Wireless Communications Assisted by Reconfigurable Intelligent Surfaces

Jiang Liu

▶ To cite this version:

Jiang Liu. Wireless Communications Assisted by Reconfigurable Intelligent Surfaces. Networking and Internet Architecture [cs.NI]. Université Paris-Saclay, 2021. English. NNT : 2021UPASG111. tel-03558602

HAL Id: tel-03558602

https://theses.hal.science/tel-03558602v1

Submitted on 4 Feb 2022

Wireless Communications Assisted by Reconfigurable Intelligent Surfaces

Communications sans fil assistées par des surfaces intelligentes reconfigurables

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 580, Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat: Réseaux, Information et Communications
Unité de recherche: Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France.
Graduate School : Informatique et sciences du numérique,
Référent : Faculté des sciences d'Orsay

Thèse présentée et soutenue à Paris-Saclay,
le 15 décembre 2021, par

Jiang LIU

Composition du jury:

Jalel BEN OTHMAN Professeur, Université Sorbonne Paris Nord, France	Président
Valeria LOSCRI CR-HDR, INRIA, France	Rapporteur & Examinatrice
Yue XIAO Professeur, UESTC, Chine	Rapporteur & Examinateur
Marwa CHAFII Professeur Associé, NYU Abu Dhabi, Émirats arabes unis	Examinatrice
Alessio ZAPPONE Professeur Associé, Université de Cassino, Italie	Examinateur
Marco DI RENZO DR CNRS, Université Paris-Saclay, France	Directeur de thèse

Thèse de doctorat

NNT: 2021UPASG111

universite
PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies
de l'information et de
la communication (STIC)

Titre: Communications sans fil assistées par des surfaces intelligentes reconfigurables

Mots clés: Communications sans fil, surfaces intelligentes reconfigurables, modulation spatiale, apprentissage en profondeur, déroulement profond.

Résumé: Récemment, l'émergence de la surface intelligente reconfigurable (RIS) a attiré l'attention de l'industrie et du monde universitaire. Un RIS est une surface plane constituée d'un grand nombre d'éléments réfléchissants passifs à faible coût. En ajustant soigneusement les déphasages des éléments réfléchissants, un RIS peut remodeler l'environnement sans fil pour une meilleure communication. Cette thèse s'articule autour de deux sujets : (i) Etudier la modélisation et l'optimisation des systèmes de communication assistés par RIS. (ii) Étudier la modulation spatiale assistée par RIS, en particulier la détection à l'aide de techniques d'apprentissage en profondeur.

Le chapitre 1 présente le concept d'environnements radio intelligents et RIS comme technique de mise en œuvre. Dans les communications futures, RIS est une technique clé pour obtenir une connectivité transparente et une consommation d'énergie moindre en même temps.

Le chapitre 2 présente les systèmes de communication assistés par RIS. Le principe de réflexion, le problème d'estimation de canal et le problème de conception du système sont présentés en détail. Les recherches de pointe sur les problèmes d'estimation de canal et de conception de système sont passées en revue.

Le chapitre 3 étudie le comportement du rapport signal sur bruit (SNR) en tant que variable aléatoire dans un système MIMO (entrées multiples et sorties multiples) assisté par RIS. L'évanouissement de Rayleigh et la propagation en visibilité directe sont considérés séparément. La dérivation théorique et la simulation numérique ont prouvé que le SNR est équivalent en distribution au produit de trois (évanouissement de Rayleigh) ou de deux (propagation en ligne de visée) variables aléatoires indépendantes.

Le chapitre 4 étudie le comportement des interférences dans un système MIMO assisté par RIS, où chaque station de base dessert un équipement utilisateur (UE) via un RIS. L'interférence au niveau d'un UE est causée par son RIS non desservant. Il est prouvé que le rapport interférence/bruit est équivalent en distribution au produit d'une variable aléatoire Chi-carré et d'une variable aléatoire qui peut être approchée avec une distribution Gamma.

Le chapitre 5 se concentre sur la modulation spatiale assistée par RIS. Tout d'abord, nous présentons la détection assistée par apprentissage en profondeur pour les systèmes MIMO. Ensuite, en généralisant les systèmes de modulation spatiale assistés par RIS en tant que cas particulier des systèmes de modulation spatiale traditionnels, nous étudions la détection basée sur l'apprentissage en profondeur pour les systèmes de modulation spatiale assistés par RIS. Les résultats numériques valident les schémas de détection d'apprentissage en profondeur basés sur des données et sur des modèles pour les systèmes de modulation spatiale assistés par RIS.

Enfin, le chapitre 6 conclut la thèse et discute des pistes de recherche futures possibles.

ÉCOLE DOCTORALE
Sciences et technologies
de l'information et de
la communication (STIC)

Title: Wireless Communications Assisted by Reconfigurable Intelligent Surfaces

Keywords: Wireless communications, reconfigurable intelligent surfaces, spatial modulation, deep learning, deep unfolding

Abstract: Recently, the emergence of reconfigurable intelligent surface (RIS) has attracted heated attention from both industry and academia. An RIS is a planar surface that consists of a large number of low-cost passive reflecting elements. By carefully adjusting the phase shifts of the reflecting elements, an RIS can reshape the wireless environment for better communication. In this thesis, we focus on two subjects: (i) To study the modeling and optimization of RIS-aided communication systems. (ii) To study RIS-aided spatial modulation, especially the detection using deep learning techniques.

Chapter 1 introduces the concept of smart radio environments and RIS as its implementation technique. In future communications, RIS is a key technique to achieve seamless connectivity and less energy consumption at the same time.

Chapter 2 introduces RIS-aided communication systems. The reflection principle, channel estimation problem and system design problem are introduced in detail. State-of-the-art research on the problems of channel estimation and system design are overviewed.

Chapter 3 investigates the distribution of the signal-to-noise ratio (SNR) as a random variable in an RIS-aided multiple-input multiple-output (MIMO) system. Rayleigh fading and line-of-sight propagation are considered separately. The theoretical derivation and numerical simulation prove that the SNR is equivalent in distribution to the product of three (Rayleigh fading) or two (line-of-sight propagation) independent random variables.

Chapter 4 studies the behavior of interference in an RIS-aided MIMO system, where each base station serves a user equipment (UE) through an RIS. The interference at a UE is caused by its non-serving RIS. It is proven that the interference-to-noise ratio is equivalent in distribution to the product of a Chi-squared random variable and a random variable which can be approximated with a Gamma distribution.

Chapter 5 focuses on RIS-aided spatial modulation. First, we introduce deep learning aided detection for MIMO systems. Then, by generalizing RIS-aided spatial modulation systems as a special case of traditional spatial modulation systems, we investigate deep learning based detection for RIS-aided spatial modulation systems. Numerical results validate the proposed data-based and model-based deep learning detection schemes for RIS-aided spatial modulation systems.

Finally, Chapter 6 concludes the thesis and discusses possible future research directions.

Acknowledgements

To my family.

Synthèse en français

La surface intelligente reconfigurable (SIR) est considérée comme l'une des techniques prometteuses pour les futures communications sans fil. Il a attiré l'attention de l'industrie et du milieu universitaire. Dans cette thèse, nous nous concentrons sur les systèmes de communication basés sur les SIF.

Dans le chapitre 1, le contexte de cette thèse est présenté. Pour répondre aux exigences de plus en plus élevées des communications futures, une solution consiste à créer un environnement radio intelligent (SRE), où le canal sans fil est une variable d'optimisation et peut être programmé. Un RIS est une surface plane composée d'un grand nombre d'éléments réfléchissants passifs à faible coût. La reconfigurabilité du RIS réside dans le fait que le coefficient de réflexion de chaque élément peut être ajusté indépendamment. En raison de sa capacité à réaliser une connectivité transparente avec une consommation d'énergie moindre, le RIS est considéré comme une technique clé pour activer le SRE.

Dans le chapitre 2, nous avons résumé les recherches récentes sur les systèmes de communication assistés par SIF. Le principe de réflexion, le problème d'estimation de canal et le problème de conception de système sont présentés en détail. Les systèmes assistés par RIS peuvent être classés en deux catégories, à savoir la transmission d'informations basée sur le SI (RBIT) où le RIS module des informations supplémentaires, et les communications sans fil assistées par RIS (RAWC) où le RIS ne code pas les informations. Dans cette thèse, pour les systèmes RAWC et RBIT, nous nous concentrons sur différents sujets. La distribution du rapport signal sur bruit (SNR) et du rapport interférence sur bruit (INR) dans les systèmes RAWC est abordée dans les chapitres 3 et 4. Pour le RBIT, nous visons à améliorer les performances de détection à l'aide de l'apprentissage en profondeur. Sur la base de l'idée de déploiement en profondeur, nous proposons au chapitre 5 un schéma de détection d'apprentissage en profondeur basé sur un modèle pour un système RBIT spécifique qui réalise une modulation d'indice via la réflexion de RIS.

Dans le chapitre 3, nous visons à étudier le comportement du SNR dans un système MIMO point à point assisté par RIS unique. L'évanouissement de Rayleigh et la propagation en visibilité directe sont considérés séparément. Le théorème central limite et la théorie des matrices aléatoires sont adoptés comme outils pour résoudre des problèmes mathématiques. La dérivation théorique et la simulation numérique prouvent que le SNR est équivalent en distribution au produit de trois (évanouissement de Rayleigh) ou de deux (propagation en ligne de visée) variables aléatoires indépendantes.

Dans le chapitre 4, nous nous intéressons au comportement de l'INR d'un RIS à un autre

RIS. Dans le système considéré, deux stations de base sont considérées, chaque station de base dessert un équipement utilisateur (UE) via un RIS. L'interférence au niveau d'un UE est causée par son RIS non desservant. Des résultats similaires par rapport au chapitre 3 sont obtenus. Il est prouvé que le rapport interférence/bruit est équivalent en distribution au produit d'une variable aléatoire Chi-carré et d'une variable aléatoire qui peut être approchée avec une distribution Gamma. De plus, il est prouvé que le RIS est sensible aux interférences dues à l'évanouissement de l'INR.

Le chapitre 5 se concentre sur les systèmes de modulation spatiale assistée par RIS (RIS-SM). Tout d'abord, nous introduisons la détection assistée par apprentissage profond pour les systèmes MIMO. Ensuite, en généralisant les systèmes de modulation spatiale assistés par RIS comme un cas particulier des systèmes de modulation spatiale traditionnels, nous étudions les détecteurs basés sur l'apprentissage profond pilotés par les données et les modèles pour les systèmes RIS-SM. Le détecteur d'apprentissage profond piloté par les données est réalisé via un réseau neuronal entièrement connecté, le détecteur d'apprentissage profond piloté par modèle est basé sur le déploiement profond, dont l'idée principale est de déployer un algorithme itératif et de traiter chaque itération comme une couche dans le réseau neuronal. Les résultats numériques montrent que : dans le cas où aucun CSI n'est disponible au niveau du récepteur, le détecteur piloté par les données proposé surpasse le détecteur glouton existant ; dans le cas où CSI est disponible au niveau du récepteur, le détecteur piloté par modèle proposé surpasse le détecteur piloté par les données proposé. En outre, le détecteur piloté par modèle proposé surpasse également les détecteurs ADMM et ISTA dépliés.

Enfin, le chapitre 6 conclut la thèse et discute des futures directions de recherche possibles.

Contents

Acronyms

ADC	Analog-to-Digital Converters
ADMM	Alternating Direction Method of Multipliers
AF	Amount of Fading
AI	Artificial Intelligence
AmBC	AambientBackscattering Communications
AO	Alternating optimization
BER	Bit-Error-Ratio
BiG-AMP	Bilinear Generalized Approximate Message Passing
BS	Base Station
BSUM	Block Successive Upper-bound Minimization
CDF	Cumulative Distribution Function
CG	Conjugate Gradient
CPU	Central Processing Unit
CR	Cognitive Radio
CSI	Channel State Information
DCCE	Direct Cascaded Channel Estimation
DFT	Discrete Fourier Transform
DL	Deep Learning
DNN	Deep Neural Network
DRL	Deep Reinforcement Learning
ELPC	Extremely Low-Power Communications

ERLLC	Extremely Reliable and Low-Latency Communications
FeMBB	Further-enhanced Mobile Broadband
GD	Greedy Detector
INR	Interference-to-Noise-Ratio
IoT	Internet-of-Things
IRS	Intelligent Reflecting Surface
ISTA	Iterative Soft Thresholding Algorithm
LOS	line-of-sight
LR	LOS fading
LS	Least Squares
MAP	Maximum A Posteriori
MC	Monte Carlo
MEC	Mobile Edge Computing
MEMS	Microelectromechanical System
MF	Matched Filter
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
NOMA	Non-Orthogonal Multiple Access
OFDM	Orthogonal Frequency Division Multiplexing
PG	Projected Gradient
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
RAWC	RIS-Aided Wireless Communication
RBIT	RIS Based Information Transmission
RCO	Reflection Coefficient based Optimization

RE	Reflective Element
RELU	Rectified Linear Unit
RF	Radio Frequency
RIS	Reconfigurable Intelligent Surface
RR	Rayleigh Fading
RV	Random Variable
SCCE	Separate Cascaded Channel Estimation
SCI	Soft Interference Cancellation
SDR	SemiDefinite Relaxation
SIMO	Single Input Multiple Output
SINR	Signal-to-Interference-plus-Noise Ratio
SM	Spatial Modulation
SNR	Signal-to-Noise Ratio
SOR	Ssuccessive Over-Relaxation
SPC	Short Packet Communications
SR	Symbiotic Radio
SRE	Smart Radio Environments
SSK	Space Shift Keying
SWIPT	Simultaneous Wireless Information and Power Transfer
THz	Terahertz
UE	User Equipment
umMTC	Ultra-massive Machine-Type Communications
VLC	Visible Light Communications
ZF	Zero-Forcing

Notation

The following notation is used throughout this thesis.

$\mathbf{x}$	a vector
$\mathbf{x}_n$	the nth vector
$\lVert\mathbf{x}\rVert$	the norm of $\mathbf{x}$
$\mathbf{X}$	a matrix
$\mathbf{X}_n$	the nth matrix
s_i	the ith symbol
$\overset{d}{=}$, $\sim$	Equivalent in distribution, distributed as
$\overset{N\gg 1}{=}$, $\overset{N\gg 1}{\approx}$, $\overset{N\gg 1}{\propto}$	Equality, approximation and scaling law if $N \gg 1$
$\mathrm{diag}(\mathbf{x})$	a diagonal matrix with diagonal given by vector $\mathbf{x}$
$\arg(x)$	the angle of the complex variable x.
$\mathbb{E}$, $\mathbb{V}$, cov	Expectation, variance, covariance
Re, Im	Real part, imaginary part
$(\cdot)^H$, $\mathbf{1}(\cdot)$	Hermitian operator, indicator function
$\lvert\cdot\rvert$, $\lVert\cdot\rVert$	Absolute value, norm of a vector
$\mathbf{0}_{A\times B}$	$A \times B$ matrix with all zero entries
$\mathbf{1}_{A\times B}$	$A \times B$ matrix with all one entries
$I_\alpha(\cdot)$, $\Gamma(\cdot)$	Bessel function of the first kind, Gamma function
$\mathrm{sinc}(x)$	Normalized sampling function ($\sin(\pi x)$ / (πx))
$\mathcal{O}(\cdot)$	Big O (asymptotic) notation
$\mathcal{N}\left(m,\sigma^2\right)$	Gaussian random variable ($\mathbb{E} = m$, $\mathbb{V} = \sigma^2$)
$\mathcal{CN}\left(m,\sigma^2\right)$	Complex Gaussian random variable ($\mathbb{E} = m$, $\mathbb{V} = \sigma^2$)
$\mathcal{B}(a,b)$	Beta random variable ($\mathbb{E} = a/(a+b)$, $\mathbb{V} = \mathbb{E}(1-\mathbb{E})/(a+b+1)$)
$\mathcal{U}(a,b)$	Uniform random variable in $[a,b]$
$\mathcal{G}(k,\theta)$	Gamma random variable ($\mathbb{E} = k\theta$, $\mathbb{V} = k\theta^2$)
$\mathcal{X}^2(k)$	Central Chi-Square random variable ($\mathbb{E} = k$, $\mathbb{V} = 2k$)
$\mathcal{X}_k^2(\lambda)$	Non-central chi-square random variable ($\mathbb{E} = k+\lambda$, $\mathbb{V} = 2(k+2\lambda)$)
$\mathcal{VM}\left(\mu,\kappa\right)$	Von Mises random variable ($\mathbb{E} = \mu$, $\mathbb{V} = 1 - I_1(\kappa)/I_0(\kappa)$)
Rayleigh(σ)	Rayleigh random variable with scale σ

List of Figures

List of Tables

1 Introduction

This chapter begins with *Section 1.1* which introduces the challenges of 6G networks and the concept of smart radio environments (SRE). *Section 1.2* introduces reconfigurable intelligent surface (RIS) and illustrates the working principle of RIS. In *Section 1.3*, we introduce merits of RIS by summarizing potential applications of RIS in 6G networks. *Section 1.4* highlights the major contributions in this thesis work and the organization of the thesis. *Section 1.5* provides lists of publications produced during the Ph.D. thesis.

1.1 Background

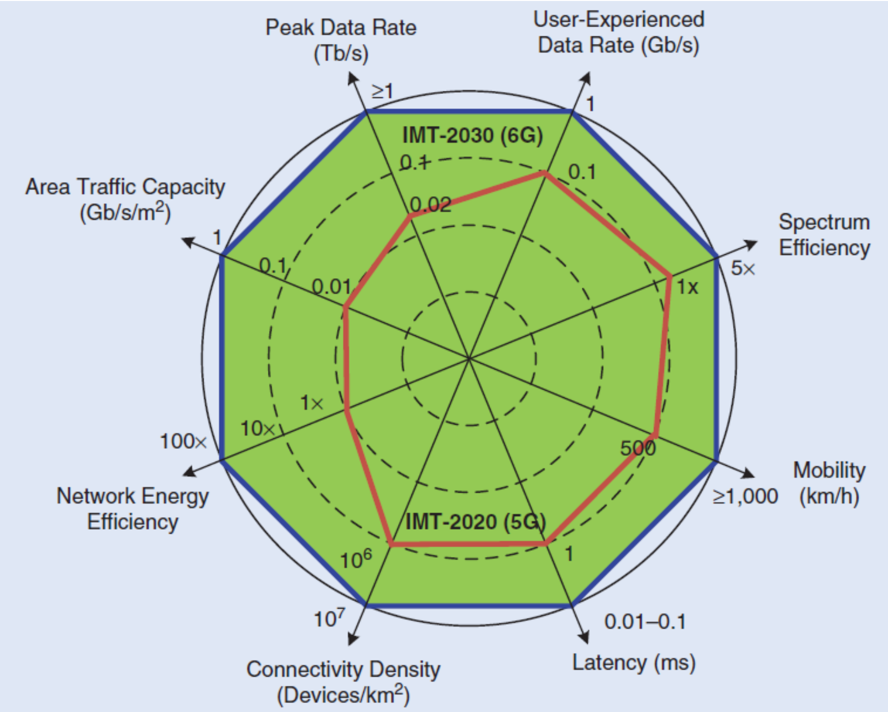


Figure 1.1 – Key capabilities of 6G networks [1].

Nowadays, with the development of technologies such as multimedia and artificial intelligence (AI), the demand for a higher wireless transmission data rate is constantly growing. As the next-generation communication, 6G communication is expected to support usage scenarios such as ultra-massive machine-type communications (umMTC), extremely reliable and low-latency communications (ERLLC), further-enhanced mobile broadband (FeMBB), extremely low-power communications (ELPC) and long-distance and high-mobility communications (LDHMC). The detailed technical objectives of 6G networks are presented in Fig. 1.1. The expected device connectivity density is larger than 10^7 devices/km^2; the reliability should be higher than 99.9999% and the latency should be smaller than 0.1 ms [5]; the peak data rate should be at least 1 Tb/s and the user-experienced data rate should be higher than 1 Gb/s; the energy efficiency should be ten or hundred times of 5G.

Various promising technologies have been proposed to fulfill these typical stringent requirements. Among the recent emerging technologies, reconfigurable intelligent surface (RIS), or intelligent reflecting surface (IRS), draws heated attention from both industry and academia for its ability to create smart radio environments (SRE) where the wireless propagation conditions are configured to enhance wireless communication [2].

For a long time in the history of wireless communication, physical objects in the environment are always regarded as a threat to communication because they usually have a negative effect on wireless communication. However, the rapid developments in meta-materials have enabled the electromagnetic components to shape how they interact with wireless signals. In other words, the scattering, absorption, reflection, and diffraction properties can be controlled by software. In this thesis, we follow the definition in [2] that an SRE is a smart reconfigurable radio environment that plays actively in transferring and processing information and that enhances the reliability of data exchange between transmitters and receivers.

1.2 Reconfigurable intelligent surface

RIS is regarded as one of the most promising techniques to achieve SRE, the research on RIS technology comes under other names, such as reconfigurable intelligent surface [2, 6, 7], intelligent walls [8, 9], smart reflectarrays [10, 11, 12], low-cost devices embedded into walls [13], software-controlled metasurfaces [14]. All of the works above aim at creating smart radio environments. The functions of the proposed RIS technology can be different. By its capability of amplifying and performing signal processing operations on the impinging radio wave, they are classified into active and passive surfaces. By its capability to alter the function of the surface, we can category them into static and dynamic/reconfigurable surfaces. In this thesis, we focus on the nearly-passive and dynamic RIS, which is a sub-wavelength metallic or dielectric planar surface that consists of a large number of low-cost passive reflecting elements, which can reconfigure the wireless communication environment by imposing the desired phase shift on the incoming signal [15, 16].

1.2.1 The working principle of RIS

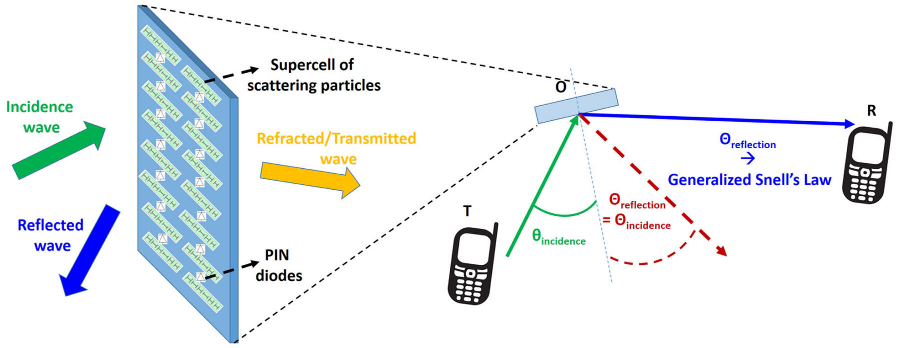
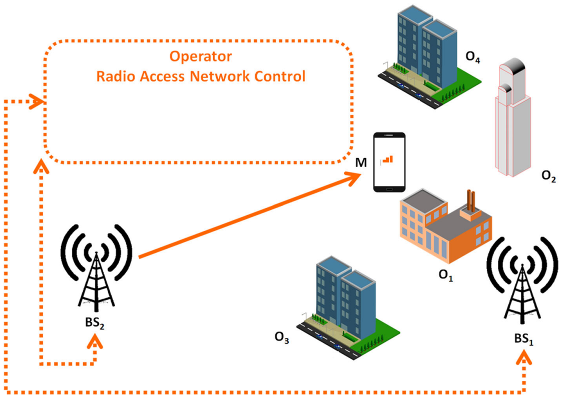
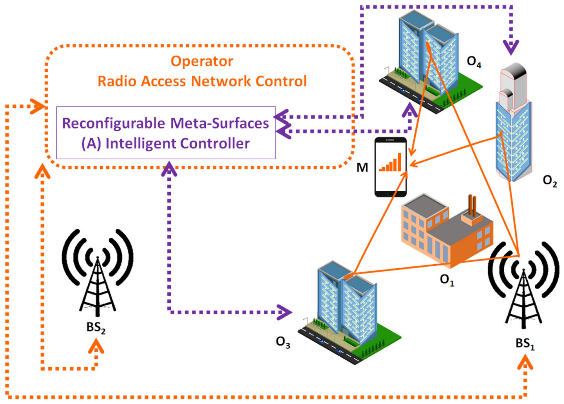


Figure 1.2 – Working principle of RIS [2].

An example of the working principle of a RIS is given in Fig.1.2. An incident radio wave is divided into two parts: the reflected and transmitted (or refracted) radio wave. The reflected and transmitted radio waves can be reconfigured with specific arrangements of the scattering particles on the RIS [17], which is achieved by using PIN diodes in this example. The RIS is capable of shaping the radio waves by the generalized Snell's laws of reflection and refraction [18], which makes the angle of reflection radio waves controllable. Typically, the intelligent reconfiguration is achieved with the aid of sensors embedded at the RIS and a feedback link connected to a CPU controller.



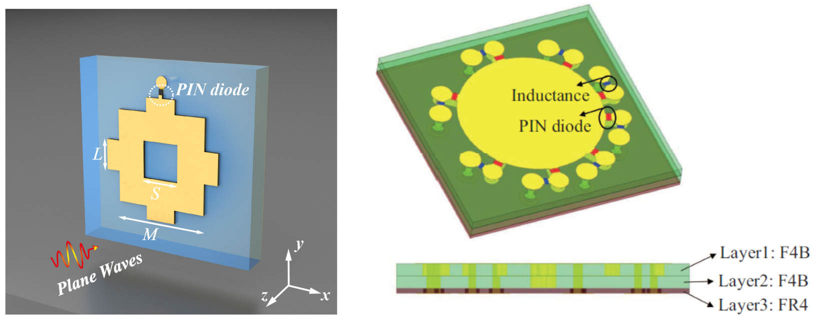
(a) Communications without RIS



(b) Communications with RIS

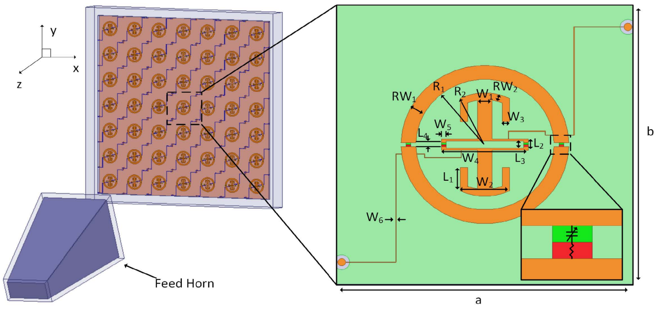
Figure 1.3 – An example of RIS [2].

Fig.1.3 shows a typical application scenario of an RIS. As shown in Fig.1.3 (a), when a user equipment (UE) M wants to connect to the Internet via its nearest base station (BS) BS1, the wireless link is blocked by the object O1. In this case, M has to choose BS2 with a weaker signal because BS2 has a longer distance from M. In Fig.1.3 (b) where RISs are installed on the surfaces of the objects O2, O3, and O4, the signals from BS1 can be transmitted to M by refraction from O2 and reflection from O3 and O4. This is because the responses of RISs are optimized to refract or reflect towards anomalous directions by the generalized Snell's laws, i.e., the refractions and reflections of the RISs are controllable. In future communications, the blockage caused by environmental objects is more severe because high-frequency signals have high penetration losses and reduced diffraction. Therefore, RIS could be a key technology in future wireless communications.

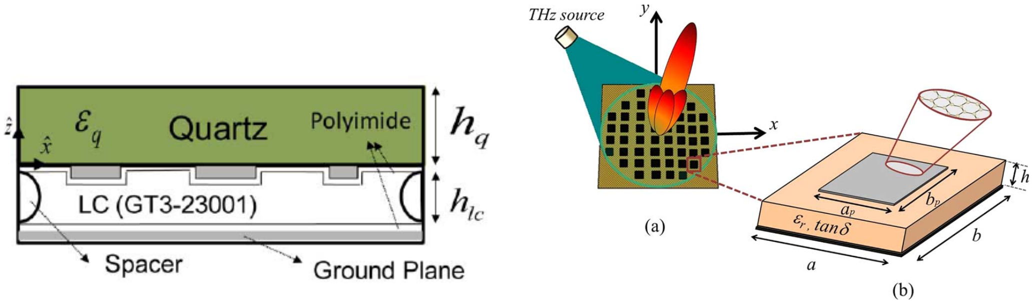


(a) Structure of the RE loaded with a PIN diode proposed in [19].

(b) Structure of the circularly polarized RE loaded with 8 PIN diodes proposed in [3].



(c) Structure of the RE loaded with four varactor diodes proposed in [20].



(d) Structure of the RE loaded with liquid crystal proposed in [21].

(e) Structure of the RE loaded with graphene proposed in [22]: Reflectarray architecture and graphene-based patch

Figure 1.4 – Structures of the REs.

1.2.2 Implementations of RIS

Some researchers are currently working to realize the RISs, different methods have been proposed to reconfigure the reflection coefficients, such as loading PIN diodes [19], varactor diodes [20], microelectromechanical system (MEMS) switches [23], liquid crystal [21], and graphene [22]. In a PIN loaded RE, the PIN can be switched on/off to reverse the reflection phase, a single PIN can achieve 1-bit phase quantization. To achieve a continuous reflection phase shift, a varactor diode was adopted in [20]. With less insertion losses, the MEMS switches outperform the PIN diodes in the millimeter-wave band [23]. Graphene [22] and liquid crystal [21] have been used to design continuously tunable reconfigurable reflectarrays at millimeter and sub-millimeter wavelengths but are not suited to designs at lower frequencies due to high losses [24, 25]. Some structures are displayed in Fig. 1.4.

1.3 Potential applications in 6G networks

To meet the stringent requirements of 6G networks, promising technologies such as Terahertz (THz) communications, short packet communications (SPC), visible light communications (VLC) and mobile edge computing (MEC) have been proposed. Moreover, RIS can be jointly designed with these technologies and brings additional degrees of freedom to enhance communication. With its ability to create a smart radio environment without increasing energy consumption and hardware cost, RIS is a perfect supplement to these technologies to realize 6G. In the following, we list some potential applications of RIS in 6G networks.

