optimizing the transmit beamforming at the BS and phase shifts at the IRS, subject to the individual Signal-to-Interference-and-Noise (SINR) constraints at all

users. In the considered model for IRS, the reflection coefficient for each reflector is not independent of its induced phase shift; instead, it is a function of it. First, the reflection coefficient for IRS reflectors is approximated as a function of the induced phase shift, specifically obtained for a SISO channel where this approximation is valid when L tends to infinity. This approximation is then applied to the transmit power minimization problem, but it requires further proof of whether such an approximation is valid for MISO channels. By applying the AO method, IRS phase shifts are optimized iteratively. Each iteration optimizes one phase shift using an approximated solution while the rest remain fixed. Based on the accuracy of the approximated function obtained for the non-linear reflection coefficient, a trust region is established. The phase shift optimization is then performed by searching the trust region or by fitting a curve within the same region. This process continues until the convergence point is reached. Yet, it is mentioned that the obtained solution is, at best, sub-optimal, and no comparison with previously suggested methods has been performed to clarify the superiority of the proposed method.

In an IRS-assisted MISO channel, transmit power minimization with discrete phase shifts is examined for single and multi-user cases in [71]. The optimization problem is reformulated for the single-user case as an Integer Linear Programming (ILP) problem. The Branch-and-Bound (BnB) method is claimed to provide the optimal solution to this problem; however, no proof is provided, which is crucial since the complexity scales exponentially with the number of bits and L. Then, a successive refinement algorithm is proposed to reduce complexity, which achieves a sub-optimal solution. The optimization problem is then formulated for the multi-user case, where each user is subject to a specific SINR constraint. It is shown that the successive refinement algorithm, which uses a Zero Forcing (ZF)-based precoder at the BS, achieves a locally optimal solution for an approximated problem.

IRS optimization to maximize the weighted sum rate of multiple users is studied

in [59] for both perfect and imperfect CSI. As suggested in [43], the MO method is used when perfect CSI is available. To reduce computational complexity compared with the MO method, a novel method is proposed, which is based on introducing a dual Lagrangian problem and applying Successive Convex Approximations (SCA) and Block Coordinate Descent (BCD) methods. Using the SCA method, the non-convex objective function is iteratively approximated with convex functions, which can be locally optimized using the BCD method. As a result of approximating the objective function, this method provides an approximated solution to the original problem. Numerical results illustrate the advantages provided by IRS using these methods, but are not compared with other methods.

2.5.3 Measurement Results

Several studies have examined the practical implementation of IRS using prototypes to validate the concept of this technology in practice, such as those listed in Table 2.2. However, to the best of our knowledge, these studies have all focused on IRS-assisted SISO channels, while none have considered IRS assistance for SIMO/MISO channels. This section discusses a few of these studies and their contributions.

The study in [80] examines several IRS prototypes for indoor and outdoor measurements. One prototype incorporates 1100 binary half-wavelength size reflectors, where each reflector can be assigned with phase shifts of $\pm\frac{\pi}{2}$. Each IRS reflector is implemented with a varactor diode. It is shown that phase shifts change in the 5.2 - 6.4 GHz range; however, they remain fixed for a bandwidth of 20 MHz. In an indoor measurement, where a 30 cm concrete wall blocks the direct Tx-Rx link, a 26 dB improvement in Rx power is reported. However, the exact distance between the transmitter and receiver is not specified. Further, two outdoor measurements were conducted at distances of 50 and 500 meters. These measurements report improvements of 27 dB and 14 dB in Rx power, respectively, compared to using a copper

plate instead of an IRS. However, the geometry of the setups in these measurements needs to be clearly described, and there are no comparative measurements between scenarios with and without IRS assistance. This lack of comparison makes it difficult to verify the reported measurements with the analytical results in [40, 41].

IRS performance in a rich scattering environment is considered in [81]. An IRS, equipped with 23 binary reflectors, is built that operates at 2.5 GHz. The phase shifter of each reflector can be switched between On/Off states, corresponding to phase shifts of π and 0, respectively. The phase shift values are shown to remain approximately constant within a bandwidth of 400 MHz. The channel impulse response (CIR) is measured for various IRS configurations to determine the channel capacity. IRS optimization is conducted using an artificial neural network (ANN) with 256 neurons. However, it does not provide measured values, e.g., receive SNR improvement, that could be validated against the analytical results.

A more detailed analytical discussion, supported by experimental results, is provided in [29, 82]. Path loss analysis for near-field and far-field distances is performed analytically without any condition on IRS size. Further, a broadcasting scenario is considered where a large IRS is located close to the transmitter. For these cases, the received power is formulated, and power scaling with the number of reflectors is discussed. Three prototype IRSs are built for measurement purposes: two large IRSs with 100×102 and 50×34 reflectors, respectively, and one smaller prototype with 8×32 reflectors. The two large IRSs are used to analyze the analytical results for broadcasting and near-field scenarios. The measurements are conducted with Tx-IRS and IRS-Rx distances of 1 meter, operating at 10.5 GHz. These measurements show a 3 dB difference between the analytical and experimental results. Also, measurements for the far-field case agree with the analytical results. However, in all these cases, only measurements and analytical results are compared, with no comparison made to demonstrate the improvement in receive SNR compared to scenarios

without IRS. Further, in this study, the analytical results are obtained for continuous phase shifts, while the measurements are based on discrete values of phase shifts. Near-field path loss is also analyzed using the free space model, but its validity is questionable.

Chapter 3

Channel Model and Problem Formulation

Let us consider an IRS-assisted SIMO (uplink) channel as shown in Fig. 3.1, consisting of a single-antenna transmitter (Tx, e.g., user equipment), an IRS equipped with L passive reflectors, and a receiver (Rx) equipped with N antennas (e.g., a base station). Following the standard discrete-time baseband model [83] [35]- [42], the signal y_n received by n-th antenna is

$$y_n = \big(h_{0,n} + \sum_{l=1}^{L} e^{j\phi_l} h_l^{Tx} h_{l,n}^{Rx}\big)x + z_n, \tag{3.1}$$

where x is the scalar transmitted (Tx) signal satisfying the Tx power constraint $\mathbb{E}\{|x|^2\} = P$, and $\mathbb{E}\{\cdot\}$ is the statistical expectation. $z_n \sim \mathcal{CN}(0, \sigma_0^2)$ is the circularly-symmetric, complex Gaussian noise, i.i.d. across antennas, of zero mean and its variance is σ_0^2 per Rx antenna; $h_{0,n}$ represents the direct (e.g., LoS) Tx-Rx path while h_l^{Tx} and $h_{l,n}^{Rx}$ represent the Tx-IRS and IRS-Rx links via l-th reflector (all including the average propagation loss and the reflection loss), and ϕ_l is the reflector-induced phase shift. This model can be compactly expressed as follows:

$$\mathbf{y} = \big(\mathbf{h}_0 + \sum_{l=1}^{L} e^{j\phi_l} \mathbf{h}_l\big)x + \mathbf{z}, \tag{3.2}$$

where $\mathbf{y} = [y_1, ..., y_N]^T \in \mathbb{C}^{N\times 1}$ is the Rx signal vector, $\mathbf{z} \in \mathbb{C}^{N\times 1}$ is the noise vector; $\mathbf{h}_0, \mathbf{h}_l \in \mathbb{C}^{N\times 1}$ are the channel vectors representing the direct Tx-Rx and the reflected Tx-IRS-Rx links (via l-th reflector), n-th entry of $\mathbf{h}_l$ is $h_{l,n} = h_l^{Tx} h_{l,n}^{Rx}$. Further, $|\mathbf{h}|$,

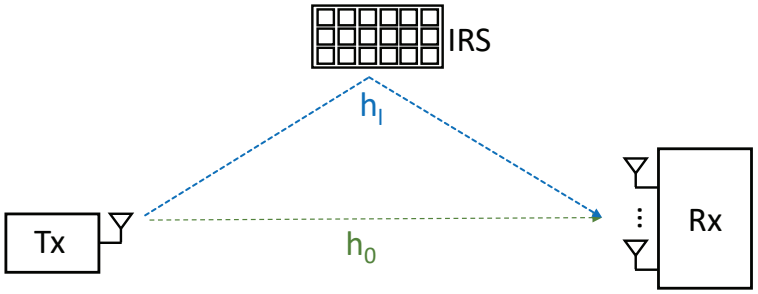


Figure 3.1: An illustration of IRS-assisted SIMO channel; $\mathbf{h}_0$ represents the direct Tx-Rx link while $\mathbf{h}_l$ - the l-th reflected link.

$\mathbf{h}^T$ and $\mathbf{h}^+$ denote Euclidean norm (length), transposition and Hermitian conjugation of column vector $\mathbf{h}$, respectively. For scalar h, h^+ denotes complex conjugation. The channel is further assumed to be static or quasi-static (stays fixed for a sufficiently long time), frequency-flat[1], with full channel state information (CSI) available to the Rx and IRS controller. The availability of CSI is adopted as a typical assumption in the literature, e.g., [38, 39, 43]. Otherwise, previously proposed channel estimation methods, such as those in [79, 84, 85], can be used to obtain such information. These results can also be extended to frequency-selective channels using an OFDM-type approach, e.g., [77, 86, 87]. It is not fundamentally impossible to optimize IRS reflectors to induce independent phase shifts at different frequencies, which is a requirement for wideband OFDM transmission. However, current hardware cannot accommodate this.

Note that this model accommodates single-IRS as well as multi-IRS settings, where L is the total number of reflectors and $\mathbf{h}_l$, $l = 1...L$, represents the respective IRS-assisted links. It is sufficiently general to address various scenarios, including deterministic, fading, LoS, non-LoS (or partially-blocked LoS), and multi-path scenarios (since we do not make any specific assumptions on $\mathbf{h}_0$, $\mathbf{h}_l$ at this point).

[1]The delay spread of the channel is much smaller than the symbol duration [56]

3.1 IRS Gain

In this section, we consider the system model introduced above and define IRS gain for that model. For the no-IRS case, the Rx SNR or power is maximized via matched filtering (MF)/beamforming (also known as MRC) and can be expressed as [56]

$$\gamma_{no-IRS} = |\mathbf{h}_0|^2\gamma_0,\ \gamma_0 = P/\sigma_0^2, \tag{3.3}$$

Absorbing the average propagation path loss into P, γ_0 becomes the Rx SNR in the unit-gain channel. Likewise, for the IRS-assisted channel in (3.2), $\mathbf{h}_{eq} \triangleq \mathbf{h}_0 + \sum_{l=1}^{L} e^{j\phi_l}\mathbf{h}_l$ is the equivalent channel vector and the MF beamforming does maximize its Rx SNR/power,

$$\gamma(\boldsymbol{\phi}) = g(\boldsymbol{\phi})\gamma_0,\ g(\boldsymbol{\phi}) = |\mathbf{h}_{eq}|^2 = \left|\mathbf{h}_0 + \sum_{l=1}^{L} e^{j\phi_l}\mathbf{h}_l\right|^2, \tag{3.4}$$

where $\boldsymbol{\phi} = [\phi_1, \cdots, \phi_L]^T$ is the vector of IRS phase shifts. The expression in (3.4) already includes MF at the Rx for SIMO channels (or at the Tx for MISO channels), which is optimal according to [56, pp. 179-180]. We emphasize that the SNR $\gamma(\boldsymbol{\phi})$ depends on $\boldsymbol{\phi}$ (to be optimized later on); $g(\boldsymbol{\phi})$ is the IRS-assisted gain for given $\boldsymbol{\phi}$; with some abuse of terminology, we call it simply "IRS gain" in the rest of this thesis.

3.2 SNR Maximization Problem

For the standard (no-IRS) SIMO channel, where the channel vector $\mathbf{h}_0$ is fixed and is out of designer's control, its induced mutual information (supported rate) is maximized via the input distribution alone and the optimal distribution is Gaussian with full available power [56]. On the other hand, the IRS-assisted channel offers additional degrees of freedom for maximizing mutual information (beyond the input

distribution) via reflector phase shifts ϕ_l. Since the equivalent IRS-assisted channel vector $\mathbf{h}_{eq}$ is independent of the input x, Gaussian input remains optimal [56]. Thus, for a given $\boldsymbol{\phi}$, the maximum achievable rate per unit bandwidth (spectral efficiency) supported by the IRS-assisted SIMO channel is

$$R(\boldsymbol{\phi}) = \log(1 + g(\boldsymbol{\phi})\gamma_0). \tag{3.5}$$

and the IRS-assisted channel capacity is

$$C_{IRS} = \max_{\boldsymbol{\phi}} R(\boldsymbol{\phi}) = \log(1 + \gamma_0 g^*), \tag{3.6}$$

$$g^* = g(\boldsymbol{\phi}^*) = \max_{\boldsymbol{\phi}} \left|\mathbf{h}_0 + \sum_{l=1}^{L} e^{j\phi_l}\mathbf{h}_l\right|^2 \tag{3.7}$$

where $g^* = g(\boldsymbol{\phi}^*) = \max_{\boldsymbol{\phi}} g(\boldsymbol{\phi})$ is the globally-optimal (maximal) IRS gain and $\boldsymbol{\phi}^*$ are the respective globally-optimal phase shifts. This SIMO channel setting can also be extended to a MISO channel via the uplink-downlink duality [56]. In particular, the column channel vectors $\mathbf{h}_0,\ \mathbf{h}_l$ are replaced by raw vectors $\mathbf{h}_0^+,\ \mathbf{h}_l^+$, and the scalar Tx signal x is replaced by a vector signal $\mathbf{x}$; the Rx signal and noise become scalars. After this transformation, (3.3)-(3.6) hold verbatim for the MISO channel as well and, hence, all our results will also apply to this channel.

The optimization problem in (3.7) is straightforward to solve for the special case of a single-antenna (SISO) channel, where h_0, h_l are scalars and the globally-optimal phase shifts are $\phi_l^* = \arg\{h_0\} - \arg\{h_l\}$ [47]. Yet, despite its apparent simplicity, no closed-form solution of this problem is known in the general SIMO/MISO case[2], i.e., for arbitrary $\mathbf{h}_0,\ \mathbf{h}_l$. No algorithm with guaranteed convergence to its global optimum is known either.

One of the key difficulties is that this problem is not convex and, hence, any stan-

[2]here and in the rest of this thesis, "solution" means *globally*-optimal solution, unless stated otherwise.

dard approach (e.g., using KKT conditions) will result in locally-optimal solutions at best, which can be far away from a globally-optimal one; the global optimality gap is not known either. To address these issues, we revisit the problem in (3.7) in the following chapters in several ways.

Chapter 4

Bounds for Globally-Optimal IRS Gain

4.1 Introduction

As previously discussed, the SNR maximization problem for an IRS-assisted SIMO channel is non-convex in general. Hence, any standard approach, including the numerical methods proposed in earlier studies [38, 39, 43, 57, 59–61], will result in locally optimal solutions at best, which can be far from a globally optimal one. In order to establish globally optimal IRS phase shifts in some cases and to assess the global optimality gap in others, this chapter establishes a number of upper and lower bounds to bound the optimality gap.

Next, we demonstrate that these bounds are tight in some cases, and for other cases, we use extensive numerical experiments to evaluate the tightness of the obtained bounds using the ratio of the upper to the lower bound. We then use these bounds to set up an initial point that can be used for any numerical method requiring initialization. Finally, we demonstrate that using this initial point sets an upper bound on the global optimality gap for any iterative method that results in a non-decreasing sequence of SNRs.

4.2 Bounds

Consider the system model defined in (3.2), where an IRS with L passive reflectors assists a MISO channel between a single-antenna Tx, and a Rx with N antennas. The channel is assumed to be fixed, as in a quasi-static channel model, where the

delay requirement is short compared to the channel coherence time [56]. To define the bounds, let

$$\mathbf{H} = [\mathbf{h}_0, \cdots, \mathbf{h}_L], \ \mathbf{w} = [1, e^{j\phi_1}, \cdots, e^{j\phi_L}]^T, \tag{4.1}$$

and $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^+$ be the singular value decomposition (SVD), where $\mathbf{U}, \mathbf{V}$ are the unitary matrices of left and right singular vectors of $\mathbf{H}$, respectively, and $\mathbf{\Sigma} = diag\{\sigma_l(\mathbf{H})\}$ is the diagonal matrix of its singular values $\sigma_l(\mathbf{H})$ sorted in decreasing order, i.e., $\sigma_1(\mathbf{H})$ is the largest one; $\mathbf{v}_l$ is the l-th column of $\mathbf{V}$ (i.e., the right singular vector corresponding to the l-th largest singular value $\sigma_l(\mathbf{H})$).

The following Proposition establishes the desired bounds using the SVD of $\mathbf{H}$, which is a time-honored tool for regular (no-IRS) MIMO channels.

Proposition 1. *The globally-optimal IRS gain $g^* = g(\boldsymbol{\phi}^*)$ for the channel in* (3.2) *is bounded as follows:*

$$g_{lb} \le g^* \le g_{ub} = \sigma_1^2(\mathbf{H})(L+1), \tag{4.2}$$

$$g_{lb} = \sigma_1^2(\mathbf{H})|\mathbf{v}_1|_1^2 + \sum_{l=2}^{r(\mathbf{H})} \sigma_l^2(\mathbf{H})|\mathbf{w}_1^+\mathbf{v}_l|^2, \tag{4.3}$$

$$w_{1l} = \exp\{j\arg(v_{1l}) - j\arg(v_{11})\}, \ l = 1...L+1, \tag{4.4}$$

where $r(\mathbf{H})$ is the rank of $\mathbf{H}$, $\mathbf{v}_1$ is the 1-st column of $\mathbf{V}$ (the right singular vector of $\mathbf{H}$ corresponding to the largest singular value) and $|\mathbf{v}_1|_1 = \sum_{l=1}^{L+1}|v_{1l}|$ is its l_1 norm; v_{1l} and w_{1l} are l-th entry of $\mathbf{v}_1$ and $\mathbf{w}_1$, respectively.

The lower bound is tight, i.e., $g_{lb} = g^*$, if $\mathbf{H}$ is rank-one, $r(\mathbf{H}) = 1$,

$$g^* = g_{lb} = \sigma_1^2(\mathbf{H})|\mathbf{v}_1|_1^2, \ \phi_l^* = \arg\{v_{1(l+1)}\} - \arg\{v_{11}\}, \tag{4.5}$$

where ϕ_l^* are globally-optimal phase shifts for this case.

If $\mathbf{v}_1$ *has equal-magnitude entries, i.e.,* $|v_{1l}| = |v_{11}|$ *for* $l = 1..L+1$*, then the upper and lower bounds coincide and are therefore tight,*

$$g^* = g_{lb} = g_{ub} = \sigma_1^2(\mathbf{H})(L+1), \tag{4.6}$$

and the globally-optimal phase shifts are as in (4.5)*.*

Proof. To prove the lower bound, we use (3.2) and (4.1) and note the following:

$$\begin{aligned}
g^\star &= \max_{\boldsymbol{\phi}} g(\boldsymbol{\phi}) && (4.7)\\
&= \max_{\boldsymbol{\phi}} |\mathbf{H}\mathbf{w}|^2,\ w_l = e^{j\phi_{l-1}},\ l = 1..L+1 && (4.8)\\
&= \max_{\boldsymbol{\phi}} \mathbf{w}^+\mathbf{H}^+\mathbf{H}\mathbf{w} && (4.9)\\
&= \max_{\boldsymbol{\phi}} \mathbf{w}^+\mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^+\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^+\mathbf{w} && (4.10)\\
&= \max_{\boldsymbol{\phi}} \mathbf{w}^+\mathbf{V}\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}\mathbf{V}^+\mathbf{w} && (4.11)\\
&= \max_{\boldsymbol{\phi}} \mathbf{q}^+\boldsymbol{\Lambda}\mathbf{q},\ \mathbf{q} = \mathbf{V}^+\mathbf{w},\ \boldsymbol{\Lambda} = \boldsymbol{\Sigma}^+\boldsymbol{\Sigma} && (4.12)\\
&= \max_{\boldsymbol{\phi}} \Big\{ \sum_{l=1}^{r(\mathbf{H})} |\mathbf{w}^+\mathbf{v}_l|^2\sigma_l^2(\mathbf{H}) \Big\} && (4.13)\\
&= \max_{\boldsymbol{\phi}} \Big\{ |\mathbf{w}^+\mathbf{v}_1|^2\sigma_1^2(\mathbf{H}) + \sum_{l=2}^{r(\mathbf{H})} |\mathbf{w}^+\mathbf{v}_l|^2\sigma_l^2(\mathbf{H}) \Big\} && (4.14)\\
&\geq |\mathbf{w}_1^+\mathbf{v}_1|^2\sigma_1^2(\mathbf{H}) + \sum_{l=2}^{r(\mathbf{H})} |\mathbf{w}_1^+\mathbf{v}_l|^2\sigma_l^2(\mathbf{H}) && (4.15)\\
&= |\mathbf{v}_1|_1^2\sigma_1^2(\mathbf{H}) + \sum_{l=2}^{r(\mathbf{H})} |\mathbf{w}_1^+\mathbf{v}_l|^2\sigma_l^2(\mathbf{H}), && (4.16)
\end{aligned}$$

where (4.10) is from the SVD $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^+$, (4.11) is due to $\mathbf{U}^+\mathbf{U} = \mathbf{I}$ and (4.13) follows from $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^+\boldsymbol{\Sigma} = diag\{\sigma_l^2(\mathbf{H})\}$; (4.15) is due to the maximization being

removed; $\mathbf{w}_1$ is defined from:

$$\max_{\boldsymbol{\phi}} |\mathbf{w}^+\mathbf{v}_1| = |\mathbf{w}_1^+\mathbf{v}_1| = |\mathbf{v}_1|_1,\ w_{1l} = e^{j\phi_l^0}, \tag{4.17}$$

where

$$\phi_l = \phi_l^0 = \arg(v_{1(l+1)}) - \arg(v_{11}), \tag{4.18}$$

i.e., $\mathbf{w}_1$ maximizes 1st term in (4.14), where (4.17) follows from:

$$\max_{\boldsymbol{\phi}} |\mathbf{w}^+\mathbf{v}_1| = \max_{\boldsymbol{\phi}} \left\{ |v_{11} + \sum_{l=1}^{L} e^{-j\phi_l} v_{1(l+1)}| \right\} \tag{4.19}$$

$$\leq |v_{11}| + \sum_{l=1}^{L} |v_{1(l+1)}| \tag{4.20}$$

$$= \sum_{l=1}^{L+1} |v_{1l}| = |\mathbf{v}_1|_1, \tag{4.21}$$

where the inequality in (4.20) is due to the triangle inequality and it holds with equality if (4.18) holds. This establishes the lower bound. Its achievability for $r(\mathbf{H}) = 1$ follows from $\sigma_l(\mathbf{H}) = 0$ for $l \geq 2$ in this case so that, from (4.14),

$$g^* = \sigma_1^2(\mathbf{H}) \max_{\boldsymbol{\phi}} |\mathbf{w}^+\mathbf{v}_1|^2 = \sigma_1^2(\mathbf{H})|\mathbf{v}_1|_1^2, \tag{4.22}$$

If $\mathbf{H}$ is consistent with wave propagation physics (e.g., Maxwell's equations), the law of energy conservation applies, and the lower bound is also consistent with wave propagation physics as well. To prove the upper bound in (4.2), note the following:

$$g(\boldsymbol{\phi}) = |\mathbf{H}\mathbf{w}|^2 \leq \sigma_1^2(\mathbf{H})|\mathbf{w}|^2 = \sigma_1^2(\mathbf{H})(L+1), \tag{4.23}$$

where the inequality is due to the singular value inequality $|\mathbf{H}\mathbf{w}| \leq \sigma_1(\mathbf{H})|\mathbf{w}|$ [88, pp. 267] and the equality is from $|\mathbf{w}|^2 = L + 1$. Now note that the right-hand side of

(4.23) is independent of $\boldsymbol{\phi}$ so that using $\max_{\boldsymbol{\phi}}$ on both sides of (4.23) results in the upper bound in (4.2). The achievability of the upper bound for channels with $\mathbf{v}_1$ having equal-magnitude entries follows by setting ϕ_l as in (4.5) so that

$$\mathbf{w} = \mathbf{w}_1 = e^{-j\arg\{v_{11}\}}\sqrt{L+1}\mathbf{v}_1, \tag{4.24}$$

and therefore

$$|\mathbf{H}\mathbf{w}|^2 = (L+1)|\mathbf{H}\mathbf{v}_1|^2 = \sigma_1^2(\mathbf{H})(L+1) = \sigma_1^2(\mathbf{H})|\mathbf{v}_1|_1^2, \tag{4.25}$$

so that the upper and lower bounds coincide and are attained, as required.