1.3.1 Massive connectivity communications

In the future 6G, how to ensure the connectivity considering a large number of Internet-of-Things (IoT) devices is a critical problem. RIS can be employed to solve this problem. In [26], the system optimization of an RIS-aided massive access system is investigated. However, the larger training overhead is still a problem. To address this issue, in [27], the sporadic traffic pattern of machine-type communications was considered for an RIS-aided massive access MISO system, where a matrix factorization based SCCE method was proposed. How to further reduce the training overhead for RIS-aided massive connectivity systems remains an interesting topic in the future.

1.3.2 THz communications

Radio spectrum is always a tightly regulated resource. Recently, THz communications operating between 100 GHz and 10 THz is considered as a key enabler for FeMBB in the next generation of wireless communication. Because high-frequency signals are much likely blocked due to high penetration loss and reduced diffraction effects, RIS is needed THz communications to solve the problem. In [28], a joint hybrid precoding strategy was proposed to maximize the sum-rate for RIS-aided THz MIMO communication systems. Then, the hierarchical search codebook design was proposed in [29] to reduce the complexity of channel estimation and data transmission. In [30], the theoretical error rate of the system was analyzed where the RIS was mounted on a satellite to enable signal propagation in low earth orbit satellite networks with THz communications.

1.3.3 Short packet communications

Short packet communications is regarded as a solution to achieve ERLLC in 6G [31]. Because of double fading cascaded channel and hardware impairments of RIS, system design of RIS-aided systems with SPC is also an interesting topic [32].

1.3.4 Visible light communications

VLC can ensure a large capacity within a limited distance, whose full potential can be realized by employing RIS. RIS is capable of mitigating intensity loss and enhance signal transmission of VLC. In [33], the reflection pattern design of RIS was investigated to extend the communication range of VLC. In [34], the RIS-assisted dualhop VLC system was investigated.

1.3.5 Mobile edge computing

MEC is a promising technology to support low-latency services with high-performance because it can enable the resource-limited devices to partially offload their computation tasks to the nearby computing nodes. However, the potential benefits of MEC systems are limited due to the long-distance offloading link. This problem can be resolved by deploying an RIS to enhance the signal transmission by passive beamforming. RIS-aided MEC systems were considered in [35, 36], where joint design of the active and passive beamformers, RIS deployment, communications, and computing resource allocation of RIS-aided MEC systems is investigated to further enhance the uplink offloading performance and system performance.

1.3.6 Air-ground communications

Sometimes the communication in an urban environments are moving vertical upward, there are more and more aerial devices being developed. In [37], the RIS was deployed into the air-ground communications with multiple aerial/terrestrial users, the RIS placement and the active and passive beamformers are jointly designed. Furthermore, in [38], the aerial user trajectory was further considered in the above system to achieve a higher spectrum and energy efficiency.

1.4 Thesis Overview and Major Contributions

In this thesis, we focus on two main subjects: (i) To study the signal-to-noise ratio (SNR) and interference-to-noise ratio (INR) of RIS-aided MIMO systems with the assumption that the phase shifts of RIS are optimized to maximize the SNR at the receiver. (ii) To investigate the deep learning aided detection for RIS-aided spatial modulation (SM) systems.

The main contributions of this thesis are the following:

i) this thesis investigates the distribution of the SNR as a random variable in RIS-aided MIMO systems. It is proven that SNR is a random variable that is equivalent in distribution to the product of three (or two) independent random variables (chapter3).

ii) this thesis investigates the distribution of the INR as a random variable in RIS-aided MIMO systems. A simplified system is considered where two user equipments (UE) communicate with two BSs separately with the aid of RISs. The INR is proven to be asymptotically equivalent in distribution to the product of a Gamma RV and a Chi-Square RV (chapter4).

iii) this thesis studies the deep learning (DL) detection for RIS-aided SM systems. Based on deep unfolding technique, a model-driven DL detector is proposed, which has a superior bit error rate (BER) performance against other detectors (chapter5).

1.5 Publications

Journal Paper The following is the journal publication produced from the research outcomes of this thesis.

(*J*1) X. Qian, M. Di Renzo, J. Liu, A. Kammoun, and M. S. Alouini,“Beamforming Through Reconfigurable Intelligent Surfaces in Single-User MIMO Systems: SNR Distribution and Scaling Laws in the Presence of Channel Fading and Phase Noise”, *IEEE Wireless Communications Letters*, 2020, 10(1), pp. 77-81.

Abstract: We consider a fading channel in which a multi-antenna transmitter communicates with a multi-antenna receiver through a reconfigurable intelligent surface (RIS) that is made of N reconfigurable passive scatterers impaired by phase noise. The beamforming vector at the transmitter, the combining vector at the receiver, and the phase shifts of the N scatterers are optimized in order to maximize the signal-to-noise-ratio (SNR) at the receiver. By assuming Rayleigh fading (or line-of-sight propagation) on the transmitter-RIS link and Rayleigh fading on the RIS-receiver link, we prove that the SNR is a random variable that is equivalent in distribution to the product of three (or two) independent random variables whose distributions are approximated by two (or one) gamma random variables and the sum of two scaled non-central chi-square random variables. The proposed analytical framework allows us to quantify the robustness of RIS-aided transmission to fading channels. For example, we prove that the amount of fading experienced on the transmitter-RIS-receiver channel linearly decreases with N. This proves that RISs of large size can be effectively employed to make fading less severe and wireless channels more reliable.

Conference Paper The following is the conference paper produced from the research outcomes of this thesis.

($C1$) J. Liu, X. Qian and M. Di Renzo,“Interference Analysis in Reconfigurable Intelligent Surface-Assisted Multiple-Input Multiple-Output Systems”, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 8067-8071.
Abstract: Reconfigurable intelligent surfaces (RISs) are regarded as an emerging technology for the next generation of wireless communications. In this paper, we consider a multiple-input multiple-output network where each base station serves a user equipment with the aid of an RIS equipped with N reconfigurable elements. We characterize the interference at one user equipment that is caused by the signal emitted by its non-serving (interfering) RIS. By assuming Rayleigh fading channels, we study the corresponding interference-to-noise- ratio (INR) under the assumption of large values of N, and we prove that the INR is the product of a Chi-Square random variable (RV) and an RV that is approximated with a Gamma distribution. In addition, we prove that the amount of fading of the INR is equal to one in the large N regime.

($C2$) J. Liu, and M. Di Renzo,“Data-driven and model-driven deep learning detection for RIS-aided Spatial Modulation”, *IEEE 5G World Forum 2021*, accepted.
Abstract: Reconfigurable intelligent surface (RIS) is regarded as a key technology for the next generation of wireless communications. Recently, the combination of RIS and spatial modulation (SM) or space shift keying (SSK) has attracted a lot of interest in the wireless communication area by achieving a trade-off between spectral and energy efficiency. In this paper, by generalizing RIS-aided SM/SSK system to a special case of conventional SM system, we investigated deep learning based detection in RIS-aided SM/SSK systems. Based on the idea of deep unfolding, we studied the model-driven deep learning detection for RIS-aided SM systems and compare the performance against the data-driven deep learning detectors.

2 Modeling and Optimization of Reconfigurable Intelligent Surface - An Introduction

Chapter 2. Modeling and Optimization of Reconfigurable Intelligent Surface - An Introduction

In this chapter, the modeling and optimization of RIS-aided systems are introduced. This chapter is organized as follows. In *Section 2.1*, we introduce the system models of RIS-aided MIMO systems and then investigate the reflection principle of RISs. In *Section 2.2*, the fundamental channel estimation problem of RIS-aided systems are introduced and the state-of-the-art methods are overviewed. In *Section 2.3* and *Section 2.4*, we review the system design problem of the RIS-aided wireless communications and RIS based information transmission systems, respectively.

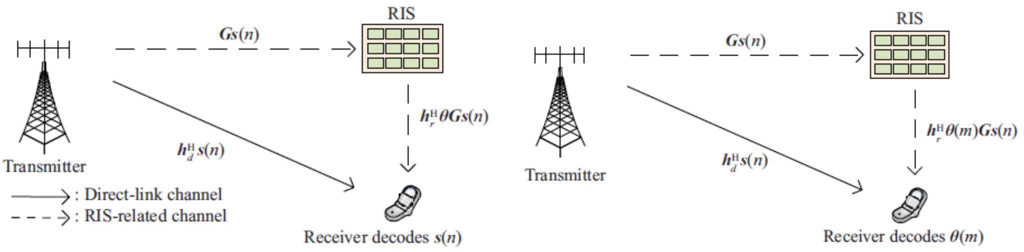
2.1 RIS-aided communication systems

In this section, first the basic system models of two communication paradigms with RISs are introduced, then we investigate the reflection principle of RISs.

2.1.1 System models

As a nearly-passive device, an RIS needs to leverage the existing active radio waves to operate its function. Based on the purpose of the RIS reflection, communication models with RISs can be classified as two categories, namely RIS-aided wireless communications (RAWC) and RIS based information transmission (RBIT) [3]. In RAWC, the RIS adjusts its reflection coefficient matrix θ according to the CSI to assist the transmitter, thus providing additional channel diversity. In RBIT, the RIS delivers its own message to the receiver by proactively changing its reflection coefficient matrix $\theta(\mathbf{m})$.

As shown in Fig. 2.1, the transmitter-receiver, transmitter-RIS, and RIS-receiver channels are denoted by $\mathbf{h}_d^H$, $\mathbf{G}$ and $\mathbf{h}_r^H$ (the superscript H denotes the conjugate transpose operation), respectively. The information message $s(n)$ is generated at the transmitter and sent to the receiver via the direct-link channel and the RIS-related channel, and the RIS varies its reflection coefficient matrix via the RIS-related channel. In Fig. 2.1(a), the messages are only embedded in $s(n)$. In Fig. 2.1(b), the messages are embedded in $s(n)$ or $\theta(\mathbf{m})$, or both.



(a) RAWC (b) RBIT

Figure 2.1 – Two communication paradigms of RIS [3].

2.1.2 Reflection principle

It is obvious that RIS-based communications rely on the design of the reflection coefficients, which is a critical issue in practical hardware implementations. To investigate the reflection from the electromagnetic perspective, we need to study the permittivity and permeability of the reflective elements (REs) and solve Maxwell's equations [39, 40, 41], which is a difficult task. Considering that the physical size of an RE is smaller than the wavelength of the incident signal, an adequate simplification of Maxwell's equations is transmission line theory [7].

The characteristic of reconfigurability lies in that the reflection coefficient of each RE can be adjusted independently by a designed control signal. The reflection coefficient, by its definition, is a parameter that describes how much of a wave is reflected by an impedance discontinuity in the transmission medium.

Denote the n-th RE as U_n, its reflection coefficient can be expressed as

$$v(n) = \beta(n)\, e^{j\theta(n)} \tag{2.1}$$

where $\beta(n)$ and $\theta(n)$ represent the controllable amplitude and phase shift of U_n, respectively. Let $\tilde{E}(n)$ and $E(n)$ be the electromagnetic (EM) wave impinging on U_n, the EM wave reflected from U_n, separately. By its definition, the reflection coefficient $v(n)$ can be expressed as

$$v(n) = \frac{\tilde{E}(n)}{E(n)} \tag{2.2}$$

Furthermore, let Z_0 and $Z(n)$ denote the impedance towards the source in the air and the equivalent load impedance of U_n, the reflection coefficient $v(n)$ can be written as [42]

$$v(n) = \frac{Z(n) - Z_0}{Z(n) + Z_0} \tag{2.3}$$

Therefore, the amplitude and the phase of the reflection coefficient $v(n)$ can be written as

$$\beta(n) = \left| \frac{Z(n) - Z_0}{Z(n) + Z_0} \right| \tag{2.4}$$

and

$$\theta(n) = \arctan\left(\frac{\mathrm{Im}\left(\frac{Z(n)-Z_0}{Z(n)+Z_0}\right)}{\mathrm{Re}\left(\frac{Z(n)-Z_0}{Z(n)+Z_0}\right)} \right) \tag{2.5}$$

Most works on RIS are mainly based on the ideal model where the phase shifts and amplitudes can be adjusted independently. However, this is not practical due to hardware constraints. In [4], the equivalent model for a reflecting element is investigated. As shown in Fig. 2.2, with L_1, L_2, $C(n)$, $R(n)$ and ω denoting the bottom/inner layer inductance, top/outer layer inductance, effective capacitance, effective resistance, and angular frequency of the incident signal, respectively, the impedance of the n-th element is a function of C_n and R_n and is given by

$$Z_n(C_n, R_n) = \frac{j\omega L_1 \left(j\omega L_2 + \frac{1}{j\omega C_n} + R_n \right)}{j\omega L_1 + j\omega L_2 + \frac{1}{j\omega C_n} + R_n} \tag{2.6}$$

From (2.4) and (2.5), theoretically the amplitude and the phase of the reflection coefficient can be carefully tuned by changing the $Z(n)$ using an external control signal. From (2.6), it is observed that $Z(n)$ is a function of the effective resistance R_n and the effective capacitance C_n, where R_n determines the amount of power dissipation due to the losses in the devices and C_n denotes the charge accumulation related to the semiconductor device. The behavior of the experimental amplitude and the phase shifts of the reflection coefficient v_n is given in Fig. 2.3. It is observed that the phase shifts and amplitudes are strongly coupled and the minimum amplitude occurs near zero phase shift and approaches unity (the maximum value) at the phase shift of π or $-\pi$.

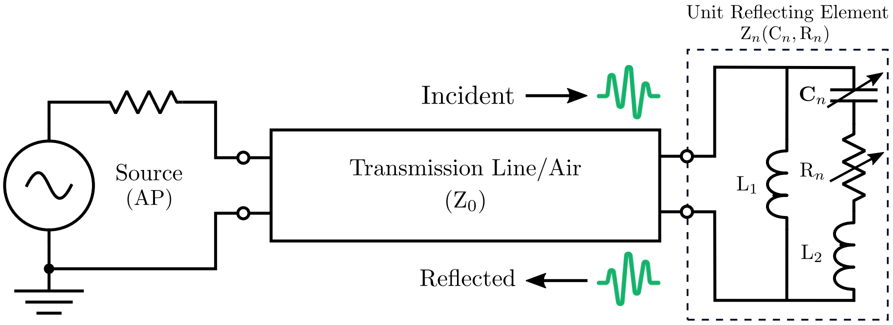
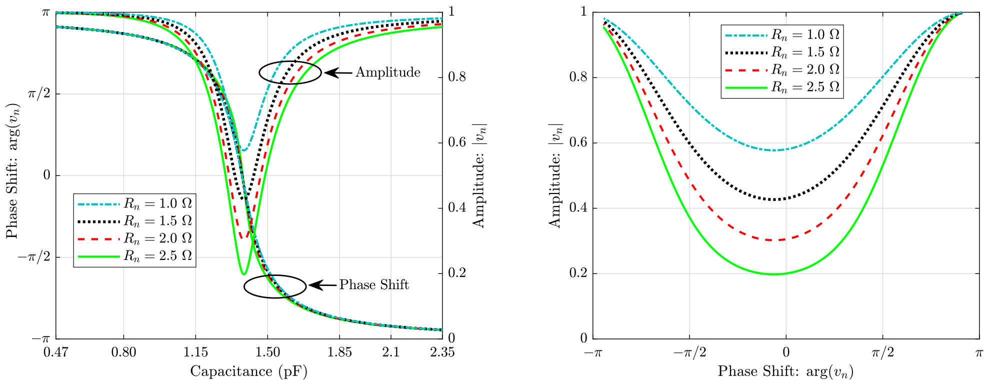


Figure 2.2 – Transmission line model of an RE [4].



(a) Phase shift and amplitude versus the effective capacitance C_n and the effective resistance R_n.

(b) Amplitude versus phase shifts.

Figure 2.3 – Reflection coefficient of an RE [4].

To obtain the ideal phase shift control, where $\beta_n = 1$ and $\theta(n) \in [-\pi, \pi)$, each element should dissipate zero energy dissipation for reflection. However, this is impossible in practical hardware. An analytical model for the phase shift which is in general applicable to a variety of semiconductor devices used for implementing IRS is proposed in [4]. The amplitude is a function of the phase shifts, which is given as

$$\beta_n(\theta_n) = (1 - \beta_{\min}) \left(\frac{\sin(\theta_n - \phi) + 1}{2} \right)^{\alpha} + \beta_{\min} \tag{2.7}$$

where $\beta_{\min} \geq 0$, $\phi \geq 0$ and $\alpha \geq 0$ are the constants related to specific circuit implementation. For more information, the reader can refer to [4].

2.2 Channel estimation for RIS-assisted systems

In this section, we introduce the channel estimation problem of RIS-assisted systems. In RIS-aided communication systems, CSI is usually needed to achieve the joint design of active beamforming at the transceiver and passive beamforming (reflection coefficient matrix) at the RIS. However, CSI acquisition is one of the most challenging tasks for RIS-assisted systems. This is because that the RIS is a nearly-passive device equipped with a large number of REs. As a passive device, we cannot perform the traditional CSI acquisition methods as in active device based systems. Consequently, it is difficult to estimate the transmitter-RIS channel and RIS-receiver channel separately because they are coupled together with the reflection coefficient matrix. Besides, the large number of REs cause a high training overhead.

For example, we consider an RIS-assisted multi-user multiple input single output (MISO) system where the base station (BS) is equipped with M antennas to serve K single-antenna users and the RIS is equipped with L REs. Denote the channel links from the BS to the RIS, from the RIS to the k-th user, and from the BS to the k-th user by $\mathbf{G} \in \mathscr{C}^{L\times M}$, $\mathbf{h}_{r,k} \in \mathscr{C}^{L}$ and $\mathbf{h}_{d,k} \in \mathscr{C}^{M}$, respectively. Then the combined channel from the BS to the k-th user is given by

$$\mathbf{h}_{d,k}^{H} + \mathbf{h}_{r,k}^{H}\text{diag}\left(\mathbf{v}\right)\mathbf{G} = \mathbf{h}_{d,k}^{H} + \mathbf{v}^{T}\text{diag}\left(\mathbf{h}_{r,k}^{H}\right)\mathbf{G} \tag{2.8}$$

where $\mathbf{v}= \left[v\left(1\right), ..., v\left(L\right)\right]$ is the reflection coefficients to be designed whose element is defined in (2.1). The CSI of the direct-link $\mathbf{h}_{d,k}^{H}$ can be estimated by using traditional multiple input multiple output (MIMO) channel estimation methods. From the right hand side of the above expression, it is observed that the design of the passive reflection coefficients $\mathbf{v}$ does not require the exact channel links of $\mathbf{h}_{r,k}$ and $\mathbf{G}$, but only requires the CSI of the direct-link $\mathbf{h}_{d,k}^{H}$ and the cascaded channel:

$$\mathbf{H}_{k} = \text{diag}\left(\mathbf{h}_{r,k}^{H}\right)\mathbf{G}. \tag{2.9}$$

To estimate the cascaded channel $\mathbf{H}_{k}$, many methods have been proposed, which can be classified into two categories: the separate cascaded channel estimation (SCCE) and the direct cascaded channel estimation (DCCE). In SCCE, the BS-RIS channel and the RIS-user channel are estimated respectively, then the cascaded channel is reconstructed. In DCCE, the cascaded channel is estimated directly with a properly designed channel estimation protocol.

2.2.1 Separate cascaded channel estimation

Depending on whether there exists active sensors on the RIS, RISs can be classified as semi-passive RISs and full-passive RISs. In this section, we introduce the separate cascaded channel estimation for these two cases separately.

SCCE for semi-passive RIS

To estimate the RIS-related channel, a natural method is to insert some active sensors among the passive REs at the RIS [43], which is defined as semi-passive RIS in this context. In [43], it is assumed that the BS-RIS channel and RIS-user channel consist of a few physical transmission paths. Using the channel sparsity, only a few training signals are needed to recover the complete CSI. However, the semi-passive RIS scheme in [43] only works on the correlated channel and the active sensors increase the hardware/energy cost and signal processing complexity. Besides, the feedback link of the CSI decreases the transmission efficiency.

SCCE for full-passive RIS

When there is no active sensors at the RIS, we define the RIS as full-passive RIS in this context. In [44, 45, 46, 47, 48, 49], SCCE methods are investigated for full-passive RIS-aided systems without active sensors. An RIS-aided single-user MIMO system is considered in [44, 45]. Specifically, in [44], in the first stage the BS-RIS channel was estimated by iteratively solving a fixed point equation, in the second stage the RIS-user channel was obtained by least-square (LS) estimation. In [45], the rank-deficient property of the BS-RIS and RIS-user channel matrix is utilized to reduce the training overhead. Particularly, in the first stage, the bi-linear generalized approximate message passing (BiG-AMP) algorithm is adopted to solve the matrix factorization problem, the BS-RIS channel and the product-matrix of the user-RIS channel and the training pilots were recovered. In the second stage, a Riemannian manifold gradient-based algorithm is introduced to solve the matrix completion problem, where the RIS-user channel was recovered.

Under the assumption of sparse, low-rank, or quasi-static channels, the SCCE methods in [46, 47, 48, 49] considered the channel estimation problems for RIS-aided multi-user MIMO systems. Specifically, in [46] a matrix-calibration based matrix factorization problem was formulated under the assumption of channel sparsity, and an iterative Bayesian inference algorithm was derived to estimate BS-RIS channel and RIS-user channels, respectively. In [47], the anchor-assisted two-phase channel estimation scheme was developed, where two anchor nodes were deployed near the RIS to help the channel estimation. Particularly, in the first phase, the two deployed anchor nodes transmit pilots successively to estimate the square of the element-wise BS-RIS channel gain. In the second phase, the obtained partial knowledge of the BS-RIS channel is further used to estimate RIS-user channels. The anchor-assisted channel estimation method is efficient for RIS-assisted massive access systems since such partial knowledge is common to all users. In [48], the parallel factor decomposition based scheme was investigated, where the parallel factor was applied to factorize the high dimensional tensor into a linear combination of multiple rank-one tensors and the alternating least squares (LS) was used to recover the unknown channels from the decomposed tensors.

2.2.2 Direct cascaded channel estimation

In this subsection, we introduce the DCCE method, which estimates the cascaded channels directly with a properly designed channel estimation protocol.

Binary reflection based DCCE

Based on sequentially activating only one RIS element for each pilot symbol, the binary reflection method was proposed in [50] for an RIS-assisted single-user system. An extra training symbol is necessary to estimate the direct BS-user channel link with all the RIS elements being deactivated. In practice, to implement the on/off switching of the massive RIS elements is costly. Besides, as only a small portion of its elements is switched on at each time, the channel estimation accuracy is degraded. To address this issue, a sub-group based binary reflection was further investigated in [51] for an RIS-assisted orthogonal frequency division multiplexing (OFDM) systems based on spatial correlation. Instead of controlling the on/off states of a single element each time, the authors applied the on/off method on the grouped RIS elements. In this scheme, REs are divided into multi sub-groups and each subgroup consisting of adjacent elements shares a common reflection coefficient.

Full reflection based DCCE

To tackle drawbacks of large training overhead and low reflected signal power in binary reflection based method, a full reflection based DCCE method was proposed in [52] where all REs are active during the whole training timeslots.The training reflection coefficients are designed based on the minimum variance unbiased estimation principle, which mimics a series of discrete Fourier transforms (DFT). In [53, 54], the method in [52] is considered together with the grouping idea in [51]. Another full reflection based DCCE method is based on deep learning, in [55, 56] a deep learning based denoising convolutional neural network was proposed based on the minimum mean-squared-error (MMSE) criterion.

Multi-user joint DCCE

In RIS-aided multi-user MIMO systems, the channel correlations can be utilized to obtain the CSI. In [57, 58, 59], the cascaded channel of an arbitrary user is used to represent the cascaded channels of the remaining users . Denote the cascaded channel of the k-th user as $\mathbf{H}_k$, we have

$$\mathbf{H}_k = \underbrace{\mathrm{diag}\left(\mathbf{h}_{r,k}^H\right)\mathrm{diag}\left(\mathbf{h}_{r,1}^H\right)^{-1}}_{\Omega_k}\underbrace{\mathrm{diag}\left(\mathbf{h}_{r,1}^H\right)\mathbf{G}}_{\mathbf{H}_1} = \Omega_k\mathbf{H}_1 \tag{2.10}$$

where Ω_k is a diagonal matrix. As can be seen from the equation above, we only need to estimate the diagonal matrix Ω_k if the common part $\mathbf{H}_1$ is given. Since the number of unknown parameters in Ω_k is much smaller than that in $\mathbf{H}_k$, the estimation of Ω_k is much easier than that of $\mathbf{H}_k$ and the training overhead can be reduced. In [57, 58], a sequential estimation method using the above channel correlation was proposed for a narrow-band communication system and a wide-band communication system, respectively.

2.3 System design of RAWC systems

In this section, the system design of the joint active beamforming at the transceiver and passive beamforming (reflection coefficient matrix) at the RIS in RAWC is overviewed.

2.3.1 System design under ideal case

In general, the optimization problems of the joint active and passive beamforming are non-convex and difficult to solve. In this subsection, we consider the system design under ideal case with perfect CSI and continuous phase shifts.

Reflection coefficient based optimization

This reflection coefficient based optimization (RCO) method is motivated by the fact that the optimal active beamformer at the transceiver has a closed-form solution when the passive beamformer (reflection coefficient matrix) is fixed. By substituting this closed-form solution into the original problem, the reformulated problem only involves the passive beamformer and is relatively easy to solve. Specifically, in [60], the RCO method was adopted to solve the transmit power minimization problem in a single-user MISO system, where the passive beamformer is then obtained using semidefinite relaxation (SDR) technique. In [61], the RCO method was extended to solve a minimum signal-to-interference-plus-noise ratio (SINR) maximization problem in a multi-user MISO system. However, the closed-form solution of the active beamformer required in the RCO method cannot be generally achieved in a complicated wireless communications system.

Alternating optimization method

To tackle the limitation in the RCO method, alternating optimization (AO) method is often adopted as an alternative to solve the formulated non-convex problem, where the active and passive beamformers are alternatively optimized in a iterative fashion. The AO method is widely used in the literature to solve the spectrum/energy efficiency maximization for various systems, e.g., multi-user MISO systems [62, 63], non-orthogonal multiple access (NOMA) systems [64, 65], cognitive radio (CR) system [66], physical layer security [67, 68, 69, 70, 71, 72, 73], active RIS-aided networks [74], wireless powered communication networks [75], and simultaneous wireless information and power transfer (SWIPT) systems [76]. The problem of AO method is that it usually requires a large number of iterations to guarantee convergence, and no closed-form expressions are given.

Machine learning

To accelerate the convergence, in [77, 78] the projected gradient method was proposed to solve the capacity maximization problem. This method optimizes all variables in each iteration, resulting in a faster convergence and lower computational complexity than those in the AO method. Besides, deep reinforcement learning (DRL) based method was adopted to solve the joint active and passive beamformer design problems in [79, 80, 81, 82].

2.3.2 System design under hardware constraints

In this subsection, we review the system design for the RAWC under practical hardware constraints, i.e., discrete phase shifts, phase-dependent amplitude variation, and transmit/receive signal distortion.

Discrete phase shifts

Continuous phase shift is an ideal assumption in literature normally for simplification. In practical, the discrete phase shifts will cause a degradation in performance [83]. However, system design considering discrete phase shifts is a mixed-integer non-linear program problem, which is NP hard and difficult. In [84], the branch-and-bound method and AO method were adopted to minimize the total transmit power by considering discrete phase shift for a single-user case and multi-user case, respectively. In [85], a new penalized Dinkelbach block successive upper-bound minimization (BSUM) method was proposed to solve the rate maximization problem for an RIS-aided single-user MISO system.

Phase-dependent amplitude variation

As introduced in previous section 2.1.2, in practice the amplitude of the reflection coefficient is a function of the phase shifts of the incident wave. The relation is given in (2.7). To take this into consideration, in [86], the RCO method and the AO method are adopted to solve the total transmit power minimization problem for single-user and multi-user MISO cases, respectively.

Transmit/receive signal distortion

The transmit/receive signal distortion due to the inevitable hardware impairments of amplifiers, oscillators, digital-to-analog converters (DACs), and analog-to-digital converters (ADCs) is another cause of performance degradation. In [87], the authors studied the received signal-to-noise ratio (SNR) maximization problem considering transmit/receive signal distortion, the RCO method and the minorization-maximization algorithm are adopted.

2.3.3 System design under statistical or imperfect CSI

In this part, we consider the system design for the RAWC under statistical or imperfect CSI. Because of the large training overhead of channel estimation, some literature considered statistical CSI for system design [88, 89, 90]. In practice, the estimated CSI is usually imperfect due to the inherent noise and limited training overhead, which will lead to certain performance degradation. The robust joint active and passive beamforming design was studied in [91] for a multi-user MISO system and in [92] for a multi-user MISO CR system. The AO method was applied to deal with the coupled active and passive beamformer and S-procedure was proposed to handle the non-convex terms stemming from the channel uncertainty. In [93], the tradeoff between the training overhead and the energy efficiency was investigated for a single-user MIMO system with a joint robust active and passive beamforming scheme. In [94, 95], the robust joint active and passive beamforming design was studied to minimize total transmission power subject to the individual outage probability constraint, where S-procedure and constrained stochastic successive convex approximation algorithm were applied to tackle the non-convex terms in the outage probability. In [96], the statistical information such as the distribution of the locations of the users and the distribution of the multi-path channels was considered, a two-phase optimization algorithm was proposed for a sum-rate maximization problem.

2.4 System design of RBIT systems

In RBIT, RIS acts as an information transmitter which modulates information message in the reflection coefficients of the REs [97, 98]. The merits of RBIT systems lie in that: (1) RBIT is achieved via passive reflection, which is energy and cost effective; (2) The large number of REs on the RIS can provide considerable spatial multiplexing or substantial diversity. In this section, we investigated the RBIT systems with the ambient RF radio wave with modulated or unmodulated signals.

2.4.1 RBIT with unmodulated signals

As shown in Fig. 2.4, the message transmitted by the RIS is denoted as w, which is transmitted to the receiver leveraging the RF carrier emitted from the active transmitter. Different from backscatter communications which only has one single RE, the RBIT system is able to serve multiple users simultaneously. In this case, the critical design problem is the mapping between the message symbol and the reflecting pattern.