□

The lower bound is close to the globally-optimal IRS gain, $g^* \approx g_{lb}$, if $\sigma_1(\mathbf{H}) \gg \sigma_2(\mathbf{H})$ and this becomes exact equality if $r(\mathbf{H}) = 1$. Note also from (4.2) and (4.3) that the globally-optimal IRS gain g^* (or Rx SNR/power) scales at least as $\sigma_1^2(\mathbf{H})$ in the general case,

$$\sigma_1^2(\mathbf{H}) \leq \sigma_1^2(\mathbf{H})|\mathbf{v}_1|_1^2 \leq g^* \leq \sigma_1^2(\mathbf{H})(L+1), \tag{4.26}$$

where 1st inequality is due to $|\mathbf{v}_1|_1 \geq 1$. This is somewhat similar to the regular MIMO channel with channel matrix $\mathbf{H}$, where the maximum SNR gain achievable with Tx/Rx beamforming is $\sigma_1^2(\mathbf{H})$ so that the globally-optimal IRS gain is at least as large (assuming both channels have the same channel matrix $\mathbf{H}$).

The bounds in (4.2) and (4.3) can be used to evaluate the global optimality gap of known numerical algorithms (for which only *locally*-optimal convergence was established; the *globally*-optimal one is out of reach due to the lack of convexity for the respective optimization problem). Further, for algorithms that require initialization, it is desirable to initialize the optimization variables with values that are closer to

the optimal solution and intelligently chosen based on the channel. Although selecting the initial point based on the upper bound in (4.2) is preferable, this upper bound may or may not be achievable depending on the channel. Further, specific phase shifts corresponding to the upper bound are generally not known, except in special cases. Thus, the phase shifts that attain the lower bound in (4.3) provide an initial point for any algorithm requiring initialization, e.g., the iterative Algorithm 1 in Chapter 6. Using this initial point results in faster convergence to a value greater or equal to the bound itself, i.e., the bounds in (4.2) not only "sandwich" the globally-optimal IRS gain but also the gain attained by the algorithm and thus provide its *global* optimality gap. This holds for any iterative algorithm that provides a non-decreasing sequence of receive SNR values, including Algorithm 1 in Chapter 6.

Next, we present different upper bounds with explicit dependence on channel vectors $\mathbf{h}_l$ (since the SVD, while being a useful tool from computational and information-theoretic perspectives, essentially "hides" such dependence).

Proposition 2. *The globally-optimal IRS gain g^* is upper bounded in the general case as follows:*

$$g^* \leq g_{UB} \triangleq |\mathbf{h}_0|^2 + 2\sum_{l=1}^{L} |\mathbf{h}_0^+ \mathbf{h}_l| + \sum_{l,k=1}^{L} |\mathbf{h}_l^+ \mathbf{h}_k| \tag{4.27}$$

$$\leq \left(|\mathbf{h}_0| + \sum_{l=1}^{L} |\mathbf{h}_l|\right)^2 \tag{4.28}$$

where both inequalities hold with equality (i.e., the upper bounds are attained) if $\mathbf{h}_l = a_l \mathbf{h}_0$ for all l and some complex a_l, and the respective globally-optimal phase shifts are $\phi_l^* = -\arg\{a_l\}$.

Proof. Using (3.4), after some manipulations, one obtains:

$$g(\boldsymbol{\phi}) = |\mathbf{h}_0 + \sum_{l=1}^{L} e^{j\phi_l}\mathbf{h}_l|^2 = \big(\mathbf{h}_0^+ + \sum_{l=1}^{L} e^{-j\phi_l}\mathbf{h}_l^+\big)\big(\mathbf{h}_0 + \sum_{k=1}^{L} e^{j\phi_k}\mathbf{h}_k\big) \tag{4.29}$$

$$= |\mathbf{h}_0|^2 + \sum_{k} e^{j\phi_k}\mathbf{h}_0^+\mathbf{h}_k + \sum_{l} e^{-j\phi_l}\mathbf{h}_l^+\mathbf{h}_0 + \sum_{l,k} e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k \tag{4.30}$$

$$= |\mathbf{h}_0|^2 + 2\mathrm{Re}\big\{\sum_{l} e^{j\phi_l}\mathbf{h}_0^+\mathbf{h}_l\big\} + \sum_{l,k} e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k \tag{4.31}$$

$$= |\mathbf{h}_0|^2 + 2\sum_{l} \mathrm{Re}\left\{e^{j\phi_l}\mathbf{h}_0^+\mathbf{h}_l\right\} + \tfrac{1}{2}\sum_{l,k}\left(e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k + e^{j(\phi_l-\phi_k)}\mathbf{h}_k^+\mathbf{h}_l\right) \tag{4.32}$$

$$= |\mathbf{h}_0|^2 + 2\sum_{l} \mathrm{Re}\left\{e^{j\phi_l}\mathbf{h}_0^+\mathbf{h}_l\right\} + \sum_{l,k} \mathrm{Re}\left\{e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k\right\} \tag{4.33}$$

$$= |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+\mathbf{h}_l|\mathrm{Re}\left\{e^{j(\phi_l+\varphi_{0l})}\right\} + \sum_{l,k} |\mathbf{h}_l^+\mathbf{h}_k|\mathrm{Re}\left\{e^{j(\phi_k-\phi_l+\varphi_{lk})}\right\} \tag{4.34}$$

$$= |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+\mathbf{h}_l|\cos(\phi_l+\varphi_{0l}) + \sum_{l,k} |\mathbf{h}_l^+\mathbf{h}_k|\cos(\phi_k-\phi_l+\varphi_{lk}) \tag{4.35}$$

$$\leq |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+\mathbf{h}_l| + \sum_{k,l} |\mathbf{h}_l^+\mathbf{h}_k| \tag{4.36}$$

$$\leq |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0||\mathbf{h}_l| + \sum_{k,l} |\mathbf{h}_l||\mathbf{h}_k| \tag{4.37}$$

$$= \big(|\mathbf{h}_0| + \sum_{l} |\mathbf{h}_l|\big)^2, \tag{4.38}$$

where $\varphi_{lk} = \arg\{\mathbf{h}_l^+\mathbf{h}_k\}$; the upper bound in (4.36) is due to $\cos(x) \leq 1$ and (4.37) is due to Cauchy-Schwarz inequality $|\mathbf{h}_l^+\mathbf{h}_k| \leq |\mathbf{h}_l||\mathbf{h}_k|$, which holds with equality if $\mathbf{h}_l = a_l\mathbf{h}_0$ for all l. Noting that the upper bounds in (4.36), (4.37) are independent of $\boldsymbol{\phi}$ and taking $\max_{\boldsymbol{\phi}}$ on (4.29), one obtains (4.27), (4.28). With $\mathbf{h}_l = a_l\mathbf{h}_0$ for all l, (4.37) holds with equality and

$$\varphi_{lk} = \arg\{\mathbf{h}_l^+\mathbf{h}_k\} = \arg\{a_k\} - \arg\{a_l\}, \tag{4.39}$$

so that setting $\phi_l = -\arg\{a_l\} = -\varphi_{0l}$,

$$\phi_l + \varphi_{0l} = 0,\ \phi_k - \phi_l + \varphi_{lk} = 0, \tag{4.40}$$

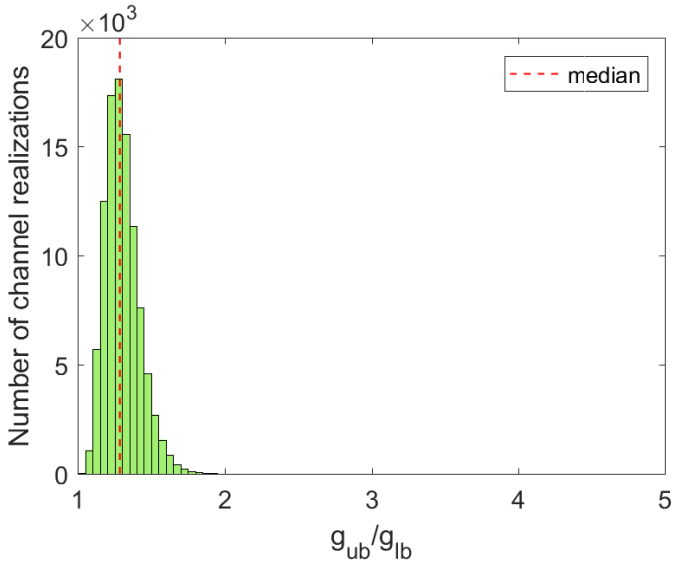
and therefore (i) (4.36) holds with equality as well, and (ii) $\phi_l = -\arg\{a_l\}$ are the globally-optimal phase shifts in this case, since they attain both upper bounds. This completes the proof. □

4.3 Numercial Experiments

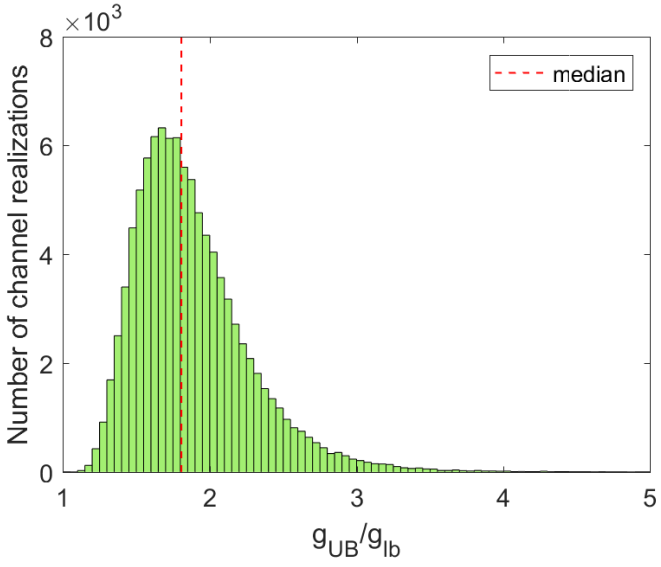
In this section, we numerically evaluate and compare the tightness of the bounds obtained in (4.2) and (4.27), particularly for cases where globally-optimal phase shifts are not known in closed-form. We provide histograms for g_{ub}/g_{lb} and g_{UB}/g_{lb} for an extensive number of channels generated under different scenarios. These numerical results have been repeated to achieve statistically stable results.

Fig. 4.1 presents the histograms of g_{ub}/g_{lb} and g_{UB}/g_{lb} both for 10^5 randomly-generated channels with $L = 10$, and $h_{l,n} \sim \mathcal{CN}(0, 1)$, all i.i.d., that is, samples from i.i.d. Rayleigh fading channel. For all channel coefficients, the average path loss is factored out into the Rx SNR definition. Under this setting, the direct Tx-Rx link and the IRS-assisted links have the same average path loss, which corresponds to a partially-blocked direct link (see e.g., [83]). From Fig. 4.1(a), g_{ub}/g_{lb} is less than 2 (i.e., 3 dB) for almost all channels, with a median value of 1.25. Meanwhile, Fig. 4.1(b) shows that g_{UB}/g_{lb} is less than 4 (i.e., 6 dB) for almost all channels, with a median value of 1.77. Based on these two figures and the median values, g_{ub}/g_{lb} offers a tighter optimality gap for this channel setup.

To examine how results are affected by a higher number of reflectors, Fig. 4.2 presents histograms for g_{ub}/g_{lb} and g_{UB}/g_{lb} using the same channel setup as in Fig. 4.1, but with $L = 100$. Fig. 4.2 shows that g_{ub}/g_{lb} is less than 2.5 (i.e., 4 dB) for



(a)



(b)

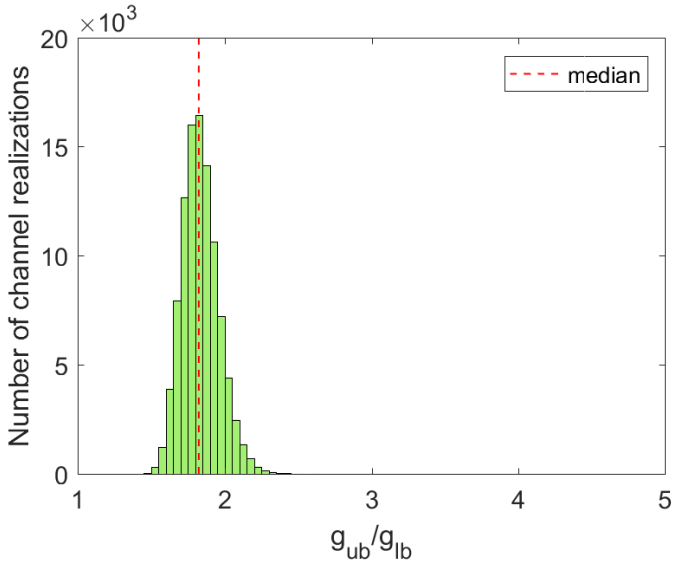
Figure 4.1: The histograms of the ratio g_{ub}/g_{lb} of the upper and lower bounds in (4.2) (a) and the ratio g_{UB}/g_{lb} of the upper and lower bounds in (4.27) (b), both for 10^5 randomly-generated channels with $h_{l,n} \sim \mathcal{CN}(0,1)$, all i.i.d., $N = 10,\ L = 10$.

almost all channels, while g_{UB}/g_{lb} is less than 3 (i.e., 4.8 dB), with median values of 1.82 and 2.19, respectively. Comparing the histograms in Fig. 4.1 and 4.2 shows that, for this specific setup, increasing the number of reflectors causes a higher median value in both cases, with a smaller increase in the median value for g_{UB}/g_{lb} compared to that for g_{ub}/g_{lb}. However, with more reflectors, g_{UB}/g_{lb} shows a narrower spread, reflecting a higher concentration around the median value. The law of large numbers can justify this, as the number of reflectors grows, g_{UB}/g_{lb} approaches the median value more closely.

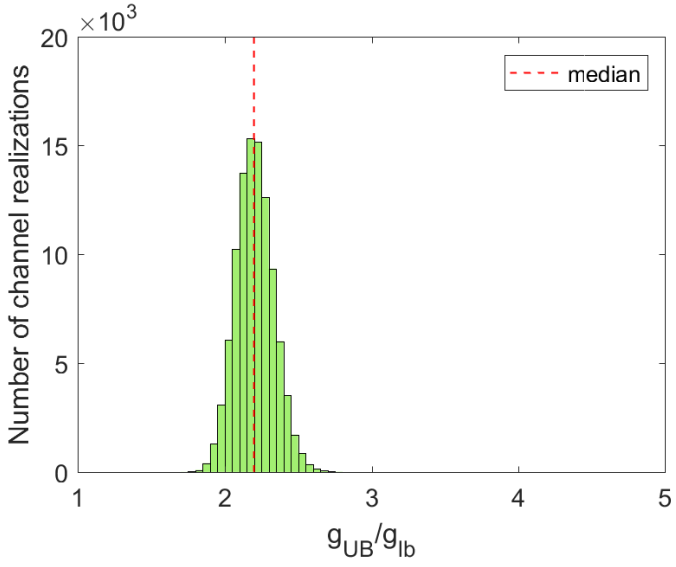
To investigate how different channel distribution parameters can affect the previous results, Fig. 4.3 illustrates histograms for g_{ub}/g_{lb} and g_{UB}/g_{lb} using the same channel setup as in Fig. 4.2, but with $h_{l,n} \sim \mathcal{CN}(0, 10)$. A comparison between Fig. 4.1 and Fig. 4.3 shows no clear differences, indicating that both g_{ub}/g_{lb} and g_{UB}/g_{lb} are very similar in both channels. This indicates that for this specific channel, i.e., samples from an i.i.d. Rayleigh fading channel, the increase in severity of fading across all channel coefficients does not significantly affect the optimality gap obtained from both bounds.

Further, Fig. 4.4 examines scenarios characterized by a strong Tx-Rx link (dominant LoS), where $|\mathbf{h}_0| \gg |\mathbf{h}_l|$. These cases were modeled using 10^5 randomly-generated channels, with $h_{0,n} = 10$ and $h_{l,n} \sim \mathcal{CN}(0, 1)$, all being i.i.d. As shown in Fig. 4.4, g_{ub}/g_{lb} is less than 3 (i.e., 4.8 dB) for almost all channels, and it is less than 2 (i.e., 3 dB) for g_{UB}/g_{lb}, with median values of 2.34 and 1.36, respectively. Based on the median values, g_{UB}/g_{lb} achieves a tighter optimality gap in cases where dominant LoS is present. This result differs from that obtained from Fig. 4.1, which was obtained for channels with partially-blocked direct links.

According to the previous numerical results, e.g., Fig. 4.2 and 4.4, it is concluded that neither the bounds in (4.2) nor (4.27) can universally provide a tighter optimality gap; thus, they are complementary.

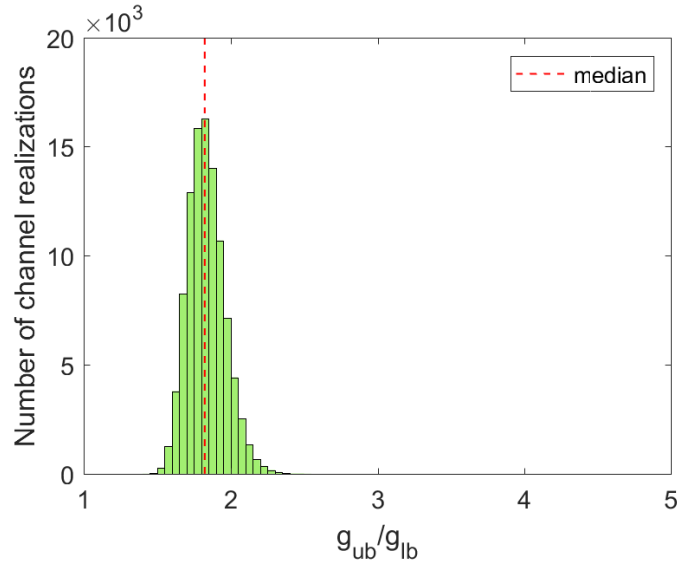


(a)

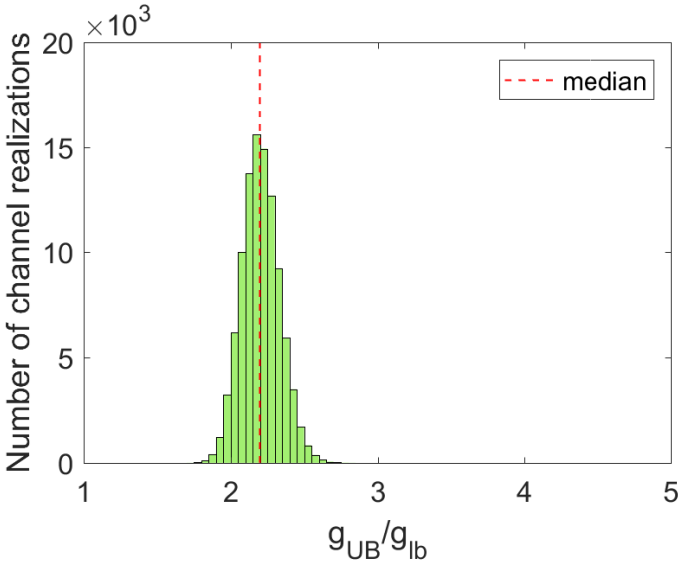


(b)

Figure 4.2: The histograms of the ratio g_{ub}/g_{lb} of the upper and lower bounds in (4.2) (a) and the ratio g_{UB}/g_{lb} of the upper and lower bounds in (4.27) (b), both for 10^5 randomly-generated channels with $h_{l,n} \sim \mathcal{CN}(0,1)$, all i.i.d., $N = 10,\ L = 100$.

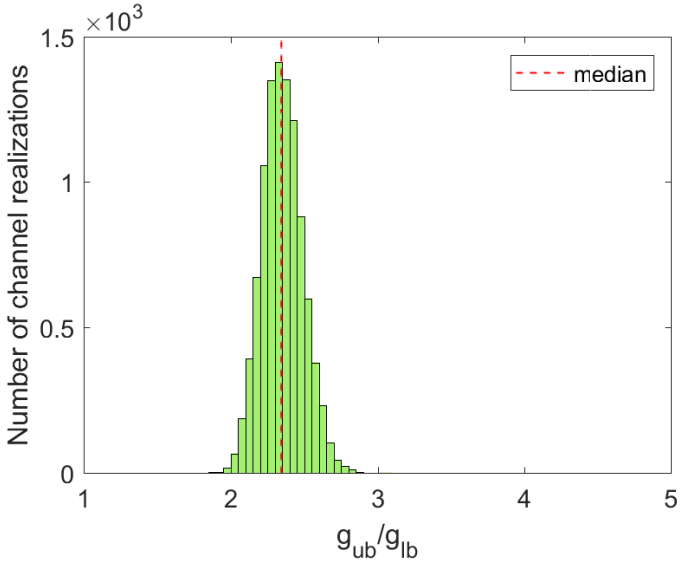


(a)

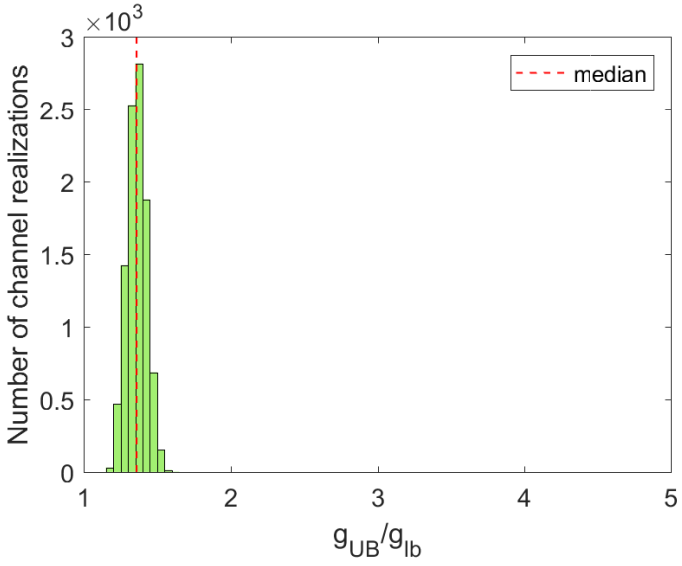


(b)

Figure 4.3: The histograms of the ratio g_{ub}/g_{lb} of the upper and lower bounds in (4.2) (a) and the ratio g_{UB}/g_{lb} of the upper and lower bounds in (4.27) (b), both for 10^5 randomly-generated channels with $h_{l,n} \sim \mathcal{CN}(0, 10)$, all i.i.d., $N = 10,\ L = 100$.



(a)

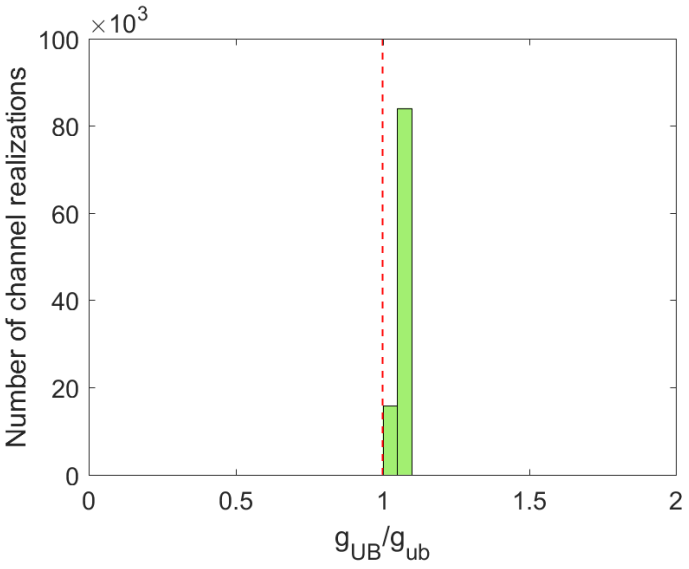


(b)

Figure 4.4: The histograms compare the ratio g_{ub}/g_{lb} of the upper and lower bounds in (4.2) (a) with the ratio g_{UB}/g_{lb} of the upper and lower bounds in (4.27) (b), for 10^5 randomly-generated channels with $h_{0,n} = 10$ and $h_{l,n} \sim \mathcal{CN}(0, 1)$, all i.i.d., and $N = 10,\ L = 100$.

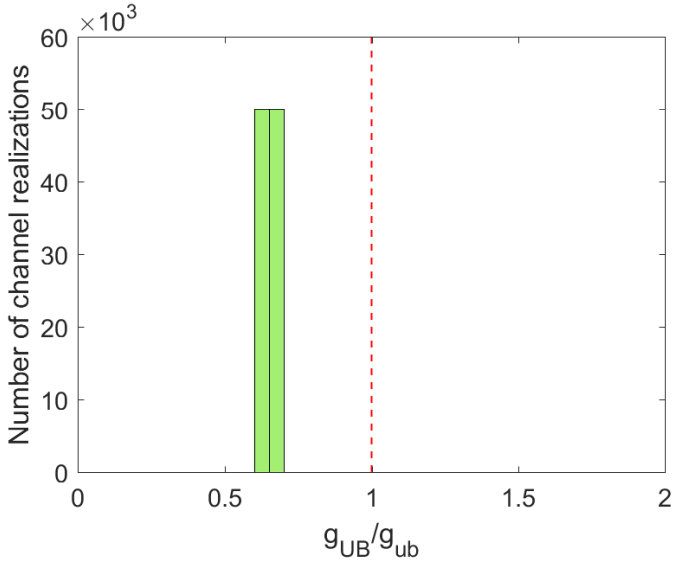
To demonstrate this result, Figs. 4.5 and 4.6 show histograms of the ratio g_{UB}/g_{ub} for the settings as in Figs. 4.2 and 4.4, respectively. As can be seen, the ratio g_{UB}/g_{ub} can take values both less than and greater than one, which shows that neither upper bound is universally tighter.

Overall, the numerical experiments in this section verify that the obtained bounds offer a reasonably accurate estimate of the globally-optimal IRS gain, for which no closed-form solution is generally known. In the next chapter, we present closed-form solutions for the problem in (3.6) and identify special cases where the upper bounds in (4.2) and (4.27) are achievable, i.e., where $g^* = g_{ub}$ or $g^* = g_{UB}$.



(a)

Figure 4.5: The histograms of the ratio g_{UB}/g_{ub} for the same settings as in Fig. 4.2. g_{ub} is better than g_{UB} when LoS is partially-blocked.



(a)

Figure 4.6: The histograms of the ratio g_{UB}/g_{ub} for the same settings as in Fig. 4.4. g_{UB} is better than g_{ub} when there is a strong LoS.

Chapter 5

Closed-Form Globally-Optimal Solutions

5.1 Introduction

As discussed in the previous chapters, a globally-optimal closed-form solution to (3.7) is not known in the general case (to the best of our knowledge, the only known closed-form solution is for scalar channels, $N = 1$). However, the bounds obtained in the previous chapter enable us to identify several cases where such solutions can be identified. In this section, we present these practically important cases, including mmWave/THz channels, single-reflector IRS, massive MIMO settings, and multi-IRS channels, and we consider variable-gain reflectors, as well.

For each case, we obtain the globally-optimal phase shifts and provide rigorous proof for their optimality. Based on the obtained results, we discuss how IRS gain can scale depending on the setup parameters, e.g., the number of reflectors (L) or the number of IRSs. We revisit the power scaling law for the IRS in these cases and identify the conditions that affect the scaling rate.

We further explore a special case of the single-reflector IRS with a closed-form globally-optimal solution. Although this case may not be practical, it serves as a foundational building block that enables us to build an iterative algorithm in Chapter 6 for the general case.

5.2 mmWave/THz channel

For mmWave/THz channels, one can argue that due to the high directivity of antennas and high diffraction/scattering loss, the absence of LoS results in a significant increase in path loss, leading to a substantial decrease in SNR at the receiver. So, the LoS becomes essential for reliable communication, including the direct Tx-Rx link and reflected links from the IRS. Measurements indicate that metamaterial-based IRS reflectors achieve relatively low reflection losses, as low as 1 dB, compared to losses greater than 10 dB observed for other surfaces [89, 90]. Therefore, we consider in this section an IRS-assisted mmWave/THz SIMO channel with LoS component, where the IRS and the Rx are both equipped with uniform linear arrays (ULAs). Following [91] [92], the respective channel vectors can be expressed as

$$\mathbf{h}_l = h \cdot \exp\{j(\varphi_l(\theta^I_{IRS}) + \phi_l)\}\mathbf{a}(\theta_{Rx}),\ l = 1..L \tag{5.1}$$

$$\varphi_n(\theta) = 2\pi d(n-1)\lambda^{-1}\sin(\theta),\ n = 1..N \tag{5.2}$$

$$a_n(\theta) = \exp\{j\varphi_n(\theta)\}, \tag{5.3}$$

where h represents large-scale path loss for the IRS links while $\mathbf{a}(\theta_{Rx})$ is the local array response vector at the Rx and $a_l(\theta_{Rx})$ is its l-th entry, θ_{Rx} is the angle of arrival (AoA) at the Rx (measured with respect to the array broadside), θ^I_{IRS} and ϕ_l are the incidence angle and the phase shift at l-th IRS reflector, d is ULA element spacing, λ is the wavelength. In this setup, IRS is assumed to be located in the far-field of both Tx and Rx, so that incident and reflected waves appear as planar wavefronts. Fig. 5.1 illustrates this channel structure. Using (5.1)-(5.3),

$$\mathbf{h}_l = e^{j\alpha_l}\mathbf{h}_1,\ \alpha_l = \varphi_l(\theta^I_{IRS}) + \phi_l. \tag{5.4}$$

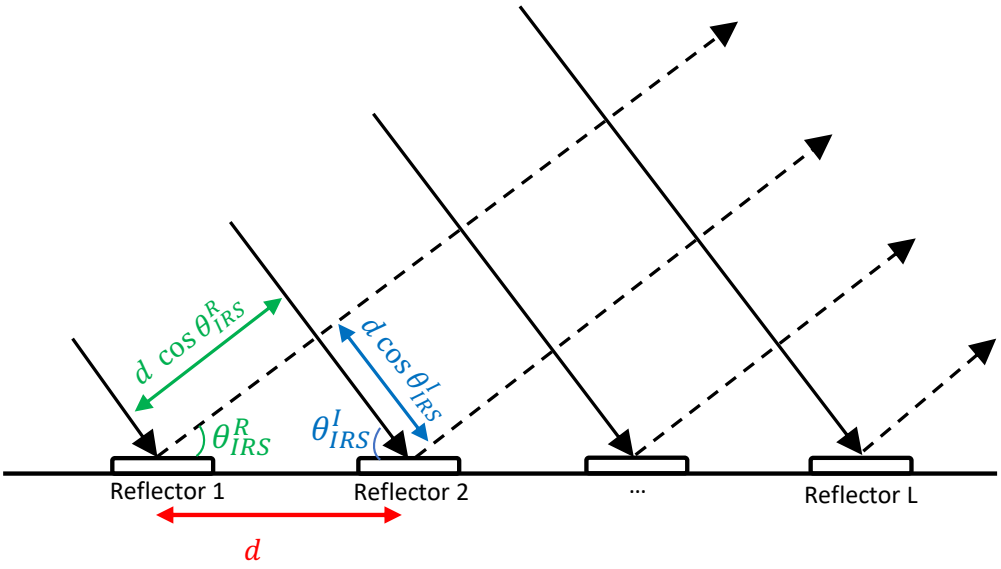


Figure 5.1: Illustration of the channel geometry for an IRS-assisted channel.

The next Proposition presents a closed-form globally-optimal solution to the problem in (3.7) for this channel.

Proposition 3. *Let the IRS-assisted channel satisfy* $\mathbf{h}_l = e^{j\alpha_l}\mathbf{h}_1,\ l = 1\cdots L$, *for some* α_l *and* $\alpha_1 = 0$, *and let* $\mathbf{h}_0$ *be arbitrary. Then, the globally-optimal IRS gain and the respective phase shifts are as follows:*

$$g^* = |\mathbf{h}_0|^2 + L^2|\mathbf{h}_1|^2 + 2L|\mathbf{h}_0^+\mathbf{h}_1| \tag{5.5}$$

$$\phi_l^* = \arg\{\mathbf{h}_1^+\mathbf{h}_0\} - \alpha_l = \arg\{\mathbf{h}_l^+\mathbf{h}_0\}, \tag{5.6}$$

where 2nd equality in (5.6) *holds if* $\mathbf{h}_1^+\mathbf{h}_0 \neq 0$. *If* $\mathbf{h}_1^+\mathbf{h}_0 = 0$, *then* $\phi_l^* = \alpha_0 - \alpha_l$, *where* α_0 *is arbitrary, i.e., the optimal phase shifts are not unique.*

Proof. From (4.35) and (4.36),

$$\begin{aligned} g(\boldsymbol{\phi}) &= |\mathbf{h}_0|^2 + 2\sum_l |\mathbf{h}_0^+\mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum_{l,k} |\mathbf{h}_l^+\mathbf{h}_k| \cos(\phi_k - \phi_l + \varphi_{lk}) \\ &\leq |\mathbf{h}_0|^2 + 2\sum_l |\mathbf{h}_0^+\mathbf{h}_l| + \sum_{k,l} |\mathbf{h}_l^+\mathbf{h}_k| \\ &= |\mathbf{h}_0|^2 + 2L|\mathbf{h}_0^+\mathbf{h}_1| + L^2|\mathbf{h}_1|^2 \end{aligned} \tag{5.7}$$

where the last equality is due to $\mathbf{h}_l = e^{j\alpha_l}\mathbf{h}_1$. Note that the upper bound is independent of $\boldsymbol{\phi}$ and it holds for any $\boldsymbol{\phi}$, including optimal one, and therefore applies to $g(\boldsymbol{\phi}^*) = g^*$ as well. It remains to show that the upper bound is attained in this case. For this particular channel structure $\mathbf{h}_l = e^{j\alpha_l}\mathbf{h}_1$ and if $\mathbf{h}_0^+\mathbf{h}_1 \neq 0$, note the following

$$\begin{aligned} &\varphi_{lk} = \arg\{\mathbf{h}_l^+\mathbf{h}_k\} = \alpha_k - \alpha_l,\ l,k = 1..L \\ &\varphi_{0l} = \arg\{\mathbf{h}_0^+\mathbf{h}_l\} = \alpha_l + \varphi_{01},\ \varphi_{01} = \arg\{\mathbf{h}_0^+\mathbf{h}_1\}, \end{aligned} \tag{5.8}$$

so that $\varphi_{0k} - \varphi_{0l} = \alpha_k - \alpha_l = \varphi_{lk}$ and, therefore, setting

$$\phi_l = -\varphi_{0l} = -\alpha_l - \varphi_{01}, \tag{5.9}$$

one obtains $\phi_l + \varphi_{0l} = 0,\ \phi_k - \phi_l + \varphi_{lk} = 0$ so that (i) the upper bound in (5.7) is attained, and (ii) the phase shifts in (5.9) are globally-optimal. If $\mathbf{h}_0^+\mathbf{h}_1 = 0$, then φ_{01} is arbitrary and therefore $\phi_l = \alpha_0 - \alpha_l$ are globally-optimal (achieve the upper bound in (5.7)) for arbitrary α_0. □

Few observations are in order based on (5.5) and (5.6). First, note that, under the condition $\mathbf{h}_l = e^{j\alpha_l}\mathbf{h}_1$ of this Proposition, $g^* = g_{UB}$ so that the upper bound in (4.27) is tight (holds with equality) in this case. Next, it follows from (5.5) that if the direct link is much weaker than the combined reflected links, i.e., if $|\mathbf{h}_0| \ll L|\mathbf{h}_1|$,

then the IRS gain is approximately

$$g^* \approx L^2 |\mathbf{h}_1|^2 \sim L^2, \tag{5.10}$$

i.e., scales as L^2 ("20 dB per decade") [1]. In the opposite regime of moderate to strong direct link,

$$g^* \approx |\mathbf{h}_0|^2 + 2L|\mathbf{h}_0^+ \mathbf{h}_1|, \tag{5.11}$$

i.e., scales as L at best ("10 dB per decade") and this scaling is achieved if $\mathbf{h}_0^+ \mathbf{h}_1 \neq 0$ and the number of reflectors is large enough,

$$L \gg \frac{|\mathbf{h}_0|^2}{2|\mathbf{h}_0^+ \mathbf{h}_1|} \;\rightarrow\; g^* \approx 2L|\mathbf{h}_0^+ \mathbf{h}_1| \sim L, \tag{5.12}$$

If this is not the case, then $g^* \approx |\mathbf{h}_0|^2$ and IRS provides negligible improvement. We further note that the above scalings in (5.5), (5.10)-(5.12) are consistent with the measurements in [31, Fig. 12(b)] [94, Fig. 18], where quadratic scaling was experimentally verified in a certain environment for large IRS. Since no specific assumptions on $\mathbf{h}_0$, $\mathbf{h}_1$ have been made in the above analysis, the difference in the propagation path loss of the direct and reflected links, as discussed in [40, 41, 83, 93], can also be included when the scalings of g^* with L are analysed as in (5.5)-(5.12). In particular, if $|\mathbf{h}_0|$ is not small enough (i.e., the direct link is not weak), the scaling of g^* with L is linear at best, which is consistent with the physically-based arguments in [40, 41] as well as with the measurements in [31, 94, 95].

[1] Note that this scaling does not hold for the asymptotic analysis with $L \to \infty$, since the condition $\mathbf{h}_l = e^{j\alpha_l}\mathbf{h}_1$ only holds in the far-field of the IRS, where this scaling is applicable. As L increases, so does the size of the IRS, and eventually the Tx and Rx will be in the near-field of the IRS so that this condition, and also the scaling, becomes invalid; additionally, due to the law of energy conservation, the total received power cannot exceed the transmit one so that this scaling will have to saturate, see [40, 41, 93] for more details. This remark also applies to all other scalings in this thesis.

To get further insights, let us consider the case when the direct and reflected links are resolvable, i.e., $\mathbf{h}_0^+\mathbf{h}_1 = 0$. In this case, (5.5) simplifies to

$$g^* = |\mathbf{h}_0|^2 + L^2|\mathbf{h}_1|^2, \tag{5.13}$$

and $\phi_l^* = \alpha_0 - \alpha_l$, where α_0 is arbitrary, i.e., the optimal phase shifts are *not unique*. IRS provides significant improvement over the direct link alone when the 2nd term of g^* exceeds the 1st one or, equivalently, if the number L of reflectors is large enough,

$$L \gg |\mathbf{h}_0||\mathbf{h}_1|^{-1} = L_{min}, \tag{5.14}$$

and, under this condition, $g^* \approx L^2|\mathbf{h}_1|^2$, i.e., the scaling with L is quadratic. If $L < L_{min}$, then IRS provides negligible improvement. The threshold L_{min} can be evaluated using the physically-based path loss models in [29, 40, 41, 83, 93], where $|\mathbf{h}_0|^{-2}$ and $|\mathbf{h}_1|^{-2}$ represent the path losses of the direct and reflected links. For example, if the path losses of Tx-Rx, Tx-IRS and IRS-Rx links are all 60 dB, then $|\mathbf{h}_0|^{-2} = 10^{-6}$, $|\mathbf{h}_1|^{-2} = 10^{-12}$ (since the combined Tx-IRS-Rx path loss is 120 dB) so that $L_{min} = 10^3$.

One can further consider the opposite case of non-resolvable direct and reflected paths, i.e., $\mathbf{h}_1 = a_1\mathbf{h}_0$ for some complex a_1. In this case, (5.5) and (5.6) simplify to

$$g^* = |\mathbf{h}_0|^2(1 + L|a_1|)^2, \ \phi_l^* = -\alpha_l - \arg\{a_1\}, \tag{5.15}$$

Note that the optimal phase shifts here are *unique* and IRS provides significant improvement if $L|a_1| \gg 1$, i.e., the same condition as in (5.14) (since $|a_1| = |\mathbf{h}_1|/|\mathbf{h}_0|$), and, in this case, $g^* \approx L^2|a_1|^2|\mathbf{h}_0|^2 = L^2|\mathbf{h}_1|^2$, i.e the same as for the resolvable paths above.

5.3 Single reflector IRS

Next, we note that the condition of Proposition 3 is always satisfied in the single-reflector case, $L = 1$, so that globally-optimal IRS gain and the respective phase shifts follow.

Corollary 1. *The globally-optimal IRS gain and phase shift for the single-reflector IRS ($L = 1$) with arbitrary $\mathbf{h}_0,\ \mathbf{h}_1$ are as follows:*

$$g^* = |\mathbf{h}_0|^2 + |\mathbf{h}_1|^2 + 2|\mathbf{h}_0^+\mathbf{h}_1|,\ \phi_1^* = \arg\{\mathbf{h}_1^+\mathbf{h}_0\}, \tag{5.16}$$

If $\mathbf{h}_0^+\mathbf{h}_1 = 0$, then ϕ_1^* is arbitrary.

While IRS offers significant advantages when equipped with a large number of reflectors (100s or 1000s), as demonstrated in (5.14) and [29,83], the closed-form solution obtained for a single-reflector IRS will be instrumental in developing an iterative optimization algorithm for the general case (any L). This algorithm demonstrates good convergence properties. More details about this proposed algorithm are provided in Chapter 6.

5.4 Massive MIMO setting

When the number N of Rx antennas is large, as in massive MIMO, and the condition known as "favorable propagation" holds, individual channel paths become resolvable and the respective channel vectors become orthogonal to each other, $\mathbf{h}_l^+\mathbf{h}_k = 0,\ l \neq k$ [54] [96] [11]. The next Proposition presents a closed-form globally-optimal solution of (3.7) in this case.

Proposition 4. *Let the channel vectors of reflected paths be mutually-orthogonal, $\mathbf{h}_l^+\mathbf{h}_k = 0,\ l \neq k$, where $l, k = 1...L$ (no such assumption is made for the LoS path).*

Then, the globally optimal phase shifts and the respective IRS gain are as follows:

$$\phi_l^* = \arg\{\mathbf{h}_l^+ \mathbf{h}_0\},\ g^* = |\mathbf{h}_0|^2 + 2\sum_{l=1}^{L} |\mathbf{h}_l^+ \mathbf{h}_0| + \sum_{l=1}^{L} |\mathbf{h}_l|^2. \tag{5.17}$$

If $\mathbf{h}_l^+ \mathbf{h}_0 = 0$*, then* ϕ_l^* *is arbitrary.*

Proof. In this case, using (4.35),

$$\begin{aligned} g(\boldsymbol{\phi}) &= |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum_{l,k} |\mathbf{h}_l^+ \mathbf{h}_k| \cos(\phi_k - \phi_l + \varphi_{lk}) \\ &= |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum_{l} |\mathbf{h}_l|^2 \qquad (5.18) \\ &\leq |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+ \mathbf{h}_l| + \sum_{l} |\mathbf{h}_l|^2, \end{aligned}$$

where (5.18) is due to $\mathbf{h}_l^+ \mathbf{h}_k = 0$ for $l \neq k$. Note that the upper bound holds for any $\boldsymbol{\phi}$, including the optimal one, is independent of $\boldsymbol{\phi}$ and is attained by $\phi_l = -\varphi_{0l} = \arg\{\mathbf{h}_l^+ \mathbf{h}_0\}$, which are therefore globally-optimal. If $\mathbf{h}_l^+ \mathbf{h}_0 = 0$, then ϕ_l^* is arbitrary since $\arg\{0\}$ is arbitrary (and yet achieves the upper bound in (5.18)). □

Note that, in this case, the upper bound in (4.27) is also tight, $g^* = g_{UB}$. Comparing (5.17) to (5.6), we note that $\phi_l^* = \arg\{\mathbf{h}_l^+ \mathbf{h}_0\}$ is globally-optimal in the cases where $\mathbf{h}_l$ are either orthogonal or parallel to each other. This feature can be used to obtain a globally-optimal solution for a multi-IRS channel.

One may argue that the mutual orthogonality condition in Proposition 4 may not hold in practice as the spacing between adjacent IRS reflectors may not suffice for the Rx array to resolve the corresponding reflected links. To overcome this, once the spacing between adjacent reflectors is insufficient for resolving the corresponding reflected links, the entire set of reflectors can be divided into a number of groups; each group consists of a number of adjacent reflectors and is treated as a single reflector (i.e., all reflectors in a group are assigned with the same phase shift). This tech-

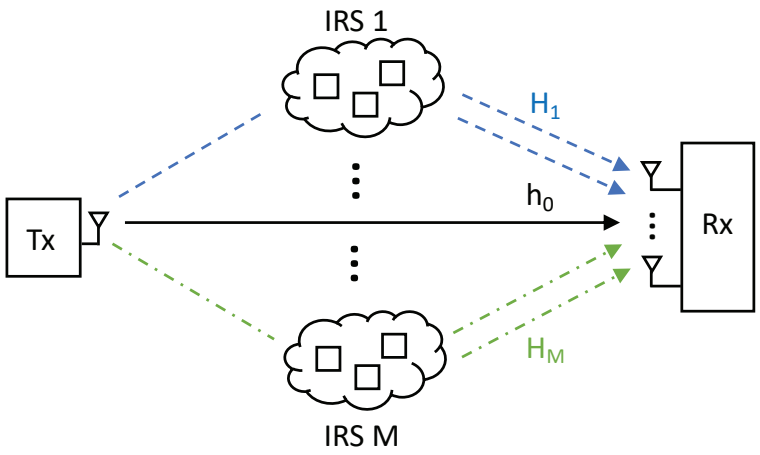


Figure 5.2: An illustration of multi-IRS SIMO channel with M IRSs.

nique, known as element grouping, has been previously discussed in the literature (e.g., [77, 97]). This method lowers the complexity of channel estimation and IRS control, which becomes more evident when the IRS is equipped with many reflectors. Applying the element grouping method increases the spacing between different groups of IRS reflectors, thus achieving favourable propagation conditions.

5.5 Multi-IRS channel

A multi-IRS channel is often considered in the literature as an inexpensive way to enhance system performance. In this section, we consider a scenario where several IRSs are spatially distributed, as in Fig. 5.2. When the number N of Rx antennas is large and, therefore, the Rx antenna array angular resolution is high [98], reflectors from different IRSs can be resolved at the Rx (since their spacing is large) so that their channel vectors are orthogonal to each other. On the other hand, reflectors from the same IRS are not resolvable since their spacing is not large enough so that their Rx angles of arrival are almost the same.

Based on this consideration, the overall multi-IRS channel matrix $\mathbf{H}$ can be block-partitioned as $\mathbf{H} = [\mathbf{h}_0\ \mathbf{H}_1\ ...\ \mathbf{H}_M]$, where each block represents the respective IRS

(collects channel vectors from that IRS):

$$\mathbf{H}_1 = [\mathbf{h}_1,\ ...,\ \mathbf{h}_{L_1}] \tag{5.19}$$

$$\mathbf{H}_m = [\mathbf{h}_{L_1+...+L_{m-1}+1},\ ...,\ \mathbf{h}_{L_1+...+L_m}], \tag{5.20}$$

where $\mathbf{H}_m$ represents m-th IRS, L_m is its number of reflectors, $m = 1...M$, and M is the number of IRSs. The total number of reflectors (in all IRSs) is $L = \sum_{m=1}^{M} L_m$. For further use, let us define the index set $\mathcal{I}_m$ of columns in $\mathbf{H}_m$:

$$\mathcal{I}_m = \{l : L_1 + ... + L_{m-1} + 1 \leq l \leq L_1 + ... + L_m\}, \tag{5.21}$$

Since reflectors from different (distant) IRSs are resolvable, their channel vectors are orthogonal to each other, so that $\mathbf{h}_l^+ \mathbf{h}_k = 0$ for $l \in \mathcal{I}_{m_1}$, $k \in \mathcal{I}_{m_2}$, $m_1 \neq m_2$, and also $\mathbf{H}_{m_1}^+ \mathbf{H}_{m_2} = \mathbf{0}$. On the other hand, since reflectors of the same IRS are not resolvable due to their proximity to each other, $r(\mathbf{H}_m) = 1$ and therefore

$$\mathbf{h}_l = \alpha_l \mathbf{u}_m \text{ for some } \alpha_l \ \forall l \in \mathcal{I}_m,\ m = 1...M, \tag{5.22}$$

where $\mathbf{u}_m$ is the unit basis vector of span$\{\mathbf{H}_m\}$; without loss of generality, we further assume that $\alpha_l \neq 0$ for all l. Since different blocks are orthogonal to each other, it follows that $\{\mathbf{u}_1 ... \mathbf{u}_m\}$ is an orthonormal set.

For this multi-IRS channel, the following Proposition provides an explicit globally-optimal solution of (3.7).