With the phase-shift ability, it is easy for the RIS to achieve the M-ary phase shift-keying (PSK) modulation [99]. In [100], a prototype of RIS-based transmitter with 8 × 32 REs and 8-PSK

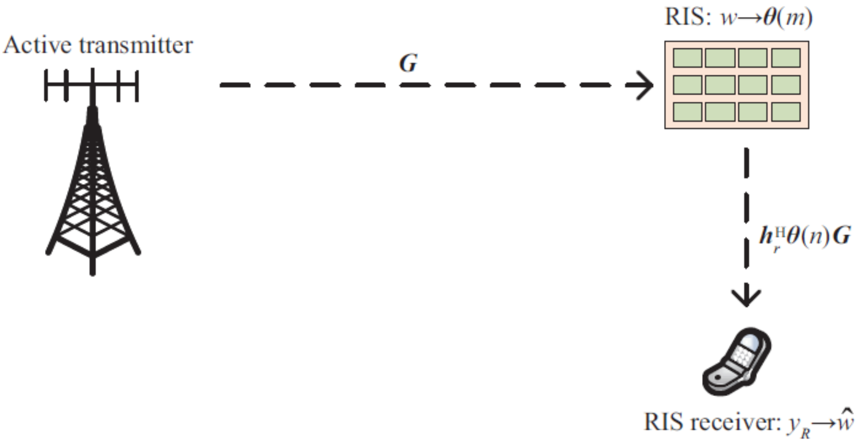


Figure 2.4 – RBIT with unmodulated signals: the active transmitter provides the RIS with the unmodulated signals as the RF carrier [3].

was proposed. In [101], combined with the concept of index modulation, RIS-aided space shift keying (SSK) system was proposed. Besides, quadrature amplitude modulation (QAM) modulation was investigated in [102, 103].

2.4.2 RBIT with modulated signals

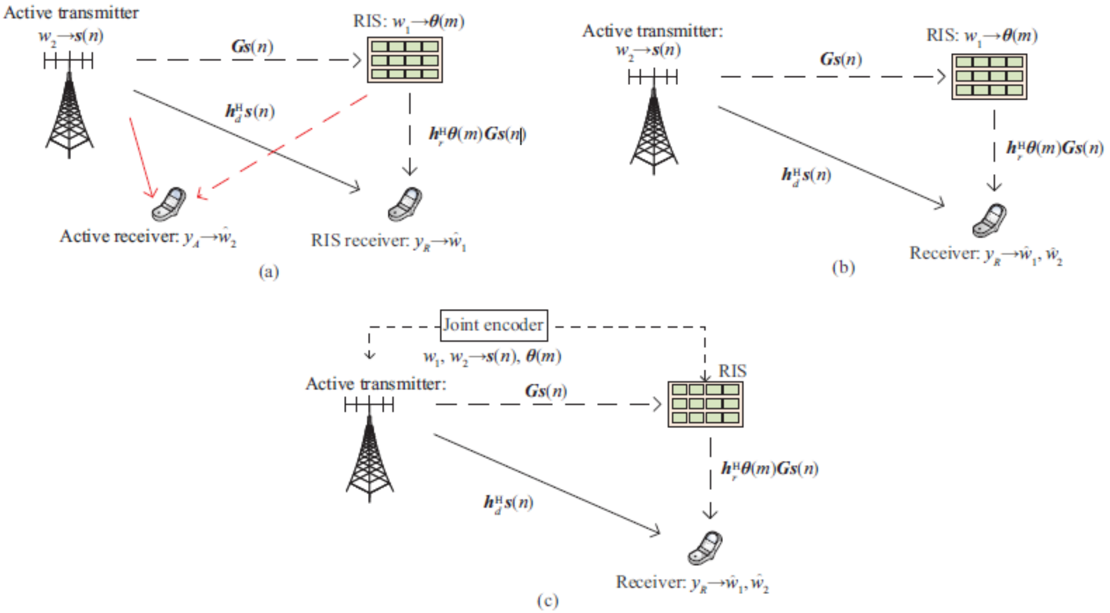


Figure 2.5 – RBIT with modulated signals: (a) No collaboration: mutual interference. (b) Symbiotic radio with joint decoding. (c) Symbiotic radio with joint encoding and decoding.[3].

Denote the message modulated in the reflection matrix of the RIS as w_1, the message embedded in the active transmitter as w_2, respectively. As shown in Fig. 2.5, the reflected signal is embedded with messages w_1 and w_2. If there exists a direct link, the receiver can receive the signal directly from the active transmitter, which can be an extra interference to the receiver if the receiver is not well designed. Collaboration between the RIS-based and active transmissions is needed to solve this problem.

If there is no collaboration between the RIS-based transmission and active transmission, separate receivers are needed, as shown in Fig. 2.5 (a). In this case, these two sub systems will cause interference to each other. To solve this problem, the authors in [104] exploited the large degree of freedoms in RIS to let a slave antenna only receive the direct-link signal, and thus by normalizing the signals at the other antenna, the receiver can cancel the unknown source symbols. Nevertheless, the modulated carrier signal is still not so ideal that the RIS-based transmission without any collaboration undergoes unsatisfactory achievement.

In [105], a technique called symbiotic radio (SR), is proposed to exploit the benefits and address the drawbacks of cognitive radio (CR) and ambient backscattering communications (AmBC), leading to mutualism spectrum sharing and highly reliable backscattering communications. The mutual interference in Fig. 2.5 (a) can be avoided via forming the SR between the RIS-based transmission and the active transmission [105, 106, 107]. As shown in Fig. 2.5 (b), the active transmitter and RIS-based transmitter send the separated messages to the same receiver. The RIS receiver is designed to have the prior knowledge on the codebooks of the two transmissions and is thus able to adopt the joint decoding to decode the message w_1 and w_2 at the same time. In [108, 109], the collaboration is further extended to joint coding and joint decoding, as shown in Fig. 2.5 (c), which exhibits better overall BER and rate performances than other aforementioned schemes.

2.5 Conclusion

In this chapter, the RIS-aided wireless communication and RIS based information transmission systems are introduced in detail, respectively. The state-of-the-art research on the system design is overviewed. With the ability to achieve a reconfigurable and controllable environment, RIS is a promising technique in the future communications.

3 SNR Analysis in RIS-assisted MIMO Systems

In this chapter, we consider a fading channel in which a multi-antenna transmitter communicates with a multi-antenna receiver through a reconfigurable intelligent surface (RIS) that is made of N reconfigurable passive scatterers impaired by phase noise. The beamforming vector at the transmitter, the combining vector at the receiver, and the phase shifts of the N scatterers are optimized in order to maximize the signal-to-noise-ratio (SNR) at the receiver. By assuming Rayleigh fading (or line-of-sight propagation) on the transmitter-RIS link and Rayleigh fading on the RIS-receiver link, we prove that the SNR is a random variable that is equivalent in distribution to the product of three (or two) independent random variables whose distributions are approximated by two (or one) gamma random variables and the sum of two scaled non-central chi-square random variables. The proposed analytical framework allows us to quantify the robustness of RIS-aided transmission to fading channels. For example, we prove that the amount of fading experienced on the transmitter-RIS-receiver channel linearly decreases with N. This proves that RISs of large size can be effectively employed to make fading less severe and wireless channels more reliable.

This chapter is organized as follows. In *Section 3.1*, we introduce the background of the analysis of SNR in RIS-aided MIMO system. The system model is introduced in *Section 3.2*. Some lemmas are summarized in *Section 3.3* as preliminaries for the analysis in *Section 3.4*. Simulation results are provided in *Section 3.5* followed by the conclusion in *Section 3.6*.

3.1 Introduction

Reconfigurable intelligent surfaces (RISs) are an emerging transmission technology for application to wireless communications [2]. RISs can be realized in different ways, which include (i) implementations based on large arrays of inexpensive antennas that are usually spaced half of the wavelength apart; and (ii) metamaterial-based planar or conformal large surfaces whose scattering elements have sizes and inter-distances much smaller than the wavelength [7]. In this thesis, we consider RISs made of scatterers that are passive, are spaced half of the wavelength apart, and are individually configured and optimized for realizing passive beamforming through the environment [110], [111]. Compared with other transmission technologies, e.g., phased arrays, multi-antenna transmitters, and relays, RISs require the largest number of scattering elements, but each of them needs to be backed by the fewest and least costly components. Also, no power amplifiers are usually needed. For these reasons, RISs constitute an emerging and promising software-defined architecture that can be realized at reduced cost, size, weight, and power (C-SWaP design) [112], [6].

Quantifying the performance of optimized RIS-empowered multi-antenna wireless systems is an open research issue. In particular, several researchers have developed algorithms for jointly optimizing the beamforming vector ($\mathbf{q}$) at the transmitter, the matrix of phase shifts at the RIS (Φ), and the combining vector ($\mathbf{w}$) at the receiver [7, Sec. V-J]. In general, however, the optimal triplet ($\mathbf{q}, \Phi, \mathbf{w}$) cannot be formulated in closed-form and can only be computed numerically. An exception is constituted by wireless systems in which the transmitter and receiver are equipped with a single antenna. For this reason, currently available analytical frameworks and scaling laws are only applicable to single-antenna transmitters and receivers. Representative contributions include [60]-[113]. In [60] and [84], in particular, the authors show that the average SNR at the receiver scales with the square of the number of tunable elements (N) of the RIS. In [99], the authors study the error probability over Rayleigh fading channels by using the central limit theorem. In [114] and [115], the authors quantify the impact of phase noise for transmission over Rayleigh and Rician fading channels, respectively. In [113], the authors analyze the impact of phase noise and hardware impairments for transmission over line-of-sight (LOS) channels.

Motivated by these considerations, we consider a fading channel in which a multi-antenna transmitter communicates with a multi-antenna receiver through an RIS whose N scattering elements are impaired by phase noise. We introduce an analytical approach for characterizing the distribution of the SNR and for determining its scaling laws as a function of N. Over Rayleigh fading or LOS channels, we prove that the SNR can be formulated, for any phase noise distribution, as the product of gamma and scaled non-central chi-square random variables. With the aid of numerical simulations, in addition, we show that the SNR can be well approximated with a gamma random variable whose parameters are formulated in closed-form. The proposed approach unveils the scaling laws of the mean, the variance, and the amount of fading (AF) [116] of the SNR as a function of N. Our analysis confirms that RISs of large size can be effectively employed to make the transmission of information over fading channels more reliable.

3.2 System model

We consider a point-to-point wireless system in which a transmitter equipped with N_T antennas and a receiver equipped with N_R antennas communicate through an RIS. The RIS is made of N antenna elements that are spaced half-wavelength apart and that apply independent phase shifts to the incident signal. The phase shift applied by the nth element is denoted by ϕ_n for $n = 1, 2, \ldots, N$. For ease of notation, the N phase shifts are collected in the $N \times N$ diagonal matrix Φ. The nth phase shift is assumed to be subject to phase noise, e.g., due to the finite resolution of the phase shifts or to phase estimation errors. The phase noise is assumed to be independent among the N phase shifts. We define $\phi_n = \phi_n^{(\text{opt})} + \delta_n$, where $\phi_n^{(\text{opt})}$ is the optimal phase shift in the absence of phase noise and δ_n is the phase noise. The distribution of δ_n is arbitrary but its mean is assumed to be zero. Examples of phase noise distributions are given in Section 3.4. The $N_T \times 1$ unit-norm beamforming vector at the transmitter is denoted by $\mathbf{q}$ and the $N_R \times 1$ unit-norm combining vector at the receiver is denoted by $\mathbf{w}$. The triplet $(\mathbf{q}, \Phi, \mathbf{w})$ is jointly optimized to maximize the receive SNR. As detailed in Section 3.3, we assume that the RIS operates in the far-field regime. Hence, N can be large but cannot tend to infinity [7, Sec. IV-D].

3.2.1 Channel model

The $N_R \times N$ channel matrix on the RIS-receiver link is denoted by $\mathbf{G}$. Its entries are assumed to be independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, i.e., $\mathbf{G} \sim \mathcal{CN}\left(\mathbf{0}_{N_R \times N}, \mathbf{I}_{N_R \times N}\right)$, i.e., Rayleigh fading is considered. This assumption is motivated by the mobility of the receiver and, hence, the difficulty of establishing an LOS link. The $N \times N_T$ channel matrix on the transmitter-RIS link is denoted by $\mathbf{H}$. Two canonical case studies are considered for this link.

Rayleigh fading

$\mathbf{H} \sim \mathcal{CN}\left(\mathbf{0}_{N \times N_T}, \mathbf{I}_{N \times N_T}\right)$, similar to the RIS-receiver link. This setup is representative of scenarios in which the RISs are randomly deployed, e.g., on spatial blockages whose locations are not under the control of the system designer, and, thus, the locations of the RISs cannot be optimized [117]. Thus, LOS propagation cannot be ensured.

Deterministic LOS

$\mathbf{H} = \sqrt{N_T N_R}\mathbf{a}_{\mathrm{RIS}}\mathbf{a}_T^H$, where $\mathbf{a}_T$ and $\mathbf{a}_{\mathrm{RIS}}$ are the array responses of the transmitter and RIS, respectively. $\mathbf{a}_T$ is an $N_T \times 1$ unit-norm (i.e., $\|\mathbf{a}_T\|^2 = 1$) vector whose generic entry is $\mathbf{a}_T(t) = \left(1/\sqrt{N_T}\right)\exp\left(-j2\pi f(t)\right)$, where $f(t)$ depends on the geometry of the transmit array, and $\mathbf{a}_{\mathrm{RIS}}$ is an $N\times 1$ unit-norm (i.e., $\|\mathbf{a}_{\mathrm{RIS}}\|^2 = 1$) vector whose generic entry is $\mathbf{a}_{\mathrm{RIS}}(t) = \left(1/\sqrt{N}\right)\exp\left(-j2\pi g(n)\right)$, where $g(n)$ depends on the geometry of the RIS. Explicit expressions for $\mathbf{a}_T$ and $\mathbf{a}_{\mathrm{RIS}}$ can be found in [113, Eq. (2)]. This setup is representative of scenarios in which the locations of the RISs can be appropriately optimized for ensuring an LOS link [111].

For ease of notation, the case study in which Rayleigh fading is assumed on both links is denoted by the subscript "RR", the setup in which LOS and Rayleigh fading are assumed on the first and second link, respectively, is denoted by the subscript "LR". The analysis of other channel models is left to future research.

3.3 Problem formulation

In [93], the authors have recently proved that, by jointly optimizing the triplet $(\mathbf{q}, \Phi, \mathbf{w})$, the SNR in single-user multi-antenna systems can be tightly approximated as follows:

$$\begin{aligned} &\mathrm{SNR}_{\mathrm{RR}} \approx \gamma_0 N^2 \max\nolimits_{l,k}\left\{\lambda_{l,\mathbf{G}}\lambda_{k,\mathbf{H}}\left|\Upsilon_{l,k}\right|^2\right\} \\ &\Upsilon_{l,k} = \sum\nolimits_{n=1}^{N}\left|\mathbf{v}_{l,\mathbf{G}}(n)\right|\left|\mathbf{u}_{k,\mathbf{H}}(n)\right|\exp\left(j\delta_n\right) \end{aligned} \tag{3.1}$$

$$\begin{aligned} &\mathrm{SNR}_{\mathrm{LR}} \approx \gamma_0 N_T N \max\nolimits_{l}\left\{\lambda_{l,\mathbf{G}}|\Psi_l|^2\right\} \\ &\Psi_l = \sum\nolimits_{n=1}^{N}\left|\mathbf{v}_{l,\mathbf{G}}(n)\right|\exp\left(j\delta_n\right) \end{aligned} \tag{3.2}$$

where: (i) γ_0 is a scaling factor that accounts for, e.g., the transmission distances, the geometric size of the RIS, the transmission bandwidth, the noise figure [118]. In this thesis, it is considered to be a constant; (ii) $\lambda_{l,\mathbf{G}}$ is the l-th non-zero eigenvalue of the matrix $\mathbf{W}_\mathbf{G} = (1/N)\,\mathbf{G}^H\mathbf{G}$ and $\lambda_{l,\mathbf{H}}$ is the l-th non-zero eigenvalue of the matrix $\mathbf{W}_\mathbf{H} = (1/N)\,\mathbf{H}\mathbf{H}^H$; (iii) $\mathbf{v}_{l,\mathbf{G}}$ and $\mathbf{u}_{l,\mathbf{G}}$ are the lth eigenvectors of $\mathbf{W}_\mathbf{G}$ and $\mathbf{W}_{\mathbf{G}^H}$, respectively, that correspond to $\lambda_{k,\mathbf{G}}$; (iv) and $\mathbf{v}_{k,\mathbf{H}}$ and $\mathbf{u}_{k,\mathbf{H}}$ are the kth eigenvectors of $\mathbf{W}_\mathbf{H}$ and $\mathbf{W}_{\mathbf{H}^H}$, respectively, that correspond to $\lambda_{k,\mathbf{H}}$. As mentioned, the SNRs in (3.1) and (3.2) are applicable in the far-field regime, as defined in [118], [119]. Thus, N can be large but it needs to be finite [7, Sec. IV-D].

3.3.1 Preliminaries

The semi-analytical expressions of the SNR in (3.1) and (3.2) are the departing point for calculating the distribution and the scaling laws of the SNR as a function of N (see Section 3.4). First, we summarize some lemmas to enable such analysis.

Lemma 3.1 *Let* $\lambda_{\mathbf{H}}^{+}$ *and* $\lambda_{\mathbf{G}}^{+}$ *be the largest eigenvalues of* $\mathbf{W}_{\mathbf{H}}$ *and* $\mathbf{W}_{\mathbf{G}}$*, respectively.* $\lambda_{\mathbf{H}}^{+}$ *and* $\lambda_{\mathbf{G}}^{+}$ *are well approximated by Gamma random variables whose mean and variance are:*

$$\begin{aligned} &\mathbb{E}\left\{\lambda_{\mathbf{X}}^{+}\right\} = \alpha_1\left(M,N\right) - \alpha_0\beta_1\left(M,N\right) \\ &\mathbb{V}\left\{\lambda_{\mathbf{X}}^{+}\right\} = \beta_0\beta_1^2\left(M,N\right) \end{aligned} \tag{3.3}$$

where $\mathbf{X} = \{\mathbf{G},\mathbf{H}\}$, $M = N_T$ *if* $\mathbf{X} = \mathbf{H}$ *and* $M = N_R$ *if* $\mathbf{X} = \mathbf{G}$, $\alpha_0 = 1.7711$, $\beta_0 = 0.8132$*, and:*

$$\begin{aligned} &\alpha_1\left(M,N\right) = \left(1+\sqrt{M/N}\right)^2 \\ &\beta_1\left(M,N\right) = N^{-2/3}\left(1+\sqrt{M/N}\right)\left(1+\sqrt{N/M}\right)^{1/3} \end{aligned} \tag{3.4}$$

Proof 3.1 *It follows from [120] and [121] by applying results on random matrix theory and by calculating numerically the mean and the variance of the Tracy-Widom distribution.*

Remark 3.1 $\lambda_{\mathbf{H}}^{+}$ *and* $\lambda_{\mathbf{G}}^{+}$ *may be approximated by a shifted Gamma random variable [121]. We consider a Gamma random variable due to its simplicity yet satisfactory accuracy.*

Lemma 3.2 *Let* $\mathbf{v}_{l,\mathbf{G}}$ *and* $\mathbf{u}_{k,\mathbf{H}}$ *be the eigenvectors in* (3.1) *and* (3.2)*. For any* N*, they are i.i.d. and uniformly distributed vectors on the* $N-1$ *sphere, i.e., on the surface of the unit* N*-ball. Thus, their distribution is equivalent to (for any* l,k*):*

$$\mathbf{v}_{l,\mathbf{G}} \overset{d}{=} \mathbf{v}/\left\|\mathbf{v}\right\| \qquad \mathbf{u}_{l,\mathbf{H}} \overset{d}{=} \mathbf{u}/\left\|\mathbf{u}\right\| \tag{3.5}$$

where $\mathbf{v}(n) \sim \mathcal{CN}(0,1)$ *and* $\mathbf{u}(n) \sim \mathcal{CN}(0,1)$ *are mutually i.i.d. random variables for* $n = 1,2,\ldots,N$*.*

Proof 3.2 *See [122].*

Remark 3.2 *From Lemma 3.2, we evince that, for every finite* N*, the eigenvectors of a Wishart matrix with zero mean complex Gaussian entries (i.e.,* $\mathbf{W}_{\mathbf{H}}$ *and* $\mathbf{W}_{\mathbf{G}}$*) do not point towards any privileged direction. If the entries are not Gaussian, this result does not hold in general [122].*

Lemma 3.3 *For any* l*, the eigenvalues* $\lambda_{l,\mathbf{X}}$ *and the eigenvectors* $\mathbf{v}_{l,\mathbf{X}}$ *or* $\mathbf{u}_{l,\mathbf{X}}$ *for* $\mathbf{X} = \{\mathbf{G},\mathbf{H}\}$ *are independent.*

Proof 3.3 *See [122].*

Lemma 3.4 *Let* $\mathbf{y}$ *be an* $N \times 1$ *vector whose entries are i.i.d. standard complex Gaussian random variables, i.e.,* $\mathbf{y}(n) \sim \mathcal{CN}(0,1)$ *for* $n = 1,2,\ldots,N$. *Define the normalized vector* $\widehat{\mathbf{y}}(n) = |\mathbf{y}(n)| \big/ \|\mathbf{y}\|$. *Then,* $\widehat{\mathbf{y}}^2(n) \sim \mathcal{B}(1, N-1)$ *and:*

$$\begin{aligned} &\mathbb{E}\left\{\widehat{\mathbf{y}}^2(n)\right\} = 1/N \\ &\mathbb{E}\left\{\widehat{\mathbf{y}}(n)\right\} = \left(\sqrt{\pi}/2\right)\left(\Gamma(N)/\Gamma(N+1/2)\right) \end{aligned} \tag{3.6}$$

Proof 3.4 *Since* $\mathbf{y}(n) \sim \mathcal{CN}(0,1)$ *for* $n = 1,2,\ldots,N$, *then* $\widehat{\mathbf{y}}^2(n) = \mathcal{Y}_1 \big/ (\mathcal{Y}_1 + \mathcal{Y}_2)$ *where* $\mathcal{Y}_1 = \widehat{\mathbf{y}}^2(n) \sim \mathcal{G}(1,1)$ *and* $\mathcal{Y}_2 = \sum_{m \neq n=1}^{N} \widehat{\mathbf{y}}^2(m) \sim \mathcal{G}(N-1,1)$ *are independent random variables. Thus,* $\mathcal{Y}_1 \big/ (\mathcal{Y}_1 + \mathcal{Y}_2) \sim \mathcal{B}(1, N-1)$, *and* (3.6) *follows from the moments of a Beta random variable.*

Lemma 3.5 *Let* $\mathbf{y}$ *be an* $N \times 1$ *vector of i.i.d. standard complex Gaussian random variables, i.e.,* $\mathbf{y}(n) \sim \mathcal{CN}(0,1)$ *for* $n = 1,2,\ldots,N$. *Define* $\widehat{\mathbf{y}}(n) = |\mathbf{y}(n)| / \|\mathbf{y}\|$. *For* $m \neq n = 1,2,\ldots,N$, *we have* $\mathbb{E}\left\{\widehat{\mathbf{y}}(n)\widehat{\mathbf{y}}(m)\right\} = \pi/(4N)$.

Proof 3.5 *Define the variable* $z = \sum_{k=1}^{N} |\mathbf{y}(k)|^2$. *By using the notable integral* $\int_0^{+\infty} e^{-zt} dt = 1/z$ *for* $z > 0$, *we obtain:*

$$\begin{aligned} \mathbb{E}\left\{\widehat{\mathbf{y}}(n)\widehat{\mathbf{y}}(m)\right\} &= \mathbb{E}\left\{|\mathbf{y}(n)|\,|\mathbf{y}(m)|\,(1/z)\right\} \\ &\overset{(a)}{=} \int_0^{+\infty} \mathbb{E}\left\{\mathcal{J}_n(t)\right\}\mathbb{E}\left\{\mathcal{J}_m(t)\right\} \prod_{k=1\neq n,m}^{N} \mathbb{E}\left\{\mathcal{J}_k(t)\right\} dt \end{aligned} \tag{3.7}$$

where (a) follows because the entries of $\mathbf{y}$ *are independent and we defined* $\mathcal{J}_n(t) = |\mathbf{y}(n)| \exp\left(-|\mathbf{y}(n)|^2 t\right)$, $\mathcal{J}_m(t) = |\mathbf{y}(m)| \exp\left(-|\mathbf{y}(m)|^2 t\right)$, $\mathcal{J}_k(t) = \exp\left(-|\mathbf{y}(k)|^2 t\right)$. *The proof follows by computing each expectation since the distribution of* $|\mathbf{y}(n)|^2$ *is known, i.e.,* $|\mathbf{y}(n)|^2 \sim \mathcal{G}(1,1)$, *and by using the notable integral* $\int_0^{+\infty} (1+t)^{1+N} dt = 1/N$.

Lemma 3.6 *Consider* $\Upsilon_{l,k}$ *and* Ψ_l *in* (3.1) *and* (3.2) *for* $l, k = 1,2,\ldots,N$. *For* $\eta = 1,2$, *let us define the moments* $\overline{\overline{m}}_{\mathcal{R}}^{(\eta)} = \mathbb{E}\left\{\left(\mathrm{Re}\left\{\Upsilon_{l,k}\right\}\right)^{\eta}\right\}$, $\overline{\overline{m}}_{\mathcal{I}}^{(\eta)} = \mathbb{E}\left\{\left(\mathrm{Im}\left\{\Upsilon_{l,k}\right\}\right)^{\eta}\right\}$, $\overline{m}_{\mathcal{R}}^{(\eta)} = \mathbb{E}\left\{(\mathrm{Re}\{\Psi_l\})^{\eta}\right\}$, *and* $\overline{m}_{\mathcal{I}}^{(\eta)} = \mathbb{E}\left\{(\mathrm{Im}\{\Psi_l\})^{\eta}\right\}$. *Then, we have:*

$$\begin{aligned} \overline{\overline{m}}_{\mathcal{R}}^{(1)} &= N(\pi/4)\left(\Gamma(N)/\Gamma(N+1/2)\right)^2 c_1 \\ \overline{\overline{m}}_{\mathcal{I}}^{(1)} &= N(\pi/4)\left(\Gamma(N)/\Gamma(N+1/2)\right)^2 s_1 \\ \overline{\overline{m}}_{\mathcal{R}}^{(2)} &= (1/N)\, c_2 + \left(\pi^2/16\right)\left((N-1)/N\right) c_1^2 \\ \overline{\overline{m}}_{\mathcal{I}}^{(2)} &= (1/N)\, s_2 + \left(\pi^2/16\right)\left((N-1)/N\right) s_1^2 \end{aligned} \tag{3.8}$$

Table 3.1 – Examples of phase noise distributions ($s_1 = 0$)

Distribution	c_1	c_2	s_2
$\delta_n = 0$	1	1	0
$\delta_n \sim \mathscr{U}(-\pi,\pi)$	0	1/2	1/2
$\delta_n \sim \mathscr{U}(-\varepsilon\pi,\varepsilon\pi)$	$\operatorname{sinc}(\varepsilon)$	$(1+\operatorname{sinc}(2\varepsilon))/2$	$(1-\operatorname{sinc}(2\varepsilon))/2$
$\delta_n \sim \mathscr{V}\mathscr{M}(0,\kappa)$	$\frac{I_1(\kappa)}{I_0(\kappa)}$	$\frac{I_1(\kappa)-\kappa I_2(\kappa)}{\kappa I_0(\kappa)}$	$\frac{I_1(\kappa)}{\kappa I_0(\kappa)}$

$$
\begin{aligned}
\overline{m}_{\mathscr{R}}^{(1)} &= N(\sqrt{\pi/4})\,(\Gamma(N)/\Gamma(N+1/2))\,c_1\\
\overline{m}_{\mathscr{I}}^{(1)} &= N(\sqrt{\pi/4})\,(\Gamma(N)/\Gamma(N+1/2))\,s_1\\
\overline{m}_{\mathscr{R}}^{(2)} &= c_2 + (\pi/4)\,(N-1)\,c_1^2\\
\overline{m}_{\mathscr{I}}^{(2)} &= s_2 + (\pi/4)\,(N-1)\,s_1^2
\end{aligned}
\tag{3.9}
$$

where $c_1 = \mathbb{E}\{\cos(\delta_n)\}$, $s_1 = \mathbb{E}\{\sin(\delta_n)\}$, $c_2 = \mathbb{E}\left\{\cos^2(\delta_n)\right\}$, *are* $s_2 = \mathbb{E}\left\{\sin^2(\delta_n)\right\}$ *are given in Table 3.1.*

Proof 3.6 *It follows by re-writing* (3.1) *and* (3.2) *by using* (3.5)*, and by computing the moments using Lemmas 3.4 and 3.5.*

Lemma 3.7 *Consider* $\Upsilon_{l,k}$ *and* Ψ_l *in* (3.1) *and* (3.2) *for* $l,k = 1,2,\ldots,N$. *The two random variables* $\operatorname{Re}\left\{\Upsilon_{l,k}\right\}$ *and* $\operatorname{Im}\left\{\Upsilon_{l,k}\right\}$ *and the two random variables* $\operatorname{Re}\{\Psi_l\}$ *and* $\operatorname{Im}\{\Psi_l\}$ *are uncorrelated for any* $l,k = 1,2,\ldots,N$ *if the distribution of* δ_n*, for every* $n = 1,2,\ldots,N$*, is symmetric around zero.*

Proof 3.7 *By definition of covariance, we have:*

$$
\begin{aligned}
\operatorname{cov}\left\{\operatorname{Re}\left\{\Upsilon_{l,k}\right\}\operatorname{Im}\left\{\Upsilon_{l,k}\right\}\right\} &= (1/N)\,\mathbb{E}\{\sin(2\delta_n)\}/2\\
&\quad + \left(\pi^2/16\right)((N-1)/N)\,c_1 s_1 - \overline{\overline{m}}_{\mathscr{R}}^{(1)}\overline{\overline{m}}_{\mathscr{I}}^{(1)}
\end{aligned}
\tag{3.10}
$$

$$
\begin{aligned}
\operatorname{cov}\{\operatorname{Re}\{\Psi_l\}\operatorname{Im}\{\Psi_l\}\} &= \mathbb{E}\{\sin(2\delta_n)\}/2\\
&\quad + (\pi/4)\,(N-1)\,c_1 s_1 - \overline{m}_{\mathscr{R}}^{(1)}\overline{m}_{\mathscr{I}}^{(1)}
\end{aligned}
\tag{3.11}
$$

The proof follows by noting that $\mathbb{E}\{\sin(2\delta_n)\} = 0$ *and* $s_1 = 0$ *if the distribution of* δ_n *is symmetric around zero.*

Remark 3.3 *Based on Table 3.1, the distribution of* δ_n *is usually symmetric around zero, and the real and imaginary parts of* $\Upsilon_{l,k}$ *and* Ψ_l *can be assumed to be uncorrelated.*

Remark 3.4 *As N grows large, we obtain* $\operatorname{cov}\left\{\operatorname{Re}\left\{\Upsilon_{l,k}\right\}\operatorname{Im}\left\{\Upsilon_{l,k}\right\}\right\} = \mathcal{O}\left(1/N\right)$*, since, for* $N \gg 1$, $\Gamma\left(N\right)/\Gamma\left(N+1/2\right) = N^{-1/2}(1+(8N)^{-1}+\mathcal{O}\left(N^{-2}\right))$. *This implies that the real and imaginary parts of $\Upsilon_{l,k}$ are asymptotically (i.e., for large values of N) uncorrelated even if the distribution of δ_n is not symmetric around zero.*

Lemma 3.8 *Assume that N grows large (i.e., $N \gg 1$). The random variables $\left|\Upsilon_{l,k}\right|^2$ and $|\Psi_l|^2$ for $l,k = 1,2,\ldots,N$ are (asymptotically) equivalent in distribution to the sum of two scaled non-central chi-square random variables:*

$$\begin{aligned}
\left|\Upsilon_{l,k}\right|^2 = \left(\operatorname{Re}\left\{\Upsilon_{l,k}\right\}\right)^2 + \left(\operatorname{Im}\left\{\Upsilon_{l,k}\right\}\right)^2 &\overset{N\gg1}{=} \overline{\overline{\sigma}}^2_{\mathscr{R}}\overline{\overline{C}}_{\mathscr{R}} + \overline{\overline{\sigma}}^2_{\mathscr{I}}\overline{\overline{C}}_{\mathscr{I}} \\
\overline{\overline{C}}_{\mathscr{R}} \sim \mathscr{X}_1^2\left(\overline{\overline{\mu}}^2_{\mathscr{R}}\right) \quad & \overline{\overline{C}}_{\mathscr{I}} \sim \mathscr{X}_1^2\left(\overline{\overline{\mu}}^2_{\mathscr{I}}\right)
\end{aligned} \tag{3.12}$$

$$\begin{aligned}
|\Psi_l|^2 = (\operatorname{Re}\{\Psi_l\})^2 + (\operatorname{Im}\{\Psi_l\})^2 &\overset{N\gg1}{=} \bar{\sigma}^2_{\mathscr{R}}\overline{C}_{\mathscr{R}} + \overline{\sigma}^2_{\mathscr{I}}\overline{C}_{\mathrm{I}} \\
\overline{C}_{\mathscr{R}} \sim \mathscr{X}_1^2\left(\overline{\mu}^2_{\mathscr{R}}\right) \quad & \overline{C}_{\mathscr{I}} \sim \mathscr{X}_1^2\left(\overline{\mu}^2_{\mathscr{I}}\right)
\end{aligned} \tag{3.13}$$

where, for $\mathscr{S} = \{\mathscr{R},\mathscr{I}\}$, $\overline{\overline{\sigma}}^2_{\mathscr{S}} = \overline{\overline{m}}^{(2)}_{\mathscr{S}} - \left(\overline{\overline{m}}^{(1)}_{\mathscr{S}}\overline{\overline{m}}^{(1)}_{\mathscr{S}}\right)$, $\overline{\overline{\mu}}_{\mathscr{S}} = \overline{\overline{m}}^{(1)}_{\mathscr{S}}\Big/\overline{\overline{\sigma}}_{\mathscr{S}}$*, and* $\overline{\sigma}^2_{\mathscr{S}} = \overline{m}^{(2)}_{\mathscr{S}} - \left(\overline{m}^{(1)}_{\mathscr{S}}\overline{m}^{(1)}_{\mathscr{S}}\right)$, $\overline{\mu}_{\mathscr{S}} = \overline{m}^{(1)}_{\mathscr{S}}\Big/\overline{\sigma}_{\mathscr{S}}$.