Proposition 5. *The globally-optimal phase shifts and the respective IRS gain of the*

multi-IRS channel described above are as follows:

$$\phi_l^* = \arg\{\mathbf{u}_m^+ \mathbf{h}_0\} - \arg\{\alpha_l\} = \arg\{\mathbf{h}_l^+ \mathbf{h}_0\} \ \forall l \in \mathcal{I}_m, \ m = 1...M \tag{5.23}$$

$$g^* = g_{UB} = |\mathbf{h}_0|^2 + 2\sum_{m=1}^{M} A_m |\mathbf{h}_0^+ \mathbf{u}_m| + \sum_{m=1}^{M} A_m^2, \tag{5.24}$$

where 2nd equality in (5.23) *holds if* $\mathbf{u}_m^+ \mathbf{h}_0 \neq 0$*;* A_m *is the combined amplitude gain of* m*-th IRS,*

$$A_m = \sum_{l \in \mathcal{I}_m} |\alpha_l| = \sum_{l \in \mathcal{I}_m} |\mathbf{h}_l|. \tag{5.25}$$

If $\mathbf{u}_m^+ \mathbf{h}_0 = 0$*, then* $\phi_l^* = \psi_m - \arg\{\alpha_l\}$ *for all* $l \in \mathcal{I}_m$ *and arbitrary* ψ_m*.*

Proof. Observe that the upper bound g_{UB} in (4.27) also applies to the multi-IRS case here, where L is the total number of reflectors. It remains to show that this upper bound is achievable by the phase shifts in (5.23). To this end, using (4.35), we obtain:

$$g(\boldsymbol{\phi}) = |\mathbf{h}_0|^2 + 2\sum_{l=1}^{L} |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum_{l,k=1}^{L} |\mathbf{h}_l^+ \mathbf{h}_k| \cos(\phi_k - \phi_l + \varphi_{lk}) \tag{5.26}$$

$$= |\mathbf{h}_0|^2 + 2\sum_{m=1}^{M} \sum_{l \in \mathcal{I}_m} |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum_{m=1}^{M} \sum_{l,k \in \mathcal{I}_m} |\mathbf{h}_l^+ \mathbf{h}_k| \cos(\phi_k - \phi_l + \varphi_{lk}) \tag{5.27}$$

$$\leq |\mathbf{h}_0|^2 + 2\sum_{m=1}^{M} \sum_{l \in \mathcal{I}_m} |\mathbf{h}_0^+ \mathbf{h}_l| + \sum_{m=1}^{M} \sum_{l,k \in \mathcal{I}_m} |\mathbf{h}_l^+ \mathbf{h}_k| = g_{UB} \tag{5.28}$$

$$= |\mathbf{h}_0|^2 + 2\sum_{m=1}^{M} \sum_{l \in \mathcal{I}_m} |\alpha_l| |\mathbf{h}_0^+ \mathbf{u}_m| + \sum_{m=1}^{M} \sum_{l,k \in \mathcal{I}_m} |\alpha_l| |\alpha_k| \tag{5.29}$$

$$= |\mathbf{h}_0|^2 + 2\sum_{m=1}^{M} A_m |\mathbf{h}_0^+ \mathbf{u}_m| + \sum_{m=1}^{M} A_m^2, \tag{5.30}$$

where (5.27) is due to $\mathbf{h}_l^+ \mathbf{h}_k = 0$ for $l \in \mathcal{I}_{m_1}$, $k \in \mathcal{I}_{m_2}$, $m_1 \neq m_2$; (5.29) is due to

$\mathbf{h}_l = \alpha_l \mathbf{u}_m$ for $l \in \mathcal{I}_m$ and (5.30) is due to (5.25). Next, assume that $\mathbf{h}_0^+ \mathbf{u}_m \neq 0$ and observe that, for this channel,

$$\begin{aligned} \varphi_{0l} &= \arg(\mathbf{h}_0^+ \mathbf{h}_l) = \arg(\alpha_l) + \arg(\mathbf{h}_0^+ \mathbf{u}_m) \ \forall l \in \mathcal{I}_m \\ \varphi_{lk} &= \arg(\mathbf{h}_l^+ \mathbf{h}_k) = \arg(\alpha_k) - \arg(\alpha_l) = \varphi_{0k} - \varphi_{0l} \ \forall k, l \in \mathcal{I}_m \end{aligned} \tag{5.31}$$

so that setting $\boldsymbol{\phi}$ as in (5.23), $\phi_l = \arg\{\mathbf{h}_l^+ \mathbf{h}_0\} = -\varphi_{0l}$, it follows that

$$\phi_l + \varphi_{0l} = 0, \ \ \phi_k - \phi_l + \varphi_{lk} = 0, \tag{5.32}$$

so that the inequality in (5.28) holds with equality and hence (i) the upper bound is attained, and (ii) this choice of ϕ_l is globally optimal. If $\mathbf{h}_0^+ \mathbf{u}_m = 0$, then $\phi_l = \psi_m - \arg(\alpha_l)$ attains the upper bound in (5.28) with arbitrary ψ_m and is therefore globally-optimal. □

It should be noted that the globally-optimal phase shifts in (5.23) are of the same form as in (5.6) and (5.17), and, for this multi-IRS channel, the upper bound in (4.27) is also tight, $g^* = g_{UB}$. (5.25) represents an equal-gain combiner (EGC) and A_m is its amplitude gain.

To get some insight, let us consider the case of absent (blocked) direct link, $\mathbf{h}_0 = 0$, for which (5.24) reduces to

$$g^* = \sum_{m=1}^{M} A_m^2 = \sum_{m=1}^{M} \big(\sum_{l \in \mathcal{I}_m} |\mathbf{h}_l| \big)^2. \tag{5.33}$$

Note that the internal summation represents amplitude-wise combining (or EGC) for each IRS across its reflectors while the external sum is a power-wise combining across different IRSs. This difference is due to the fact that the reflectors within the same IRS are not resolvable (so that their channel vectors are parallel to each other) while the reflectors of different IRSs are resolvable (so that their channel vectors are

orthogonal to each other).

To reveal the scaling of the globally-optimal IRS gain g^* with the number M of IRSs and the numbers L_m of reflectors in each IRSs, let us consider the case when all IRSs are identical and have similar channels to their reflectors so that $L_m = L_1$, $|\mathbf{u}_m^+\mathbf{h}_0| = |\mathbf{u}_1^+\mathbf{h}_0|$ for all m, and $|\mathbf{h}_l| = |\mathbf{h}_1|$ for all l. In this case, (5.24) reduces to

$$g^* = |\mathbf{h}_0|^2 + 2|\mathbf{h}_1^+\mathbf{h}_0|ML_1 + |\mathbf{h}_1|^2ML_1^2. \tag{5.34}$$

In the case of weak or absent direct link, the last term dominates so that

$$g^* \approx |\mathbf{h}_1|^2ML_1^2 \;\text{ if }\; \frac{|\mathbf{h}_0|^2}{ML_1} + 2|\mathbf{h}_1^+\mathbf{h}_0| \ll |\mathbf{h}_1|^2L_1. \tag{5.35}$$

i.e., g^* scales quadratically with L_1 but only linearly with M and this scaling holds provided the number L_1 of reflectors per IRS is large enough. If the opposite is true,

$$g^* \approx |\mathbf{h}_0|^2 + 2|\mathbf{h}_1^+\mathbf{h}_0|ML_1 \;\text{ if }\; \frac{|\mathbf{h}_0|^2}{ML_1} + 2|\mathbf{h}_1^+\mathbf{h}_0| \gg |\mathbf{h}_1|^2L_1. \tag{5.36}$$

i.e., the scaling with M, L_1 is linear at best, and this holds provided the direct link is not too strong,

$$g^* \approx 2|\mathbf{h}_1^+\mathbf{h}_0|ML_1 \;\text{ if }\; |\mathbf{h}_0|^2 \ll 2|\mathbf{h}_1^+\mathbf{h}_0|ML_1. \tag{5.37}$$

If the opposite is true, then $g^* \approx |\mathbf{h}_0|^2$, i.e., IRSs have negligible impact. Thus, of all three cases considered, the first one is most favorable in terms of IRSs impact, i.e., the stronger the direct link, the smaller the impact of the IRSs. For IRSs to be effective, either (5.35) or (5.37) has to hold, which can be used as design guidelines as to (i) how many reflectors per IRS or (ii) how many IRSs are needed to make a significant impact.

Finally, one can consider the case when the direct and reflected paths are resolv-

able so that their channel vectors are orthogonal to each other, $\mathbf{h}_1^+\mathbf{h}_0 = 0$. In this case, (5.34) reduces to

$$g^* = |\mathbf{h}_0|^2 + |\mathbf{h}_1|^2 ML_1^2, \tag{5.38}$$

and IRSs have significant impact if 2nd term is dominant, i.e., $ML_1^2 \gg |\mathbf{h}_0|^2|\mathbf{h}_1|^{-2}$.

5.6 Variable-gain reflectors

One can further consider a more general setting where IRS reflectors, while being passive, have variable gains β_l, where $\beta_l \leq 1$ reflects their passive nature. In this setting, the IRS gain is

$$g(\boldsymbol{\phi}, \boldsymbol{\beta}) = |\mathbf{h}_0 + \sum_l \beta_l e^{j\phi_l}\mathbf{h}_l|^2, \tag{5.39}$$

and it can be jointly optimized over $\boldsymbol{\phi}$ and $\boldsymbol{\beta}$:

$$g^* = \max_{\boldsymbol{\phi},\boldsymbol{\beta}} \left|\mathbf{h}_0 + \sum_{l=1}^{L} \beta_l e^{j\phi_l}\mathbf{h}_l\right|^2 \text{ s.t. } \beta_l \leq 1,\ l = 1...L. \tag{5.40}$$

It follows that using the largest possible gains $\beta_l = 1$ is optimal in many cases provided phase shifts are also optimized and, therefore, there is no loss of optimality in assuming $\beta_l = 1$, as the following Proposition shows.

Proposition 6. *Consider the optimization problem in* (5.40)*. It is optimal to use the largest possible gains, i.e., $\beta_l^* = 1$, for the settings in Propositions 3-5 so that their respective optimal phase shifts and IRS gains apply to this problem as well.*

Proof. One can use the proofs of Propositions 3-5 with $\mathbf{h}_l$ replaced by $\beta_l\mathbf{h}_l$. In

particular, for Proposition 3, (5.7) is replaced by

$$\begin{aligned} g(\boldsymbol{\phi},\boldsymbol{\beta}) &= |\mathbf{h}_0|^2 + 2\sum_l \beta_l |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum_{l,k} \beta_l \beta_k |\mathbf{h}_l^+ \mathbf{h}_k| \cos(\phi_k - \phi_l + \varphi_{lk}) \\ &\leq |\mathbf{h}_0|^2 + 2\sum_l |\mathbf{h}_0^+ \mathbf{h}_l| + \sum_{k,l} |\mathbf{h}_l^+ \mathbf{h}_k| \\ &= |\mathbf{h}_0|^2 + 2L|\mathbf{h}_0^+ \mathbf{h}_1| + L^2 |\mathbf{h}_1|^2, \end{aligned} \tag{5.41}$$

so that $\beta_l^* = 1$ follows, where the inequality is due to $\beta_l \leq 1$. Likewise, for Proposition 4, (5.18) is replaced by

$$\begin{aligned} g(\boldsymbol{\phi},\boldsymbol{\beta}) &= |\mathbf{h}_0|^2 + 2\sum_l \beta_l |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum_l \beta_l^2 |\mathbf{h}_l|^2 \\ &\leq |\mathbf{h}_0|^2 + 2\sum_l |\mathbf{h}_0^+ \mathbf{h}_l| + \sum_l |\mathbf{h}_l|^2, \end{aligned} \tag{5.42}$$

so that $\beta_l = 1$ is optimal (since it attains the upper bound with optimal phase shifts). In a similar way, the proof of Proposition 5 can also be extended (in particular, the upper bound in (5.28) still holds and is attained by $\beta_l = 1$ in combination with optimal phase shifts). □

It should be noted that this result also applies to the case of discrete gains, e.g., using controllers with a finite number of bits to adjust reflection amplitudes, in which case the maximum possible gain should be used. However, it can be shown that $\beta_l = 1$ are not necessarily optimal if phase shifts are not optimized as well, i.e., the largest possible gains are optimal for the joint optimization only. For example, consider the following 2-reflector IRS setup with phase shifts $\phi_1 = 0$ and $\phi_2 = \pi$,

where setting $\beta_l = 1$ is not optimal, as shown below:

$$\begin{aligned} &\mathbf{h}_0 = \mathbf{h}_1 = \mathbf{h}_2 = \mathbf{h} \\ &g(\boldsymbol{\phi}, \boldsymbol{\beta}) = |\mathbf{h}_0 + \beta_1 \mathbf{h}_1 - \beta_2 \mathbf{h}_2| = (1 + \beta_1 - \beta_2)|\mathbf{1}| \\ &g(\boldsymbol{\phi}, \beta_1 = 1, \beta_2 = 1) = |\mathbf{h}| < g(\boldsymbol{\phi}, \beta_1 = 1, \beta_2 = 0) = 2|\mathbf{h}|, \end{aligned} \tag{5.43}$$

where $\mathbf{h}$ is any non-zero vector, $|\mathbf{h}| \neq 0$.

Chapter 6

Alternating Optimization Algorithm

6.1 Introduction

In this chapter, we introduce the Alternating Optimization Algorithm (AOA), an iterative algorithm based on the closed-form, globally optimal solution obtained for the single-reflector IRS in the previous chapter. We provide explanations and a block diagram to illustrate how this algorithm iteratively deals with the general case for any L, and include the proof of its convergence.

However, convergence to a global optimum cannot be guaranteed due to the non-convexity of the optimization problem; therefore, it is beneficial to run the algorithm for multiple initial points. To achieve this, we have developed the multi-start AOA. This approach enables searching over different initial points and identifies the one leading to the highest IRS gain. Additionally, we suggest an initial point based on the bounds established earlier in this thesis. We demonstrate that such an initialization of AOA ensures a guaranteed global optimality gap for the general SIMO/MISO channel. Further, we conduct numerical experiments to show the impact of different initial points on AOA's performance. These experiments compare algorithm performance under two scenarios: one using the initial point derived from the bounds, and the other using an arbitrary initial point.

We examine the special cases for which closed-form solutions were found in the previous chapter and show that AOA converges to these solutions, thus achieving global optimality. For cases without known closed-form solutions, we conduct an

extensive number of numerical experiments to evaluate AOA. These numerical results demonstrate its superior performance compared to other numerical methods in the literature.

6.2 Alternating Optimization Algorithm

While a number of closed-form solutions to the IRS optimization problem in (3.7) have been presented in Chapter 5, the general case remains an open problem. No algorithm to solve this problem numerically for a global optimum is known either. To partially address this problem, we use the closed-form globally-optimal solution for single-reflector IRS, see Corollary 1 in Chapter 5. While this Corollary applies to the $L = 1$ case only, it can be used as a building block to construct a semi-analytical iterative algorithm for any number of reflectors; since such algorithm is gradient-free, it is efficient, even for a large number of reflectors. The key idea is to optimize a single reflector phase at each iteration using the closed-form solution in (5.16) while keeping all other phases fixed. This can be done sequentially for all reflectors. In the optimization literature, this is known as alternating optimization (optimizing only single variable at a time). Let us illustrate this idea for $L = 2$:

- Step 1: optimize ϕ_1 using (5.16) with $\mathbf{h}_0 + e^{j\phi_2}\mathbf{h}_2$ in place of $\mathbf{h}_0$,

$$\phi_1 = \arg\{\mathbf{h}_1^+(\mathbf{h}_0 + e^{j\phi_2}\mathbf{h}_2)\}. \tag{6.1}$$

- Step 2: optimize ϕ_2 using (5.16) with $\mathbf{h}_0 + e^{j\phi_1}\mathbf{h}_1$ in place of $\mathbf{h}_0$ and $\mathbf{h}_2$ in place of $\mathbf{h}_1$,

$$\phi_2 = \arg\{\mathbf{h}_2^+(\mathbf{h}_0 + e^{j\phi_1}\mathbf{h}_1)\}. \tag{6.2}$$

These iterations can be repeated as many times as necessary, using some initial

value of ϕ_2 at step 1 of the very first iteration. Since this algorithm generates non-decreasing sequences of SNRs and since this sequence is bounded, it will converge, which is a welcome property.

In the general case (arbitrary L), ϕ_l is optimized at step l while all other phases are kept constant:

- Step l: optimize ϕ_l using (5.16) with $\mathbf{h}_0 + \sum_{k \neq l} e^{j\phi_k}\mathbf{h}_k$ in place of $\mathbf{h}_0$ and $\mathbf{h}_l$ in place of $\mathbf{h}_1$,

$$\phi_l = \arg\left\{\mathbf{h}_l^+(\mathbf{h}_0 + \sum_{k \neq l} e^{j\phi_k}\mathbf{h}_k)\right\}. \tag{6.3}$$

This is summarized in Algorithm 1, and the corresponding flowchart is shown in Fig. 6.1.

Phase optimization is performed alternatingly in the inner loop, one phase at a time, using the closed-form solution in (5.17) (which is globally-optimal for that particular step), as in Step l above. Since one-by-one optimization is not necessarily globally-optimal (even though each step is), multiple iterations are needed, which are performed by the outer loop. Step 4 of Algorithm 1 is needed to account for a rare (but possible) case $\mathbf{h}_0^{(l)} = 0$, for which ϕ_l is arbitrary so it is randomly generated. $\boldsymbol{\phi}_0$ is an initial point (phase shift vector), Δg is a convergence tolerance, i_0 is the number of outer iterations over which the increase in the gain is evaluated in the termination criterion; $g^{(i)} = 0$ for $i < 0$, $g^{(i,l)}$ is the IRS gain after l-th inner loop iteration of i-th outer loop iteration, while $g^{(i)}$ is the gain after the complete i-th outer loop iteration. The stopping criteria is to terminate the Algorithm if the increase in the gain over the last i_0 outer iteration does not exceed the threshold Δg. It should be pointed out that this Algorithm is different from those presented in [39] and [43]; its superior performance is demonstrated below in Proposition 8 and in Section 6.5.

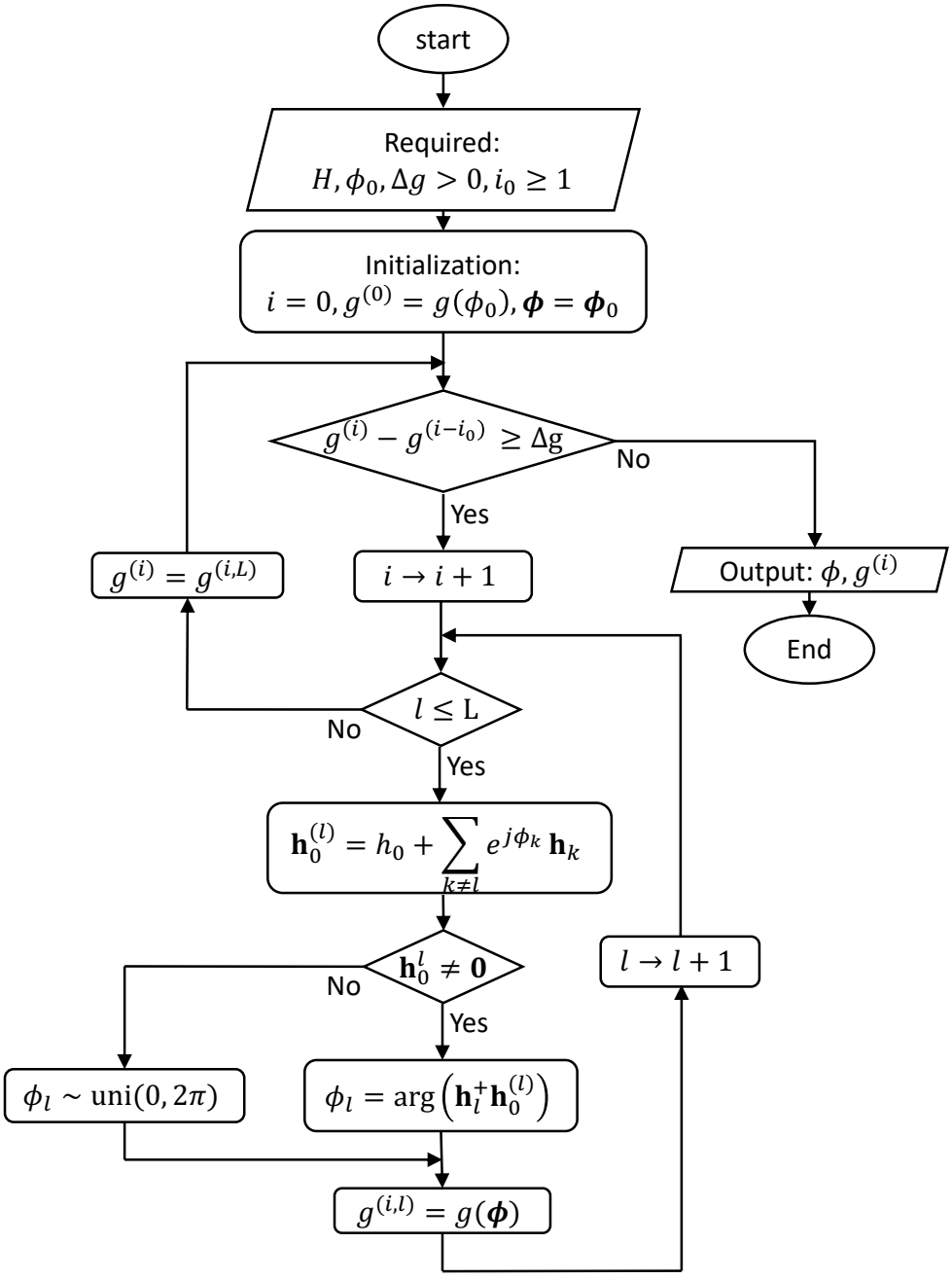


Figure 6.1: Flowchart for Alternating Optimization Algorithm (AOA).

Algorithm 1 Alternating Optimization Algorithm (AOA)

Required: H, $\boldsymbol{\phi}_0$, $\Delta g > 0$, $i_0 \geq 1$
Initialization: $i = 0$, $g^{(0)} = g(\boldsymbol{\phi}_0)$, $\boldsymbol{\phi} = \boldsymbol{\phi}_0$.
repeat (outer loop)
  1. Update $i \to i + 1$.
  for $l = 1$ to L **do** (inner loop)
    2. Set $\mathbf{h}_0^{(l)} = \mathbf{h}_0 + \sum_{k \neq l} e^{j\phi_k} \mathbf{h}_k$.
    if $\mathbf{h}_0^{(l)} \neq \mathbf{0}$
      3. Set $\phi_l = \arg(\mathbf{h}_l^{+} \mathbf{h}_0^{(l)})$.
    else
      4. Set $\phi_l \sim \text{uni}(0, 2\pi)$.
    end if
    5. Set $g^{(i,l)} = g(\boldsymbol{\phi})$.
  end for
  6. $g^{(i)} = g^{(i,L)}$.
until $g^{(i)} - g^{(i-i_0)} \leq \Delta g$.
Output: $\boldsymbol{\phi}$, $g^{(i)}$

Since a closed-form solution is used in each iteration of the inner loop, no gradients or Hessians are necessary and no numerical optimization is used. Hence, Algorithm 1 is computationally-efficient. As Proposition 7 shows, the sequence $g^{(i)}$ of IRS gains generated by this Algorithm is non-decreasing and bounded and therefore converges, and so is the Algorithm.

Proposition 7. *Algorithm 1 generates a non-decreasing and bounded sequence of IRS gains $g^{(i)}$ and therefore converges. Its convergence point is a local maximum for the problem in* (3.7).

Proof. To see that Algorithm 1 generates non-decreasing sequence $g^{(i)}$, let $\boldsymbol{\phi}^{(i,l)}$ be the phase vector after step l of the inner loop has been completed at step i of the

outer loop, and observe the following:

$$g^{(i)} = g^{(i,L)} \tag{6.4}$$

$$\leq \max_{\phi_1} g(\boldsymbol{\phi}^{(i,L)}) = g(\boldsymbol{\phi}^{(i+1,1)}) = g^{(i+1,1)} \tag{6.5}$$

$$\leq \max_{\phi_2} g(\boldsymbol{\phi}^{(i+1,1)}) = g^{(i+1,2)} \tag{6.6}$$

$$\leq \cdots \leq \max_{\phi_L} g(\boldsymbol{\phi}^{(i+1,L-1)})) = g^{(i+1)}, \tag{6.7}$$

where (6.4) is the IRS gain after step i of the outer loop has been completed; (6.5), (6.6) and (6.7) represent steps 1, 2 and L of the inner loop at step $i+1$ of the outer loop. Thus, $g^{(i)} \leq g^{(i+1)}$, i.e., $g^{(i)}$ is a non-decreasing sequence. Intuitively, this is so because, at each step of the inner loop, the IRS gain cannot decrease since the respective phase is optimal, i.e., maximizes the gain, at that step. This sequence is bounded, as has been established in Propositions 1 and 2. Therefore, it converges and hence the termination criterion in Algorithm 1, i.e., $g^{(i)} - g^{(i-i_0)} \leq \Delta g$, will be eventually satisfied, for any $\Delta g > 0$, any $i_0 > 0$ and any initial point $\boldsymbol{\phi}_0$. To see that this convergence point is a local maximum, observe that it cannot be a local minimum or an inflection point since the latter would mean that at least one ϕ_l is not (locally) optimal and hence it can be improved at step l of the inner loop (since Step 3 of Algorithm 1 sets optimal ϕ_l while all other phases are fixed). □

6.3 Multi-start AOA

Since the problem in (3.7) is not convex, its local maximum obtained by Algorithm 1 is not necessarily a global one (this is a common property for non-convex problems in general). Further, Algorithm 1 requires initialization, and because the optimization problem in (3.7) is non-convex, its convergence point is not independent of the initial point. A multi-start implementation, referred to as multi-start AOA, addresses this issue. It executes Algorithm 1 using numerous randomly-generated initial points.

The best outcome, characterized by the highest IRS gain, is then selected and reported. Executing Algorithm 1 for a number of initial points leads to a linear increase in computational complexity. However, as Algorithm 1 is based on a closed-form solution and is computationally efficient, the overall complexity of the multi-start Algorithm remains reasonable. This search across different initial points, while adding to the complexity of the multi-start AOA, enables achieving higher IRS gain values, thus creating a trade-off between performance and complexity. Algorithm 2 summarizes multi-start AOA, and Fig. 6.3 shows its flowchart.

Algorithm 2 Multi-start AOA

Required: $\mathbf{H}$, $\Delta g > 0$, $i_0 \geq 1, m_{\max}$
Initialization: $m = 0, G = 0$.
repeat (initial point search)
1. Update $m \to m + 1$.
2. Set $\boldsymbol{\phi} \sim \text{uni}^L(0, 2\pi)$.
3. Set $[\boldsymbol{\phi}_m, g_m] = \mathbf{AOA}[\mathbf{H}, \boldsymbol{\phi}, \Delta g, i_0]$.
if $g_m > G$ **then**
4. Set $\boldsymbol{\phi}^* = \boldsymbol{\phi}_m$.
5. Set $G = g_m$.
end if
until $(m \leq m_{\max})$
Output: $\boldsymbol{\phi}^*, G$

The multi-start algorithm performs AOA iteratively for a number of different initial points. In step 2 of Algorithm 2, an L-dimensional initial point is generated due to $\boldsymbol{\phi} \sim \text{uni}^L(0, 2\pi)$, with i.i.d. elements. During the m-th iteration of the multi-start algorithm, the AOA algorithm is performed for the same channel and parameters but for a different initial point that is generated specifically for this iteration; It then returns the optimized phase shifts (i.e., $\boldsymbol{\phi}_m$) and IRS gain (i.e., g_m), which are the outputs from performing the AOA algorithm for the m-th initial point. Next, g_m is compared to the highest IRS gain obtained in the previous iterations; If

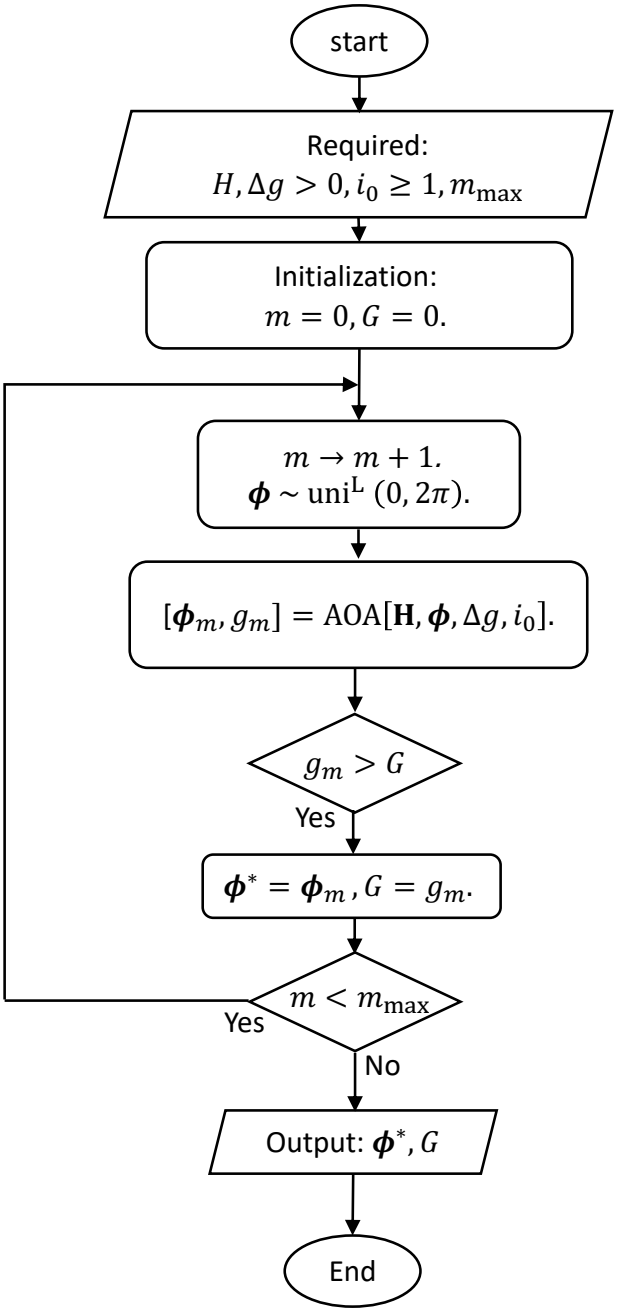


Figure 6.2: Flowchart for multi-start AOA.

it is greater, the outputs for the multi-start algorithm are updated. Otherwise, AOA is performed for another initial point. This process continues until it completes the search across the specified number of initial points, denoted as $m_{\max}$.

6.4 Initial Point Based on Bounds

In the previous section, we demonstrated how using the multi-start Algorithm allows for searching across various initial points to identify the one that leads to the highest IRS gain. However, numerical experiments show that, in many cases, Algorithm 1 does achieve a gain close to the upper bounds established in Chapter 4 and hence close to a global optimum. In fact, as Proposition 8 below outlines, under certain conditions, Algorithm 1 does converge to a global optimum, provided that its initial point $\boldsymbol{\phi}_0$ is properly selected.

To ensure a guaranteed gap to the global optimum in the general case, we set its initial point $\boldsymbol{\phi}_0$ as in (4.5),

$$\phi_{0l} = \arg(v_{1(l+1)}) - \arg(v_{11}), \tag{6.8}$$

This will guarantee that the Algorithm's output gain g_{out} satisfies

$$g_{lb} \leq g_{out} \leq g^* \leq g_{ub}, \tag{6.9}$$

since it generates a non-decreasing sequence of IRS gains $g^{(i)}$ so that the global optimality gap $g^* - g_{out}$ does not exceed $\Delta g = g_{ub} - g_{out}$, i.e.,

$$g^* - \Delta g \leq g_{out} \leq g^*. \tag{6.10}$$

This is illustrated in Fig. 6.3. To the best of our knowledge, this is the first time a numerical algorithm for IRS optimization has a guaranteed *global* optimality gap

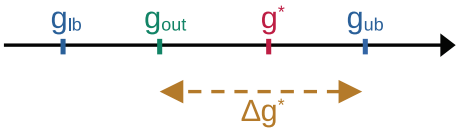


Figure 6.3: Bounding the globally-optimal IRS gain g^* using the bounds in (4.2) and the sub-optimality gap Δg of Algorithm 1.

for the general SIMO/MISO channel.

Next Proposition shows that, under certain conditions, a convergence point of Algorithm 1 is indeed a global optimum.

Proposition 8. *Let the initial point* $\boldsymbol{\phi}_0$ *be set as* $\phi_{0l} = \arg\{\mathbf{h}_l^+ \mathbf{h}_0\}$, $l = 1...L$. *Then, Algorithm 1 converges to a global optimum* g^* *of* (3.7) *in the following cases:*

1. *the setting of Proposition 3 (see Chapter 5, mmWave channel)*
2. *the setting of Proposition 4 (see Chapter 5, mMIMO channel)*
3. *the setting of Proposition 5 (see Chapter 5, multi-IRS channel)*
4. *for the* $L = 1$, *any* N *case*
5. *for the* $N = 1$, *any* L *case*

Proof. For all considered cases, note the following:

$$g^* = g(\boldsymbol{\phi}^*) = g(\boldsymbol{\phi}_0) = g^{(0)} \tag{6.11}$$

$$\leq g^{(1)} \leq \ldots \leq g^{(c)} = g(\boldsymbol{\phi}_c) \leq g^*, \tag{6.12}$$

where $\boldsymbol{\phi}_c$, $g^{(c)}$ is the output of Algorithm 1 at its convergence point (the respective phase shifts and IRS gain); (6.11) follows from (5.6), (5.16), (5.17), (5.23), which imply that $\boldsymbol{\phi}_0 = \boldsymbol{\phi}^*$ for the considered cases; (6.12) holds since, from Proposition 7, $g^{(i)}$ is a non-decreasing sequence (see (6.4)-(6.7)); the last inequality in (6.12) is from the definition of global optimum: $g^* \geq g(\boldsymbol{\phi})$ for any $\boldsymbol{\phi}$. Thus, all inequalities

in (6.12) hold with equality and therefore $g(\boldsymbol{\phi}_c) = g^*$, as required. Note, however, that the latter equality does not imply that $\boldsymbol{\phi}_c = \boldsymbol{\phi}_0$ since globally-optimal point is not necessarily unique (in general). □

We believe this is the first time that an algorithm for IRS optimization has provable convergence to a *global* optimum. It should be pointed out that the choice of an initial point is important here (as well as for any algorithm operating on a non-convex problem in general), due to the following:

- it affects convergence to a global optimum (an algorithm may converge to global optimum for some initial points but not for others)
- it affects the speed of convergence

Since the objective $g(\boldsymbol{\phi})$ in (3.7) is a continuous and bounded function, one can expect that global convergence property will also hold for other settings that are not "far away" from the cases of Proposition 8 or for other initial points that are not far away from the considered one. The multi-start version of Algorithm 1 is implemented to improve its chances of global convergence for other cases. However, the initial point in (6.8) which was obtained based on the lower bound in (4.3), along with some others, remains a good choice.

The next Section presents the results of extensive numerical experiments and compares Algorithm 1 with other known algorithms, demonstrating its superior performance and its ability to converge to a point in close proximity of a global optimum.

6.5 Numerical Experiments

To validate the analytical results above, we present here a number of numerical experiments comparing these results with Algorithm 1 (AOA) above as well as with the fixed point iteration (FPI) method in [43] and the semi-definite relation (SDR) method in [39].

In order to evaluate the tightness of the bounds in (4.2) and (4.27), histograms of g_{ub}/g_{lb} and g_{UB}/g_{lb} for an extensive number of randomly-generated channels under different setups are presented in Figs. 4.1-4.4. These figures demonstrate that $\min\{g_{ub}, g_{UB}\}/g_{lb}$ is less than 4 (i.e., 4.8 dB) for almost all channels in all considered scenarios. This implies that the global optimality gap of Algorithm 1, when initialized as in (6.8), does not exceed these values, as indicated by (6.9),

$$1 \leq g^*/g_{out} \leq g_{ub}/g_{lb}, \tag{6.13}$$

Fig. 6.4 shows the performance of Algorithm 1 (AOA) along with the FPI [43] and the SDR [39] algorithms in terms of the normalized IRS gain $g(\boldsymbol{\phi})/|\mathbf{h}_0|^2$ (i.e., the IRS-assisted gain over that with the direct link alone). Further, the figure includes upper and lower bounds based on (4.2) and (4.27), denoted as UB and LB. For UB, we select either g_{ub} or g_{UB} based on whichever one provides a tighter global optimality gap. To obtain representative results for a wide range of scenarios, 10^4 channels were randomly generated with $h_{l,n} \sim \mathcal{CN}(0,1)$, all i.i.d., samples from i.i.d. Rayleigh fading; the average path loss is factored into the Rx SNR definition for all channel coefficients. This setup represents scenarios where the direct Tx-Rx link and the IRS-assisted links have the same average path loss, i.e., partially-blocked direct link [56]. Each channel is individually optimized using the above algorithms; after this, the best (highest normalized IRS gain) and worst (smallest normalized IRS gain) cases were selected, which are shown in Fig. 6.4(a) and 6.4(b). AOA-1 is Algorithm 1 with the initial point as in (6.8) while AOA-2 uses the all-zero initial point for comparison. Note that the performance of AOA-1 is in-between the lower and upper bounds in (4.2), thus validating these bounds, while AOA-2 does not satisfy the lower bound (since its initial point is not as in (6.8)) but always satisfies the upper bound, as expected. Note also that AOA-1 and AOA-2 outperform the other algorithms both in terms of the achieved IRS gain as well as convergence speed, for the best

and worst-case scenarios. The global optimality gap of Algorithm 1 does not exceed 0.2 dB in both scenarios, while this gap is noticeably larger for the other algorithms. A comparison of AOA-1 and AOA-2 shows the impact of an initial point on the algorithm's performance (a good initial point, e.g., as in AOA-1, can significantly speed up the convergence). Note the differences in the convergence points of different algorithms, which are due to the different phase shift update rules as well as the non-convexity of the optimization problem so that the FPI and SDR algorithms converge to locally-optimal points, which are located away from the globally-optimal one, for both the best and worst channel realizations, while Algorithm 1 convergence point is closer to the global optimum (which does not exceed UB). It is noted that the high values of IRS gain in this scenario may not be achievable in practice due to the law of conservation of energy. However, the obtained results are based on optimizing the objective function in (3.6) without taking into account other limitations.

To investigate the performance of the AOA algorithm in scenarios where a strong direct link exists, Fig. 6.5 compares the AOA algorithm's performance with that of the FPI and SDR algorithms, same as in Fig. 6.4. Fig. 6.5(a) and 6.5(b) show the highest and lowest IRS gains among 10^4 randomly generated channels, where $h_{0,n} = 10$ and $h_{l,n} \sim \mathcal{CN}(0,1)$, all i.i.d.. In both cases, the AOA algorithm outperforms the FPI and SDR algorithms in terms of achieved IRS gain and convergence speed. The global optimality gap does not exceed 0.2 dB, consistent with the results shown in Fig. 6.4. In this scenario, the lower values of achieved IRS gains, compared to those in Fig. 6.4, is due to the limited IRS gain when a strong LoS exists, as discussed in Chapter 5.

To validate Propositions 3 and 8, Fig. 6.6 shows the respective results for the settings of Proposition 3, where 10^4 random channel realizations where generated according to $h_{0,n},\ h_{1,n} \sim \mathcal{CN}(0,1)$, all i.i.d., and $\alpha_l \sim \text{uni}[0, 2\pi]$, also i.i.d.. AOA-3 corresponds to Algorithm 1 with the initial point as in Proposition 3. Note that, for

this initial point, Algorithm 1 converges to global optimum, confirming Propositions 3 and 8. In addition, global convergence also holds for the other initial points (AOA-1 and AOA-2), for both the best and worst-case channels. Based on this observation, we conjecture that Proposition 8 also holds for other initial points. Note also that, for both channels, the upper bound coincides with the global optimum, confirming its tightness.

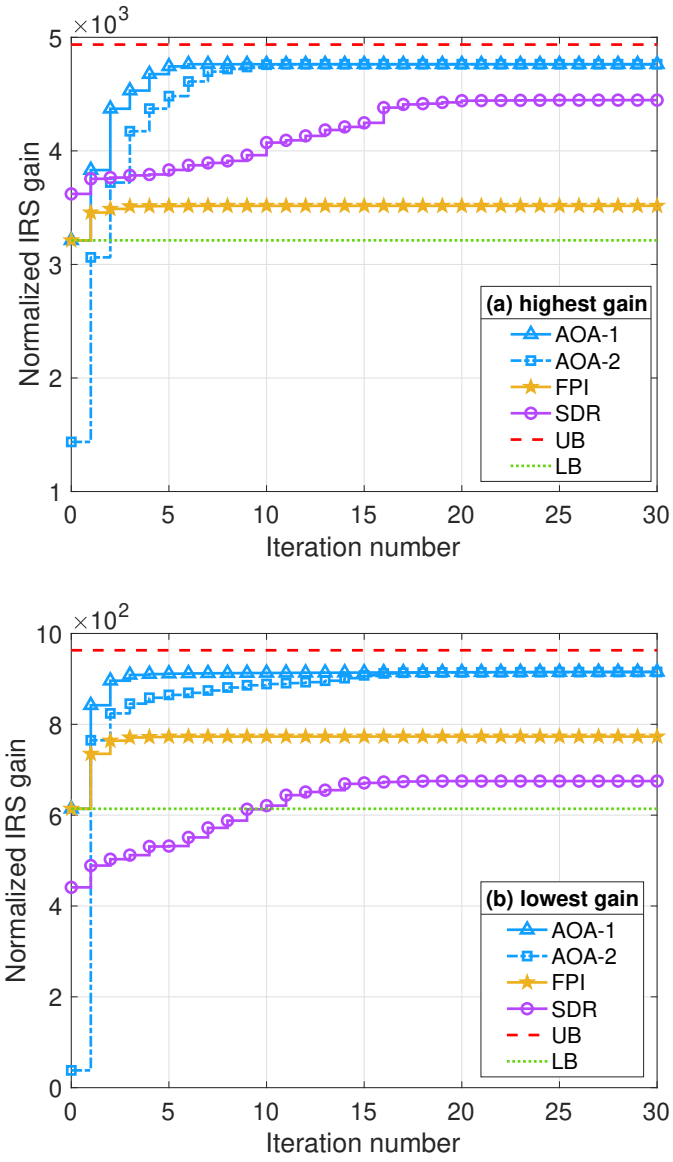


Figure 6.4: Highest (a) and lowest (b) normalized IRS gain scenarios among 10^4 randomly-generated channels with $h_{l,n} \sim \mathcal{CN}(0,1)$, i.i.d. (partially-blocked direct link), $N = 10, L = 100$.

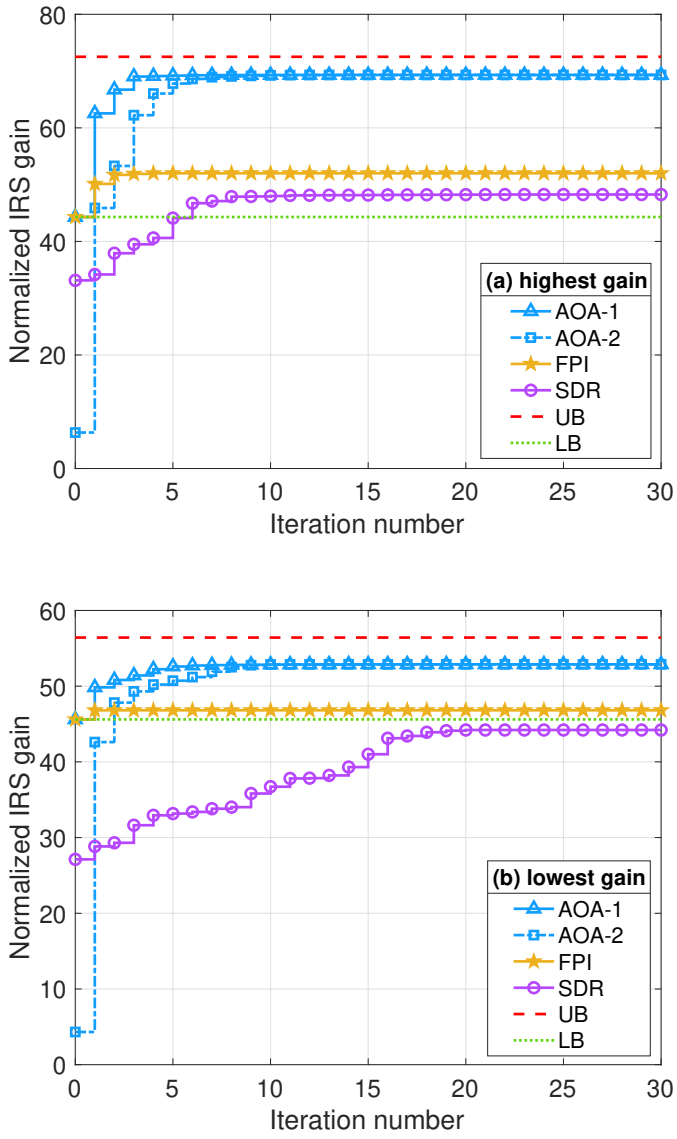


Figure 6.5: Highest (a) and lowest (b) normalized IRS gain scenarios among 10^4 randomly-generated channels with $h_{0,n} = 10$ and $h_{l,n} \sim \mathcal{CN}(0,1)$, i.i.d. (strong direct link), $N = 10, L = 100$.

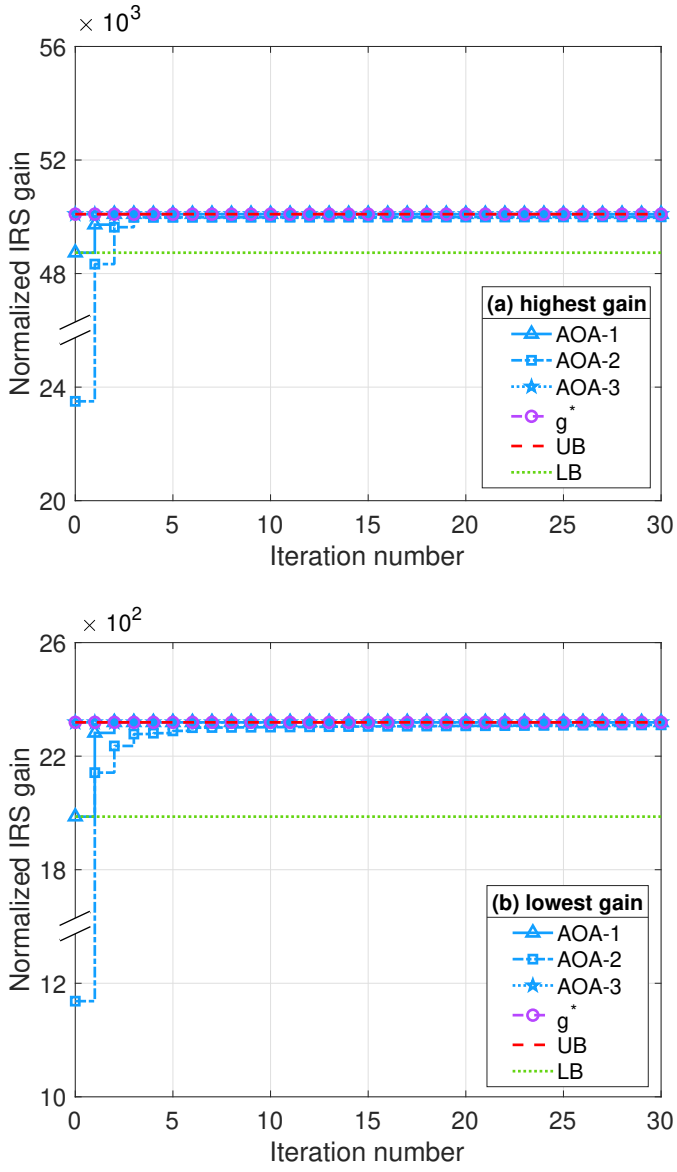


Figure 6.6: Highest (a) and lowest (b) normalized IRS gain scenarios among 10^4 randomly-generated channels for the setting of Proposition 3; $N = 10, L = 100$.

Chapter 7

Small Phase Dispersion Regime

7.1 Introduction

As investigated in the previous chapters, the globally-optimal phase shifts for maximizing IRS gain for SIMO/MISO channels are not known in the general case. Further, the global optimality gap is not known or bounded, either. This chapter partially addresses these issues and proposes a new approach that is based on a concave lower bound to the IRS-assisted channel gain. It is demonstrated that once this lower bound gets maximized, it coincides with the global optimum of the original problem in many cases and, in some additional cases, gives a tight approximation. For the latter case, we introduce a new concept of phase dispersion, consider a small phase dispersion regime, and show that, under this regime, the maximized lower bound is a tight approximation of the global optimum. We further provide numerical experiments for an extensive number of channel realizations to support our results.

Overall the proposed approach is amenable not only to numerical implementations but also to the analysis resulting in closed-form solutions in some cases. In particular, we establish a global optimality gap for IRS gain using discrete 1-bit phase shifters, which offer a low-complexity alternative compared to the full-range continuous phase shifters.

7.2 Global Optimum via Maximized Lower Bound

We adopt here a novel approach based on a concave lower bound, whose maximization is a convex problem that can be solved for a global optimum in a closed form. Next, we demonstrate that the maximized lower bound is tight in many cases, i.e., the IRS phase shifts maximizing the lower bound also maximize the original IRS gain and hence are globally-optimal for the original problem.

To this end, the following Lemma presents a desired lower bound for the IRS gain $g(\boldsymbol{\phi})$ valid for any $\boldsymbol{\phi}$, $\mathbf{h}_0, \mathbf{h}_l$ and tight in many cases after being maximized.