Proof 3.8 *It follows from the central limit theorem if $N \gg 1$:* $\operatorname{Re}\left\{\Upsilon_{l,k}\right\} \sim \mathcal{N}\left(\overline{\overline{m}}^{(1)}_{\mathscr{R}},\overline{\overline{\sigma}}^2_{\mathscr{R}}\right)$, $\operatorname{Im}\left\{\Upsilon_{l,k}\right\} \sim \mathcal{N}\left(\overline{\overline{m}}^{(1)}_{\mathscr{I}},\overline{\overline{\sigma}}^2_{\mathscr{I}}\right)$, $\operatorname{Re}\{\Psi_l\} \sim \mathcal{N}\left(\overline{m}^{(1)}_{\mathscr{R}},\overline{\sigma}^2_{\mathscr{R}}\right)$*, and* $\operatorname{Im}\{\Psi_l\} \sim \mathcal{N}\left(\overline{m}^{(1)}_{\mathscr{I}},\overline{\sigma}^2_{\mathscr{I}}\right)$.

Remark 3.5 *If the distribution of δ_n is symmetric around zero, the non-central chi-square random variables in* (3.12) *and* (3.13) *are independent. This originates from Remark 3, and because the real and imaginary parts of $\Upsilon_{l,k}$ and Ψ_l converge, asymptotically (i.e., $N \gg 1$), to Gaussian random variables.*

Remark 3.6 *Readers are referred to [123] for the sum of independent scaled non-central chi-square random variables.*

3.4 Analysis of the Signal-to-Noise-Ratio

In this section, we analyze the distribution, the mean, the variance, and the AF of the SNR, as well as the corresponding scaling laws as a function of N. The AF of the SNR is, in particular, a unified statistical measure that quantifies the severity of fading and, correspondingly, the robustness of transmission technologies against channel fading. Some results are applicable to arbitrary values of N, while others apply only for large values of N. This is elaborated in further text.

3.4.1 Equivalent in Distribution Representation

For an arbitrary N, the following theorem yields an equivalent in distribution representation of the SNRs in (3.1) and (3.2).

Theorem 3.1 *Consider the SNRs in* (3.1) *and* (3.2). *The following equivalent in distribution representations hold true:*

$$\mathrm{SNR}_{\mathrm{RR}} \overset{d}{=} \gamma_0 N^2 \lambda_{\mathbf{G}}^{+} \lambda_{\mathbf{H}}^{+} \left| \sum_{n=1}^{N} \hat{\mathbf{v}}(n) \hat{\mathbf{u}}(n) \exp\left(j\delta_n\right) \right|^2 \tag{3.14}$$

$$\mathrm{SNR}_{\mathrm{LR}} \overset{d}{=} \gamma_0 N_T N \lambda_{\mathbf{G}}^{+} \left| \sum_{n=1}^{N} \hat{\mathbf{v}}(n) \exp\left(j\delta_n\right) \right|^2 \tag{3.15}$$

where $\hat{\mathbf{v}}(n) = |\mathbf{v}(n)| / \|\mathbf{v}\|$, $\hat{\mathbf{u}}(n) = |\mathbf{u}(n)| / \|\mathbf{u}\|$, and $\mathbf{v}(n) \sim \mathcal{CN}(0,1)$, $\mathbf{u}(n) \sim \mathcal{CN}(0,1)$ are i.i.d. for $n = 1, 2, \ldots, N$.

Proof 3.9 *From Lemma 3.2 and Remark 2, the eigenvectors of a Wishart matrix with zero mean complex Gaussian entries are equal in distribution, and, thus, the maximization in* (3.1) *and* (3.2) *is determined only by the distribution of the (largest) eigenvalues. From Lemma 3.3, the eigenvectors and the eigenvalues of a Wishart matrix with zero mean complex Gaussian entries are independent. This concludes the proof.*

Theorem 3.1 provides us with a general tool for the analysis of RIS-aided wireless systems. Let us consider, e.g., $\mathrm{SNR}_{\mathrm{RR}}$. The same comments apply to $\mathrm{SNR}_{\mathrm{LR}}$. Equation (3.14) holds true for any N and it brings to our attention that the SNR is equivalent in distribution to the product of three independent random variables. There exist different approaches for computing the distribution of the product of independent random variables, e.g., [124]. For example, the distribution of the square absolute value of the sum in (3.14) may be obtained by first computing the Laplace transform of the sum of independent random variables, which is equal to the product of Laplace transforms of the individual random variables. In this thesis, we do not purse this line of research, since the resulting analytical expressions are likely not to be sufficiently tractable to gain insights for system design. In the next two sub-sections, on the other hand, we focus our attention on the case study in which $N \gg 1$, which is relevant for RIS-aided wireless systems.

3.4.2 Channel Model: Rayleigh Fading – Rayleigh Fading

In this section, we analyze the statistics of $\mathrm{SNR}_{\mathrm{RR}}$ in (3.14) under the assumption that N is large, i.e., $N \gg 1$.

Table 3.2 – SNR scaling laws as a function of N ($s_1 = 0$). $\overline{\overline{\zeta}}_{v1} = -6 + \beta_0 \left(N_T^{-1/3} + N_R^{-1/3}\right)$ and $\overline{\zeta}_{v1} = -5 + \beta_0 N_R^{-1/3}$.

$\mathrm{SNR}_{\mathrm{RR}}$	$\mathrm{SNR}_{\mathrm{LR}}$
$o_{e0} = 1$ $o_{e1} = \left(\pi^2/16\right) c_1^2$	$o_{e0} = N_T$ $o_{e1} = (\pi/4)\, N_T c_1^2$
$o_{v0} = 2\left(c_2^2 + s_2^2\right)$ $o_{v1} = \left(\pi^2/4\right) c_1^2 c_2$ $+ \left(\pi^4/256\right) \overline{\overline{\zeta}}_{v1} c_1^4$	$o_{v0} = 2N_T^2 \left(c_2^2 + s_2^2\right)$ $o_{v1} = \pi N_T^2 c_1^2 c_2$ $+ \left(\pi^4/16\right) N_T^2 \overline{\zeta}_{v1} c_1^4$

Theorem 3.2 *Let us assume the same notation and definitions as in Section 3.3. If $N \gg 1$, the following holds true:*

$$\mathrm{SNR}_{\mathrm{RR}} \overset{N \gg 1}{=} \gamma_0 N^2 P_{\mathbf{G}} P_{\mathbf{H}} \left(\overline{\overline{\sigma}}_{\mathscr{R}}^2 \overline{\overline{C}}_{\mathscr{R}} + \overline{\overline{\sigma}}_{\mathrm{I}}^2 \overline{\overline{C}}_{\mathscr{I}} \right) \tag{3.16}$$

where $P_{\mathbf{X}} \sim \mathscr{G}\left(\left(\mathbb{E}\left\{\lambda_{\mathbf{X}}^{+}\right\}\right)^2 \big/ \mathbb{V}\left\{\lambda_{\mathbf{X}}^{+}\right\}, \mathbb{V}\left\{\lambda_{\mathbf{X}}^{+}\right\} / \mathbb{E}\left\{\lambda_{\mathbf{X}}^{+}\right\} \right)$ for $\mathbf{X} = \{\mathbf{G}, \mathbf{H}\}$ and $\overline{\overline{C}}_{\mathscr{S}} \sim \mathscr{X}_1^2 \left(\overline{\overline{\mu}}_{\mathscr{S}}^2 \right)$ for $\mathscr{S} = \{\mathscr{R}, \mathscr{I}\}$ are four mutually independent random variables.

Proof 3.10 *It follows from* (3.14) *by using Lemmas 3.1 and 3.8.*

Proposition 3.1 *Let us assume $N \gg 1$. The mean and the variance of* $\mathrm{SNR}_{\mathrm{RR}}$ *in* (3.14) *can be formulated as follows:*

$$\begin{aligned} \mathbb{E}\left\{\mathrm{SNR}_{\mathrm{RR}}\right\} &\overset{N \gg 1}{=} \gamma_0 N^2 \mathscr{M}_{\mathbf{G}} \mathscr{M}_{\mathbf{H}} \overline{\overline{\mathscr{M}}} \\ \mathbb{V}\left\{\mathrm{SNR}_{\mathrm{RR}}\right\} &\overset{N \gg 1}{=} \gamma_0^2 N^4 \left(\mathscr{T}_{\mathbf{G}} \mathscr{T}_{\mathbf{H}} \overline{\overline{\mathscr{T}}} - \mathscr{M}_{\mathbf{G}}^2 \mathscr{M}_{\mathbf{H}}^2 \overline{\overline{\mathscr{M}}}^2 \right) \end{aligned} \tag{3.17}$$

where $\mathscr{M}_{\mathbf{X}} = \mathbb{E}\left\{\lambda_{\mathbf{X}}^{+}\right\}$ and $\mathscr{T}_{\mathbf{X}} = \mathbb{V}\left\{\lambda_{\mathbf{X}}^{+}\right\} + \left(\mathbb{E}\left\{\lambda_{\mathbf{X}}^{+}\right\}\right)^2$ for $\mathbf{X} = \{\mathbf{G}, \mathbf{H}\}$, and $\overline{\overline{\mathscr{M}}} = \left(\overline{\overline{\sigma}}_{\mathscr{R}}^2 \left(1 + \overline{\overline{\mu}}_{\mathscr{R}}^2\right) + \overline{\overline{\sigma}}_{\mathscr{I}}^2 \left(1 + \overline{\overline{\mu}}_{\mathscr{I}}^2\right) \right)$, $\overline{\overline{\mathscr{T}}} = \left(2\overline{\overline{\sigma}}_{\mathscr{R}}^4 \left(1 + 2\overline{\overline{\mu}}_{\mathscr{R}}^2\right) + 2\overline{\overline{\sigma}}_{\mathscr{I}}^4 \left(1 + 2\overline{\overline{\mu}}_{\mathscr{I}}^2\right) \right)$.

Proof 3.11 *It follows from the independence of the random variables in* (3.16) *and by using Lemma 3.1 and Lemma 3.8.*

3.4.3 Channel Model: Line-of-Sight – Rayleigh Fading

In this section, we analyze the statistics of $\mathrm{SNR}_{\mathrm{LR}}$ in (3.15) under the assumption that N is large, i.e., $N \gg 1$.

Theorem 3.3 *Let us assume the same notation and definitions as in Section 3.3. If $N \gg 1$, the following holds true:*

$$\mathrm{SNR}_{\mathrm{LR}} \overset{N \gg 1}{=} \gamma_0 N_T N P_{\mathbf{G}} \left(\overline{\sigma}_{\mathscr{R}}^2 \overline{C}_{\mathscr{R}} + \overline{\sigma}_{\mathrm{I}}^2 \overline{C}_{\mathscr{I}} \right) \tag{3.18}$$

where $P_{\mathbf{G}} \sim \mathcal{G}\left(\left(\mathbb{E}\left\{\lambda_{\mathbf{G}}^{+}\right\}\right)^2 \big/ \mathbb{V}\left\{\lambda_{\mathbf{G}}^{+}\right\}, \mathbb{V}\left\{\lambda_{\mathbf{G}}^{+}\right\} / \mathbb{E}\left\{\lambda_{\mathbf{G}}^{+}\right\}\right)$ *and* $\overline{C}_{\mathscr{S}} \sim \mathscr{X}_1^2\left(\overline{\overline{\mu}}_{\mathscr{S}}^2\right)$ *for* $\mathscr{S} = \{\mathscr{R}, \mathscr{I}\}$ *are three mutually independent random variables.*

Proof 3.12 *It is similar to the proof of Theorem 3.2.*

Proposition 3.2 *Let us assume* $N \gg 1$. *The mean and the variance of* $\mathrm{SNR}_{\mathrm{LR}}$ *in* (3.15) *can be formulated as follows:*

$$\begin{aligned} \mathbb{E}\left\{\mathrm{SNR}_{\mathrm{LR}}\right\} &\overset{N\gg 1}{=} \gamma_0 N_T N \mathscr{M}_{\mathbf{G}} \overline{\mathscr{M}} \\ \mathbb{V}\left\{\mathrm{SNR}_{\mathrm{LR}}\right\} &\overset{N\gg 1}{=} \gamma_0^2 N_T^2 N^2 \left(\mathscr{T}_{\mathbf{G}} \overline{\mathscr{T}} - \mathscr{M}_{\mathbf{G}}^2 \overline{\mathscr{M}}^2\right) \end{aligned} \tag{3.19}$$

where $\mathscr{M}_{\mathbf{G}} = \mathbb{E}\left\{\lambda_{\mathbf{G}}^{+}\right\}$, $\mathscr{T}_{\mathbf{G}} = \mathbb{V}\left\{\lambda_{\mathbf{G}}^{+}\right\} + \left(\mathbb{E}\left\{\lambda_{\mathbf{G}}^{+}\right\}\right)^2$, $\overline{\mathscr{M}} = \left(\overline{\sigma}_{\mathscr{R}}^2\left(1 + \overline{\mu}_{\mathscr{R}}^2\right) + \overline{\sigma}_{\mathscr{I}}^2\left(1 + \overline{\mu}_{\mathscr{I}}^2\right)\right)$, *and* $\overline{\mathscr{T}} = \left(2\overline{\sigma}_{\mathscr{R}}^4\left(1 + 2\overline{\mu}_{\mathscr{R}}^2\right) + 2\overline{\sigma}_{\mathscr{I}}^4\left(1 + 2\overline{\mu}_{\mathscr{I}}^2\right)\right)$.

Proof 3.13 *It is similar to the proof of Proposition 3.1.*

3.4.4 Scaling Laws and Insights

From Propositions 3.1 and 3.2, explicit analytical expressions for the mean and the variance of the SNR can be obtained. The resulting formulas are, however, not tractable enough to gain insights for system design. Therefore, we analyze the dominant terms (scaling laws) in the asymptotic regime $N \gg 1$.

Proposition 3.3 *Define* $\mathrm{D} = \{\mathrm{RR}, \mathrm{LR}\}$, *and assume* $s_1 = 0$ *and* $N \gg 1$. *Let* $\mathrm{AF}_{\mathrm{SNR}_{\mathrm{D}}} = \mathbb{V}\left\{\mathrm{SNR}_{\mathrm{D}}\right\} / \left(\mathbb{E}\left\{\mathrm{SNR}_{\mathrm{D}}\right\}\right)^2$ *be the AF of* $\mathrm{SNR}_{\mathrm{D}}$. *As a function of* N, *while keeping the other system parameters fixed, the following scaling laws hold true:*

$$\begin{aligned} \mathbb{E}\left\{\mathrm{SNR}_{\mathrm{D}}\right\} &\overset{N\gg 1}{\propto} o_{e0} N^1 \mathbf{1}\left(c_1 = 0\right) + o_{e1} N^2 \mathbf{1}\left(c_1 \neq 0\right) \\ \mathbb{V}\left\{\mathrm{SNR}_{\mathrm{D}}\right\} &\overset{N\gg 1}{\propto} o_{v0} N^2 \mathbf{1}\left(c_1 = 0\right) + o_{v1} N^3 \mathbf{1}\left(c_1 \neq 0\right) \\ \mathrm{AF}_{\mathrm{SNR}_{\mathrm{D}}} &\overset{N\gg 1}{\propto} \frac{o_{v0}}{o_{e0}^2} N^0 \mathbf{1}\left(c_1 = 0\right) + \frac{o_{v1}}{o_{e1}^2} N^{-1} \mathbf{1}\left(c_1 \neq 0\right) \end{aligned} \tag{3.20}$$

where o_{e0}, o_{e1}, o_{v0}, *and* o_{v1} *are defined in Table 3.2.*

Proof 3.14 *It follows from* (3.17), (3.19), (3.4) *since* $\alpha_1(M, N) = 1 + \mathscr{O}\left(N^{-1/2}\right)$, $\beta_1(M, N) = M^{-1/6} N^{-1/2} + \mathscr{O}\left(N^{-1}\right)$.

From Proposition 3.3, we can draw the following conclusions on the scaling laws of the SNR as a function of N.

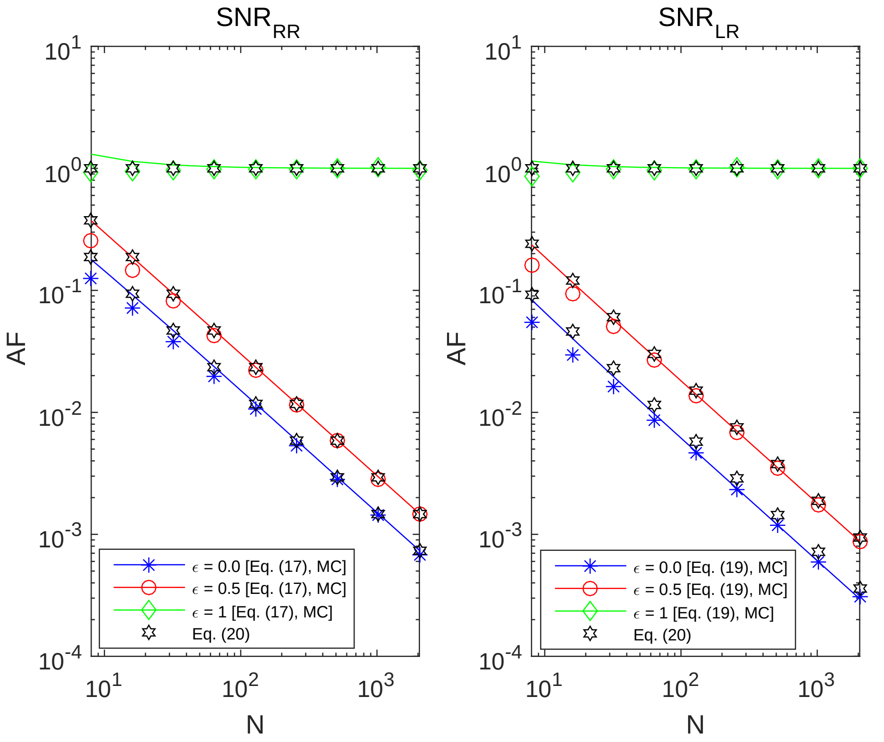


Figure 3.1 – AF of the SNR. ($\gamma_0 = 1$, $N_T = N_R = 4$, $\delta_n \sim \mathcal{U}(-\varepsilon\pi, \varepsilon\pi)$).

- The scaling laws highly depend on whether $c_1 = 0$ or $c_1 \neq 0$. From Table 3.1, e.g., the condition $c_1 = 0$ corresponds to the case study of totally random phase noise. Also, the condition $c_1 = 0$ can be thought of as representative of a scenario with no controllable RIS, in which the RIS is a conventional wall whose phase response is unknown and cannot be optimized. If $c_1 = 0$, in particular, the AF is constant with N, since the RIS is not capable of customizing the radio waves. If $c_1 \neq 0$, on the other hand, the AF decays linearly with N. This unveils the capability of RISs of reducing the fading severity and, as a result, making the transmission of information more robust.
- The robustness of RISs against the phase noise can be quantified by studying the ratios $o_{e1}(c_1)/o_{e1}(c_1 = 1)$ and $o_{v1}(c_1, c_2)/o_{v1}(c_1 = 1, c_2 = 1)$ defined in Table 3.2, since $c_1 = c_2 = 1$ for the benchmark setup with no phase noise (see Table 3.1). This provides a simple tool for quantifying, e.g., the discretization of the phase shifts that yields a suitable trade-off between performance and implementation complexity.

3.5 Numerical Results

Figure 3.1 shows the AF obtained from Monte Carlo (MC) simulations [93, Proposition 1] (markers), and compares it against the analytical frameworks in (3.17) and (3.19) (solid lines), and the scaling laws in (3.20). Figure 3.2 shows the cumulative distribution function (CDF) of

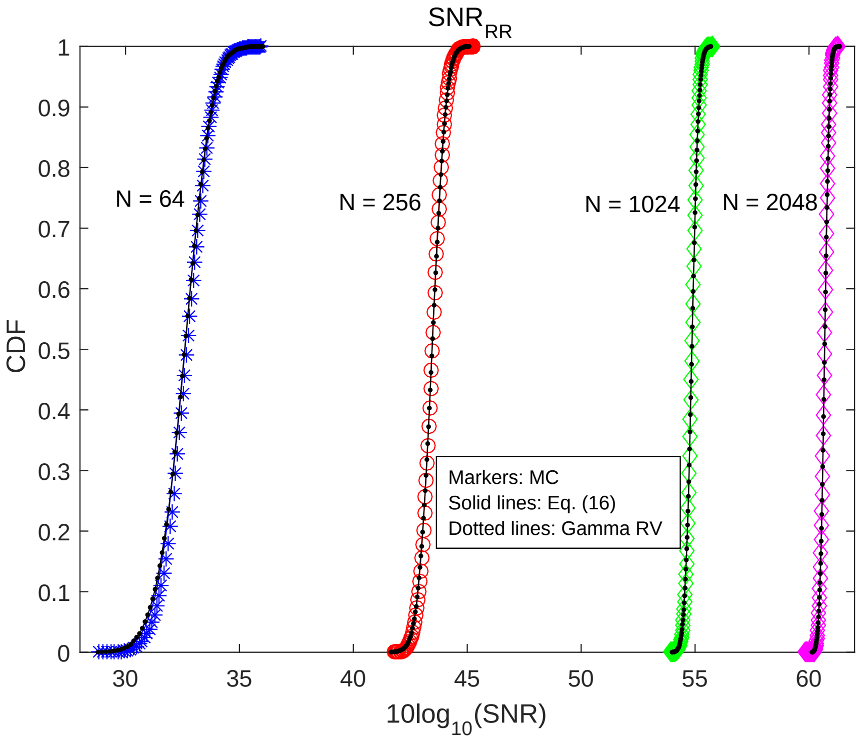


Figure 3.2 – CDF of $\mathrm{SNR}_{\mathrm{RR}}$ ($\gamma_0 = 1$, $N_T = N_R = 4$).

the SNR obtained from Monte Carlo simulations [93, Proposition 1], and compares it against the distributions in (3.16), and a Gamma-based approximation for the SNR, i.e., $\mathrm{P}_{\mathrm{SNR}_{\mathrm{RR}}} \sim \mathscr{G}\left((\mathbb{E}\{\mathrm{SNR}_{\mathrm{RR}}\})^2/\mathbb{V}\{\mathrm{SNR}_{\mathrm{RR}}\}, \mathbb{V}\{\mathrm{SNR}_{\mathrm{RR}}\}/\mathbb{E}\{\mathrm{SNR}_{\mathrm{RR}}\}\right)$. The proposed analytical approach is in good agreement with the simulations and confirm our findings.

3.6 Conclusion

We have introduced an analytical framework to quantify the performance of RIS-aided multi-antenna transmission. If $N \gg 1$, we have proved that the AF of the SNR linearly decreases with N. Also, we have shown that the distribution of the SNR can be well approximated with a Gamma random variable. The proposed approach can be generalized to multiple scenarios, e.g., the analysis of multi-user and multi-RIS transmission.

4 INR Analysis in RIS-assisted MIMO Systems

In this chapter, we consider a multiple-input multiple-output network where each base station serves a user equipment with the aid of an RIS equipped with N reconfigurable elements. We characterize the interference at one user equipment that is caused by the signal emitted by its non-serving (interfering) RIS. By assuming Rayleigh fading channels, we study the corresponding interference-to-noise-ratio (INR) under the assumption of large values of N, and we prove that the INR is the product of a Chi-Square random variable (RV) and an RV that is approximated with a Gamma distribution. In addition, we prove that the amount of fading of the INR is equal to one in the large N regime.

This chapter is organized as follows. In *Section 4.1*, we introduce the background of the analysis of INR in the simplified RIS-aided two-user MIMO system. The system model is introduced in *Section 4.2*. With the aid of the lemmas derived in *Chapter 2*, we analyze the behavior of INR as a random variable in *Section 4.3*. Simulation results are provided in *Section 4.4* followed by the conclusion in *Section 4.5*.

4.1 Introduction

A reconfigurable intelligent surface (RIS) is an emerging technology for increasing the reliability of wireless communications at a reduced hardware cost and energy consumption [2], [7]. An RIS is capable of realizing passive beamforming by adding customized phase shifts to the incident waves, thus appropriately shaping the scattered waves towards specified locations [118], [125]. In contrast with traditional relays, an RIS does not need power amplifiers, introduces no additive noise, and no self-interference in full-duplex implementations [6].

Modeling and analyzing the performance of RIS-aided wireless systems is an open research issue. As outlined in a recent survey and tutorial paper [7], several authors have investigated the joint optimization of the beamforming vector at the transmitter, the matrix of phase shifts at the RIS, and the combining vector at the receiver. However, very few papers have tackled the evaluation of the performance of multiple-input multiple-output (MIMO) RIS-aided systems in fading channels [126]. In [126], the authors have recently characterized the distribution of the signal-to-noise-ratio (SNR) by using tools from random matrix theory and by capitalizing on the optimization framework introduced in [93]. The existing works, however, consider a single transmitter, a single RIS, and a single receiver. Thus, they ignore the interference that is generated by the presence of multiple RISs that are optimized for serving their respective transmitters and receivers.

Motivated by these considerations, we analyze a two-user interference channel in the presence of two RISs. Each MIMO transmitter communicates with its intended MIMO receiver with the aid of an RIS. Each RIS is equipped with N reconfigurable elements that are optimized to maximize the received SNR of each intended user. Therefore, each RIS constitutes a source of interference for the non-intended user. We introduce an analytical approach for characterizing the distribution of the interference-to-noise-ratio (INR) at each receiver, and to study its scaling laws as a function of N in terms of mean, variance, and amount of fading (AF) [116]. Numerical results are illustrated to validate the analysis.

4.2 System model

We consider the RIS-aided communication system depicted in Fig. 4.1, which is made of two BSs, two UEs, and two RISs. BS_i communicates with UE_i via RIS_i, $i = 1,2$, where BS_i is equipped with N_T transmit antennas, UE_i is equipped with N_R receive antennas, and RIS_i is made of N reconfigurable elements that apply independent phase shifts to the incident signal. We assume that perfect channel state information (CSI) is known at the BSs, the UEs, and the RISs. The impact of channel estimation and configuration can be taken into account as recently described in [93], but it is here not considered for simplicity.