Lemma 1. *The IRS gain $g(\boldsymbol{\phi})$ in (3.4) can be lower bounded, for any $\boldsymbol{\phi}$ and any $\mathbf{h}_0, \mathbf{h}_l$, $l = 1...L$, as follows:*

$$g(\boldsymbol{\phi}) \geq g_a(\boldsymbol{\phi}) \triangleq -\boldsymbol{\phi}^T \mathbf{W} \boldsymbol{\phi} - 2\boldsymbol{\phi}^T \mathbf{b} + c, \tag{7.1}$$

where $\mathbf{W}, \mathbf{b}$, and c are defined as follows:

$$\mathbf{W} = \begin{bmatrix} |\mathbf{h}_1^+ \mathbf{h}_0| + \sum_{l=1, l\neq 1}^{L} |\mathbf{h}_1^+ \mathbf{h}_l| & -|\mathbf{h}_1^+ \mathbf{h}_2| & \cdots & -|\mathbf{h}_1^+ \mathbf{h}_L| \\ -|\mathbf{h}_2^+ \mathbf{h}_1| & |\mathbf{h}_2^+ \mathbf{h}_0| + \sum_{l=1, l\neq 2}^{L} |\mathbf{h}_2^+ \mathbf{h}_l| & \cdots & -|\mathbf{h}_2^+ \mathbf{h}_L| \\ \vdots & \ddots & & \vdots \\ -|\mathbf{h}_L^+ \mathbf{h}_1| & -|\mathbf{h}_L^+ \mathbf{h}_2| & \cdots & |\mathbf{h}_L^+ \mathbf{h}_0| + \sum_{l=1, l\neq L}^{L} |\mathbf{h}_L^+ \mathbf{h}_l| \end{bmatrix}$$

$$\mathbf{b} = \left[\sum_{\substack{m=0 \\ m\neq 1}}^{L} |\mathbf{h}_1^+ \mathbf{h}_m| \varphi_{m1}, \quad \cdots, \quad \sum_{\substack{m=0 \\ m\neq L}}^{L} |\mathbf{h}_L^+ \mathbf{h}_m| \varphi_{mL} \right]^T, \ \varphi_{ml} = \arg(\mathbf{h}_m^+ \mathbf{h}_l) \tag{7.2}$$

$$c = |\mathbf{h}_0|^2 + \sum_{l=1}^{L} |\mathbf{h}_0^+ \mathbf{h}_l| (2 - \varphi_{0l}^2) + \sum_{k,l=1}^{L} |\mathbf{h}_l^+ \mathbf{h}_k| (1 - \varphi_{lk}^2/2).$$

Proof. Using (3.4), after some manipulations, one obtains

$$g(\boldsymbol{\phi}) = |\mathbf{h}_0 + \sum\nolimits_{l=1}^{L} e^{j\phi_l}\mathbf{h}_l|^2 \tag{7.3}$$

$$= \big(\mathbf{h}_0^+ + \sum\nolimits_{l=1}^{L} e^{-j\phi_l}\mathbf{h}_l^+\big)\big(\mathbf{h}_0 + \sum\nolimits_{k=1}^{L} e^{j\phi_k}\mathbf{h}_k\big) \tag{7.4}$$

$$= |\mathbf{h}_0|^2 + \sum\nolimits_k e^{j\phi_k}\mathbf{h}_0^+\mathbf{h}_k + \sum\nolimits_l e^{-j\phi_l}\mathbf{h}_l^+\mathbf{h}_0 + \sum\nolimits_{l,k} e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k \tag{7.5}$$

$$= |\mathbf{h}_0|^2 + 2\mathrm{Re}\big\{\sum\nolimits_l e^{j\phi_l}\mathbf{h}_0^+\mathbf{h}_l\big\} + \sum\nolimits_{l,k} e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k$$

$$= |\mathbf{h}_0|^2 + 2\sum\nolimits_l \mathrm{Re}\left\{e^{j\phi_l}\mathbf{h}_0^+\mathbf{h}_l\right\} + \tfrac{1}{2}\sum\nolimits_{l,k}\left(e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k + e^{j(\phi_l-\phi_k)}\mathbf{h}_k^+\mathbf{h}_l\right) \tag{7.6}$$

$$= |\mathbf{h}_0|^2 + 2\sum\nolimits_l \mathrm{Re}\big\{e^{j\phi_l}\mathbf{h}_0^+\mathbf{h}_l\big\} + \sum\nolimits_{l,k} \mathrm{Re}\big\{e^{j(\phi_k-\phi_l)}\mathbf{h}_l^+\mathbf{h}_k\big\} \tag{7.7}$$

$$= |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+\mathbf{h}_l|\mathrm{Re}\big\{e^{j(\phi_l+\varphi_{0l})}\big\} + \sum\nolimits_{l,k} |\mathbf{h}_l^+\mathbf{h}_k|\mathrm{Re}\big\{e^{j(\phi_k-\phi_l+\varphi_{lk})}\big\} \tag{7.8}$$

$$= |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+\mathbf{h}_l|\cos(\phi_l+\varphi_{0l}) + \sum\nolimits_{l,k} |\mathbf{h}_l^+\mathbf{h}_k|\cos(\phi_k-\phi_l+\varphi_{lk}) \tag{7.9}$$

$$\geq |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+\mathbf{h}_l|\left(1 - \tfrac{1}{2}(\phi_l+\varphi_{0l})^2\right)$$

$$+ \sum\nolimits_{k,l} |\mathbf{h}_l^+\mathbf{h}_k|\left(1 - \tfrac{1}{2}(\phi_k-\phi_l+\varphi_{lk})^2\right) \tag{7.10}$$

$$= -\boldsymbol{\phi}^T\mathbf{W}\boldsymbol{\phi} - 2\boldsymbol{\phi}^T\mathbf{b} + c = g_a(\boldsymbol{\phi}), \tag{7.11}$$

where $\varphi_{kl} = \arg(\mathbf{h}_k^+\mathbf{h}_l)$; the lower bound in (7.10) is due to

$$\cos(x) \geq 1 - x^2/2, \tag{7.12}$$

which holds for any x; (7.11) follows from (7.10), (7.2) after some manipulations. □

To proceed further, we need the following technical result.

Lemma 2. *$g_a(\boldsymbol{\phi})$ in (7.1) is a concave function and $\mathbf{W}$ is positive semi-definite, $\mathbf{W} \geq 0$. If $\mathbf{h}_0^+\mathbf{h}_l \neq 0$ for all l, the concavity is strict and $\mathbf{W}$ is strictly positive definite, $\mathbf{W} > 0$ (and thus non-singular).*

Proof. To prove 1st part, it is sufficient to show that $\mathbf{W}$ is positive semi-definite, $\mathbf{W} \geq 0$, i.e., $\mathbf{z}^+\mathbf{W}\mathbf{z} \geq 0$ for all $\mathbf{z} \in \mathbb{C}^L$ (see e.g., [99]). To this end, any $\mathbf{z}$ can be

expressed as $\mathbf{z} = \mathbf{D}\mathbf{y}$, where

$$\mathbf{D} = \text{diag}\{e^{j\arg(z_l)}\},\ y_l = |z_l|,\ 1 \le l \le L \tag{7.13}$$

so that

$$\begin{aligned}
\mathbf{z}^+\mathbf{W}\mathbf{z} &= \mathbf{y}^T\mathbf{D}^+\mathbf{W}\mathbf{D}\mathbf{y} && (7.14)\\
&= \mathbf{y}^T\mathbf{Q}\mathbf{y},\ \mathbf{Q} = \mathbf{D}^+\mathbf{W}\mathbf{D} && (7.15)\\
&= \sum\nolimits_{l=1}^{L} q_{ll}y_l^2 + \sum\nolimits_{\substack{l,k=1\\k\neq l}}^{L} q_{lk}y_l y_k && (7.16)\\
&= \sum\nolimits_{l=1}^{L} q_{ll}y_l^2 + \sum\nolimits_{\substack{l,k=1\\k<l}}^{L} (q_{lk}+q_{kl})y_l y_k && (7.17)\\
&= \sum\nolimits_{l=1}^{L} q_{ll}y_l^2 + \sum\nolimits_{\substack{l,k=1\\k<l}}^{L} 2\text{Re}\{q_{lk}\}y_l y_k && (7.18)\\
&\ge \sum\nolimits_{l=1}^{L} q_{ll}y_l^2 - \sum\nolimits_{\substack{l,k=1\\k<l}}^{L} 2|q_{lk}|y_l y_k && (7.19)\\
&\ge \sum\nolimits_{\substack{l,k=1\\k\neq l}}^{L} |q_{lk}|y_l^2 - \sum\nolimits_{\substack{l,k=1\\k<l}}^{L} 2|q_{lk}|y_l y_k && (7.20)\\
&= \sum\nolimits_{\substack{l,k=1\\k<l}}^{L} |q_{lk}|(y_l^2 + y_k^2 - 2y_l y_k) && (7.21)\\
&= \sum\nolimits_{\substack{l,k=1\\k<l}}^{L} |q_{lk}|(y_l - y_k)^2 \ge 0, && (7.22)
\end{aligned}$$

where (7.18) follows from $2\text{Re}\{z\} = z + z^+$ for complex z and $q_{lk} = q_{kl}^+$ since $\mathbf{Q}$ is a Hermitian matrix:

$$\mathbf{Q}^+ = (\mathbf{D}^+\mathbf{W}\mathbf{D})^+ = \mathbf{D}^+\mathbf{W}^+\mathbf{D} = \mathbf{D}^+\mathbf{W}\mathbf{D} = \mathbf{Q}, \tag{7.23}$$

where $\mathbf{W}^+ = \mathbf{W}$ since it is real and symmetric; (7.19) follows from $-|z| \le \text{Re}\{z\}$; (7.20) follows from

$$q_{ll} \ge \sum_{\substack{k=1\\k\neq l}}^{L} |q_{lk}|, \tag{7.24}$$

which is due to the following:

$$\begin{aligned} q_{ll} - \sum_{\substack{k=1 \\ k\neq l}}^{L} |q_{lk}| &= w_{ll} - \sum_{\substack{k=1 \\ k\neq l}}^{L} |w_{lk}| \\ &= |\mathbf{h}_l^+ \mathbf{h}_0| + \sum_{\substack{k=1 \\ k\neq l}}^{L} |\mathbf{h}_l^+ \mathbf{h}_k| - \sum_{\substack{k=1 \\ k\neq l}}^{L} |\mathbf{h}_l^+ \mathbf{h}_k| \\ &= |\mathbf{h}_l^+ \mathbf{h}_0| \geq 0, \end{aligned} \tag{7.25}$$

where, from (7.15), $q_{lk} = w_{lk} e^{j \arg(z_k) - j \arg(z_l)}$ and therefore $q_{ll} = w_{ll}$, $|q_{lk}| = |w_{lk}|$. Thus, $\mathbf{W} \geq 0$ and therefore $g_a(\boldsymbol{\phi})$ is concave. To prove strict concavity, let $\mathbf{z} \neq 0$ and note that the inequality in (7.25) and thus in (7.24) are strict if $\mathbf{h}_l^+ \mathbf{h}_0 \neq 0$ for all l. Therefore, the inequality in (7.20) is also strict, $\mathbf{W}$ is strictly positive-definite, $\mathbf{W} > 0$, and thus non-singular and the concavity is strict as well. □

Using the lower bound in (7.1), the globally-optimal IRS gain g^* can be lower-bounded as follows:

$$g^* = \max_{\boldsymbol{\phi}} g(\boldsymbol{\phi}) \tag{P1}$$

$$\geq \max_{\boldsymbol{\phi}} g_a(\boldsymbol{\phi}) = g_a^*. \tag{P2}$$

While the original problem (P1) is difficult to solve for global maximum, either analytically or numerically, due to its non-convex nature, the maximized lower bound problem (P2) is much easier to deal with since it is a convex quadratic problem:

$$\max_{\boldsymbol{\phi}} g_a(\boldsymbol{\phi}) = -\min_{\boldsymbol{\phi}} (\boldsymbol{\phi}^T \mathbf{W} \boldsymbol{\phi} + 2\boldsymbol{\phi}^T \mathbf{b}) + c. \tag{P2}$$

The following Proposition presents its closed-form solution in the general case and Proposition 10 shows that its global optimum is a tight bound for the original problem (P1) in many cases.

Proposition 9. *The globally-optimal phase shifts $\boldsymbol{\phi}_a^*$ and the globally-optimal IRS gain g_a^* for the problem (P2) are as follows:*

$$\boldsymbol{\phi}_a^* = -\mathbf{W}^\dagger \mathbf{b},\ g_a^* = \mathbf{b}^T \mathbf{W}^\dagger \mathbf{b} + c, \tag{7.26}$$

where $\mathbf{W}^\dagger$ is the Moore-Penrose pseudo-inverse of $\mathbf{W}$.

Proof. Since (P2) is an unconstrained convex (quadratic) problem, its stationarity condition is sufficient for global optimality [99], which can be solved in a closed form:

$$\nabla_{\boldsymbol{\phi}} g_a(\boldsymbol{\phi}) = 2(\mathbf{W}\boldsymbol{\phi} + \mathbf{b}) = 0 \Rightarrow\ \boldsymbol{\phi}_a^* = -\mathbf{W}^\dagger \mathbf{b}, \tag{7.27}$$

where $\nabla_{\boldsymbol{\phi}}$ is the gradient with respect to $\boldsymbol{\phi}$. Using (7.27), (7.26) follows (note that pseudo-inverse is needed here since $\mathbf{W}$ can be singular). □

Since $\mathbf{W}$ is positive semi-definite, its pseudo-inverse can be computed via its eigenvalue decomposition [99] [100]:

$$\mathbf{W} = \mathbf{U} \begin{bmatrix} \boldsymbol{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^T \ \Rightarrow\ \mathbf{W}^\dagger = \mathbf{U} \begin{bmatrix} \boldsymbol{\Lambda}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^T, \tag{7.28}$$

where $\mathbf{U}$ is a unitary matrix whose columns are the eigenvectors of $\mathbf{W}$ and the diagonal matrix $\boldsymbol{\Lambda}$ collects its strictly-positive eigenvalues, so that $\boldsymbol{\phi}_a^*$ and g_a^* can be expressed as

$$\boldsymbol{\phi}_a^* = \sum_{i:\lambda_i>0} \mathbf{u}_i \frac{\mathbf{u}_i^T \mathbf{b}}{\lambda_i},\ g_a^* = \sum_{i:\lambda_i>0} \frac{|\mathbf{u}_i^T \mathbf{b}|^2}{\lambda_i} + c, \tag{7.29}$$

where $\mathbf{u}_i$ is i-th eigenvector of $\mathbf{W}$ and λ_i is the respective eigenvalue, and the summation is over all strictly-positive eigenmodes. Next note that if $\mathbf{W}$ is non-singular, then $\mathbf{W}^\dagger = \mathbf{W}^{-1}$. In our setting, it is not difficult to see that $\mathbf{W}$ becomes singular

when the LoS link is resolvable from all reflected links, $\mathbf{h}_l^+\mathbf{h}_0 = 0\ \forall l \geq 1$, which may be the case when the Rx is equipped with a large number of antennas (e.g., massive MIMO under favorable propagation [54]; see also [98]), so that pseudo-inverse is indeed needed here.

Next, we demonstrate that the globally-optimal solutions of the exact and approximate problems (P1) and (P2) coincide in many cases, hence demonstrating tightness of the maximized lower bound.

Proposition 10. *The globally-optimal solutions (phase shifts and the respective IRS gains) of the exact and approximate problems (P1) and (P2) coincide in the following cases:*

1. *$L = 1$, any N:*

$$\begin{aligned} &\phi_{a1}^* = \arg\{\mathbf{h}_1^+\mathbf{h}_0\} = \phi_1^*, \\ &g_a^* = g^* = |\mathbf{h}_0|^2 + |\mathbf{h}_1|^2 + 2|\mathbf{h}_0^+\mathbf{h}_1|. \end{aligned} \tag{7.30}$$

2. *$N = 1$, any L:*

$$\begin{aligned} &\phi_{al}^* = \arg\{h_0\} - \arg\{h_l\} = \phi_l^*, \\ &g_a^* = g^* = \big(|h_0| + \sum\nolimits_{l=1}^{L} |h_l|\big)^2. \end{aligned} \tag{7.31}$$

3. *If $\mathbf{h}_l = e^{\alpha_l}\mathbf{h}_1$, $l = 1...L$, for some (real) α_l, and arbitrary $\mathbf{h}_0$ (this is typical for mmWave/THz channels):*

$$\begin{aligned} &\phi_{al}^* = \arg\{\mathbf{h}_l^+\mathbf{h}_0\} = \arg\{\mathbf{h}_1^+\mathbf{h}_0\} - \alpha_l = \phi_l^*, \\ &g_a^* = g^* = |\mathbf{h}_0|^2 + 2L|\mathbf{h}_0^+\mathbf{h}_1| + L^2|\mathbf{h}_1|^2. \end{aligned} \tag{7.32}$$

4. *If $|\mathbf{h}_l^+\mathbf{h}_0| \gg \sum_{k=1}^{L} |\mathbf{h}_l^+\mathbf{h}_k|\ \forall l \geq 1$ (i.e., dominant LoS, non-resolvable from*

IRS-supported links):

$$\begin{aligned}\phi_{al}^* &\approx \arg\{\mathbf{h}_l^+\mathbf{h}_0\} \approx \phi_l^* \\ g_a^* &\approx g^* \approx |\mathbf{h}_0|^2 + 2\sum\nolimits_{l=1}^{L} |\mathbf{h}_0^+\mathbf{h}_l|. \end{aligned} \tag{7.33}$$

Proof. For Case 1, notice that $\mathbf{W} = |\mathbf{h}_1^+\mathbf{h}_0|,\ \mathbf{b} = |\mathbf{h}_1^+\mathbf{h}_0|\varphi_{01}$, i.e., both are scalars, so that

$$\phi_{a1}^* = -\frac{|\mathbf{h}_1^+\mathbf{h}_0|\varphi_{01}}{|\mathbf{h}_1^+\mathbf{h}_0|} = -\varphi_{01} = \arg\{\mathbf{h}_1^+\mathbf{h}_0\} \tag{7.34}$$

$$g_a^* = |\mathbf{h}_0|^2 + |\mathbf{h}_1|^2 + 2|\mathbf{h}_0^+\mathbf{h}_1|. \tag{7.35}$$

On the other hand, from (7.9),

$$\begin{aligned} g(\boldsymbol{\phi}) &= |\mathbf{h}_0|^2 + 2|\mathbf{h}_0^+\mathbf{h}_1|\cos(\phi_1 + \varphi_{01}) + |\mathbf{h}_1|^2 \\ &\leq |\mathbf{h}_0|^2 + 2|\mathbf{h}_0^+\mathbf{h}_1| + |\mathbf{h}_1|^2, \end{aligned} \tag{7.36}$$

where the inequality is due to $\cos x \leq 1$. The upper bound is global (applies to any $\boldsymbol{\phi}$) and is attained with $\phi_1 = -\varphi_{01}$, which is therefore the globally-optimal phase shift in this case. Thus,

$$\phi_{a1}^* = \phi_1^*,\ g_a^* = g^*, \tag{7.37}$$

i.e., there is no loss of optimality in using the lower bound in this case, as required.

For Case 2, use the following upper bound

$$g_a(\boldsymbol{\phi}) \leq g(\boldsymbol{\phi}) \leq \left(|h_0| + \sum\nolimits_{l=1}^{L} |h_l|\right)^2, \tag{7.38}$$

where 2nd inequality is due to the triangle inequality. Note that the upper bound is

independent of $\boldsymbol{\phi}$, i.e., holds for any $\boldsymbol{\phi}$, including globally-optimal one, and further observe that both inequalities become equalities if $\phi_l = \arg\{h_l^+ h_0\} = -\varphi_{0l}$, since

$$\begin{aligned}\varphi_{lk} &= \arg\{h_l^+ h_k\} = \arg\{h_k\} - \arg\{h_l\} \\ &= \arg\{h_0^+ h_k\} - \arg\{h_0^+ h_l\} = \varphi_{0k} - \varphi_{0l}.\end{aligned} \tag{7.39}$$

so that, for $\phi_l = -\varphi_{0l}$,

$$\phi_l + \varphi_{0l} = 0, \;\; \phi_k - \phi_l + \varphi_{lk} = 0, \tag{7.40}$$

and, from (7.9) and (7.10),

$$\begin{aligned}g(\boldsymbol{\phi}) &= |h_0|^2 + 2\sum\nolimits_l |h_0||h_l| + \sum\nolimits_{l,k} |h_l||h_k| \\ &= g_a(\boldsymbol{\phi}) = \big(|h_0| + \sum\nolimits_l |h_l|\big)^2.\end{aligned} \tag{7.41}$$

Therefore, $\phi_l = -\varphi_{0l} = \arg\{h_0\} - \arg\{h_l\}$ are globally-optimal for both (P1) and (P2) in this case since they attain the upper bound in (7.38).

For Case 3, use the following upper bound, which follows from (7.9) and $\mathbf{h}_l = e^{j\alpha_l}\mathbf{h}_1$,

$$\begin{aligned}g_a(\boldsymbol{\phi}) \le g(\boldsymbol{\phi}) &= |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) + \sum\nolimits_{l,k} |\mathbf{h}_l^+ \mathbf{h}_k| \cos(\phi_k - \phi_l + \varphi_{lk}) \\ &\le |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+ \mathbf{h}_l| + \sum\nolimits_{k,l} |\mathbf{h}_l^+ \mathbf{h}_k| \\ &= |\mathbf{h}_0|^2 + 2L|\mathbf{h}_0^+ \mathbf{h}_1| + L^2|\mathbf{h}_1|^2,\end{aligned} \tag{7.42}$$

and observe that, in this case,

$$\begin{aligned}\varphi_{lk} &= \arg\{\mathbf{h}_l^+ \mathbf{h}_k\} = \alpha_k - \alpha_l = \varphi_{0k} - \varphi_{0l} \\ \varphi_{0l} &= \arg\{\mathbf{h}_0^+ \mathbf{h}_l\} = \alpha_l + \arg\{\mathbf{h}_0^+ \mathbf{h}_1\} = \alpha_l + \varphi_{01},\end{aligned} \tag{7.43}$$

so that (7.40) holds if $\phi_l = -\varphi_{0l}$ and hence both inequalities in (7.42) hold with equality, indicating global optimality of this choice of ϕ_l for both (P1) and (P2) and, furthermore, $g_a^* = g^*$, i.e., no loss of optimality in using the maximized lower bound.

For Case 4, $g(\boldsymbol{\phi})$ can be upper-bounded and approximated as follows:

$$\begin{aligned} g(\boldsymbol{\phi}) &\leq |\mathbf{h}_0|^2 + \sum\nolimits_l \left(2|\mathbf{h}_0^+ \mathbf{h}_l| + \sum\nolimits_k |\mathbf{h}_l^+ \mathbf{h}_k|\right) \\ &\approx |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+ \mathbf{h}_l|. \end{aligned} \tag{7.44}$$

This approximated upper bound is attained if $\phi_l = -\varphi_{0l} = \arg\{\mathbf{h}_l^+ \mathbf{h}_0\}$, indicating approximate global optimality of this choice of ϕ_l for (P1). It is not difficult to see that the same argument also applies to $g_a(\boldsymbol{\phi})$, indicating that these phase shifts are also approximately optimal for (P2). □

7.3 Small Phase Dispersion Regime

Here we identify another scenario where the solution of (P2) is a tight approximation for the solution of (P1). To this end, observe that the lower bound in (P2) holds due to $\cos x \geq 1 - x^2/2$ (see (7.10)) and the latter bound is tight for small x:

$$\cos x \approx 1 - x^2/2 \text{ if } |x| < 1. \tag{7.45}$$

Fig. 7.1 illustrates this: the difference does not exceed 10% if $|x| \leq \pi/3$. Therefore, using (7.9) and (7.10) when all phase terms are not large, the lower bound is

tight,

$$\begin{aligned} g(\boldsymbol{\phi}) &= |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\phi_l + \varphi_{0l}) \\ &\quad + \sum\nolimits_{l,k} |\mathbf{h}_l^+ \mathbf{h}_k| \cos(\phi_k - \phi_l + \varphi_{lk}) \\ &\approx |\mathbf{h}_0|^2 + 2\sum\nolimits_l |\mathbf{h}_0^+ \mathbf{h}_l| \left(1 - \tfrac{1}{2}(\phi_l + \varphi_{0l})^2\right) \\ &\quad + \sum\nolimits_{k,l} |\mathbf{h}_l^+ \mathbf{h}_k| \left(1 - \tfrac{1}{2}(\phi_k - \phi_l + \varphi_{lk})^2\right) \\ &= g_a(\boldsymbol{\phi}), \end{aligned} \tag{7.46}$$

where $\varphi_{lk} = \arg\{\mathbf{h}_l^+ \mathbf{h}_k\}$, i.e., $g(\boldsymbol{\phi}) \approx g_a(\boldsymbol{\phi})$, and, therefore, one can expect that (P2) is a good approximation of (P1).

To quantify this intuition, we introduce the concept of phase dispersion. To this end, let

$$\Delta\theta_{nl} = \arg\{h_{nl}\} - \arg\{h_{n0}\}, \tag{7.47}$$

i.e., $\Delta\theta_{nl}$ is the phase difference of the IRS-reflected and LoS links for n-th antenna and l-th reflector. We call Θ the *phase dispersion* if

$$|\Delta\theta_{nl}| \leq \Theta \ \forall l, n \tag{7.48}$$

and this holds with equality for some l, n. Equivalently,

$$\Theta = \max_{n,l} |\Delta\theta_{nl}|. \tag{7.49}$$

That is, Θ is the largest phase difference between the reflected and LoS links, across all antennas and reflectors. The following Lemma shows that small Θ implies that all the phase terms $\varphi_{kl} = \arg(\mathbf{h}_k^+ \mathbf{h}_l)$ are also small.