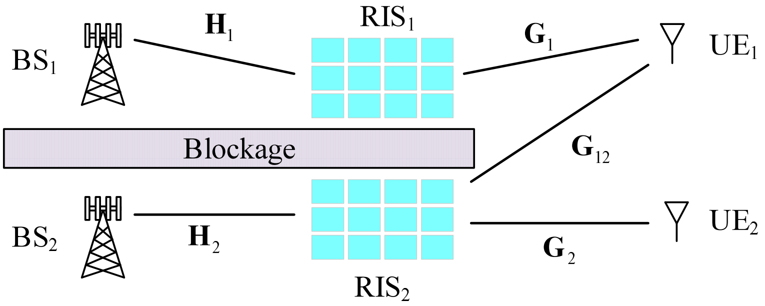


Figure 4.1 – RIS-aided MIMO communication system.

The $N_T \times 1$ unit-norm beamforming vector at BS_i is denoted by $\mathbf{q}_i$ and the $N_R \times 1$ unit-norm combining vector at the receiver is denoted by $\mathbf{w}_i$. The phase shift applied by the nth element of RIS_i is denoted as $\phi_{i,n}$. For ease of notation, the N phase shifts of RIS_i are collected in the $N \times N$ diagonal matrix Φ_i, $i = 1,2$, that is defined as $\Phi_i = \mathrm{diag}\left(e^{j\phi_{i,1}}, e^{j\phi_{i,2}}, \ldots, e^{j\phi_{i,N}}\right)$. To simplify the analysis, we assume that the channel link from BS_i to RIS_j $(i \neq j)$ is blocked. This implies that the interference is available only at the UEs. Our goal is to study the interference at UE_1 caused by RIS_2 (a similar analysis can be applied at UE_2).

Assuming a unit noise power and normalizing with respect to the path-loss, the received signal-to-noise-ratio (SNR) at UE_i can be written as follows [93]:

$$\mathrm{SNR}_i = \left|\mathbf{w}_i^H \mathbf{G}_i \Phi_i \mathbf{H}_i \mathbf{q}_i\right|^2 \tag{4.1}$$

where $\mathbf{H}_i \sim \mathscr{CN}\left(\mathbf{0}_{N \times N_T}, \mathbf{I}_{N \times N_T}\right)$ is the channel from BS_i to RIS_i and $\mathbf{G}_i \sim \mathscr{CN}\left(\mathbf{0}_{N_R \times N}, \mathbf{I}_{N_R \times N}\right)$ is the channel from RIS_i to UE_i, $i = 1,2$. Also, $\mathbf{G}_{12} \sim \mathscr{CN}\left(\mathbf{0}_{N_R \times N}, \mathbf{I}_{N_R \times N}\right)$ is the channel link from RIS_2 to UE_1.

The receiver combining vector at UE_i, the transmit beamforming vector at BS_i, and the phase shifts matrix at RIS_i are jointly optimized to maximize SNR_i in (4.1). In [93], the authors have recently proved that the maximum SNR_i that is obtained by optimizing the triplet $(\mathbf{q}_i, \Phi_i, \mathbf{w}_i)$ can be formulated as:

$$\begin{aligned} &\mathrm{SNR}_i \approx \gamma_0 N^2 \max_{l,k}\left\{\lambda_{k,\mathbf{G}_i}\lambda_{l,\mathbf{H}_i}\left|\Upsilon_{l,k}\right|^2\right\} \\ &\Upsilon_{l,k} = \sum\nolimits_{n=1}^{N}\left|\mathbf{v}_{k,\mathbf{G}}(n)\right|\left|\mathbf{u}_{l,\mathbf{H}}(n)\right| \end{aligned} \tag{4.2}$$

where: (i) $\lambda_{k,\mathbf{G}_i}$ is the kth non-zero eigenvalue of the matrix $\mathbf{W}_{\mathbf{G}_i} = \frac{1}{N}\mathbf{G}_i^H\mathbf{G}_i$ and $\lambda_{l,\mathbf{H}_i}$ is the lth non-zero eigenvalue of the matrix $\mathbf{W}_{\mathbf{H}_i} = \frac{1}{N}\mathbf{H}_i^H\mathbf{H}_i$; (ii) $\mathbf{v}_{k,\mathbf{G}_i}$ and $\mathbf{u}_{k,\mathbf{G}_i}$ are the kth eigenvectors of $\mathbf{W}_{\mathbf{G}_i}$ and $\mathbf{W}_{\mathbf{G}_i^H}$; and (iii) $\mathbf{v}_{l,\mathbf{H}_i}$ and $\mathbf{u}_{l,\mathbf{H}_i}$ are the lth eigenvectors of $\mathbf{W}_{\mathbf{H}_i}$ and $\mathbf{W}_{\mathbf{H}_i^H}$, respectively.

Denoting the optimal indices (k,l) that maximize SNR_i in (4.2) by (k_i,l_i), the triplet $(\mathbf{q}_i,\Phi_i,\mathbf{w}_i)$ that maximizes SNR_i is:

$$\begin{aligned} &\mathbf{w}_i^H=\mathbf{u}_{k_i,\mathbf{G}_i}^H, \quad \mathbf{q}_i=\mathbf{v}_{l_i,\mathbf{H}_i} \\ &\phi_{i,n}=-\arg\left\{\mathbf{v}_{k_i,\mathbf{G}_i}^*(n)\,\mathbf{u}_{l_i,\mathbf{H}_i}(n)\right\} \end{aligned} \tag{4.3}$$

The distribution of SNR_i in (4.2) is characterized in [126]. In this chapter, on the other hand, we focus our attention on the INR at UE_1 that is obtained by considering the optimized triplet $(\mathbf{q}_2,\Phi_2,\mathbf{w}_1)$ in (4.3). In particular, the INR at UE_1 can be written as follows:

$$\mathrm{INR}=\left|\mathbf{w}_1^H\mathbf{G}_{12}\Phi_2\mathbf{H}_2\mathbf{q}_2\right|^2 \tag{4.4}$$

Let $\mathbf{H}_i=\sum_{r=1}^{r_{\mathbf{H}_i}}\sqrt{N\lambda_{r,\mathbf{H}_i}}\mathbf{u}_{r,\mathbf{H}_i}\mathbf{v}_{r,\mathbf{H}_i}^H$ be the singular value decomposition of $\mathbf{H}_i$ where $r_{\mathbf{H}_i}=\mathrm{rank}(\mathbf{H}_i)$ is the rank of $\mathbf{H}_i$. From (4.3), we have $\mathbf{H}_i\mathbf{q}_i=\sqrt{N\lambda_{l_i,\mathbf{H}_i}}\mathbf{u}_{l_i,\mathbf{H}_i}$, which implies:

$$\mathrm{INR}=N\lambda_{l_2,\mathbf{H}_2}\left|\mathbf{u}_{k_1,\mathbf{G}_1}^H\mathbf{G}_{12}\Phi_2\mathbf{u}_{l_2,\mathbf{H}_2}\right|^2 \tag{4.5}$$

By denoting $\kappa=\mathbf{u}_{k_1,\mathbf{G}_1}^H\mathbf{G}_{12}\Phi_2\mathbf{u}_{l_2,\mathbf{H}_2}$, (4.5) can be rewritten as follows:

$$\mathrm{INR}=N\lambda_{l_2,\mathbf{H}_2}|\kappa|^2 \tag{4.6}$$

Based on Lemma 3.2, $\mathbf{u}_{k_1,\mathbf{G}_1}^H$, $\mathbf{u}_{l_2,\mathbf{H}_2}$, $\mathbf{v}_{k_2,\mathbf{G}_2}$ are equivalent in distribution to the following random variables:

$$\begin{aligned} &\mathbf{u}_{k_1,\mathbf{G}_1}^H\overset{d}{=}\hat{\mathbf{u}}_1=\mathbf{u}_1/\left\|\mathbf{u}_1\right\| \\ &\mathbf{u}_{l_2,\mathbf{H}_2}\overset{d}{=}\hat{\mathbf{u}}_2=\mathbf{u}_2/\left\|\mathbf{u}_2\right\| \\ &\mathbf{v}_{k_2,\mathbf{G}_2}\overset{d}{=}\hat{\mathbf{v}}_2=\mathbf{v}_2/\left\|\mathbf{v}_2\right\| \end{aligned} \tag{4.7}$$

where $\mathbf{u}_1\sim\mathcal{CN}\left(\mathbf{0}_{1\times N_R},\mathbf{1}_{1\times N_R}\right)$, $\mathbf{u}_2\sim\mathcal{CN}\left(\mathbf{0}_{1\times N},\mathbf{1}_{1\times N}\right)$, and $\mathbf{v}_2\sim\mathcal{CN}\left(\mathbf{0}_{1\times N},\mathbf{1}_{1\times N}\right)$ are mutually independent.

By denoting $\mathbf{p}=\hat{\mathbf{u}}_1\mathbf{G}_{12}$, we have $\kappa=\mathbf{p}\Phi_2\hat{\mathbf{u}}_2$. From the definition of Φ_2, we can have that $\kappa=\sum_{n=1}^{N}\mathbf{p}(n)\,\hat{\mathbf{u}}_2(n)\exp\left(j\phi_{2,n}\right)$. Therefore, κ simplifies to:

$$\kappa=\sum\nolimits_{n=1}^{N}\left|\mathbf{p}(n)\right|\left|\hat{\mathbf{u}}_2(n)\right|\exp(j\delta(n)) \tag{4.8}$$

where, from (4.3), we have:

$$\begin{aligned}\delta(n) &= \arg\left(\mathbf{p}(n)\,\hat{\mathbf{u}}_2(n)\right) - \arg\left\{\mathbf{v}^*_{k_2,\mathbf{G}_2}(n)\,\hat{\mathbf{u}}_2(n)\right\}\\ &= \arg\left(\mathbf{p}(n)\right) + \arg\left(\mathbf{v}_{k_2,\mathbf{G}_2}(n)\right)\end{aligned} \tag{4.9}$$

Since $\arg(\mathbf{v}_{k_2,\mathbf{G}_2}(n)) \overset{d}{=} \arg(\mathbf{u}_{l_2,\mathbf{H}_2}(n)) \overset{d}{=} \arg(\hat{\mathbf{u}}_2(n))$, then κ in (4.8) is equivalent in distribution to the following:

$$\kappa \overset{d}{=} \sum\nolimits_{n=1}^{N} \hat{\mathbf{u}}_2(n)\,\mathbf{p}(n) \tag{4.10}$$

Based on (4.10), we evince that the phase shift $\phi_{2,n}$ introduced by RIS_2 does not change the distribution of κ. From (4.6), the INR can be formulated as follows:

$$\text{INR} \overset{d}{=} N\lambda_{l_2,\mathbf{H}_2}\left|\sum\nolimits_{n=1}^{N} \hat{\mathbf{u}}_2(n)\,\mathbf{p}(n)\right|^2 \tag{4.11}$$

Next, we characterize the distribution of the INR in (4.11).

4.3 Analysis of the Interference-to-Noise-Ratio

In this section, we analyze the distribution, the mean, the variance, and the AF of the INR in (4.11), as well as the corresponding scaling laws as a function of N.

Theorem 4.1 *Consider the INR in* (4.11). *The following approximate equivalent in distribution representation holds true:*

$$\text{INR} \overset{d}{\approx} N\lambda^+_{\mathbf{H}}\left|\sum\nolimits_{n=1}^{N} \hat{\mathbf{u}}_2(n)\,\mathbf{p}(n)\right|^2 \tag{4.12}$$

where $\hat{\mathbf{u}}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$, $\hat{\mathbf{u}}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$, $\mathbf{u}_1 \sim \mathcal{CN}\left(\mathbf{0}_{1\times N_R}, \mathbf{1}_{1\times N_R}\right)$, $\mathbf{u}_2 \sim \mathcal{CN}\left(\mathbf{0}_{1\times N}, \mathbf{1}_{1\times N}\right)$, $\mathbf{p}(n) = \sum_{m=1}^{N_R} \hat{\mathbf{u}}_1(m)\,\mathbf{G}_{12}(m,n)$, $\mathbf{G}_{12} \sim \mathcal{CN}\left(\mathbf{0}_{N_R\times N}, \mathbf{1}_{N_R\times N}\right)$, *and* $\lambda^+_{\mathbf{H}}$ *is a Gamma RV, whose mean and variance are defined as follows:*

$$\begin{aligned}&\mathbb{E}\left(\lambda^+_{\mathbf{H}}\right) = \alpha_1(N_T,N) - \alpha_0\beta_1(N_T,N)\\ &\mathbb{V}\left(\lambda^+_{\mathbf{H}}\right) = \beta_0\beta_1^2(N_T,N)\end{aligned} \tag{4.13}$$

where $\alpha_0 = 1.7711$, $\beta_0 = 0.8132$ *and:*

$$\begin{aligned}&\alpha_1(N_T,N) = \left(1+\sqrt{N_T/N}\right)^2\\ &\beta_1(N_T,N) = N^{-2/3}\left(1+\sqrt{N_T/N}\right)\left(1+\sqrt{N/N_T}\right)^{1/3}\end{aligned} \tag{4.14}$$

Proof 4.1 *From Lemma 3.3, the eigenvectors and the eigenvalues of a Wishart matrix with* *zero mean complex Gaussian entries are independent. Also,* $\lambda_{l_2,\mathbf{H}_2}$ *in* (4.11) *is the maximum eigenvalue of the* $\mathbf{W}_{\mathbf{H}_2} = \frac{1}{N}\mathbf{H}_2^H\mathbf{H}_2$ *whose distribution can be well approximated with a Gamma RV as stated in Lemma 3.1 based on [121, 120]. This concludes the proof.*

Theorem 4.2 *If* $N \gg 1$*, the following holds true:*

$$\mathrm{INR} \overset{N\gg 1}{=} \frac{N}{2}\lambda_{\mathbf{H}}^{+} Q \tag{4.15}$$

where $\lambda_{\mathbf{H}}^{+} \sim \mathcal{G}((\mathbb{E}(\lambda_{\mathbf{H}}^{+}))^2/\mathbb{V}(\lambda_{\mathbf{H}}^{+}), \mathbb{V}(\lambda_{\mathbf{H}}^{+})/\mathbb{E}(\lambda_{\mathbf{H}}^{+}))$ *is a Gamma RV with* $\mathbb{E}(\lambda_{\mathbf{H}}^{+})$ *and* $\mathbb{V}(\lambda_{\mathbf{H}}^{+})$ *given in* (4.13)*, and* $Q \sim \chi^2(2)$ *is a Chi-Squared RV with two degrees of freedom.*

Proof 4.2 *From Theorem 4.1,* $\lambda_{\mathbf{H}}^{+}$ *can be approximated as a Gamma random variable as follows:*

$$\lambda_{\mathbf{H}}^{+} \sim \mathcal{G}\left(\left(\mathbb{E}\left(\lambda_{\mathbf{H}}^{+}\right)\right)^2/\mathbb{V}\left(\lambda_{\mathbf{H}}^{+}\right), \mathbb{V}\left(\lambda_{\mathbf{H}}^{+}\right)/\mathbb{E}\left(\lambda_{\mathbf{H}}^{+}\right)\right) \tag{4.16}$$

For $N \gg 1$*, by virtue of the central limit theorem,* $\kappa = \sum_{n=1}^{N} \hat{\mathbf{u}}_2(n)\,\mathbf{p}(n)$ *in* (4.12) *can be approximated with a Gaussian RV. By definition,* $\mathbf{p} = \hat{\mathbf{u}}_1\mathbf{G}_{12}$*, denote* $\vartheta(n,m) = \hat{\mathbf{u}}_1(m)\,\hat{\mathbf{u}}_2(n)\,\mathbf{G}_{12}(m,n)$*, we have:*

$$\begin{aligned} \mathbb{E}(\kappa) &= \mathbb{E}\left(\sum\nolimits_{n=1}^{N} \hat{\mathbf{u}}_2(n)\,\mathbf{p}(n)\right) \\ &= \sum\nolimits_{n=1}^{N}\sum\nolimits_{m=1}^{N_R} \mathbb{E}(\vartheta(n,m)) \end{aligned} \tag{4.17}$$

Since $\mathbf{G}_{12}$*,* $\hat{\mathbf{u}}_1$ *and* $\hat{\mathbf{u}}_2$ *are mutually independent and* $\mathbb{E}(\mathbf{G}_{12}(m,n)) = 0$*, we have* $\mathbb{E}(\vartheta(n,m)) = 0$ *and* $\mathbb{E}(\kappa) = 0$*. Thus,* $\mathbb{V}(\kappa) = \mathbb{E}((\kappa - \mathbb{E}(\kappa))(\kappa - \mathbb{E}(\kappa))^*) = \mathbb{E}(\kappa\kappa^*)$ *is:*

$$\mathbb{V}(\kappa) = \mathbb{E}\left(\sum_{n_1=1}^{N}\sum_{m_1=1}^{N_R}\sum_{n_2=1}^{N}\sum_{m_2=1}^{N_R} \vartheta(n_1,m_1)\,\vartheta^*(n_2,m_2)\right) \tag{4.18}$$

If $(n_1,m_1) \neq (n_2,m_2)$*,* $\mathbf{G}_{12}(m_1,n_1)$ *and* $\mathbf{G}_{12}(m_2,n_2)$ *are independent, then we can have that* $\mathbb{E}\left(\mathbf{G}_{12}(m_1,n_1)\,\mathbf{G}_{12}^*(m_2,n_2)\right) = \mathbb{E}(\mathbf{G}_{12}(m_1,n_1))\,\mathbb{E}\left(\mathbf{G}_{12}^*(m_2,n_2)\right) = 0$*. Therefore,* (4.18) *is non-zero only for* $(n_1,m_1) = (n_2,m_2)$*, and it simplifies to:*

$$\mathbb{V}(\kappa) = \sum\nolimits_{n_1=1}^{N}\sum\nolimits_{m_1=1}^{N_R} \mathbb{E}\left(|\vartheta(n_1,m_1)|^2\right) \tag{4.19}$$

From Lemma 3.4, we obtain $\mathbb{E}(|\hat{\mathbf{u}}_1(m_1)|^2) = \frac{1}{N_R}$*,* $\mathbb{E}(|\hat{\mathbf{u}}_2(n_1)|^2) = \frac{1}{N}$*. Also,* $\mathbb{E}(|\mathbf{G}_{12}(m_1,n_1)|^2) = 1$ *because* $\mathbf{G}_{12}(m_1,n_1) \sim \mathcal{CN}(0,1)$*. Therefore,* (4.19) *is equal to:*

$$\mathbb{V}(\kappa) = 1 \tag{4.20}$$

From the central limit theorem, we have $\kappa \sim \mathscr{C}N(0,1)$*. Consequently,* $\sqrt{2}\Re(\kappa), \sqrt{2}\Im(\kappa) \sim \mathscr{N}(0,1)$ *and* $Q = 2|\kappa|^2$ *can be written as* $Q = (\sqrt{2}\Re(\kappa))^2 + (\sqrt{2}\Im(\kappa))^2 \sim \chi^2(2)$*. This concludes the proof.*

Based on Theorem 4.2, we evince that, for $N \gg 1$, the INR is the product of an (approximate) Gamma RV and a Chi-Square RV with two degrees of freedom. Also, the two RVs are mutually independent. In the following propositions, we characterize the mean, the variance, and the AF of the INR based on Theorem 4.2.

Proposition 4.1 *Assume* $N \gg 1$*. The mean and the variance of the* INR *in* (4.15) *can be formulated as follows:*

$$\begin{aligned} &\mathbb{E}\left(\mathrm{INR}\right) \overset{N\gg1}{=} N\mathbb{E}\left(\lambda_{\mathbf{H}}^{+}\right) \\ &\mathbb{V}\left(\mathrm{INR}\right) \overset{N\gg1}{=} 2N^2\mathbb{V}\left(\lambda_{\mathbf{H}}^{+}\right) + N^2\left(\mathbb{E}\left(\lambda_{\mathbf{H}}^{+}\right)\right)^2 \end{aligned} \tag{4.21}$$

Proof 4.3 *From* (4.15)*, we obtain:*

$$\begin{aligned} \mathbb{E}(\mathrm{INR}) &\overset{N\gg1}{=} (N/2)\mathbb{E}(\lambda_{\mathbf{H}}^{+})\mathbb{E}(Q) \overset{(a)}{=} N\mathbb{E}(\lambda_{\mathbf{H}}^{+}) \\ \mathbb{E}(\mathrm{INR}^2) &\overset{N\gg1}{=} (N/2)^2\mathbb{E}((\lambda_{\mathbf{H}}^{+})^2)\mathbb{E}(Q^2) \\ &\overset{(b)}{=} 2N^2\mathbb{E}((\lambda_{\mathbf{H}}^{+})^2) \end{aligned} \tag{4.22}$$

where (a) *follows from* $\mathbb{E}(Q) = 2$ *and* (b) *follows from* $\mathbb{E}(Q^2) = \mathbb{V}(Q) + (\mathbb{E}(Q))^2 = 8$*. Using* $\mathbb{E}((\lambda_{\mathbf{H}}^{+})^2) = \mathbb{V}(\lambda_{\mathbf{H}}^{+}) + (\mathbb{E}(\lambda_{\mathbf{H}}^{+}))^2$ *and* $\mathbb{V}(\mathrm{INR}) = \mathbb{E}(\mathrm{INR}^2) - (\mathbb{E}(\mathrm{INR}))^2$*, we obtain* (4.21)*. This concludes the proof.*

From Proposition 4.1, we can infer the scaling laws of the mean, variance, and AF as a function of N by taking into account the distribution of $\lambda_{\mathbf{H}}^{+}$ in Theorem 4.1. The final result is formally stated in the following proposition.

Proposition 4.2 *Assuming* $N \gg 1$ *and let the AF of the* INR *be* $\mathrm{AF}_{\mathrm{INR}} = \mathbb{V}\left(\mathrm{INR}\right) / \left(\mathbb{E}\left(\mathrm{INR}\right)\right)^2$*. As a function of N, while keeping the other parameters fixed, the following scaling laws hold:*

$$\begin{aligned} &\mathbb{E}\left(\mathrm{INR}\right) \overset{N\gg1}{\propto} N \\ &\mathbb{V}\left(\mathrm{INR}\right) \overset{N\gg1}{\propto} N^2 \\ &\mathrm{AF}_{\mathrm{INR}} \overset{N\gg1}{\propto} 1 \end{aligned} \tag{4.23}$$

Proof 4.4 *Consider the definitions in* (4.12) *and* (4.13)*. When* $N \gg 1$*, we have* $(1+\sqrt{N_T/N}) \overset{N\gg1}{\propto} 1$ *and* $(1+\sqrt{N/N_T}) \overset{N\gg1}{\propto} \sqrt{N/N_T}$*. Thus, we obtain* $\alpha_1\left(N_T, N\right) \overset{N\gg1}{\propto} 1$ *and* $\beta_1(N_T, N) \overset{N\gg1}{\propto} N_T{}^{-1/6}N^{-1/2}$,

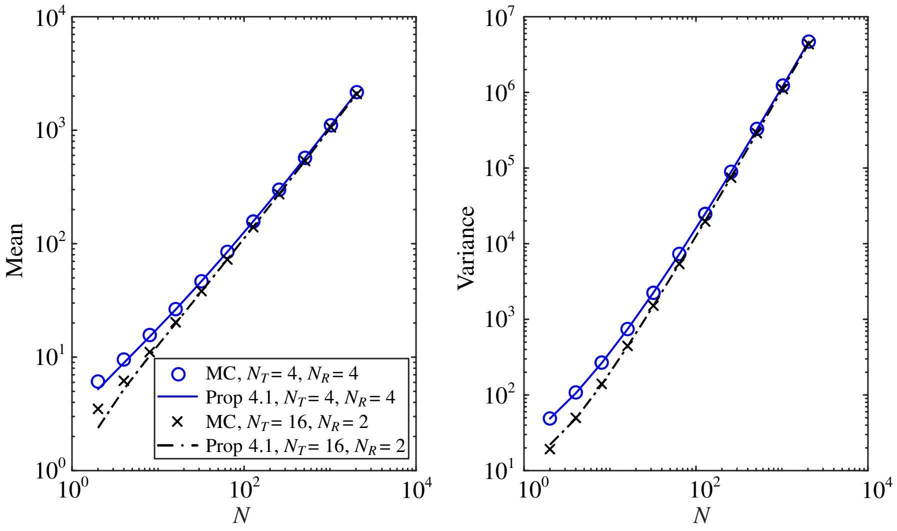


Figure 4.2 – Mean and variance of the INR.

and, from (4.13)*, we have:*

$$\begin{aligned} &\mathbb{E}\left(\lambda_{\mathbf{H}}^{+}\right) = \alpha_1\left(N_T, N\right) - \alpha_0\beta_1\left(N_T, N\right) \\ &\overset{N\gg 1}{\propto} 1 - \alpha_0 {N_T}^{-1/6} N^{-1/2} \overset{N\gg 1}{\propto} 1 \\ &\mathbb{V}\left(\lambda_{\mathbf{H}}^{+}\right) = \beta_0\beta_1^2\left(N_T, N\right) \overset{N\gg 1}{\propto} \beta_0 {N_T}^{-1/3} N^{-1} \end{aligned} \tag{4.24}$$

Therefore, from (4.21)*, we obtain:*

$$\begin{aligned} &\mathbb{E}\left(\mathrm{INR}\right) \overset{N\gg 1}{=} N\mathbb{E}\left(\lambda_{\mathbf{H}}^{+}\right) \overset{N\gg 1}{\propto} N \\ &\mathbb{V}\left(\mathrm{INR}\right) \overset{N\gg 1}{=} 2N^2\mathbb{V}\left(\lambda_{\mathbf{H}}^{+}\right) + N^2\left(\mathbb{E}\left(\lambda_{\mathbf{H}}^{+}\right)\right)^2 \\ &\overset{N\gg 1}{\propto} 2\beta_0 {N_T}^{-1/3} N + N^2 \overset{N\gg 1}{\propto} N^2 \end{aligned} \tag{4.25}$$

Finally, we have $\mathrm{AF}_{\mathrm{INR}} = \mathbb{V}\left(\mathrm{INR}\right) / \left(\mathbb{E}\left(\mathrm{INR}\right)\right)^2 \overset{N\gg 1}{\propto} 1$. *This concludes the proof.*

4.4 Numerical Results

In this section, we illustrate some numerical results in order to substantiate the obtained analytical findings against Monte Carlo (MC) simulations. Figure 4.2 shows the mean and variance of the INR, and compares MC simulations and Proposition 4.1. Figure 4.3 compares the AF obtained from MC simulations against the analytical definition of the AF obtained from (4.21), and the scaling laws in (4.23). In both figures, we obtain a good agreement for $N \gg 1$. Finally, Fig. 4.4 reports the cumulative distribution function (CDF) of the INR obtained from MC simulations, and compares it against the corresponding CDF obtained from Theorems 4.1 and 4.2. Also in this case, we obtain a good agreement between analysis and simulations.

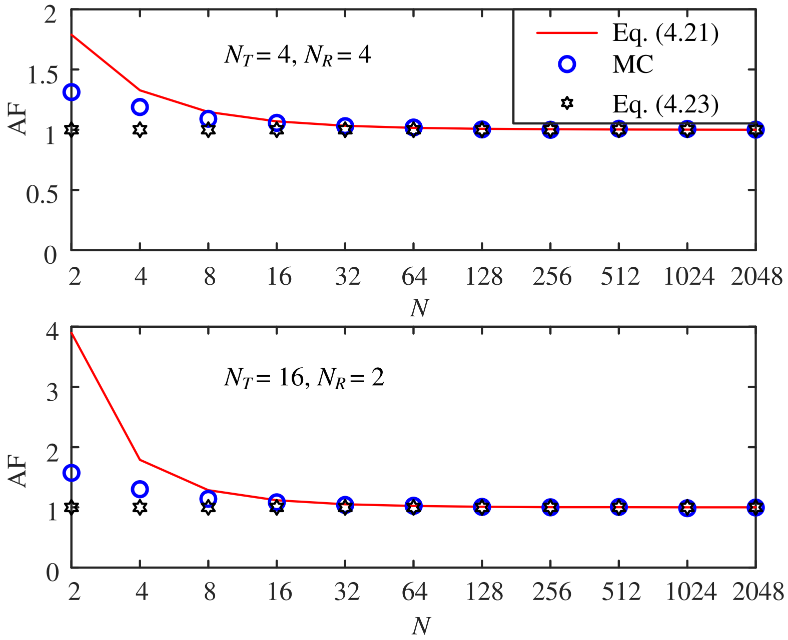


Figure 4.3 – AF of the INR.

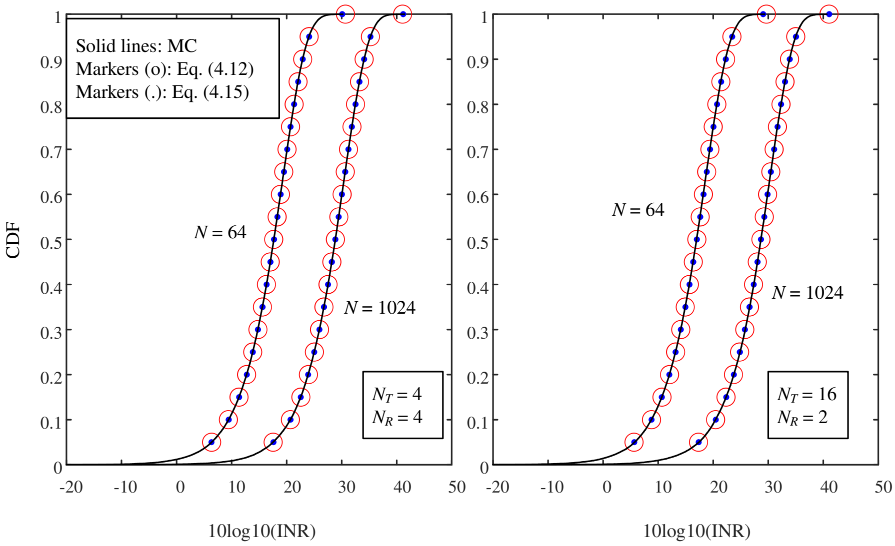


Figure 4.4 – CDF of the INR: Simulations vs. Theorems 4.1 and 4.2.

By comparing the distribution of the INR with that of the SNR obtained in [126], we evince that an RIS is capable of reducing the AF of the intended link linearly with N, while this scaling law is not observed for the INR (i.e., the AF is constant with N). The study conducted in this chapter highlights that the interference-aware optimization of RISs may be needed to achieve a better robustness to the interference.

4.5 Conclusion

We have introduced an analytical framework to characterize the distribution of the INR in RIS-aided systems. Over Rayleigh fading, in particular, we have proved that the INR is asymptotically equivalent in distribution to the product of a Gamma RV and a Chi-Square RV. Also, the AF is proved to be independent of the number N of reconfigurable elements.

Generalizations of the results reported in this chapter include the analysis and optimization of RIS-assisted communications based on electromagnetic-compliant communication models, such as that recently introduced in [39] and [40].

5 Data-driven and model-driven deep learning detection for RIS-aided Spatial Modulation

Reconfigurable intelligent surface (RIS) is regarded as a key technology for the next generation of wireless communications. Recently, the combination of RIS and spatial modulation (SM) or space shift keying (SSK) has attracted a lot of interest in the wireless communication area by achieving a trade-off between spectral and energy efficiency. In this chapter, we first introduce how to utilize deep unfolding in MIMO detection. Then, by generalizing RIS-aided SM/SSK system to a special case of conventional SM system, we investigate deep learning based detection in RIS-aided SM/SSK systems. Based on the idea of deep unfolding, we propose a model-driven deep learning detection for RIS-aided SM systems and compare the performance against the data-driven deep learning detectors.

This chapter is organized as follows. In *Section 5.1*, we introduce the deep learning and RIS-aided SM/SSK system. A review of MIMO detectors including conventional model-based algorithms and deep learning is given in *Section 5.2.* The system model and the traditional detection algorithms for RIS-SSK/SM systems are introduced in *Section 5.3.* The data-driven and the model-driven DL detectors are presented in *Section 5.4.* Simulation results are provided in *Section 5.5* followed by the conclusion in *Section 5.6.*

5.1 Introduction

5.1.1 RIS-SSK/SM

Future wireless communication is expected to provide reliable connections between a large amount of devices with limited energy consumption. To overcome the unreliable wireless communication environment, reconfigurable intelligent surface (RIS) is considered as an important approach to enhance the classical communication channels [2, 7]. By imposing customized phase shifts to the incident waves, an RIS can appropriately shape the scattered waves towards specified locations [118, 125]. Compared to traditional relays, an RIS does not need power amplifiers, thus reducing the interference and power consumption [6].

Recently, the idea of combing RIS and spatial modulation (SM) is getting heated discussion. In [101], Basar et al. proposed to utilize the RIS to maximize the received signal to noise ratio (SNR) on a chosen antenna of the receiver. In this chapter, by generalizing RIS-SSK/SM to a special case of conventional SM system, we aim to investigate deep learning detectors for RIS-SSK/SM.

5.1.2 Deep Unfolding

In deep learning, the structure design of a neural network is critical. To integrate the model expert knowledge into the network structure and design model-driven deep learning (DL) detectors, instead of using a conventional DNN, deep unfolding [127] takes an iterative algorithm with a fixed number of iterations, unfolds its structure, and introduces a number of trainable parameters. The basic idea of deep unfolding is to treat an iteration as a layer in a neural network. This model-driven DL approach can achieve or exceed the performance of corresponding iterative algorithms since the advantages of model-driven and data-driven approach are effectively complementary to each other. Deep unfolding can mitigate the negative impact of time-varying channels, reduce the number of trainable parameter in the neural network and accelerate the convergence.

5.1.3 Literature Study

By unfolding the model-based iterative detection algorithms into a neural network, there has a surge of works on deep unfolding of MIMO detection in recent years. [128] theoretically proved that deep unfolding will guarantee a better performance against the original iterative algorithm. It is also proved that data-driven DL detector can approach the optimal MAP detector with enough training, while the performance of the model-driven DL detector depends on the original iterative algorithm.

Generic iterative algorithms

The MIMO detectors based on Richardson iteration [129, 130], successive over-relaxation (SOR) [131, 132] and Jacobi iteration [133] all have a similar iterative structure. Exploiting the idea of Deep Unfolding, the parameters in the detectors can be set random initially and trained using a DNN. The deep learning based detector exhibits a better performance than the iterative detectors in [129, 130, 131, 132, 133].

Projected Gradient (PG)

Based on PG, [134, 135] proposed a deep learning detector named as DetNet, where extra dimensions are added to enhance the performance of MIMO detection. [136, 137] proposed to train the step-size and projection parameter in PG iterations. In massive overloaded MIMO system where the number of transmit antennas is much greater than that of receive antennas, the unfolded PG-based detector exhibited a good performance.

Conjugate Gradient (CG)

To accelerate the convergence rate of gradient descent while avoiding the high computational cost (for example, in Newton Method the inverse of the Hessian matrix is needed), CG is proposed in [138]. In [139], a MIMO detector based on CG is proposed. In [140], the authors proposed to unfold CG iteartions and train the step-sizes in each iteration.

Proximal Gradient Descent

For the regularized least squares estimation, proximal gradient methods such as iterative soft thresholding algorithm (ISTA), approximate message passing (AMP) and alternating direction method of multipliers (ADMM) [141] are standard approaches to circumvent the non-smoothness of the regularizer while simplifying the optimization problem into a sequence of computationally efficient operations. [142] proposed to learn the ISTA. The idea of unfolding ISTA and AMP is further discussed in [143]. A deep detector based on unfolding ADMM is proposed in [144].

5.2 Deep Unfolding in MIMO detection

We consider a MIMO system with N_T transmit antennas and N_R receive antennas, the input-output relation of data transmission over a frequency-flat MIMO channel is as follows:

$$\mathbf{y}_c = \mathbf{H}_c \mathbf{x}_c + \mathbf{n}_c \tag{5.1}$$

where the index c stands for complex-valued, $\mathbf{y}_c \in \mathbb{C}^{N_R \times 1}$ is the complex-valued receive vector, $\mathbf{H}_c \in \mathbb{C}^{N_T \times N_R}$ is the complex-valued MIMO channel matrix whose elements follow complex

Gaussian distribution, i.e., $\mathbf{H}_c \sim \mathscr{CN}\left(0, \mathbf{I}_{N_T \times N_R}\right)$, $\mathbf{x}_c \in \mathbb{C}^{N_T \times 1}$ is the vector of transmit symbols, and $\mathbf{n}_c \in \mathbb{C}^{N_R \times 1}$ is the noise vector whose elements follows the complex Gaussian distribution $\mathscr{CN}\left(0, \sigma_n^2 \mathbf{I}_{N_R \times 1}\right)$, where σ_n is the noise power.

We convert the complex-valued model into the corresponding real-valued one for ease of representation and computation. The system model becomes real-valued as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{5.2}$$

where $T = 2N_T$, $R = 2N_R$ and

$$\mathbf{y} = \begin{bmatrix} \mathrm{Re}\left(\mathbf{y}_c\right) \\ \mathrm{Im}\left(\mathbf{y}_c\right) \end{bmatrix} \in \mathbb{R}^{R \times 1} \tag{5.3}$$

$$\mathbf{H} = \begin{bmatrix} \mathrm{Re}\left(\mathbf{H}_c\right) & -\mathrm{Im}\left(\mathbf{H}_c\right) \\ \mathrm{Im}\left(\mathbf{H}_c\right) & \mathrm{Re}\left(\mathbf{H}_c\right) \end{bmatrix} \in \mathbb{R}^{T \times R} \tag{5.4}$$

$$\mathbf{x} = \begin{bmatrix} \mathrm{Re}\left(\mathbf{x}_c\right) \\ \mathrm{Im}\left(\mathbf{x}_c\right) \end{bmatrix} \in \mathbb{R}^{T \times 1} \tag{5.5}$$

$$\mathbf{w} = \begin{bmatrix} \mathrm{Re}\left(\mathbf{w}_c\right) \\ \mathrm{Im}\left(\mathbf{w}_c\right) \end{bmatrix} \in \mathbb{R}^{R \times 1} \tag{5.6}$$

5.2.1 Conventional Detectors

The optimal solution to Eq. (5.2) is the maximum likelihood (ML) detection, which requires a huge calculation complexity.

$$\hat{\mathbf{x}}_{ML} = \arg\min_{\mathbf{x}} \left\|\mathbf{y} - \mathbf{H}\mathbf{x}\right\|_F^2 \tag{5.7}$$

When the number of antennas is huge, the conventional linear detection methods such as matched filter (MF), zero-forcing (ZF), minimum mean square error (MMSE) are more practical in implementation.

$$\begin{aligned} \hat{\mathbf{x}}_{\mathrm{MF}} &= \mathbf{H}^T \mathbf{y} && (5.8) \\ \hat{\mathbf{x}}_{\mathrm{ZF}} &= \left(\mathbf{H}^T \mathbf{H}\right)^{-1} \mathbf{H}^T \mathbf{y} && (5.9) \\ \hat{\mathbf{x}}_{\mathrm{MMSE}} &= \left(\mathbf{H}^T \mathbf{H} + \sigma_n^2 \mathbf{I}_T\right)^{-1} \mathbf{H}^T \mathbf{y} && (5.10) \end{aligned}$$

To tackle the complexity in matrix inversion

We can rewrite (5.10) as

$$\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{b} \tag{5.11}$$

where $\mathbf{A} = \mathbf{H}^T\mathbf{H} + \sigma^2\mathbf{I}_T$ and $\mathbf{b} = \mathbf{H}^T\mathbf{y}$. In massive MIMO systems where the number of antennas are huge, the matrix inversion in Eq. (5.11) is a problem. To approximate or avoid the matrix inversion, many algorithms have been proposed, many of them adopt an iterative fashion, which can be combined with deep learning with the aid of deep unfolding.

5.2.2 Unfold Projected Gradient Descent

With the loss function defined to evaluate the performance as $L(\mathbf{x}) = \left\|\mathbf{y} - \mathbf{H}\mathbf{x}\right\|_F^2$, the well-known gradient descent algorithm allows us to obtain the solution via an iterative fashion:

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{x}}_t - \delta_t \frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} \tag{5.12}$$

which leads to the following iteration:

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{x}}_t - \delta_t \mathbf{H}^T\mathbf{y} + \delta_t \mathbf{H}^T\mathbf{H}\hat{\mathbf{x}}_t \tag{5.13}$$

where δ_t corresponds to the step-size parameter which controls the convergence behavior such as the convergence to a fixed point and the convergence speed.

Unfold the Gradient Descent Detector

Intuitively, we can build a L-layer neural network using the iteration in Eq. (5.13). In t-th layer, the operation of the neural network is expressed as:

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{x}}_t - \mathbf{A}_t\mathbf{H}^T\mathbf{y} + \mathbf{B}_t\mathbf{H}^T\mathbf{H}\hat{\mathbf{x}}_t, t = 1, 2, ..., L \tag{5.14}$$

where the $T \times T$ matrices $\{\mathbf{A}_t, \mathbf{B}_t\}_{t=1}^L$ are considered as trainable parameters, $\hat{\mathbf{x}}_{L+1}$ is the final output.

DetNet

To enhance the performance of the unfolded detector, [134] and [135] proposed a deep detector named as “DetNet" which introduces extra dimensions with the following iteration:

$$\mathbf{z}_t = \mathrm{ReLU}\left(\mathbf{W}_{1t}\begin{bmatrix} \hat{\mathbf{x}}_t - \delta_{1t}\mathbf{H}^T\mathbf{y} - \delta_{2t}\mathbf{H}^T\mathbf{H}\hat{\mathbf{x}}_t \\ \mathbf{v}_t \end{bmatrix} + \mathbf{b}_{1t}\right), \tag{5.15}$$

$$\hat{\mathbf{x}}_{t+1} = \sigma\left(\mathbf{W}_{2t}\mathbf{z}_t + \mathbf{b}_{2t}\right), \tag{5.16}$$

$$\hat{\mathbf{v}}_{t+1} = \mathbf{W}_{3t}\mathbf{z}_t + \mathbf{b}_{3t}, \tag{5.17}$$

where $\sigma(\cdot)$ is a logistic sigmoid. The set of trainable parameters is

$$\{\mathbf{W}_{1t}, \mathbf{W}_{2t}, \mathbf{W}_{3t}, \mathbf{b}_{1t}, \mathbf{b}_{2t}, \mathbf{b}_{3t}, \delta_{1t}, \delta_{2t} : t = 0, ..., T-1\} \tag{5.18}$$

Compared to Eq. (5.13), DetNet introduces extra dimensions by adding a new parameter $\mathbf{v}_t$.

To address the problems in deep network training such as vanishing gradients, saturation of the activation functions, sensitivity to initialization, etc., inspired by [145], DetNet adopted a loss function that takes into account the outputs of all of the layers, which is given by

$$l\left(\mathbf{x}, \hat{\mathbf{x}}_L\right) = \sum_{l=1}^{L} \log(l) \left\|\mathbf{x} - \hat{\mathbf{x}}_L\right\|^2 \tag{5.19}$$

To further enhance the performance, a residual feature from [146] where the output of each layer is a weighted average with the output of the previous layer is added.

$$\hat{\mathbf{x}}_l = \alpha\hat{\mathbf{x}}_{l-1} + (1-\alpha)\hat{\mathbf{x}}_l \tag{5.20}$$

where α is chosen as 0.9 in the implementation.

Overloaded MIMO

In [136, 137], the authors proposed a massive overloaded MIMO detector based on unfolding PG algorithm. The considered PG algorithm is given by

$$\mathbf{r}_t = \hat{\mathbf{x}}_t - \delta_t\mathbf{H}^T\mathbf{y} + \delta_t\mathbf{H}^T\mathbf{H}\hat{\mathbf{x}}_t \tag{5.21}$$

$$\hat{\mathbf{x}}_{t+1} = \tanh\left(\xi\mathbf{r}_t\right) \tag{5.22}$$

Compared to Eq. (5.13), in the projection step Eq. (5.22) a soft projection function tanh(·) is applied to $\mathbf{r}_t$ to obtain the estimate $\hat{\mathbf{x}}_{t+1}$ of the tth iteration. The soft projection ensures that the estimate takes a real value close to ±1. In this projection step, the parameter ξ adjusts the softness of the soft projection.

By unfolding the iterative process in Eq. (5.21) and (5.22), the authors proposed a trainable algorithm given by

$$\begin{aligned} \mathbf{r}_t &= \hat{\mathbf{x}}_t - \delta_t \mathbf{W}\big(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_t\big) && (5.23) \\ \hat{\mathbf{x}}_{t+1} &= \tanh\left(|\xi_t|\,\mathbf{r}_t\right) && (5.24) \end{aligned}$$

where $\mathbf{W} = \mathbf{H}^T\big(\mathbf{H}\mathbf{H}^T\big)^{-1}$ or $\mathbf{W} = \mathbf{H}^T\big(\mathbf{H}\mathbf{H}^T + \sigma_n^2 \mathbf{I}_T\big)^{-1}$. Compared to Eq. (5.21) and Eq. (5.22), $\mathbf{H}^T$ is replaced by $\mathbf{W}$ because this modification improves the BER performance [147]. It is proved that in [137] that the former $\mathbf{W}$ has a poor detection performance when the ratio R/T is relatively large.

5.2.3 Conjugate Gradient Descent

In [139], the authors proposed to use CG in MIMO detection. The estimated signal $\hat{\mathbf{x}}$ for Eq. (5.11) can be obtained using

$$\hat{\mathbf{x}}^{(n+1)} = \hat{\mathbf{x}}^{(n)} + \alpha^{(n)} \mathbf{p}^{(n)} \tag{5.25}$$

where $\alpha^{(n)}$ is a scalar parameter and $\mathbf{p}^{(n)}$ is the conjugate direction with respect to $\mathbf{A}$, i.e.,

$$(\mathbf{p}^{(n)})^H \mathbf{A} \mathbf{p}^{(m)} = 0 \quad \text{for} \quad n \neq m \tag{5.26}$$

More specific, the iteration of the CG detector can be expressed as

1. Initialization: $i = 0, \hat{\mathbf{x}}_0 = \mathbf{0}, \hat{\mathbf{d}}_0 = \hat{\mathbf{r}}_0 = \mathbf{b}$
2. Update the step-size: $\alpha_i = \big(\mathbf{r}_i^H \mathbf{r}_i\big) / \big(\mathbf{r}_i^H \mathbf{A} \mathbf{d}_i\big)$
3. Update the estimate: $\hat{\mathbf{x}}_{i+1} = \hat{\mathbf{x}}_i + \alpha_i \mathbf{d}_i$
4. Update the residual: $\mathbf{r}_i = -f'(\hat{\mathbf{x}}_i) = \mathbf{r}_{i-1} - \alpha_{i-1} \mathbf{A} \mathbf{d}_{i-1}$
5. Update the direction step-size: $\beta_i = \big(\mathbf{r}_{i+1}^H \mathbf{r}_{i+1}\big) / \big(\mathbf{r}_i^H \mathbf{r}_i\big)$
6. Update the direction: $\mathbf{d}_i = \mathbf{r}_i + \beta_{i-1} \mathbf{d}_{i-1}$
7. $i = i + 1$

In [140], the authors proposed to use deep unfolding to improve the CG detector. In the proposed learned CG detector, the update of step-sizes in Eq. (5.27) is implemented using deep learning.

$$\begin{aligned} \alpha_i &= \big(\mathbf{r}_i^H \mathbf{r}_i\big) / \big(\mathbf{r}_i^H \mathbf{A} \mathbf{d}_i\big) \\ \beta_i &= \big(\mathbf{r}_{i+1}^H \mathbf{r}_{i+1}\big) / \big(\mathbf{r}_i^H \mathbf{r}_i\big) \end{aligned} \tag{5.27}$$

In the deep network, the performance is further improved using vector step-sizes, where the trainable variables are:

$$\Theta_v^{(i)} = \{\boldsymbol{\alpha}_r^{(i)} \in {}^{2N_t\times 1}, \boldsymbol{\beta}_r^{(i)} \in {}^{2N_t\times 1}\} \tag{5.28}$$

5.2.4 Deep Unfolding the generic iterative algorithms

The MIMO detectors based on Richardson iteration [129, 130], successive over-relaxation (SOR) [131, 132] and Jacobi iteration [133] all have a similar iterative structure. Exploiting the idea of Deep Unfolding, the parameters in the detectors can be set random initially and trained using a DNN. The deep learning based detector exhibits a better performance than the iterative detectors in [129, 130, 131, 132, 133].

(1) Generic Iterative Detectors

Richardson Iteration In [129, 130], an iterative MIMO detector based on the Richardson iteration is given by

$$\hat{\mathbf{x}}_i = \hat{\mathbf{x}}_{i-1} + \omega\{\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}_{i-1}\} \tag{5.29}$$

where ω is a constant scale factor.

SOR In [131, 132], the iterative algorithm based on SOR can be expressed as

$$((1/\omega)\mathbf{D} + \mathbf{L})\hat{\mathbf{x}}_i = \mathbf{b} + \left\{(1/\omega - 1)\mathbf{D} - \mathbf{L}^H\right\}\hat{\mathbf{x}}_{i-1} \tag{5.30}$$

where ω is a constant scale factor, $\mathbf{D}$ is the main diagonal of the MMSE filtering matrix $\mathbf{A}$ in Eq. (5.11), $\mathbf{L}$ is the lower triangular matrix of $\mathbf{A}$.

Jacobi Iteration In [133], the iterative detector based on Jacobi iteration is given by

$$\hat{\mathbf{x}}_i = \omega\left(-\mathbf{D}^{-1}\mathbf{E}\hat{\mathbf{x}}_{i-1} + \mathbf{D}^{-1}\mathbf{b}\right) + (1-\omega)\hat{\mathbf{x}}_{i-1} \tag{5.31}$$

where $\mathbf{E}$ is the off-diagonal of $\mathbf{A}$, i.e., $\mathbf{E} = \mathbf{A} - \mathbf{D}$.

(2) Unfold the Iterations

The above iterative detectors all share the same structure:

$$\hat{\mathbf{x}}_i = \mathbf{A}\hat{\mathbf{x}}_{i-1} + \mathbf{B}\mathbf{b}, \tag{5.32}$$

which inspires a T-layer neural network whose each layer is defined as

$$\hat{\mathbf{x}}_{t+1} = \mathbf{A}_t \hat{\mathbf{x}}_t + \mathbf{B}_t \mathbf{b} \tag{5.33}$$

where $\{\mathbf{A}_t, \mathbf{B}_t\}_{t=1}^{T}$ are trainable parameters.

5.2.5 Proximal Gradient Descent

Instead of solving the unconstrained least squares estimation in Eq. (5.11), a popular method is to solve the regularized least squares estimation by adding a penalty term. The added regularization function penalizes values that are less likely or desirable as solutions, forces $\hat{\mathbf{x}}$ to be close to the mean value $\mathbb{E}(\mathbf{x})$. The well-known LASSO regression [148] is given by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 + \lambda \|\mathbf{x}\|_1 \tag{5.34}$$

To solve the regularized least squares estimation problem, proximal gradient methods such as iterative ISTA, approximate message passing (AMP) and ADMM are standard approaches which avoid the over-fitting problem.

ISTA

To solve (5.34), a simple solution is ISTA [149]:

$$\begin{aligned} \mathbf{v}_t &= \mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_t \\ \hat{\mathbf{x}}_{t+1} &= \eta_{st}\left(\hat{\mathbf{x}}_t + \mathbf{H}^T \mathbf{v}_t; \lambda\right) \end{aligned} \tag{5.35}$$

where η_{st} is the soft thresholding shrinkage function, which is defined component-wise as

$$\eta_{st}(x, \lambda) = \mathrm{sign}(\mathrm{x}) \cdot \max(|\mathrm{x}| - \lambda, 0) \tag{5.36}$$

With the aid of deep unfolding, the t-th layer of the unfolded ISTA network [143] can be expressed as

$$\begin{aligned} \mathbf{v}_t &= \mathbf{y} - \mathbf{A}_t \hat{\mathbf{x}}_t, \\ \hat{\mathbf{x}}_{t+1} &= \eta_{st}(\hat{\mathbf{x}}_t + \mathbf{B}_t \mathbf{v}_t; \lambda_t) \end{aligned} \tag{5.37}$$

where $\{\mathbf{A}_t, \mathbf{B}_t, \lambda_t\}_{t=1}^{T}$ are trainable parameters.

AMP

To enhance the performance of ISTA, [150] proposed an AMP algorithm, which can be expressed as

$$\begin{aligned} \mathbf{v}_t &= \mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_t + b_t \mathbf{v}_{t-1}, \\ \hat{\mathbf{x}}_{t+1} &= \eta_{st}\left(\hat{\mathbf{x}}_t + \mathbf{H}^T \mathbf{v}_t; \lambda_t\right) \end{aligned} \tag{5.38}$$

where $b_t = \|\hat{\mathbf{x}}_t\|_0 / R$, $\lambda_t = \alpha \|\mathbf{v}_t\|_2 / \sqrt{R}$, α is the tuning parameter. Compared to (5.35), the added term $b_t \mathbf{v}_{t-1}$ is called Onsager correction, which aims to decouple the errors across iterations. Similarly, the t-th layer of the unfolded AMP network [143] can be written as

$$\begin{aligned} \mathbf{v}_{t+1} &= \mathbf{y} - \mathbf{A}_t \hat{\mathbf{x}}_{t+1} + b_{t+1} \mathbf{v}_t, \\ \hat{\mathbf{x}}_{t+1} &= \eta_{st}\left(\hat{\mathbf{x}}_t + \mathbf{B}_t \mathbf{v}_t; \alpha_t \|\mathbf{v}_t\|_2 / \sqrt{N_R}\right) \end{aligned} \tag{5.39}$$

where $\{\mathbf{A}_t, \mathbf{B}_t, \alpha_t\}_{t=1}^T$ are trainable parameters.

ADMM

The alternating direction method of multipliers (ADMM) [141] is an algorithm that solves convex optimization problems. More specific, ADMM is for problems with an objective function that is the sum of two separate convex functions with linear constraints. In [144], the authors proposed to unfold ADMM. The MIMO detector based on ADMM can be expressed as

$$\begin{aligned} \mathbf{z}^{(k)} &= \operatorname{sign}\left(\mathbf{x}^{(k-1)} - \mathbf{u}^{(k-1)}\right), \\ \mathbf{u}^{(k)} &= \mathbf{z}^{(k)} - \mathbf{x}^{(k-1)} + \mathbf{u}^{(k-1)}, \\ \mathbf{x}^{(k)} &= \left(\operatorname{diag}(\lambda) + \mathbf{H}^T \mathbf{H}\right)^{-1} \left(\mathbf{H}^T \mathbf{y} + \lambda \odot \left(\mathbf{z}^{(k)} + \mathbf{u}^{(k)}\right)\right) \end{aligned} \tag{5.40}$$

where $\lambda = [\lambda_1, ..., \lambda_T]^T$ is the non-negative penalty parameter and $\odot$ denotes the Hardmard product.

To unfold the iterations, firstly the sign function of each iteration is replaced by a projection function given by

$$\Upsilon_{t,\beta}(\mathbf{x}) = -\beta + \frac{\operatorname{ReLU}\left(\mathbf{x} + \beta \odot \mathbf{t}\right)}{\mathbf{t}} - \frac{\operatorname{ReLU}\left(\mathbf{x} - \beta \odot \mathbf{t}\right)}{\mathbf{t}} \tag{5.41}$$

The operation of the k-th layer of the unfolded network ($k = 1, ..., L$) can be expressed as

$$\begin{aligned} \mathbf{z}^{(k)} &= \Upsilon_{t_{k-1}, \beta_k}\left(\mathbf{x}^{(k-1)} - \mathbf{u}^{(k-1)}\right), \\ \mathbf{u}^{(k)} &= \mathbf{z}^{(k)} - \mathbf{x}^{(k-1)} + \mathbf{u}^{(k-1)}, \\ \mathbf{x}^{(k)} &= \mathbf{D}\left(\mathbf{H}^T \mathbf{y} + \left(\mathbf{p} \odot \omega_k\right) \odot \left(\mathbf{z}^{(k)} + \mathbf{u}^{(k)}\right)\right) \end{aligned} \tag{5.42}$$

where $D = \left(\text{diag}\left(\omega_0 \odot \mathbf{p}\right) + \mathbf{H}^T\mathbf{H}\right)^{-1}$ and $\mathbf{x}^{(0)} = \mathbf{0}, \mathbf{u}^{(0)} = \mathbf{0}, \mathbf{p} = \text{diag}\left(\mathbf{H}^T\mathbf{H}\right)$. The final output is $\hat{\mathbf{x}}(\theta) = \Upsilon_{t_L,1}\left(\mathbf{x}^{(L)}\right)$.

The trainable parameters are $\theta = \left\{\{\omega_l, \mathbf{t}_l\}_{l=0}^{L}, \{\beta_k\}_{k=1}^{L}\right\}$ and the loss function is given by

$$l\left(\mathbf{s};\hat{\mathbf{x}}(\theta)\right) = \sum_{n=1}^{N} \|\mathbf{h}_n\|_2^2 \left(x_n - \hat{x}_n(\theta)\right)^2 \tag{5.43}$$

5.2.6 Performance Comparison

In this section, we simulated all the unfolded detectors to compare the performance. In all the experiments, the unfolded models are trained using TensorFlow [151]. The learning rate is set as 0.0001. The channel matrix is fixed and constant. The training SNR is a fixed high SNR (e.g., 30dB).

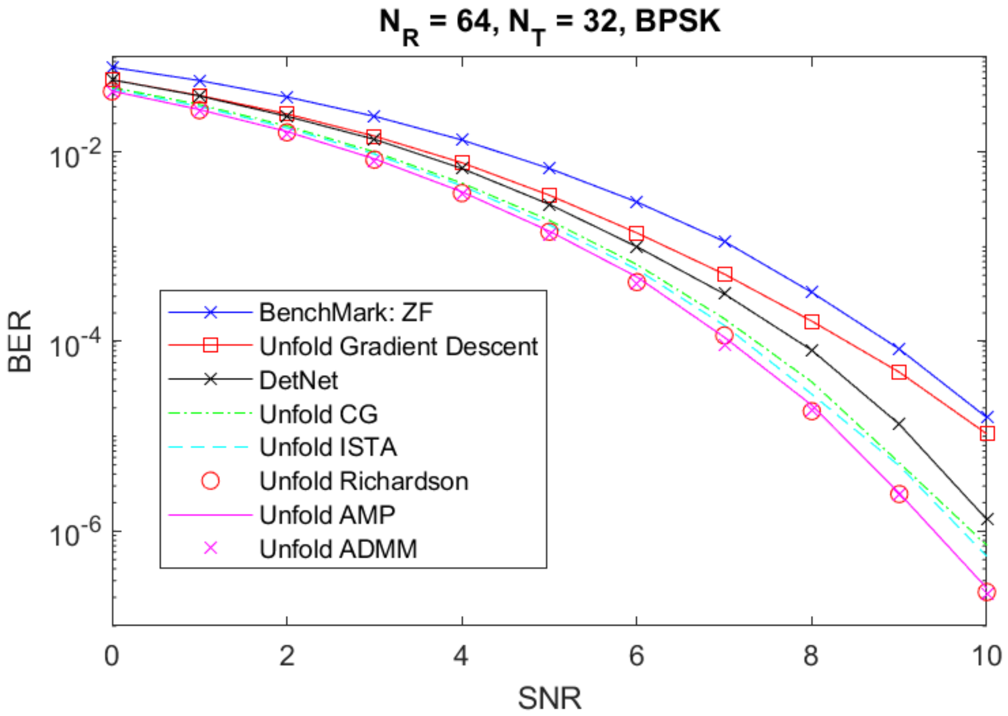


Figure 5.1 – $N_T = 32, N_R = 64$, BPSK. BER performance comparison of unfolded detectors.

The BER performance comparison of different unfolded detectors is given in Fig. 5.1. We can observe that, of all the unfolded detectors we investigated, unfolded Richardson, AMP and ADMM performs better performance.