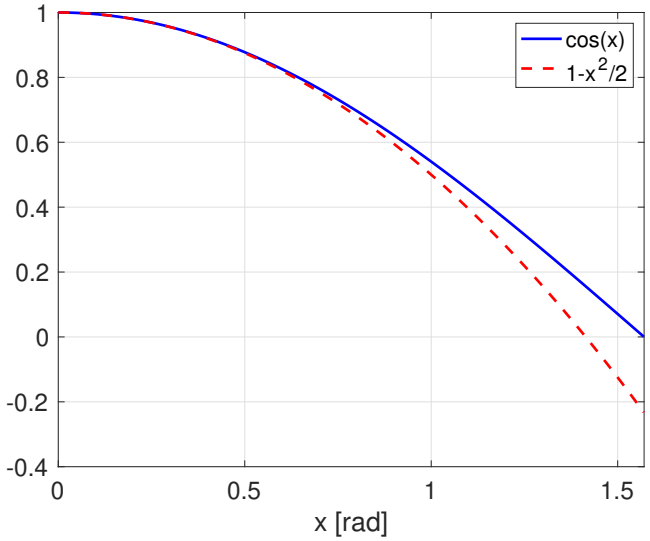


Figure 7.1: Comparison of $\cos(x)$ and its lower bound $1 - x^2/2$.

Lemma 3. *If* (7.48) *holds with* $\Theta \leq \pi/4$*, then*

$$|\varphi_{kl}| = |\arg(\mathbf{h}_k^+ \mathbf{h}_l)| \leq 2\Theta \ \forall \ k, l. \tag{7.50}$$

Proof. To simplify the notations, let us use the following:

$$\begin{aligned} &a_{nl} = |h_{nl}|, \ c_{kl}^n = a_{nl} a_{nk}, \ \theta_{nl} = \arg\{h_{nl}\} \\ &\Delta\vartheta_{kl}^n = \Delta\theta_{nl} - \Delta\theta_{nk}, \ \Delta\vartheta_{kl} = \max_n |\Delta\vartheta_{kl}^n| \\ &\beta_n = \frac{c_{kl}^n \cos(\Delta\vartheta_{kl}^n)}{\sum_n c_{kl}^n \cos(\Delta\vartheta_{kl}^n)} \ \rightarrow \ \sum_n \beta_n = 1. \end{aligned} \tag{7.51}$$

Using the triangle inequality,

$$\begin{aligned} |\Delta\vartheta_{kl}^n| &= |\Delta\theta_{nl} - \Delta\theta_{nk}| \\ &\leq |\Delta\theta_{nl}| + |\Delta\theta_{nk}| \leq 2\Theta \leq \pi/2, \end{aligned} \tag{7.52}$$

so that $\Delta\vartheta_{kl} \leq 2\Theta \leq \pi/2$, $\beta_n \geq 0$ and therefore

$$\begin{aligned}
\arg\{\mathbf{h}_k^+ \mathbf{h}_l\} &= \arg\Big\{ \sum\nolimits_{n=1}^{N} a_{nl} a_{nk} e^{j(\theta_{nl}-\theta_{nk})} \Big\} && (7.53)\\
&= \arg\Big\{ \sum\nolimits_{n=1}^{N} c_{kl}^n e^{j\Delta\vartheta_{kl}^n} \Big\} && (7.54)\\
&= \arg\Big\{ \sum_n c_{kl}^n \cos(\Delta\vartheta_{kl}^n) + j \sum_n c_{kl}^n \sin(\Delta\vartheta_{kl}^n) \Big\} \\
&= \tan^{-1}\left(\frac{\sum_n c_{kl}^n \sin(\Delta\vartheta_{kl}^n)}{\sum_n c_{kl}^n \cos(\Delta\vartheta_{kl}^n)} \right) && (7.55)\\
&= \tan^{-1}\Big(\sum_n \frac{c_{kl}^n \cos(\Delta\vartheta_{kl}^n)}{\sum_n c_{kl}^n \cos(\Delta\vartheta_{kl}^n)} \tan(\Delta\vartheta_{kl}^n) \Big) \\
&= \tan^{-1}\Big(\sum_n \beta_n \tan(\Delta\vartheta_{kl}^n) \Big) && (7.56)\\
&\leq \tan^{-1}\Big(\tan(\Delta\vartheta_{kl}) \sum_n \beta_n \Big) && (7.57)\\
&= \tan^{-1}(\tan(\Delta\vartheta_{kl})) && (7.58)\\
&= \Delta\vartheta_{kl} \leq 2\Theta, && (7.59)
\end{aligned}$$

where (7.55) is due to $\arg(z) = \tan^{-1}(\mathrm{Im}(z)/\mathrm{Re}(z))$ if $\mathrm{Re}(z) \geq 0$ and the latter is ensured by $|\Delta\vartheta_{kl}^n| \leq \pi/2$ (from (7.52)) so that $\sum_n c_{lk}^n \cos(\Delta\vartheta_{kl}^n) \geq 0$; (7.57) follows from $\tan(x)$ being a strictly-increasing function for $|x| \leq \pi/2$ and $\tan^{-1}(\cdot)$ being a strictly-increasing function so that $|\Delta\vartheta_{kl}^n| \leq \Delta\vartheta_{kl} \leq \pi/2$ and $\beta_n \geq 0$ imply $\tan(\Delta\vartheta_{kl}^n) \leq \tan(\Delta\vartheta_{kl})$ and hence (7.57); (7.58) follows from $\sum_n \beta_n = 1$; (7.59) follows from $\tan^{-1}(\tan(x)) = x$ for $|x| \leq \pi/2$. In a similar way, one can establish that $\arg\{\mathbf{h}_k^+ \mathbf{h}_l\} \geq -\Delta\vartheta_{kl} \geq -2\Theta$, from which (7.50) follows. □

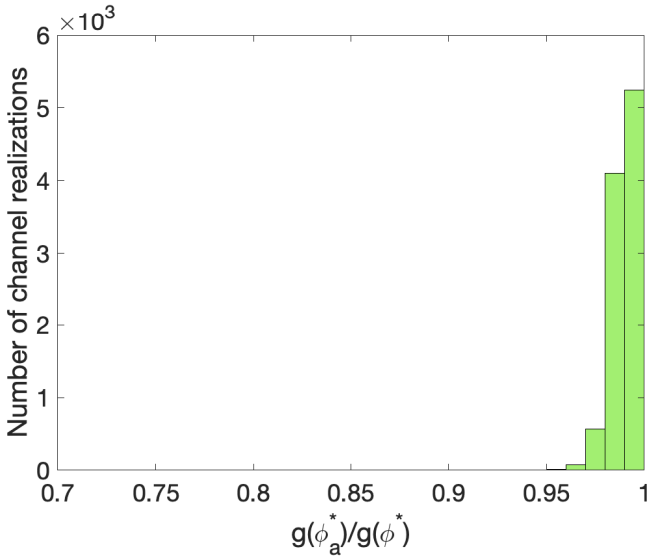
It can be shown (by examples) that no tighter bound is possible in general (since (7.50) holds with equality in some cases). Thus, from this Lemma, $\varphi_{kl} = \arg(\mathbf{h}_k^+ \mathbf{h}_l)$ are small provided that Θ is small, i.e., in the small phase dispersion regime. Keeping in mind that $g(\boldsymbol{\phi})$ is a continuous function and the purpose of IRS phase shifts optimization is to reduce phase imbalance across all links (i.e., to reduce $|\phi_l + \varphi_{0l}|$, $|\phi_k - \phi_l + \varphi_{lk}|$) so that, from (7.46), IRS gain increases, it is natural

to expect that, under optimal phase shifts and for small Θ, all phase terms in (7.46) are also small so that the lower bound in (7.1) is tight and therefore the solution of (P2) is sufficiently close to that of (P1),

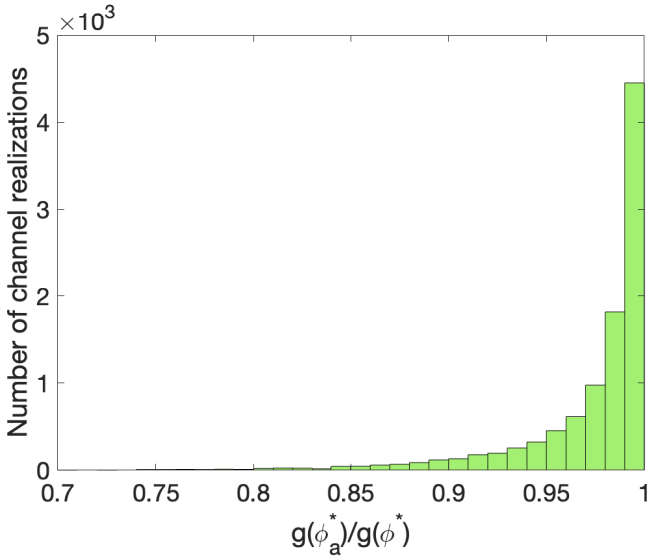
$$\boldsymbol{\phi}^* \approx \boldsymbol{\phi}_a^*,\ g(\boldsymbol{\phi}_a^*) \approx g(\boldsymbol{\phi}^*) = g^*. \tag{7.60}$$

To validate this intuition and thus the tightness of $\boldsymbol{\phi}_a^*$ in the small phase dispersion regime, extensive numerical experiments were carried out. Figs. 7.2 and 7.3 show some representative results in terms of the normalized IRS gain, $g(\boldsymbol{\phi}_a^*)/g(\boldsymbol{\phi}^*)$, for 10^4 randomly-generated channels satisfying (7.48), (7.49), as described earlier. In both figures, the AOA method is used to approximate $g(\boldsymbol{\phi}^*)$ in the small phase dispersion regime, due to Assumption 1, as using exhaustive search becomes infeasible while the number of reflectors, L, grows; e.g., searching for each phase shift at 100 equally spaced points in $[0, 2\pi]$ for only $L = 10$ reflectors results in searching over 10^{20} phase shifts, which is computationally hard, if not impossible. Fig. 7.2 shows the normalized IRS gain for $\Theta = \pi/8$ and $\Theta = \pi/4$, both for $L = 10$. It is shown that for $\Theta = \pi/8$, $g(\boldsymbol{\phi}_a^*)$ is within 5% (or 0.2 dB) of the global maximum $g(\boldsymbol{\phi}^*)$ for most channel realizations, where for larger $\Theta = \pi/4$, this becomes 20% (or 1 dB).

Fig. 7.3 presents the normalized IRS gain for the same settings as in Fig. 7.2 but for $L = 100$. As shown in this figure, for $\Theta = \pi/8$, $g(\boldsymbol{\phi}_a^*)$ is within 5% (or 0.2 dB) of the global maximum for most channel realizations, and for $\Theta = \pi/4$, this becomes 15% (or 0.7 dB). Comparing results from both figures, $g(\boldsymbol{\phi}_a^*)$ exhibits almost the same tightness to global optimal for $\Theta = \pi/8$. However, as the number of reflectors grows for $\Theta = \pi/4$, $g(\boldsymbol{\phi}_a^*)$ provides a tighter approximation of $g(\boldsymbol{\phi}^*)$, which can be justified due to the law of large numbers.

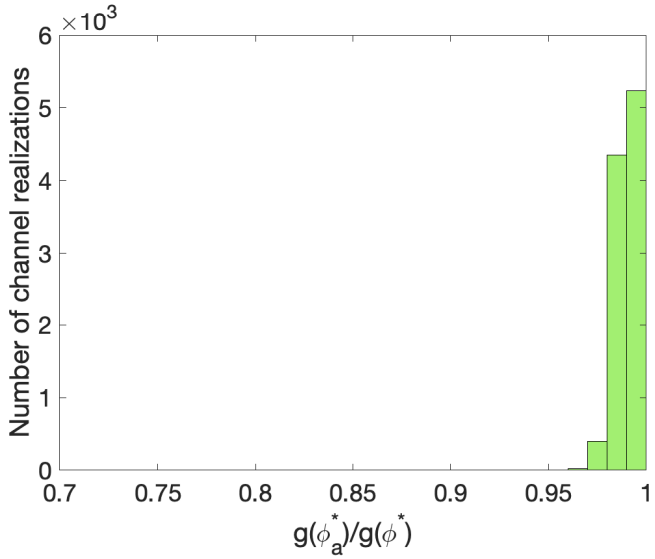


(a) $\Theta = \pi/8$

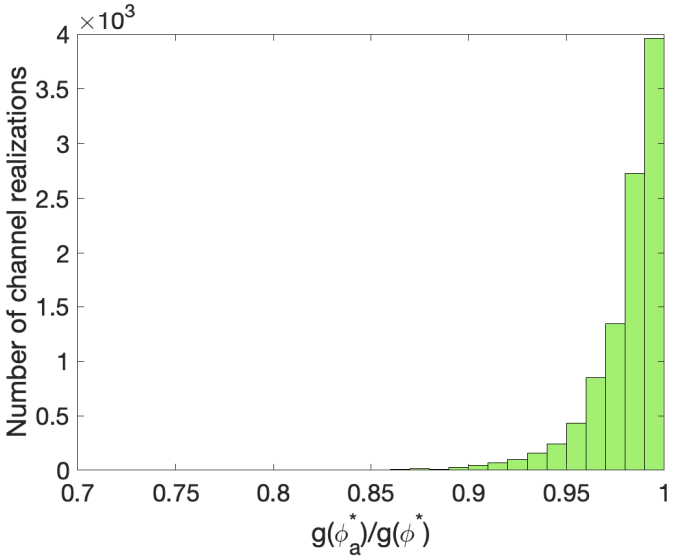


(b) $\Theta = \pi/4$

Figure 7.2: The histogram of the ratio of $g(\boldsymbol{\phi}_a^*)/g(\boldsymbol{\phi}^*)$ for 10^4 randomly-generated channels realizations with $N = 10,\ L = 10$.



(a) $\Theta = \pi/8$



(b) $\Theta = \pi/4$

Figure 7.3: The histogram of the ratio of $g(\boldsymbol{\phi}_a^*)/g(\boldsymbol{\phi}^*)$ for 10^4 randomly-generated channels realizations with $N = 10,\ L = 100$.

To address global optimality, let us consider the cases where the following assumption holds:

Assumption 1. *Let globally optimal phase shifts satisfy*

$$|\phi_l^* + \varphi_{0l}|, |\phi_k^* - \phi_l^* + \varphi_{lk}| \leq \pi/2 \ \ \forall\ k, l \tag{7.61}$$

Using this assumption, we add, without loss of generality, the following constraint to the problem in (3.6):

$$|\phi_l + \varphi_{0l}|, |\phi_k - \phi_l + \varphi_{lk}| \leq \pi/2 \ \ \forall\ k, l \tag{7.62}$$

so that (P1) is equivalent to the constrained problem (P3) below:

$$\text{(P1)} \ \max_{\boldsymbol{\phi}} g(\boldsymbol{\phi}) \ \longleftrightarrow \ \max_{\boldsymbol{\phi}} g(\boldsymbol{\phi}) \text{ s.t. (7.62) (P3).} \tag{7.63}$$

Both (P1) and (P3) have the same optimal phase shifts and respective IRS gains; this is because applying the constraint does not exclude the globally optimal solution, as follows from Assumption 1. The problem in (P3) represents a convex optimization problem; its objective function is concave due to (7.9) and (7.62), and $\cos(x)$ is a concave function for $|x| \leq \pi/2$, and the constraint in (7.62) defines a convex set. The convexity of the problem in (P3) ensures that any locally optimal solution is also globally optimal. Therefore, if the constraint in (7.62) is satisfied at the convergence point of Algorithm 1, i.e., a locally optimal solution, then this convergence point becomes globally optimal. This enables us to optimize IRS globally using the AOA method.

While we acknowledge the absence of formal proof for Assumption 1 in general, it holds in many special cases. For instance, all the cases in Proposition 10, where closed-form globally optimal phase shifts are obtained, satisfy Assumption 1, as the

following conditions hold in all these cases:

$$|\phi_l^* + \varphi_{0l}| = |\phi_k^* - \phi_l^* + \varphi_{lk}| = 0 \ \ \forall\ k, l \tag{7.64}$$

These special cases are listed as follows:

- $L = 1$, any N.
- $N = 1$, any L.
- the setting of Case 3 (see Proposition 10, mmWave/THz channel).
- the setting of Case 4 (see Proposition 10, dominant LoS).

Further, we examined the validity of Assumption 1 in the small phase dispersion regime through extensive numerical experiments. We evaluated Assumption 1 across 10^5 randomly-generated channels satisfying (7.48), (7.49). In particular, complex-valued channel coefficients $h_{nl} = |h_{nl}|e^{j\theta_{nl}}$ were generated with $|h_{nl}| \sim \text{uni}(0, 1)$, all i.i.d., $\theta_{nl} = \theta_{n0} + \Delta\theta_{nl}$ with $\Delta\theta_{nl} \sim \text{uni}(-\Theta, \Theta)$ (to satisfy (7.48), (7.49)), $\Theta \sim \text{uni}(0, \pi/4)$, and $\theta_{n0} \sim \text{uni}(-\pi, \pi)$, all i.i.d., and $\Delta\theta_{n0} = 0$. The AOA method was used for each channel to approximate $\boldsymbol{\phi}^*$, which was then evaluated to verify in (7.62). No violations of Assumption 1 were observed in any channel.

It is noted that the enumerated special cases, along with the conducted numerical experiments, do not constitute definitive proof of the validity of Assumption 1. Instead, they provide strong empirical support, suggesting a high likelihood that the assumption generally holds. However, if Assumption 1 does not hold in general, the feasible set for the problem in (P3), which has an extra constraint compared to the original problem in (P1), implies that any solution to (P3) is also feasible for (P1), and thus, (P3) serves as a lower bound for the original problem:

$$\text{(P3)} \ \max_{\boldsymbol{\phi}} g(\boldsymbol{\phi}) \text{ s.t. (7.62)} \ \leq \ \max_{\boldsymbol{\phi}} g(\boldsymbol{\phi}) \text{ (P1).} \tag{7.65}$$

7.4 Global Optimality Gap

To get further insight into IRS performance in the small phase dispersion regime, the following Proposition presents bounds on the globally-optimum IRS gain in this regime, where lower bound is attained without variable phase shifters (and, therefore, with 1-bit phase shifters as well).

Proposition 11. *Consider the IRS-assisted channel in the small phase dispersion regime as in* (7.48) *with* $\Theta \leq \pi/4$. *Its globally-optimum IRS gain* g^* *can be bounded as follows:*

$$|\mathbf{h}_0|^2(1-\cos(2\Theta)) + g_{UB}\cos(2\Theta) \leq g(0) \leq g(\boldsymbol{\phi}_a^*) \leq g^* \leq g_{UB}, \tag{7.66}$$

where $g_{UB} = |\mathbf{h}_0|^2 + 2\sum_{l=1}^{L} |\mathbf{h}_0^+ \mathbf{h}_l| + \sum_{l,k=1}^{L} |\mathbf{h}_l^+ \mathbf{h}_k|$.

Proof. The upper bound has been already established in (7.42). To establish the lower bounds, observe that $g^* \geq g(0)$ and use (7.9) with $\phi_l = 0$ so that, from (7.50),

$$\cos(\varphi_{0l}) \geq \cos(2\Theta), \ \cos(\varphi_{lk}) \geq \cos(2\Theta), \tag{7.67}$$

and therefore

$$\begin{aligned} g(0) &= |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+ \mathbf{h}_l| \cos(\varphi_{0l}) + \sum_{l,k} |\mathbf{h}_l^+ \mathbf{h}_k| \cos(\varphi_{lk}) \\ &\geq |\mathbf{h}_0|^2 + 2\sum_{l} |\mathbf{h}_0^+ \mathbf{h}_l| \cos(2\Theta) + \sum_{l,k} |\mathbf{h}_l^+ \mathbf{h}_k| \cos(2\Theta) \\ &= |\mathbf{h}_0|^2(1-\cos(2\Theta)) + g_{UB}\cos(2\Theta), \end{aligned} \tag{7.68}$$

as required. □

While the lower bound in (7.66) is not tight in general, it becomes progressively tighter as Θ decreases and the lower and upper bounds converge as $\Theta \to 0$, in which case $\boldsymbol{\phi} = 0$ becomes optimal. For non-zero but small Θ, the sub-optimality gap of using $\boldsymbol{\phi} = 0$ can be bounded, from (7.66), as follows:

$$1 \geq \frac{g(\boldsymbol{\phi}_a^*)}{g^*} \geq \frac{g(0)}{g^*} \geq \cos(2\Theta) + \frac{|\mathbf{h}_0|^2(1-\cos(2\Theta))}{g_{UB}} \geq \cos(2\Theta), \tag{7.69}$$

so that, with $\Theta = \pi/6$, the gap does not exceed 3 dB. Since the IRS gain within these bounds is attained by $\boldsymbol{\phi} = 0$, this may represent a low complexity alternative to optimized IRS for small Θ as no optimization is needed here, and the IRS gain within this gap to global optimum is achieved by $\boldsymbol{\phi} = 0$ (no variable phase shifters are needed at all). When there is flexibility in IRS location, it can be selected to induce small phase dispersion, if possible, so that the complexity of implementation is low.

Further, the results in (7.69) are observed to be valid for Figs. 7.2(a) and 7.3(a), which are obtained for $\Theta = \pi/8$. As observed in both cases, the normalized IRS gain is lower bounded by $\cos(\pi/4)$, i.e., 0.7, but numerical results in Figs. 7.2 and 7.3 show that it is higher than 0.95 for $\Theta = \pi/8$, and at 0.8 for $\Theta = \pi/4$.

The bounds in (7.69) can also be applied to an IRS with 1-bit phase shifters, since its gain g_{1-bit} is lower bounded by $g(0)$, so that

$$1 \geq \frac{g_{1-bit}}{g^*} \geq \frac{g(0)}{g^*} \geq \cos(2\Theta), \tag{7.70}$$

for $\Theta \leq \pi/4$. Therefore, with $\Theta = \pi/6$ phase dispersion, a low-complexity IRS using 1-bit phase shifters attains a gain within 3 dB of the global optimum (where achieving the latter requires infinite-bit phase shifters). To the best of our knowledge, this is the only bound on the global optimality gap of IRS with 1-bit phase shifters

available in the literature.

Chapter 8

Conclusion

8.1 Summary

Although the 5G wireless network is still being deployed worldwide, both academia and industry are already working toward the next generations of wireless networks to meet the forthcoming requirements, such as ultra-high data rates, improved EE/SE, and low latency. However, these requirements may not be achievable with current technologies. Recently, the concept of a smart radio environment has introduced a new vision by aiming to gain some control over the radio environment, which remained unexploited in previous generations of wireless networks. Among the technologies enabling this vision is the IRS. This technology typically consists of a surface equipped with numerous passive reflectors that can induce individual phase shifts to incident signals. Numerous studies have explored various optimization problems related to IRS. For this thesis, our focus is on phase shift optimization for SIMO/MISO channels, a trivial-but not easy case for which no globally optimal solution has yet been found.

In this thesis, after emphasizing the importance of IRS and its state of the art in Chapter 1, we explored the main functionalities and challenges of this technology in Chapter 2. Further, we conducted a comprehensive review of existing literature on the phase shift optimization problem in SIMO/MISO channels, discussing those studies most relevant to our focus in greater detail.

In Chapter 3, we introduced an IRS-assisted SIMO channel model where the

receiver is equipped with N antennas, and IRS is equipped with L passive reflectors. IRS reflectors are passive and ideal; each can induce a phase shift on its incident signal. To detect recieve signal, we employed matched filter beamforming, which allows us to maximize receive SNR for given IRS phase shifts so that channel capacity becomes a function of IRS phase shifts. Then, we defined the IRS gain as our performance metric based on channel gain compared to the no-IRS case.

In Chapter 4, we established bounds on the IRS gain. Using these bounds, we demonstrated that global optimality is achievable in some cases, and in others, these bounds help to assess the global optimality gap. Our extensive numerical results indicate that these bounds can approximate the global optimality gap within 6 dB in most scenarios. Further, we derived an initial point for IRS phase shifts based on these bounds, useful for iterative methods requiring initialization, effectively bounding the optimality gap.

In Chapter 5, we present some practically important cases where the phase shift optimization problem can be solved with globally optimal closed-form solutions. These cases include mmWave/THz channels, single-reflector IRS, massive MIMO settings, and multi-IRS channels. In each case, we provide rigorous proof and present the closed-form solution. Based on the obtained results, we revisited the power scaling law for IRS and showed that scaling as L^2 holds when a direct link is much weaker than the combined reflected links. However, in the presence of a moderate to strong direct link, it scales as L at best. Once dominant LoS exists, the IRS advantage is negligible, and scaling saturates.

In Chapter 6, we propose an iterative algorithm, the AOA method, which we developed using the obtained closed-form globally optimal solution for a single-reflector IRS. Since updates are based on a closed-form solution, this method is computationally efficient. We further provide proof of the algorithm's convergence and, to reduce sensitivity to the initial point, develop the multi-start AOA. This approach locally

solves the optimization problem from multiple initial points and reports the highest gain among them. Our numerical results demonstrate that AOA outperforms known methods and approximates the established upper bounds more closely.

In Chapter 7, we approximate the original problem by maximizing a concave lower bound that has a closed-form solution. We discuss the conditions under which this approximation is accurate. The concept of the small phase dispersion regime is introduced, and it is shown that under this regime, the maximized lower bound closely approximates the global optimal solution. This discussion is supported by extensive numerical results. Additionally, we establish a lower bound for this approximation based on maximum phase dispersion and demonstrate that the numerical results not only support this but often surpass it.

8.2 Future directions

Despite extensive studies on IRS optimization, there are still unresolved issues. In this section, we explore several possibilities for extending the current work.

Analytical proof for Assumption 1: Although we were unable to prove Assumption 1 in Chapter 7 analytically, we examined it through an extensive number of numerical experiments, and none of our results showed a violation of this assumption. In this respect, further attempts to prove this assumption analytically would be valuable and contribute to solving the IRS phase shift optimization problem for SIMO/MISO channels to achieve a global optimum.

Imperfect CSI: Once the IRS and the receiver are equipped with a large number of reflectors and antennas, respectively, acquiring channel information becomes challenging, if not impossible. Therefore, considering a system model that takes into account imperfect CSI is important.

Low complexity solutions: In the literature, the IRS is typically equipped with a large number of reflectors. This demands solutions capable of optimizing the IRS

with reasonable complexity. One approach in this scenario is to group the total IRS reflectors into subarrays, each comprising several adjacent reflectors that can be assigned the same phase shift. However, this approach introduces an optimization problem, leading to a trade-off between complexity and performance, which presents a significant issue in practical implementations.

Discrete phase shifts: In practical systems, the phase shift values can be continuous or discrete, primarily due to reflector design considerations. This thesis proposes several approaches to solve the IRS phase shift optimization problem for SIMO/MISO channels with continuous phase shift values. One natural extension of this problem is the requirement for discrete phase shift values, which raises the interesting question: Is it optimal to solve the problem for continuous phase shifts and then quantize the solution to achieve discrete values, or can solutions with reasonable complexity (polynomial rather than exponential) be found to directly solve the problem for discrete phase shifts?

Manifold optimization methods: The IRS phase shift optimization problem for SIMO/MISO channels can be reformulated as an optimization problem over a product of Stiefel manifolds, e.g., as in [43]. Given the rich existing literature in this field and the availability of tools, a potential future direction involves investigating whether alternative methods can be developed to guarantee global optimality in more cases. Further, analyzing the presence of local optimum points in different scenarios is important.

IRS optimization for multi-user setups: in the cellular design for 6G and next generations of wireless networks, supporting a higher number of devices is essential [9, 11]. Thus, IRS optimization for a multiuser setup, where each user is subject to interference from other users, presents a challenging problem. Problems such as sum rate maximization or transmit power minimization, where IRS assists several users, and each user is subject to a certain threshold SINR, have been considered in [12,57,

59]. However, the methods proposed in these studies typically relax some constraints or provide approximated solutions. Offering analytical or numerical solutions that outperform known methods by achieving global or local optimal solutions would be a valuable contribution, especially since the optimal solution for the simplest case, i.e., where IRS assists two users, has not yet been found. Developing bounds that can tightly bind the optimal solution would also be useful.

Phase shift optimization for OFDM: different optimization problems for OFDM transmission via IRS have been studied so far, such as in [77, 79]. Yet, the existing prototypes show that the induced phase shifts are not independent of the frequency for the current technology for IRS reflectors [57]. Therefore, the phase shifts set for reflectors would not remain the same for the entire bandwidth (e.g., in broadband communication systems). This presents a practical problem that, if solved, can help for higher practicality for this technology.

Phase shift optimization for MIMO channels: the phase shifts optimization in IRS-assisted channels has been addressed with global optimality only for SISO channels so far. This thesis discusses the special cases where globally optimal phase shift optimization for SIMO/MISO channels can be obtained in closed-form. For the other cases, we propose algorithms and approximations, and using established bounds, we show that we can tightly approximate a global optimum, if not achieving it. However, obtaining a locally or globally optimal solution for the IRS optimization problem in MIMO channels remains an open issue. Further, obtaining bounds that provide a tight optimality gap would be valuable for IRS-assisted MIMO channels.

REFERENCES

[1] M. Shafi *et al.*, "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," *IEEE Journal Sel. Areas Comm.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[2] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 2002.

[3] Cisco Annual Internet Report, Mar. 2023, https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report.

[4] Ericsson Mobility Report, Jun. 2023, https://www.ericsson.com/en/reports-and-papers.

[5] Ericsson Traffic and Market Data Report, Nov. 2011, https://www.ericsson.com/en/reports-and-papers.

[6] P. Marsch *et al.*, "5G Radio Access Network Architecture: Design Guidelines and Key Considerations," *IEEE Comm. Mag.*, vol. 54, no. 11, pp. 24–32, Nov. 2016.

[7] C. D. Alwis *et al.*, "Survey on 6G Frontiers: Trends, Applications, Requirements, Technologies and Future Research," *IEEE Open Journal Comm. Soc.*, vol. 2, pp. 836–886, Apr. 2021.

[8] B. Ji *et al.*, "Several Key Technologies for 6G: Challenges and Opportunities," *IEEE Comm. Standards Mag.*, vol. 5, no. 2, pp. 44–51, Jun. 2021.

[9] M. Series, "Minimum requirements related to technical performance for IMT-2020 radio interface (s)," ITU, Geneva, Switzerland, Rep. M.2410-0, 2017.

[10] Z. Zhang *et al.*, "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies," *IEEE Veh. Tech. Mag.*, vol. 14, no. 3, pp. 28–41, Jul. 2019.

[11] C.-X. Wang *et al.*, "On the Road to 6G: Visions, Requirements, Key Technologies, and Testbeds," *IEEE Comm. Surveys Tut.*, vol. 25, no. 2, pp. 905–974, Mar. 2023.

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

[13] C. Liaskos *et al.*, "Realizing Wireless Communication Through Software-Defined HyperSurface Environments," in *IEEE Int. Symp. on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, Chania, Greece, pp. 14-15, Jun. 2018.

[14] L. Yashvanth and C. R. Murthy, "Performance Analysis of Intelligent Reflecting Surface Assisted Opportunistic Communications," *IEEE Trans. Signal Process.*, vol. 71, pp. 2056–2070, Apr. 2023.

[15] M. D. Renzo *et al.*, "Smart Radio Environments Empowered by Reconfigurable AI Meta-Surfaces: an Idea Whose Time Has Come," *EURASIP Journal Wireless Comm. Netw.*, vol. 2019, no. 1, p. 129, May 2019.

[16] M. Di Renzo *et al.*, "Smart Radio Environments Empowered by Reconfigurable Intelligent Surfaces: How It Works, State of Research, and The Road Ahead," *IEEE Journal Sel. Areas Comm.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[17] Q. Wu *et al.*, "Intelligent Reflecting Surface-Aided Wireless Communications: A Tutorial," *IEEE Trans. Comm.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[18] R. Alghamdi *et al.*, "Intelligent Surfaces for 6G Wireless Networks: A Survey of Optimization and Performance Analysis Techniques," *IEEE Access*, vol. 8, pp. 202 795–202 818, Jan. 2020.

[19] E. Björnson *et al.*, "Reconfigurable Intelligent Surfaces: A signal processing perspective with wireless applications," *IEEE Signal Process. Mag.*, vol. 39, no. 2, pp. 135–158, Mar. 2022.

[20] S. Gong *et al.*, "Toward Smart Wireless Communications via Intelligent Reflecting Surfaces: A Contemporary Survey," *IEEE Comm. Surveys Tut.*, vol. 22, no. 4, pp. 2283–2314, Sep. 2020.

[21] Y. Liu *et al.*, "Reconfigurable Intelligent Surfaces: Principles and Opportunities," *IEEE Comm. Surveys Tut.*, vol. 23, no. 3, pp. 1546–1577, Jun. 2021.

[22] B. Zheng *et al.*, "A Survey on Channel Estimation and Practical Passive Beamforming Design for Intelligent Reflecting Surface Aided Wireless Communications," *IEEE Comm. Surveys Tut.*, vol. 24, no. 2, pp. 1035–1071, Apr. 2022.

[23] Z. Chen *et al.*, "Towards Intelligent Reflecting Surface Empowered 6G Terahertz Communications: A Survey," *IEEE China Comm.*, vol. 18, no. 5, pp. 93–119, May 2021.

[24] "Special Issue: Reconfigurable Intelligent Surfaces," *Proceedings of the IEEE*, vol. 110, no. 9, Sep. 2022.

[25] M. M. Amri, N. M. Tran, and K. W. Choi, "Reconfigurable Intelligent Surface-Aided Wireless Communications: Adaptive Beamforming and Experimental Validations," *IEEE Access*, vol. 9, pp. 147 442–147 457, Nov. 2021.

[26] L. Dai *et al.*, "Reconfigurable Intelligent Surface-Based Wireless Communications: Antenna Design, Prototyping, and Experimental Results," *IEEE Access*, vol. 8, pp. 45 913–45 923, Mar. 2020.

[27] M. Ouyang *et al.*, "Computer Vision-Aided Reconfigurable Intelligent Surface-Based Beam Tracking: Prototyping and Experimental Results," *IEEE Trans. Wireless Commu.*, vol. 22, no. 12, pp. 8681–8693, Feb. 2023.

[28] J. Rao *et al.*, "An Active Reconfigurable Intelligent Surface Utilizing Phase-Reconfigurable Reflection Amplifiers," *IEEE Trans. Microwave Theory Tech.*, vol. 71, no. 7, pp. 3189–3202, Apr. 2023.

[29] W. Tang *et al.*, "Wireless Communications With Reconfigurable Intelligent Surface: Path Loss Modeling and Experimental Measurement," *IEEE Trans. Wireless Comm.*, vol. 20, no. 1, pp. 421–439, Feb. 2021.

[30] H. Zhang and B. Di, "Intelligent Omni-Surfaces: Simultaneous Refraction and Reflection for Full-Dimensional Wireless Communications," *IEEE Comm. Surveys Tut.*, vol. 24, no. 4, pp. 1997–2028, Jun. 2022.

[31] V. Arun and H. Balakrishnan, "RFocus: Beamforming Using Thousands of Passive Antennas," in *17th USENIX Symp. Netw. Sys. Design Implementation (NSDI 20)*, Santa Clara, CA, pp. 1047-1061, Feb. 2020.

[32] R. Liu *et al.*, "A Path to Smart Radio Environments: An Industrial Viewpoint on Reconfigurable Intelligent Surfaces," *IEEE Wireless Comm.*, vol. 29, no. 1, pp. 202–208, Feb. 2022.

[33] L. Subrt and P. Pechac, "Intelligent Walls as Autonomous Parts of Smart Indoor Environments," *IET Comm.*, vol. 6, no. 8, pp. 1004–1010, May 2012.

[34] X. Tan *et al.*, "Enabling Indoor Mobile Millimeter-wave Networks Based on Smart Reflect-arrays," in *IEEE Int. Conf. Computer Comm. (INFOCOM)*, Honolulu, HI, USA, pp. 270-278, Apr. 2018.

[35] C. Huang *et al.*, "Reconfigurable Intelligent Surfaces for Energy Efficiency in Wireless Communication," *IEEE Trans. Wireless Comm.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[36] Y.-C. Liang *et al.*, "Large Intelligent Surface/Antennas (LISA): Making Reflective Radios Smart," *Journal Comm. Inf. Netw.*, vol. 4, no. 2, pp. 40–50, Jun. 2019.

[37] X. Mu *et al.*, "Simultaneously Transmitting and Reflecting (STAR) RIS Aided Wireless Communications," *IEEE Trans. Wireless Comm.*, vol. 21, no. 5, pp. 3083–3098, May 2022.

[38] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network: Joint Active and Passive Beamforming Design," in *IEEE Global Comm. Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, pp. 1-6, Dec. 2018.

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

[40] E. Björnson and L. Sanguinetti, "Demystifying the Power Scaling Law of Intelligent Reflecting Surfaces and Metasurfaces," in *IEEE 8th Int. Workshop Computation. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, Le gosier, Guadeloupe, France, pp. 549-553, Dec. 2019.

[41] E. Björnson and L. Sanguinetti, "Power Scaling Laws and Near-Field Behaviors of Massive MIMO and Intelligent Reflecting Surfaces," *IEEE Open Journal Comm. Soc.*, vol. 1, pp. 1306–1324, Sep. 2020.

[42] T. Jiang and W. Yu, "Interference Nulling Using Reconfigurable Intelligent Surface," *IEEE Journal Sel. Areas in Commu.*, vol. 40, no. 5, pp. 1392–1406, May 2022.

[43] X. Yu, D. Xu, and R. Schober, "MISO Wireless Communication Systems via Intelligent Reflecting Surfaces : (Invited Paper)," in *IEEE Int. Conf. Comm. China (ICCC)*, Changchun, China, pp. 735-740, Aug. 2019.

[44] D. Muirhead, M. A. Imran, and K. Arshad, "A Survey of the Challenges, Opportunities and Use of Multiple Antennas in Current and Future 5G Small Cell Base Stations," *IEEE Access*, vol. 4, pp. 2952–2964, May 2016.

[45] H. A. U. Mustafa *et al.*, "Separation Framework: An Enabler for Cooperative and D2D Communication for Future 5G Networks," *IEEE Comm. Surveys Tut.*, vol. 18, no. 1, pp. 419–445, Jul. 2016.

[46] F. S. Shaikh and R. Wismüller, "Routing in Multi-Hop Cellular Device-to-Device (D2D) Networks: A Survey," *IEEE Comm. Surveys Tut.*, vol. 20, no. 4, pp. 2622–2657, Jun. 2018.

[47] E. Basar *et al.*, "Wireless Communications Through Reconfigurable Intelligent Surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Dec. 2019.

[48] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?" *IEEE Trans. Comm.*, vol. 67, no. 10, pp. 7331–7376, Oct. 2019.

[49] M. Nemati, J. Park, and J. Choi, "RIS-Assisted Coverage Enhancement in Millimeter-Wave Cellular Networks," *IEEE Access*, vol. 8, pp. 188 171–188 185, Oct. 2020.

[50] D. Li, "How Many Reflecting Elements Are Needed for Energy- and Spectral-Efficient Intelligent Reflecting Surface-Assisted Communication," *IEEE Trans. Comm.*, vol. 70, no. 2, pp. 1320–1331, Feb. 2022.

[51] J.-S. Jung *et al.*, "Intelligent Reflecting Surface for Spectral Efficiency Maximization in the Multi-User MISO Communication Systems," *IEEE Access*, vol. 9, pp. 134 695–134 702, Oct. 2021.

[52] Z. Huang, B. Zheng, and R. Zhang, "Transforming Fading Channel From Fast to Slow: Intelligent Refracting Surface Aided High-Mobility Communication," *IEEE Trans. Wireless Comm.*, vol. 21, no. 7, pp. 4989–5003, Jul. 2021.

[53] Özdogan, E. Björnson, and E. G. Larsson, "Using Intelligent Reflecting Surfaces for Rank Improvement in MIMO Communications," in *IEEE Int. Conf. Acoustics, Speech Signal Process. (ICASSP)*, Barcelona, Spain, pp. 9160-9164, May 2020.

[54] T. L. Marzetta *et al.*, *Fundamentals of Massive MIMO*. Cambridge University Press, Nov. 2016.

[55] E. Björnson and L. Sanguinetti, "Rayleigh Fading Modeling and Channel Hardening for Reconfigurable Intelligent Surfaces," *IEEE Wireless Comm. Letters*, vol. 10, no. 4, pp. 830–834, Apr. 2021.

[56] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge: Cambridge University Press, May 2005.

[57] S. Abeywickrama *et al.*, "Intelligent Reflecting Surface: Practical Phase Shift Model and Beamforming Optimization," *IEEE Trans. Comm.*, vol. 68, no. 9, pp. 5849–5863, Sep. 2020.

[58] B. O. Zhu, J. Zhao, and Y. Feng, "Active Impedance Metasurface with Full 360° Reflection Phase Tuning," *Nature Sci. Rep.*, vol. 3, no. 1, p. 3059, Oct. 2013.

[59] H. Guo *et al.*, "Weighted Sum-Rate Maximization for Reconfigurable Intelligent Surface Aided Wireless Networks," *IEEE Trans. Wireless Comm.*, vol. 19, no. 5, pp. 3064–3076, May 2020.

[60] H. Xie, J. Xu, and Y.-F. Liu, "Max-Min Fairness in IRS-Aided Multi-Cell MISO Systems With Joint Transmit and Reflective Beamforming," *IEEE Trans. Wireless Comm.*, vol. 20, no. 2, pp. 1379–1393, Feb. 2021.

[61] J.-C. Chen, "Beamforming Optimization for Intelligent Reflecting Surface-Aided MISO Communication Systems," *IEEE Trans. Veh. Tech.*, vol. 70, no. 1, pp. 504–513, Jan. 2021.

[62] Z. Ding, R. Schober, and H. V. Poor, "On the Impact of Phase Shifting Designs on IRS-NOMA," *IEEE Wireless Comm. Letters*, vol. 9, no. 10, pp. 1596–1600, Oct. 2020.

[63] T. Hou *et al.*, "Reconfigurable Intelligent Surface Aided NOMA Networks," *IEEE Journal Sel. Areas Comm.*, vol. 38, no. 11, pp. 2575–2588, Nov. 2020.

[64] J. Lyu and R. Zhang, "Spatial Throughput Characterization for Intelligent Reflecting Surface Aided Multiuser System," *IEEE Wireless Comm. Letters*, vol. 9, no. 6, pp. 834–838, Jun. 2020.

[65] A. U. Makarfi *et al.*, "Reconfigurable Intelligent Surface Enabled IoT Networks in Generalized Fading Channels," in *IEEE Int. Conf. Comm. (ICC)*, Dublin, Ireland, pp. 1-6, Jun. 2020.

[66] Q. Tao *et al.*, "Intelligent Reflecting Surface Aided Multicasting With Random Passive Beamforming," *IEEE Wireless Comm. Letters*, vol. 10, no. 1, pp. 92–96, Jan. 2021.

[67] H. Zhang *et al.*, "Reconfigurable Intelligent Surfaces Assisted Communications With Limited Phase Shifts: How Many Phase Shifts Are Enough?" *IEEE Trans. Veh. Tech.*, vol. 69, no. 4, pp. 4498–4502, Apr. 2020.

[68] C. You, B. Zheng, and R. Zhang, "Channel Estimation and Passive Beamforming for Intelligent Reflecting Surface: Discrete Phase Shift and Progressive Refinement," *IEEE Journal Sel. Areas Comm.*, vol. 38, no. 11, pp. 2604–2620, Nov. 2020.

[69] Y. Zhang *et al.*, "Configuring Intelligent Reflecting Surface With Performance Guarantees: Optimal Beamforming," *IEEE Journal Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 967–979, Aug. 2022.

[70] Y. Han *et al.*, "Large Intelligent Surface-Assisted Wireless Communication Exploiting Statistical CSI," *IEEE Trans. Veh. Tech.*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.

[71] Q. Wu and R. Zhang, "Beamforming Optimization for Wireless Network Aided by Intelligent Reflecting Surface With Discrete Phase Shifts," *IEEE Trans. Comm.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.

[72] F. Olver, "Asymptotics and Special Functions," CRC Press, 1997.

[73] A. Goldsmith, "Wireless Communications," Cambridge University Press, 2005.

[74] E. Björnson, Özdogan, and E. G. Larsson, "Intelligent Reflecting Surface Versus Decode-and-Forward: How Large Surfaces are Needed to Beat Relaying?" *IEEE Wireless Comm. Letters*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

[75] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming," Sep. 2013, http://cvxr.com/cvx.

[76] M. Grant and S. Boyd, "Graph Implementations for Nonsmooth Convex Programs," in *Recent Advances in Learning and Control*, Springer-Verlag, pp. 95-110, 2008.

[77] Y. Yang *et al.*, "Intelligent Reflecting Surface Meets OFDM: Protocol Design and Rate Maximization," *IEEE Trans. Comm.*, vol. 68, no. 7, pp. 4522–4535, Jul. 2020.

[78] S. Zhang and R. Zhang, "Capacity Characterization for Intelligent Reflecting Surface Aided MIMO Communication," *IEEE Journal Sel. Areas Comm.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.

[79] B. Zheng and R. Zhang, "Intelligent Reflecting Surface-Enhanced OFDM: Channel Estimation and Reflection Optimization," *IEEE Wireless Comm. Letters*, vol. 9, no. 4, pp. 518–522, Apr. 2020.

[80] X. Pei *et al.*, "RIS-Aided Wireless Communications: Prototyping, Adaptive Beamforming, and Indoor/Outdoor Field Trials," *IEEE Trans. Comm.*, vol. 69, no. 12, pp. 8627–8640, Dec. 2021.

[81] G. C. Alexandropoulos, N. Shlezinger, and P. del Hougne, "Reconfigurable Intelligent Surfaces for Rich Scattering Wireless Communications: Recent Ex-

periments, Challenges, and Opportunities," *IEEE Comm. Mag.*, vol. 59, no. 6, pp. 28–34, Jun. 2021.

[82] W. Tang *et al.*, "Path Loss Modeling and Measurements for Reconfigurable Intelligent Surfaces in the Millimeter-Wave Frequency Band," *IEEE Trans. Comm.*, vol. 70, no. 9, pp. 6259–6276, Sep. 2022.

[83] M. Najafi *et al.*, "Physics-Based Modeling and Scalable Optimization of Large Intelligent Reflecting Surfaces," *IEEE Trans. Comm.*, vol. 69, no. 4, pp. 2673–2691, Apr. 2021.

[84] Y. Lin *et al.*, "Tensor-Based Algebraic Channel Estimation for Hybrid IRS-Assisted MIMO-OFDM," *IEEE Trans. Wireless Commu.*, vol. 20, no. 6, pp. 3770–3784, Jun. 2021.

[85] C. Hu *et al.*, "Two-Timescale Channel Estimation for Reconfigurable Intelligent Surface Aided Wireless Communications," *IEEE Trans. Comm.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021.

[86] H. Li *et al.*, "Intelligent Reflecting Surface Enhanced Wideband MIMO-OFDM Communications: From Practical Model to Reflection Optimization," *IEEE Trans. Comm.*, vol. 69, no. 7, pp. 4807–4820, Jul. 2021.

[87] B. Zheng, C. You, and R. Zhang, "Intelligent Reflecting Surface Assisted Multi-User OFDMA: Channel Estimation and Training Design," *IEEE Trans. Wireless Comm.*, vol. 19, no. 12, pp. 8315–8329, Dec. 2020.

[88] F. Zhang, "Matrix theory," Springer, 2011.

[89] E. Vassos *et al.*, "Air-Bridged Schottky Diodes for Dynamically Tunable millimeter-wave Metamaterial Phase Shifters," *Nature Sci. Rep.*, vol. 11, no. 1, pp. 1–10, Mar. 2021.

[90] M. Wang *et al.*, "Reflection Characteristics Measurements of Indoor Wireless Link in D-Band," *Sensors*, vol. 22, no. 18, pp. 1–16, Sep. 2022.

[91] A. Alkhateeb *et al.*, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems," *IEEE Journal Sel. Topics Signal Process.*, vol. 8, no. 5, p. 831–846, Oct. 2014.

[92] H. Do, N. Lee, and A. Lozano, "Reconfigurable ULAs for Line-of-Sight MIMO Transmission," *IEEE Trans. Wireless Comm.*, vol. 20, no. 5, pp. 2933–2947, May 2021.

[93] Özdogan, E. Björnson, and E. G. Larsson, "Intelligent Reflecting Surfaces: Physics, Propagation, and Pathloss Modeling," *IEEE Wireless Comm. Letters*, vol. 9, no. 5, pp. 581–585, May 2020.

[94] M. Rossanese *et al.*, "Designing, Building, and Characterizing RF Switch-Based Reconfigurable Intelligent Surfaces," in *16th ACM Workshop Wireless Netw. Testbeds, Exp. Eval. Characterization (WiNTECH)*, Sydney, NSW, Australia, pp. 69-76, Oct. 2022.

[95] W. Tang *et al.*, "MIMO Transmission Through Reconfigurable Intelligent Surface: System Design, Analysis, and Implementation," *IEEE Journal Sel. Areas Comm.*, vol. 38, no. 11, pp. 2683–2699, Nov. 2020.

[96] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of Favorable Propagation in Massive MIMO," in *22nd Eur. Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, pp. 76-80, Sep. 2014.

[97] Z. Mao *et al.*, "Element-Grouping Intelligent Reflecting Surface: Electromagnetic-Compliant Model and Geometry-Based Optimization," *IEEE Trans. Wireless Comm.*, vol. 21, no. 7, pp. 5362–5376, Jul. 2022.

[98] H. L. Van Trees, "Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory," Wiley, 2002.

[99] S. Boyd and L. Vandenberghe, "Convex Optimization," Cambridge University Press, 2004.

[100] R. A. Horn and C. R. Johnson, "Matrix Analysis," Cambridge University Press, 1990.