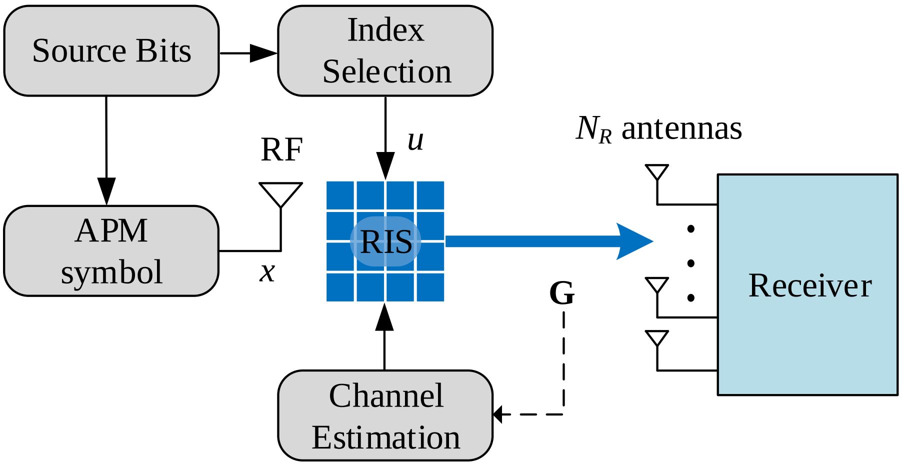


Figure 5.2 – The framework of an RIS-aided SM system.

5.3 RIS-SSK/SM

5.3.1 System Model

Based on the concept of RIS-AP introduced in [152], RIS-SSK/SM scheme is proposed in [101]. A schematic diagram of RIS-aided SM system is given in Fig. 5.2. The RIS is a part of the transmitter, it reflects the signals generated by a near RF source in a deliberate manner to convey information bits. We consider a RIS consists of N passive and low-cost reflector elements (reconfigurable meta-surfaces) and a receiver with N_R antennas, The receiver lies in the far-field of the RIS and does not receive transmission from the RF source. The wireless fading channel between the l-th receive antenna of the receiver and i-th reflector element is characterized by $g_{l,i} = \beta_{l,i} e^{-j\psi_{l,i}}$ for $l = 1, \cdots, N_R$ and $i = 1, \cdots, N$, and follows $\mathcal{CN}(0,1)$ distribution under the assumption of flat Rayleigh fading channels, where $\beta_{l,i}$ and $\psi_{l,i}$ are the absolute value and phase of $g_{l,i}$, respectively. For intelligent reflection, the RIS has the knowledge of channel phases $\psi_{l,i}$ for all l and i. We also assume that all wireless channels are uncorrelated and the receiver has perfect channel state information (CSI).

In each transmission of RIS-SSK, $\log_2(N_R)$ bits are utilized to choose one out of N_R receive antennas. After the antenna selection, the RIS phase shifts are adjusted to maximize the SNR at the target receive antenna. Assuming the chosen antenna index is u, the phase shifts of the RIS are adjusted as

$$\Psi = [e^{j\phi_{u,1}}, ..., e^{j\phi_{u,N}}]^T \in \mathbb{C}^{N\times 1} \tag{5.44}$$

where $\phi_{u,i} = -\arg(g_{u,i}) = \psi_{u,i}$ for $i = 1, \cdots, N$.

To further improve the spectral efficiency, an RIS-SM system conveys extra $\log_2 M$ bits in each transmission to modulate a M-PSK/QAM symbol, which is transmitted from the RF source to the RIS.

Denote $\mathbf{G} = \left\{g_{l,i}, l = 1, \cdots, N_R, i = 1, \cdots, N\right\}$ as the channel matrix between the RIS and the base station (BS) with N_R receiver antennas, the received signal $\bar{\mathbf{y}} \in \mathbb{C}^{N_R \times 1}$ can be expressed as

$$\text{RIS-SSK:} \quad \bar{\mathbf{y}} = \mathbf{G}\Psi + \bar{\mathbf{n}} \tag{5.45}$$

$$\text{RIS-SM:} \quad \bar{\mathbf{y}} = \mathbf{G}\Psi x + \bar{\mathbf{n}} \tag{5.46}$$

where $\bar{\mathbf{n}} \in \mathbb{C}^{N_R \times 1}$ is the additive white Gaussian noise (AWGN), which follows $\mathscr{CN}\left(\mathbf{0}_{N_R \times 1}, N_0 \mathbf{I}_{N_R \times 1}\right)$ with N_0 denoting the noise power. In RIS-SM, $x \in \mathbb{S}$ is the modulated M-PSK/QAM signal drawn from a discrete constellation $\mathbb{S}$.

5.3.2 ML and Greedy Detection

The maximum likelihood (ML) detector for RIS-SSK/SM [101] is given as

$$\hat{u}_{\text{ML}}^{\text{RIS-SSK}} = \arg \min_{u \in \{1,\ldots,N_R\}} \left\| \bar{\mathbf{y}} - \mathbf{G}\left[e^{j\phi_{u,1}}, \ldots, e^{j\phi_{u,N}}\right]^T \right\|_2^2 \tag{5.47}$$

$$\begin{aligned} &\left(\hat{u}_{\text{ML}}^{\text{RIS-SM}}, \hat{x}_{\text{ML}}^{\text{RIS-SM}}\right) \\ &= \arg \min_{u \in \{1,\ldots,N_R\}, x \in \mathbb{S}} \left\| \bar{\mathbf{y}} - \mathbf{G}\left[e^{j\phi_{u,1}}, \ldots, e^{j\phi_{u,N}}\right]^T x \right\|_2^2 \end{aligned} \tag{5.48}$$

The authors in [101] also proposed a greedy detector (GD) with reduced complexity to estimate the activated antenna index, which is given as

$$\hat{u}_{\text{GD}}^{\text{RIS-SSK/SM}} = \arg\max_i \left|y_i\right| \tag{5.49}$$

where y_i is the i-th element of the received $\bar{\mathbf{y}}$. After the estimation of u using (5.49), an estimation of x in RIS-SM can be obtained as

$$\hat{x}_{\text{GD}}^{\text{RIS-SM}} = \arg\min_x \left| y_u - x \sum_{i=1}^{N} \beta_{u,i} \right|^2 \tag{5.50}$$

5.3.3 RIS-SM: A special case of conventional SM

Let $\mathbf{B}$ be a $N \times N_R$ matrix whose elements $\left\{b_{l,i}\right\}$ are described as

$$b_{l,i} = g_{l,i}^* / \left|g_{l,i}\right| = e^{j\psi_{l,i}}, \tag{5.51}$$

then we can rewrite (5.45) and (5.46) as

$$\bar{\mathbf{y}} = \mathbf{G}\mathbf{B}\bar{\mathbf{c}} + \bar{\mathbf{n}}, \tag{5.52}$$

where $\bar{\mathbf{c}}$ is defined for RIS-SSK and RIS-SM separately as

$$\begin{aligned} &\text{RIS-SSK:} \quad \bar{\mathbf{c}} = \left[0, \ldots, \underbrace{1}_{u-\text{th}}, \ldots, 0\right]^T \\ &\text{RIS-SM:} \quad \bar{\mathbf{c}} = \left[0, \ldots, \underbrace{x}_{u-\text{th}}, \ldots, 0\right]^T \end{aligned} \tag{5.53}$$

Thus, (5.45) and (5.46) can be written as

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}_{eq}\bar{\mathbf{c}} + \bar{\mathbf{n}}, \tag{5.54}$$

where $\bar{\mathbf{H}}_{eq}$ is the equivalent channel matrix defined as

$$\bar{\mathbf{H}}_{eq} = \mathbf{G}\mathbf{B}. \tag{5.55}$$

In fact, under the assumption that $\mathbf{G}$ being flat Rayleigh fading channel, when $N \gg 1$, for $i, j = 1, \cdots, N_R$ and $j \neq i$, we can obtain the distribution (see Appendix 5.7.1) of the diagonal and non-diagonal elements of $\bar{\mathbf{H}}_{eq}$ as

$$\begin{aligned} &h_{ii} \sim \mathcal{N}\left(\frac{N}{2}\sqrt{\pi}, \frac{N}{4}(4-\pi)\right), \\ &h_{ij} \sim \mathcal{CN}(0, N), \end{aligned} \tag{5.56}$$

Therefore, when $N \gg 1$, the distortion of $\bar{\mathbf{H}}_{eq}$ mainly comes from the diagonal element, which explains the advantage to adopt greedy detection in RIS-SSK/SM.

From the expression in (5.54), we can observe that the RIS-SM system is equivalent to the conventional SM system with the difference in the channel matrix. Therefore, the detection of RIS-SM can utilize the variety of detection algorithms in conventional SM systems.

5.4 Detection based on deep learning

In this section, we introduce data-driven and model-driven DL detectors separately. To avoid handling complex values in detection, (5.54) is re-parameterized into a real-valued signal model as

$$\mathbf{y} = \mathbf{H}\mathbf{c} + \mathbf{n}, \tag{5.57}$$

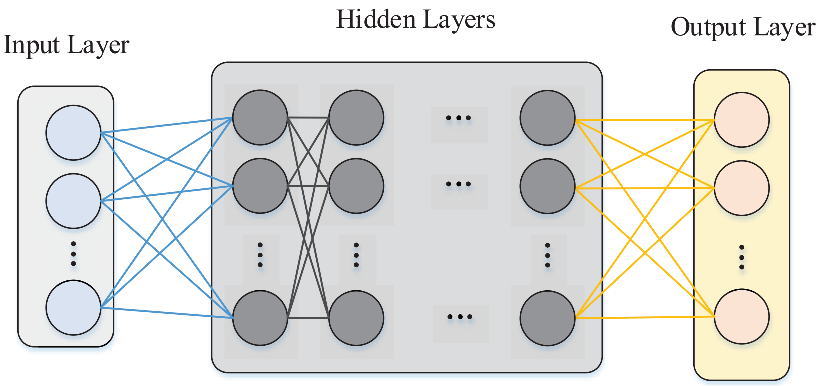


Figure 5.3 – Network structures for the data-driven DL detectors.

where

$$\mathbf{y} = \begin{bmatrix} \mathrm{Re}\left(\bar{\mathbf{y}}\right) \\ \mathrm{Im}\left(\bar{\mathbf{y}}\right) \end{bmatrix} \in \mathbb{R}^{2N_R \times 1} \tag{5.58}$$

$$\mathbf{c} = \begin{bmatrix} \mathrm{Re}\left(\bar{\mathbf{c}}\right) \\ \mathrm{Im}\left(\bar{\mathbf{c}}\right) \end{bmatrix} \in \mathbb{R}^{2N_R \times 1} \tag{5.59}$$

$$\mathbf{n} = \begin{bmatrix} \mathrm{Re}\left(\bar{\mathbf{n}}\right) \\ \mathrm{Im}\left(\bar{\mathbf{n}}\right) \end{bmatrix} \in \mathbb{R}^{2N_R \times 1} \tag{5.60}$$

and

$$\mathbf{H} = \begin{bmatrix} \mathrm{Re}\left(\bar{\mathbf{H}}_{eq}\right) & -\mathrm{Im}\left(\bar{\mathbf{H}}_{eq}\right) \\ \mathrm{Im}\left(\bar{\mathbf{H}}_{eq}\right) & \mathrm{Re}\left(\bar{\mathbf{H}}_{eq}\right) \end{bmatrix} \in \mathbb{R}^{2N_R \times 2N_R} \tag{5.61}$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary parts separately. In the following, we introduce the data-driven and model-driven DL detectors for RIS-SSK/SM, respectively.

5.4.1 Data-driven DL Detector

We consider a data-driven DL detector with a fully-connected ReLU DNN, the network structure is illustrated in Fig. 5.3. The deep neural network (DNN) of the detector consists of input and output layers, $l \in \mathbb{N}$ hidden layers and neuron assignment $\mathbf{d} = (d_0, d_1, ..., d_l, d_{l+1}) \in \mathbb{N}^{l+1}$.

The number of nodes in the input layer is determined by the knowledge of CSI. Denote the output of the DNN as $\mathbf{b}$, and the input of the DNN is $\mathbf{a}$. When CSI is unknown, the input of the DNN is $\mathbf{a} = \{\mathbf{y}\}$, thus we have $d_0 = 2N_R$ for the input layer of RIS-SSK/SM. When CSI is available, the input of the DNN is $\mathbf{a} = \{\mathbf{y}, vec(\mathbf{H})\}$, where $vec(\cdot)$ stands for vectorization. In this case, we have $d_0 = 2N_R + 4N_R^2$ for the input layer of RIS-SSK/SM.

The binary cross entropy (BCE) between the output $\mathbf{b}$ of the DNN and the one-hot encoding of $\bar{\mathbf{c}}$ is adopted as the loss function, which can be expressed as

$$f_{loss} = \mathrm{BCE}\left(\mathbf{b}, \bar{\mathbf{b}}\right) \tag{5.62}$$

where $\bar{\mathbf{b}}$ is the one-hot encoding of $\bar{\mathbf{c}}$ and is defined as

$$\bar{\mathbf{b}} = \left[0, \ldots, \underbrace{1}_{v-\mathrm{th}}, \ldots, 0\right]^T \tag{5.63}$$

In RIS-SSK, v is the chosen antenna index u. In RIS-SM, when the modulated signal x is the m-th constellation in $\mathbb{S}$, we have $v = M(u-1) + m$. Therefore, we have $d_{l+1} = N_R$ for RIS-SSK and $d_{l+1} = MN_R$ for RIS-SM.

The set of all trainable parameters of the DNN can be expressed as

$$\Theta = \{\mathrm{vec}\left(\mathbf{W}_i\right), \mathbf{b}_i\}_{i=0}^{l}, \tag{5.64}$$

where $\mathbf{W}_i \in \mathbb{R}^{d_{i+1} \times d_i}$ is the weight matrix connecting the i-th layer to the $(i+1)$-th layer, and $\mathbf{b}_i \in \mathbb{R}^{d_{i+1}}$ is the bias vector of the $(i+1)$-th layer for $i \in \{0, 1, \ldots, l\}$. The total number of trainable parameters is $\sum_{i=0}^{l} d_{i+1}\left(d_i + 1\right)$.

For the activation function, we choose rectified linear units (ReLU) for the hidden layers and softmax for the output layer. Finally, the DNN can be expressed as

$$\mathbf{b} = \psi_{d_{l+1}}\left(\mathscr{A}_l\left(\varphi_{d_l}\left(\mathscr{A}_{l-1}\left(\varphi_{d_{l-1}}\left(\cdots\left(\varphi_{d_1}\left(\mathscr{A}_0\left(\mathbf{a}\right)\right)\right)\right)\right)\right)\right)\right) \tag{5.65}$$

where $\psi_{d_{l+1}}$: $\mathbb{R}^{d_{l+1}} \to \mathbb{R}^{d_{l+1}}$ is the entry-wise softmax function, $\mathscr{A}_i$: $\mathbb{R}^{d_i} \to \mathbb{R}^{d_{i+1}}$ is the affine transformation with weight $\mathbf{W}_i$ and bias $\mathbf{b}_i$ and φ_{d_i}: $\mathbb{R}^{d_i} \to \mathbb{R}^{d_i}$ is the entry-wise rectified linear units (ReLU) activation function for $i \in \{0, \ldots, l\}$.

For ease of notation, we denote the data-driven fully connected DL detector without/with CSI as $\mathscr{D}_1$ and $\mathscr{D}_2$, respectively.

5.4.2 Model-driven DL Detector with CSI

In this section we introduce the model-driven DL detector based on deep unfolding. Based on the idea of well-known gradient descent, we can obtain the estimated $\hat{\mathbf{c}}$ in an iterative fashion

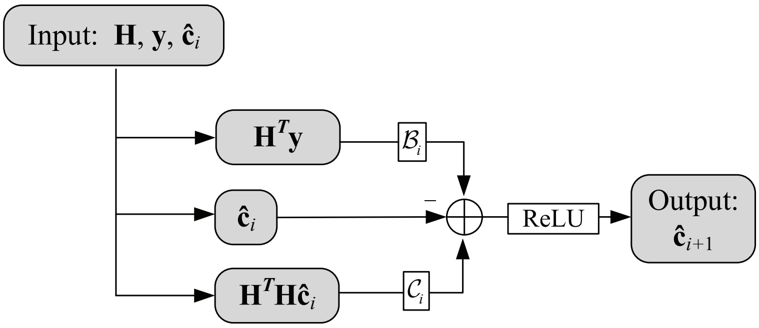


Figure 5.4 – Schematic diagram of the i-th iteration of $\mathscr{D}_3$.

as:

$$\hat{\mathbf{c}}_{i+1} = \hat{\mathbf{c}}_i - \delta_i \frac{\partial \left\| \mathbf{y} - \mathbf{H}\mathbf{c} \right\|^2}{\partial \mathbf{c}} \Big|_{\mathbf{c}=\hat{\mathbf{c}}_i} \tag{5.66}$$

where δ_i is the step-size of iteration, $\mathbf{c}_1$ can be initialized as zeros. Thus we obtain

$$\hat{\mathbf{c}}_{i+1} = \hat{\mathbf{c}}_i + 2\delta_i \mathbf{H}^T\mathbf{y} - 2\delta_i \mathbf{H}^T\mathbf{H}\hat{\mathbf{c}}_i \tag{5.67}$$

Consequently, following the idea of deep unfolding, we can design a L-layer neural network by unfolding the iterations in (5.67) as:

$$\hat{\mathbf{c}}_{i+1} = \varphi\left(\hat{\mathbf{c}}_i - \mathscr{B}_i\left(\mathbf{H}^T\mathbf{y}\right) + \mathscr{C}_i\left(\mathbf{H}^T\mathbf{H}\hat{\mathbf{c}}_i\right)\right) \tag{5.68}$$

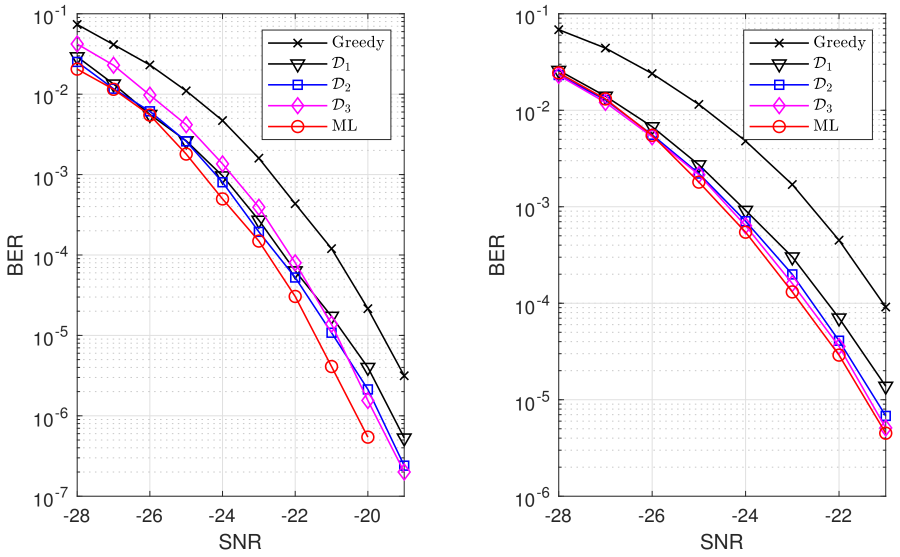
where φ: $\mathbb{R}^{2N_R} \to \mathbb{R}^{2N_R}$ is the entry-wise ReLU activation function for $i \in \{0, ..., l\}$, $\mathscr{B}_i, \mathscr{C}_i$: $\mathbb{R}^{2N_R} \to \mathbb{R}^{2N_R}$ are the affine transformations with weight $\mathbf{W}_{1,i}, \mathbf{W}_{2,i}$ and bias $\mathbf{b}_{1,i}, \mathbf{b}_{2,i}$, respectively. For simplicity, we denote the model-driven detector as $\mathscr{D}_3$. In Fig. 5.4, we plot the i-th iteration of $\mathscr{D}_3$, the whole signal-flow graph of $\mathscr{D}_3$ can be obtained by stack multiple Fig. 5.4.

The set of trainable parameters of $\mathscr{D}_3$ is given as

$$\Theta = \left\{\text{vec}\left(\mathbf{W}_{1,i}\right), \mathbf{b}_{1,i}, \text{vec}\left(\mathbf{W}_{2,i}\right), \mathbf{b}_{2,i}\right\}_{i=1}^{L} \tag{5.69}$$

The total number of trainable parameters is $2N_R L(2N_R + 1)$. The loss function is defined as the mean squared error (MSE) of the estimation $\hat{\mathbf{c}}$ and the original vector $\mathbf{c}$, which is given as

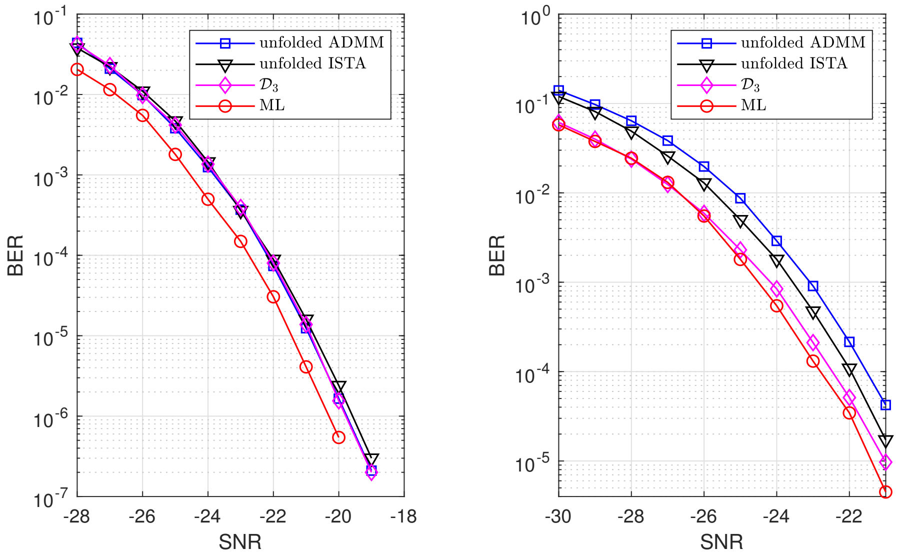
$$f_{loss} = \text{MSE}\left(\mathbf{c}, \hat{\mathbf{c}}\right) \tag{5.70}$$



(a) Over time-invariant channel.

(b) Over time-varing channel.

Figure 5.5 – The BER performance of the DL detectors and conventional ML and greedy detectors for RIS-SSK system with $N_R = 4$ and $N = 64$.



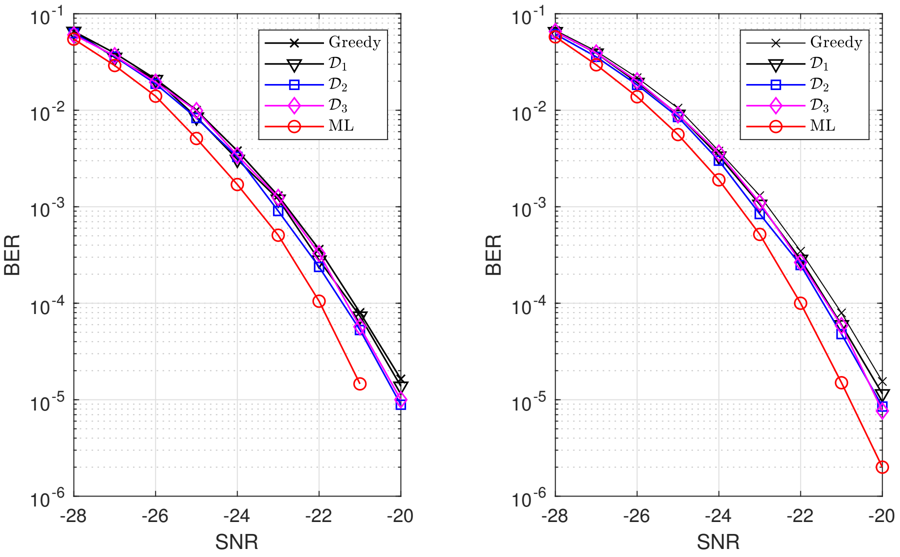
(a) Over time-invariant channel.

(b) Over time-varing channel.

Figure 5.6 – The BER performance of the model-driven DL detectors for RIS-SSK system with $N_R = 4$ and $N = 64$.

5.5 Numerical results

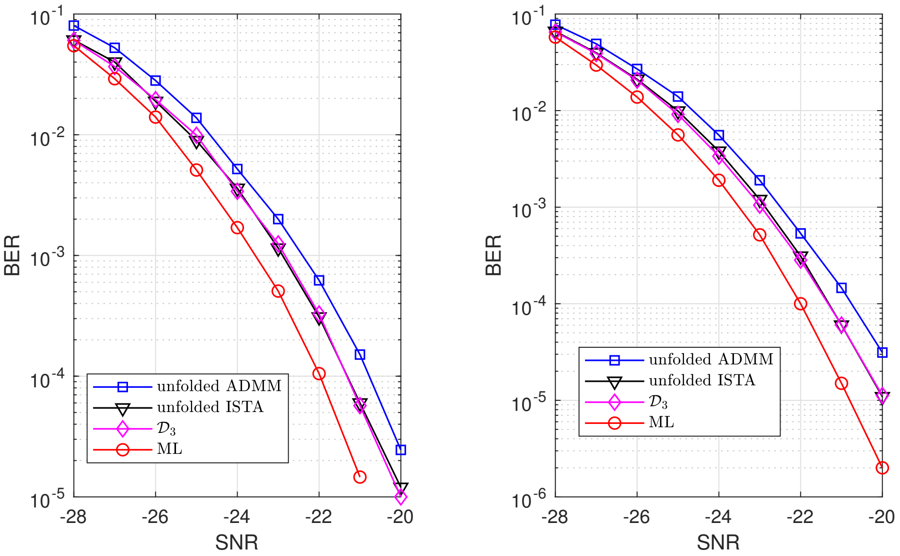
In this section, the performance of the DL detectors is evaluated by computer experiments. Time-invariant and time-varying Rayleigh fading channels are investigated. For all the DL



(a) Over time-invariant channel.

(b) Over time-varing channel.

Figure 5.7 – The BER performance of the DL detectors and conventional ML and greedy detectors for RIS-SM system with $N_R = 4$ and $N = 64$.



(a) Over time-invariant channel.

(b) Over time-varing channel.

Figure 5.8 – The BER performance of the model-driven DL detectors for RIS-SM system with $N_R = 4$ and $N = 64$.

detectors, we choose Adam as the optimizer [153]. The learning rate is 0.001 for all the DL detectors. The data-driven DNNs ($\mathscr{D}_1$ and $\mathscr{D}_2$) have 3 hidden layers. The model-driven DNN

($\mathscr{D}_3$) has 4 layers (or iterations).

In Fig. 5.5, we compare the BER performance in RIS-SSK systems over time-invariant and time-varying channels. The model-based greedy, ML detectors and data/model-driven DL detectors are investigated and compared in a RIS-SSK system with $N_R = 4$ and $N = 64$. We have $(d_i, d_1, d_2, d_3, d_4) = (8, 72, 144, 72, 4)$ for $\mathscr{D}_1$, $(d_i, d_1, d_2, d_3, d_4) = (72, 72, 144, 72, 4)$ for $\mathscr{D}_2$.

From Fig. 5.5, we can observe that the proposed DL detectors outperform the greedy detector. With CSI, $\mathscr{D}_2$ exhibits a better performance than $\mathscr{D}_1$. The performance of $\mathscr{D}_3$ is better than $\mathscr{D}_1$ and $\mathscr{D}_2$ with fewer training samples.

The numbers of all the trainable parameters are 21892, 26500 and 288 for $\mathscr{D}_1$, $\mathscr{D}_2$ and $\mathscr{D}_3$, respectively. Because that $\mathscr{D}_3$ has a smaller set of trainable parameter, $\mathscr{D}_1$ and $\mathscr{D}_2$ need much more training batches to converge, in the simulation the number of training batches is 20000 for $\mathscr{D}_1$ and $\mathscr{D}_2$ and 10000 for $\mathscr{D}_3$. Each training batch consists of 5000 channel realizations.

Fig. 5.6 gives a BER comparison of different model-driven DL detectors for RIS-SSK over time-invariant and time-varying channels. The performance of unfolded ISTA detector [154] and unfolded ADMM detector [144] are investigated. We can observe that our proposed model-driven DL detector has better performance against the other model-driven DL detectors. In Fig. 5.7, similar performance comparisons are given for RIS-SM system over time-invariant and time-varying channels. We have $(d_i, d_1, d_2, d_3, d_4) = (8, 144, 288, 72, 4)$ for $\mathscr{D}_1$, $(d_i, d_1, d_2, d_3, d_4) = (72, 144, 288, 72, 4)$ for $\mathscr{D}_2$. In Fig. 5.8, we compare the BER performance of different model-driven DL detectors for RIS-SM with $N_R = 4$ and $N = 64$. Compared with unfolded ISTA and ADMM detector, $\mathscr{D}_3$ has a better performance.

5.6 Conclusion

In this chapter, we proposed to detect the RIS-SSK/SM systems with the aid of data-driven and model-driven deep learning. The numerical results show that our proposed data-driven DL detector and model-driven DL detectors have better performance against conventional greedy detection. Besides, with a simple design, our proposed model-driven DL detector exhibits a promising performance against the unfolded ISTA and ADMM detectors.

5.7 appendices

5.7.1 Distribution of $\mathbf{H}_{eq}$ when $N \gg 1$

Denote $g_{i,j}, h_{i,j}$ as the (i, j) element of $\mathbf{G}$, $\mathbf{H}_{eq}$, respectively. From (5.55), we have

$$h_{ii} = \sum_{k=1}^{N} |g_{ik}| \tag{5.71}$$

$$h_{ij} = \sum_{k=1}^{N} \frac{g_{ik} g_{jk}^{*}}{|g_{jk}|} \tag{5.72}$$

To better analyze the equivalent channel matrix, we give the following lemmas.

Lemma 5.1 *Given* $g \sim \mathcal{CN}(0,1)$, *we have* $|g| \sim \text{Rayleigh}\left(\sqrt{1/2}\right)$, *whose moments are given as* $\mathbb{E}\left(|g|^{n}\right) = \Gamma(1 + n/2)$.

Proof 5.1 *Omitted.*

Lemma 5.2 *If* $N \gg 1$, $h_{ii} \sim \mathcal{N}\left(\frac{N}{2}\sqrt{\pi}, \frac{N}{4}(4-\pi)\right)$.

Proof 5.2 *From 5.1, we have*

$$\mathbb{E}(h_{ii}) = \mathbb{E}\left(\sum_{k=1}^{N} |g_{ik}|\right) = \sum_{k=1}^{N} \mathbb{E}|g_{ik}| = \frac{N}{2}\sqrt{\pi}. \tag{5.73}$$

Because

$$\begin{aligned} h_{ii}^{2} &= \sum_{k=1}^{N} \sum_{l=1}^{N} |g_{ik}||g_{il}| \\ &= \sum_{k=1}^{N} |g_{ik}|^{2} + \sum_{k=1}^{N} \sum_{l \neq k}^{N} |g_{ik}||g_{il}|, \end{aligned} \tag{5.74}$$

we have

$$\begin{aligned} &\mathbb{E}\left(h_{ii}^{2}\right) = \mathbb{E}\left(\sum_{k=1}^{N} |g_{ik}|^{2}\right) + \mathbb{E}\left(\sum_{k=1}^{N} \sum_{l \neq k}^{N} |g_{ik}||g_{il}|\right) \\ &= \sum_{k=1}^{N} \mathbb{E}\left(|g_{ik}|^{2}\right) + \sum_{k=1}^{N} \sum_{l \neq k}^{N} \mathbb{E}\left(|g_{ik}|\right) E\left(|g_{il}|\right) \\ &= N + N(N-1)\frac{\pi}{4} \end{aligned} \tag{5.75}$$

$$\mathbb{V}(h_{ii}) = \mathbb{E}\left(h_{ii}^{2}\right) - [\mathbb{E}(h_{ii})]^{2} = N - \frac{N}{4}\pi \tag{5.76}$$

From central limit theorem, when $N \gg 1$, *we have* $h_{ii} \sim \mathcal{N}\left(\frac{N}{2}\sqrt{\pi}, \frac{N}{4}(4-\pi)\right)$.

Lemma 5.3 *If* $N \gg 1$, $h_{ij} \sim \mathcal{CN}(0, N)$ *for* $i \neq j$.

Proof 5.3 $\mathbb{E}\left(h_{ij}\right) = \mathbb{E}\left(\sum_{k=1}^{N} \frac{g_{ik} g_{jk}^*}{|g_{jk}|}\right) = \sum_{k=1}^{N} \mathbb{E}\left(g_{ik}\right) \mathbb{E}\left(g_{jk}^* / \left|g_{jk}\right|\right) = 0$

$$\begin{aligned} h_{ij} h_{ij}^* &= \sum_{k=1}^{N} \frac{g_{ik} g_{jk}^*}{\left|g_{jk}\right|} \sum_{l=1}^{N} \frac{g_{il}^* g_{jl}}{\left|g_{jl}\right|} \\ &= \sum_{k=1}^{N} \left|g_{ik}\right|^2 + \sum_{k=1}^{N} \sum_{l \neq k}^{N} \frac{g_{ik} g_{jk}^*}{\left|g_{jk}\right|} \frac{g_{il}^* g_{jl}}{\left|g_{jl}\right|} \end{aligned} \tag{5.77}$$

Therefore, $\mathbb{V}\left(h_{ij}\right) = \mathbb{E}\left(h_{ij} h_{ij}^*\right) = N$. *From central limit theorem, when* $N \gg 1$, *we have* $h_{ij} \sim \mathcal{CN}(0, N)$.

6 Conclusion

Chapter 6. Conclusion

This chapter highlights conclusions of the thesis and discusses possible directions for future work.

6.1 Summary

In general, this thesis provides contributions on: (i) the distributions of the SNR and INR in RIS-aided MIMO systems, (ii) deep learning based detection methods in RIS-aided spatial modulation systems.

In Chapter 1, the RIS is introduced as a promising technique to realize SRE. The state of research is introduced.

In Chapter 2, the RIS-aided communication systems are summarized and introduced in detail. The reflection principle, channel estimation problem, and system design problem are overviewed.

Chapter 3 introduces the SNR analysis of the RIS-aided MIMO system where the phase shifts are already optimized to maximize the received SNR.

Chapter 4 considers a two-RIS aided MIMO system. The interference is investigated by characterizing the behavior of the INR as a random variable.

Chapter 5 studies the detection for RIS-aided SM systems. First, we introduce deep unfolding in MIMO detection. Then, by generalizing RIS-aided SM systems as a special case of traditional SM systems, we propose data-driven and model-driven deep learning detectors for RIS-aided SM systems.

6.2 Future Work

6.2.1 SNR/SINR Analysis in more complicated systems

We have analyzed the SNR in a wireless communication system with a single RIS and a single user. The analysis of the SNR/SINR in more general systems with multiple users or multiple RISs is still an open issue.

6.2.2 Research on joint channel estimation and data detection for RIS-SM

In RIS-SM system, channel estimation is required at the RIS to adjust the phase shifts. The system design of joint channel estimation and data detection based on deep learning is a interesting topic.

6.2.3 Research on RIS-based differential spatial modulation and generalized spatial modulation systems

Research on the variants of RIS-SM systems such as differential spatial modulation and generalized spatial modulation is another future direction. How to design the reflection coefficient which could realize the index modulation remains a challenging task.

Bibliography

[1] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019.

[2] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–20, 2019.

[3] Y.-C. Liang, J. Chen, R. Long, Z.-Q. He, X. Lin, C. Huang, S. Liu, X. S. Shen, and M. Di Renzo, "Reconfigurable intelligent surfaces for smart wireless environments: channel estimation, system design and applications in 6G networks," *Science China Information Sciences*, vol. 64, no. 10, pp. 1–21, 2021.

[4] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5849–5863, 2020.

[5] S. Nayak and R. Patgiri, "6G communication: Envisioning the key issues and challenges," *arXiv preprint arXiv:2004.04024*, 2020.

[6] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 798–807, 2020.

[7] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. De Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[8] L. Subrt and P. Pechac, "Controlling propagation environments using intelligent walls," in *2012 6th European Conference on Antennas and Propagation (EUCAP)*. IEEE, 2012, pp. 1–5.

[9] ——, “Intelligent walls as autonomous parts of smart indoor environments,” *IET communications*, vol. 6, no. 8, pp. 1004–1010, 2012.

[10] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, “Increasing indoor spectrum sharing capacity using smart reflect-array,” in *2016 IEEE International Conference on Communications (ICC)*. IEEE, 2016, pp. 1–6.

[11] O. Abari, D. Bharadia, A. Duffield, and D. Katabi, “Enabling high-quality untethered virtual reality,” in *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17)*, 2017, pp. 531–544.

[12] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, “Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays,” in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 270–278.

[13] A. Welkie, L. Shangguan, J. Gummeson, W. Hu, and K. Jamieson, “Programmable radio environments for smart spaces,” in *Proceedings of the 16th ACM Workshop on Hot Topics in Networks*, 2017, pp. 36–42.

[14] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, “A new wireless communication paradigm through software-controlled metasurfaces,” *IEEE Communications Magazine*, vol. 56, no. 9, pp. 162–169, 2018.

[15] K. Achouri, M. A. Salem, and C. Caloz, “General metasurface synthesis based on susceptibility tensors,” *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 7, pp. 2977–2991, 2015.

[16] Y. Vahabzadeh, N. Chamanara, K. Achouri, and C. Caloz, “Computational analysis of metasurfaces,” *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 3, pp. 37–49, 2018.

[17] G. Lavigne, K. Achouri, V. S. Asadchy, S. A. Tretyakov, and C. Caloz, “Susceptibility derivation and experimental demonstration of refracting metasurfaces without spurious diffraction,” *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 3, pp. 1321–1330, 2018.

[18] N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, “Light propagation with phase discontinuities: generalized laws of reflection and refraction,” *science*, vol. 334, no. 6054, pp. 333–337, 2011.

[19] D. Wang, L.-Z. Yin, T.-J. Huang, F.-Y. Han, Z.-W. Zhang, Y.-H. Tan, and P.-K. Liu, “Design of a 1 bit broadband space-time-coding digital metasurface element,” *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 4, pp. 611–615, 2020.

[20] M. E. Trampler, R. E. Lovato, and X. Gong, “Dual-resonance continuously beam-scanning x-band reflectarray antenna,” *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 8, pp. 6080–6087, 2020.

[21] G. Perez-Palomino, P. Baine, R. Dickie, M. Bain, J. A. Encinar, R. Cahill, M. Barba, and G. Toso, "Design and experimental validation of liquid crystal-based reconfigurable reflectarray elements with improved bandwidth in f-band," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1704–1713, 2013.

[22] E. Carrasco and J. Perruisseau-Carrier, "Reflectarray antenna at terahertz using graphene," *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 253–256, 2013.

[23] X. Yang, S. Xu, F. Yang, and M. Li, "Design of a 2-bit reconfigurable reflectarray element using two MEMS switches," in *2015 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting.* IEEE, 2015, pp. 2167–2168.

[24] J. Perruisseau-Carrier, M. Tamagnone, J. S. Gomez-Diaz, and E. Carrasco, "Graphene antennas: Can integration and reconfigurability compensate for the loss?" in *2013 European Microwave Conference.* Ieee, 2013, pp. 369–372.

[25] W. Hu, M. Ismail, R. Cahill, J. Encinar, V. Fusco, H. Gamble, R. Dickie, D. Linton, N. Grant, and S. Rea, "Electronically reconfigurable monopulse reflectarray antenna with liquid crystal substrate," 2007.

[26] P. Mursia, V. Sciancalepore, A. Garcia-Saavedra, L. Cottatellucci, X. C. Pérez, and D. Gesbert, "Risma: Reconfigurable intelligent surfaces enabling beamforming for IoT massive access," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1072–1085, 2020.

[27] S. Xia and Y. Shi, "Intelligent reflecting surface for massive device connectivity: Joint activity detection and channel estimation," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2020, pp. 5175–5179.

[28] Y. Lu and L. Dai, "Reconfigurable intelligent surface based hybrid precoding for THz communications," *arXiv preprint arXiv:2012.06261*, 2020.

[29] B. Ning, Z. Chen, W. Chen, and Y. Du, "Channel estimation and transmission for intelligent reflecting surface assisted THz communications," in *ICC 2020-2020 IEEE International Conference on Communications (ICC).* IEEE, 2020, pp. 1–7.

[30] K. Tekbıyık, G. Karabulut Kurt, A. Rıza Ekti, A. Görçin, and H. Yanikomeroglu, "Reconfigurable intelligent surface empowered terahertz communication for LEO satellite networks," *arXiv e-prints*, pp. arXiv–2007, 2020.

[31] J. Östman, G. Durisi, E. G. Ström, M. C. Coşkun, and G. Liva, "Short packets over block-memoryless fading channels: Pilot-assisted or noncoherent transmission?" *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1521–1536, 2018.

[32] R. Hashemi, S. Ali, N. H. Mahmood, and M. Latva-aho, "Average rate and error probability analysis in short packet communications over ris-aided URLLC systems," *arXiv preprint arXiv:2102.13363*, 2021.

[33] A. R. Ndjiongue, T. M. Ngatched, O. A. Dobre, and H. Haas, “Re-configurable intelligent surface-based VLC receivers using tunable liquid-crystals: The concept,” *Journal of Lightwave Technology*, vol. 39, no. 10, pp. 3193–3200, 2021.

[34] L. Yang, X. Yan, D. B. Da Costa, T. A. Tsiftsis, H.-C. Yang, and M.-S. Alouini, “Indoor mixed dual-hop VLC/RF systems through reconfigurable intelligent surfaces,” *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 1995–1999, 2020.

[35] T. Bai, C. Pan, C. Han, and L. Hanzo, “Empowering mobile edge computing by exploiting reconfigurable intelligent surface,” *arXiv preprint arXiv:2102.02569*, 2021.

[36] X. Hu, C. Masouros, and K.-K. Wong, “Reconfigurable intelligent surface aided mobile edge computing: From optimization-based to location-only learning-based solutions,” *IEEE Transactions on Communications*, 2021.

[37] H. Hashida, Y. Kawamoto, and N. Kato, “Intelligent reflecting surface placement optimization in air-ground communication networks toward 6G,” *IEEE Wireless Communications*, vol. 27, no. 6, pp. 146–151, 2020.

[38] L. Ge, P. Dong, H. Zhang, J.-B. Wang, and X. You, “Joint beamforming and trajectory optimization for intelligent reflecting surfaces-assisted UAV communications,” *IEEE Access*, vol. 8, pp. 78 702–78 712, 2020.

[39] G. Gradoni and M. Di Renzo, “End-to-end mutual coupling aware communication model for reconfigurable intelligent surfaces: An electromagnetic-compliant approach based on mutual impedances,” *IEEE Wireless Communications Letters*, vol. 10, no. 5, pp. 938–942, 2021.

[40] X. Qian and M. Di Renzo, “Mutual coupling and unit cell aware optimization for reconfigurable intelligent surfaces,” *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1183–1187, 2021.

[41] A. Abrardo, D. Dardari, M. Di Renzo, and X. Qian, “MIMO interference channels assisted by reconfigurable intelligent surfaces: Mutual coupling aware sum-rate optimization based on a mutual impedance channel model,” *arXiv preprint arXiv:2102.07155*, 2021.

[42] D. M. Pozar, *Microwave engineering.* John wiley & sons, 2011.

[43] A. Taha, M. Alrabeiah, and A. Alkhateeb, “Enabling large intelligent surfaces with compressive sensing and deep learning,” *IEEE Access*, vol. 9, pp. 44 304–44 321, 2021.

[44] J. Zhang, C. Qi, P. Li, and P. Lu, “Channel estimation for reconfigurable intelligent surface aided massive MIMO system,” in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5.

[45] Z.-Q. He and X. Yuan, “Cascaded channel estimation for large intelligent metasurface assisted massive mimo,” *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 210–214, 2019.

[46] H. Liu, X. Yuan, and Y.-J. A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2621–2636, 2020.

[47] X. Guan, Q. Wu, and R. Zhang, "Anchor-assisted intelligent reflecting surface channel estimation for multiuser communications," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.

[48] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for ris-empowered multi-user MISO wireless communications," *IEEE Transactions on Communications*, 2021.

[49] G. T. de Araújo and A. L. de Almeida, "PARAFAC-based channel estimation for intelligent reflective surface assisted mimo system," in *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE, 2020, pp. 1–5.

[50] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 4659–4663.

[51] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4522–4535, 2020.

[52] T. L. Jensen and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 5000–5004.

[53] S. Sun and H. Yan, "Channel estimation for reconfigurable intelligent surface-assisted wireless communications considering doppler effect," *IEEE Wireless Communications Letters*, vol. 10, no. 4, pp. 790–794, 2020.

[54] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 518–522, 2019.

[55] N. K. Kundu and M. R. McKay, "A deep learning-based channel estimation approach for MISO communications with large intelligent surfaces," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, 2020, pp. 1–6.

[56] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep residual network empowered channel estimation for IRS-assisted multi-user communication systems," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–7.

[57] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6607–6620, 2020.

[58] B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface assisted multi-user OFDMA: Channel estimation and training design," *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 8315–8329, 2020.

[59] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," *arXiv preprint arXiv:1912.03619*, 2019.

[60] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[61] A. Kammoun, A. Chaaban, M. Debbah, M.-S. Alouini *et al.*, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 7748–7764, 2020.

[62] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3064–3076, 2020.

[63] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

[64] G. Yang, X. Xu, and Y.-C. Liang, "Intelligent reflecting surface assisted non-orthogonal multiple access," in *2020 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2020, pp. 1–6.

[65] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, "Exploiting intelligent reflecting surfaces in NOMA networks: Joint beamforming optimization," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6884–6898, 2020.

[66] L. Zhang, Y. Wang, W. Tao, Z. Jia, T. Song, and C. Pan, "Intelligent reflecting surface aided MIMO cognitive radio systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 445–11 457, 2020.

[67] J. Chen, Y.-C. Liang, Y. Pei, and H. Guo, "Intelligent reflecting surface: A programmable wireless environment for physical layer security," *IEEE Access*, vol. 7, pp. 82 599–82 612, 2019.

[68] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1410–1414, 2019.

[69] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, “Robust and secure wireless communications via intelligent reflecting surfaces,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2637–2652, 2020.

[70] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, “Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications,” *IEEE Communications Letters*, vol. 23, no. 9, pp. 1488–1492, 2019.

[71] S. Hu, Z. Wei, Y. Cai, C. Liu, D. W. K. Ng, and J. Yuan, “Robust and secure sum-rate maximization for multiuser MISO downlink systems with self-sustainable IRS,” *arXiv preprint arXiv:2101.10549*, 2021.

[72] Z. Chu, W. Hao, P. Xiao, and J. Shi, “Intelligent reflecting surface aided multi-antenna secure transmission,” *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 108–112, 2019.

[73] S. Hong, C. Pan, H. Ren, K. Wang, and A. Nallanathan, “Artificial-noise-aided secure MIMO wireless communications via intelligent reflecting surface,” *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7851–7866, 2020.

[74] R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson, “Active reconfigurable intelligent surface aided wireless communications,” *IEEE Transactions on Wireless Communications*, 2021.

[75] B. Lyu, D. T. Hoang, S. Gong, and Z. Yang, “Intelligent reflecting surface assisted wireless powered communication networks,” in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 2020, pp. 1–6.

[76] Q. Wu and R. Zhang, “Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1735–1748, 2020.

[77] N. S. Perović, L.-N. Tran, M. Di Renzo, and M. F. Flanagan, “Achievable rate optimization for MIMO systems with reconfigurable intelligent surfaces,” *IEEE Transactions on Wireless Communications*, 2021.

[78] ——, “Optimization of RIS-aided MIMO systems via the cutoff rate,” *IEEE Wireless Communications Letters*, 2021.

[79] C. Huang, R. Mo, and C. Yuen, “Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1839–1850, 2020.

[80] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, “Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 375–388, 2020.

[81] G. Lee, M. Jung, A. T. Z. Kasgari, W. Saad, and M. Bennis, “Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces,” in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.

[82] C. Huang, Z. Yang, G. C. Alexandropoulos, K. Xiong, L. Wei, C. Yuen, Z. Zhang, and M. Debbah, “Multi-hop ris-empowered terahertz communications: A DRL-based hybrid beamforming design,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1663–1677, 2021.

[83] D. Li, “Ergodic capacity of intelligent reflecting surface-assisted communication systems with phase errors,” *IEEE Communications Letters*, vol. 24, no. 8, pp. 1646–1650, 2020.

[84] Q. Wu and R. Zhang, “Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts,” *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1838–1851, 2019.

[85] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, “IRS-aided wireless communication with imperfect CSI: Is amplitude control helpful or not?” in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.

[86] M. Jung, W. Saad, M. Debbah, and C. S. Hong, “On the optimality of reconfigurable intelligent surfaces (RISs): Passive beamforming, modulation, and resource allocation,” *IEEE Transactions on Wireless Communications*, 2021.

[87] H. Shen, W. Xu, S. Gong, C. Zhao, and D. W. K. Ng, “Beamforming optimization for irs-aided communications with transceiver hardware impairments,” *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 1214–1227, 2020.

[88] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, “Large intelligent surface-assisted wireless communication exploiting statistical CSI,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8238–8242, 2019.

[89] H. Guo, Y.-C. Liang, and S. Xiao, “Model-free optimization for reconfigurable intelligent surface with statistical CSI,” *arXiv preprint arXiv:1912.10913*, 2019.

[90] J. Zhang, J. Liu, S. Ma, C.-K. Wen, and S. Jin, “Transmitter design for large intelligent surface-assisted MIMO wireless communication with statistical CSI,” in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2020, pp. 1–5.

[91] G. Zhou, C. Pan, H. Ren, K. Wang, M. Di Renzo, and A. Nallanathan, “Robust beamforming design for intelligent reflecting surface aided MISO communication systems,” *IEEE wireless communications letters*, vol. 9, no. 10, pp. 1658–1662, 2020.

[92] J. Yuan, Y.-C. Liang, J. Joung, G. Feng, and E. G. Larsson, “Intelligent reflecting surface (IRS)-enhanced cognitive radio system,” in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.

[93] A. Zappone, M. Di Renzo, F. Shams, X. Qian, and M. Debbah, "Overhead-aware design of reconfigurable intelligent surfaces in smart radio environments," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 126–141, 2020.

[94] J. Wang, Y.-C. Liang, S. Han, and Y. Pei, "Robust beamforming and phase shift design for IRS-enhanced multi-user MISO downlink communication," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.

[95] M.-M. Zhao, A. Liu, and R. Zhang, "Outage-constrained robust beamforming for intelligent reflecting surface aided wireless communication," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1301–1316, 2021.

[96] A. Abrardo, D. Dardari, and M. Di Renzo, "Intelligent reflecting surfaces: Sum-rate optimization based on statistical position information," *IEEE Transactions on Communications*, 2021.

[97] Q. Li, M. Wen, and M. Di Renzo, "Single-RF MIMO: From spatial modulation to metasurface-based modulation," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 88–95, 2021.

[98] S. Lin, B. Zheng, G. C. Alexandropoulos, M. Wen, M. Di Renzo, and F. Chen, "Reconfigurable intelligent surfaces with reflection pattern modulation: Beamforming design and performance analysis," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 741–754, 2020.

[99] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE access*, vol. 7, pp. 116 753–116 773, 2019.

[100] W. Tang, J. Y. Dai, M. Chen, X. Li, Q. Cheng, S. Jin, K.-K. Wong, and T. J. Cui, "Programmable metasurface-based rf chain-free 8PSK wireless transmitter," *Electronics Letters*, vol. 55, no. 7, pp. 417–420, 2019.

[101] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3187–3196, 2020.

[102] J. Y. Dai, W. Tang, L. X. Yang, X. Li, M. Z. Chen, J. C. Ke, Q. Cheng, S. Jin, and T. J. Cui, "Realization of multi-modulation schemes for wireless communication by time-domain digital coding metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 3, pp. 1618–1627, 2019.

[103] W. Tang, J. Y. Dai, M. Z. Chen, K.-K. Wong, X. Li, X. Zhao, S. Jin, Q. Cheng, and T. J. Cui, "MIMO transmission through reconfigurable intelligent surface: System design, analysis, and implementation," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2683–2699, 2020.

[104] H. Zhao, Y. Shuang, M. Wei, T. J. Cui, P. Del Hougne, and L. Li, “Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals,” *Nature communications*, vol. 11, no. 1, pp. 1–10, 2020.

[105] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, “Symbiotic radio: Cognitive backscattering communications for future wireless networks,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1242–1255, 2020.

[106] Q. Zhang, Y.-C. Liang, and H. V. Poor, “Large intelligent surface/antennas (LISA) assisted symbiotic radio for iot communications,” *arXiv e-prints*, pp. arXiv–2002, 2020.

[107] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, “Symbiotic radio: A new communication paradigm for passive internet of things,” *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1350–1363, 2019.

[108] S. Guo, S. Lv, H. Zhang, J. Ye, and P. Zhang, “Reflecting modulation,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2548–2561, 2020.

[109] R. Karasik, O. Simeone, M. Di Renzo, and S. Shamai, “Single-RF multi-user communication through reconfigurable intelligent surfaces: An information-theoretic analysis,” *arXiv preprint arXiv:2101.07556*, 2021.

[110] V. Arun and H. Balakrishnan, “RFocus: Beamforming using thousands of passive antennas,” in *17th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 20)*, 2020, pp. 1047–1061.

[111] M. Dunna, C. Zhang, D. Sievenpiper, and D. Bharadia, “ScatterMIMO: Enabling virtual MIMO with smart surfaces,” in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1–14.

[112] Pivotal, “White papers,” [EB/OL], https://pivotalcommware.com/technology/ Accessed July 15, 2021.

[113] S. Zhou, W. Xu, K. Wang, M. Di Renzo, and M.-S. Alouini, “Spectral and energy efficiency of IRS-assisted MISO communication with hardware impairments,” *IEEE wireless communications letters*, vol. 9, no. 9, pp. 1366–1369, 2020.

[114] M.-A. Badiu and J. P. Coon, “Communication through a large reflecting surface with phase errors,” *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 184–188, 2019.

[115] H. Zhang, B. Di, L. Song, and Z. Han, “Reconfigurable intelligent surfaces assisted communications with limited phase shifts: How many phase shifts are enough?” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4498–4502, 2020.

[116] B. Holter and G. E. Oien, “On the amount of fading in MIMO diversity systems,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2498–2507, 2005.

[117] M. Di Renzo and J. Song, “Reflection probability in wireless networks with metasurface-coated environmental objects: An approach based on random spatial processes,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–15, 2019.

[118] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, “Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 421–439, 2020.

[119] M. Di Renzo, F. H. Danufane, X. Xi, J. De Rosny, and S. Tretyakov, “Analytical modeling of the path-loss for reconfigurable intelligent surfaces–anomalous mirror or scatterer?” in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5.

[120] L. Wei and O. Tirkkonen, “Analysis of scaled largest eigenvalue based detection for spectrum sensing,” in *2011 IEEE International Conference on Communications (ICC)*. IEEE, 2011, pp. 1–5.

[121] M. Chiani, “Distribution of the largest eigenvalue for real wishart and gaussian random matrices and a simple approximation for the tracy–widom distribution,” *Journal of Multivariate Analysis*, vol. 129, pp. 69–81, 2014.

[122] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.

[123] A. Castaño-Martínez and F. López-Blázquez, “Distribution of a sum of weighted noncentral chi-square variables,” *Test*, vol. 14, no. 2, pp. 397–415, 2005.

[124] B. D. Carter and M. D. Springer, “The distribution of products, quotients and powers of independent h-function variates,” *SIAM Journal on Applied Mathematics*, vol. 33, no. 4, pp. 542–558, 1977.

[125] F. H. Danufane, M. Di Renzo, J. De Rosny, and S. Tretyakov, “On the path-loss of reconfigurable intelligent surfaces: An approach based on green’s theorem applied to vector fields,” *IEEE Transactions on Communications*, 2021.

[126] X. Qian, M. Di Renzo, J. Liu, A. Kammoun, and M.-S. Alouini, “Beamforming through reconfigurable intelligent surfaces in single-user MIMOsystems: SNR distribution and scaling laws in the presence of channel fading and phase noise,” *IEEE wireless communications letters*, vol. 10, no. 1, pp. 77–81, 2020.

[127] J. R. Hershey, J. L. Roux, and F. Weninger, “Deep unfolding: Model-based inspiration of novel deep architectures,” *arXiv preprint arXiv:1409.2574*, 2014.

[128] Q. Hu, F. Gao, H. Zhang, G. Y. Li, and Z. Xu, “Understanding deep MIMO detection,” *arXiv preprint arXiv:2105.05044*, 2021.

[129] X. Gao, L. Dai, Y. Ma, and Z. Wang, “Low-complexity near-optimal signal detection for uplink large-scale MIMO systems,” *Electronics letters*, vol. 50, no. 18, pp. 1326–1328, 2014.

[130] X. Gao, L. Dai, C. Yuen, and Y. Zhang, “Low-complexity MMSE signal detection based on richardson method for large-scale MIMO systems,” in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*. IEEE, 2014, pp. 1–5.

[131] X. Gao, L. Dai, Y. Hu, Z. Wang, and Z. Wang, “Matrix inversion-less signal detection using SOR method for uplink large-scale MIMO systems,” in *2014 IEEE Global Communications Conference*. IEEE, 2014, pp. 3291–3295.

[132] L. Dai, X. Gao, X. Su, S. Han, C. I, and Z. Wang, “Low-complexity soft-output signal detection based on Gauss–Seidel method for uplink multiuser large-scale MIMO systems,” *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4839–4845, 2015.

[133] J. Zhou, Y. Ye, and J. Hu, “Biased MMSE soft-output detection based on Jacobi method in massive MIMO,” in *2014 IEEE International Conference on Communiction Problem-solving*. IEEE, 2014, pp. 442–445.

[134] N. Samuel, T. Diskin, and A. Wiesel, “Deep MIMO detection,” in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2017, pp. 1–5.

[135] ——, “Learning to detect,” *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2554–2564, 2019.

[136] S. Takabe, M. Imanishi, T. Wadayama, and K. Hayashi, “Deep learning-aided projected gradient detector for massive overloaded MIMO channels,” in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–6.

[137] S. Takabe, M. Imanishi, T. Wadayama, R. Hayakawa, and K. Hayashi, “Trainable projected gradient detector for massive overloaded MIMO channels: Data-driven tuning approach,” *IEEE Access*, vol. 7, pp. 93 326–93 338, 2019.

[138] M. R. Hestenes, E. Stiefel *et al.*, “Methods of conjugate gradients for solving linear systems,” *Journal of research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.

[139] Y. Hu, Z. Wang, X. Gaol, and J. Ning, “Low-complexity signal detection using CG method for uplink large-scale MIMO systems,” in *2014 IEEE International Conference on Communication Systems*, 2014, pp. 477–481.

[140] Y. Wei, M.-M. Zhao, M.-j. Zhao, and M. Lei, “Learned conjugate gradient descent network for massive MIMO detection,” *arXiv preprint arXiv:1906.03814*, 2019.

[141] S. Boyd, N. Parikh, and E. Chu, *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.

[142] U. S. Kamilov and H. Mansour, "Learning optimal nonlinearities for iterative thresholding algorithms," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 747–751, 2016.

[143] M. Borgerding, P. Schniter, and S. Rangan, "AMP-inspired deep networks for sparse linear inverse problems," *IEEE Transactions on Signal Processing*, vol. 65, no. 16, pp. 4293–4308, 2017.

[144] M. Un, M. Shao, W. Ma, and P. C. Ching, "Deep Mimo detection using ADMM unfolding," in *2019 IEEE Data Science Workshop (DSW)*, 2019, pp. 333–337.

[145] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[146] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[147] J. Ma and L. Ping, "Orthogonal AMP," *IEEE Access*, vol. 5, pp. 2020–2033, 2017.

[148] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[149] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[150] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, 2009.

[151] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," 2016.

[152] E. Basar, "Transmission through large intelligent surfaces: A new frontier in wireless communications," in *2019 European Conference on Networks and Communications (EuCNC)*, 2019, pp. 112–117.

[153] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[154] D. Ito, S. Takabe, and T. Wadayama, "Trainable ISTA for sparse signal recovery," *IEEE Transactions on Signal Processing*, vol. 67, no. 12, pp. 3113–3125, 2019.