Physically Consistent Modeling and Optimization of Stacked Intelligent Metasurfaces for Wideband Multi-User Communication

by
Muhammad Ibrahim

A Thesis submitted to The Faculty of Graduate and Postdoctoral Studies of
The University of Manitoba
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg

October 2026

Copyright © Muhammad Ibrahim

The important thing is not to stop questioning. Curiosity has its own reason for existing.

Albert Einstein

Abstract

Stacked intelligent metasurfaces (SIMs) have emerged as an alternative to single-layer reconfigurable intelligent surfaces (RISs). A SIM consists of multiple programmable metasurface layers placed in front of the base station (BS) antennas and performs beamforming in the electromagnetic wave domain. Its wideband operation introduces several challenges. Closely spaced elements and layers are coupled, the resulting response is frequency dependent, and a common phase configuration must serve multiple users and subbands. This thesis develops circuit-theoretic models and optimization methods for SIM-assisted wideband and multi-band multi-user downlink transmission.

First, we develop a coupling-aware circuit-theoretic model for a SIM-assisted wideband multi-user multiple-input multiple-output (MIMO) system. The model accounts for frequency-dependent mutual coupling, inter-layer reflections, and frequency-selective channels across subbands. Based on this model, the BS precoder and SIM phase shifts are jointly optimized, with power allocation across subbands providing carrier aggregation. We also investigate a partially reconfigurable SIM in which only the last layer is updated while the first layer remains fixed, reducing the control overhead. For a fixed total number of elements, the comparison between SIM and single-layer RIS shows that their relative performance depends on the operating conditions. The RIS performs better in narrowband and low signal-to-noise ratio (SNR) regimes, with fewer users, and when power allocation is performed at the BS. In contrast, the SIM provides higher spectral efficiency in wideband and high-SNR regimes with more users. Moreover, accounting for mutual coupling improves the modeled spectral efficiency, while partial reconfiguration can achieve near-optimal performance with fewer control bits.

We then consider a multi-band multi-user multiple-input single-output (MU-MISO) downlink with inter-band carrier aggregation, where one SIM configuration serves a low

band and a high band with different bandwidths and noise levels. Since conventional phase updates need many iterations, we propose a multi-band deep-unfolding network (MBDU-Net), which unrolls projected-gradient phase updates into a fixed number of stages and learns band-specific step sizes and momentum parameters. For the same number of iterations, MBDU-Net achieves a higher sum rate than projected gradient ascent and basic deep unfolding, and generalizes to numbers of subcarriers not seen during training.

Keywords: Stacked intelligent metasurface (SIM), reconfigurable intelligent surface (RIS), wideband MIMO, carrier aggregation, mutual coupling, circuit-theoretic model, joint active and passive beamforming, deep unfolding.

Acknowledgment

I would like to sincerely thank everyone who supported me throughout my graduate studies and contributed to the completion of this thesis.

First, I would like to thank my supervisor, Professor Ekram Hossain, for giving me the opportunity to join his research group and for his guidance and support throughout my studies. I am grateful for his trust, encouragement, and valuable feedback during this work.

I would also like to thank Professor Amine Mezghani for his technical guidance and for the many helpful discussions and feedback that contributed to this research and our joint publications.

I gratefully acknowledge the financial support provided by the University of Manitoba through the University of Manitoba Graduate Fellowship.

I am especially grateful to my parents and family for their constant love, prayers, and support throughout this journey. Their encouragement has been a constant source of strength.

Finally, I would like to thank my friends at the University of Manitoba and back home for their support and companionship during my graduate studies.

Table of Contents

List of Figures **x**

List of Tables **xii**

List of Abbreviations **xiii**

1 Introduction **1**

1.1 Overview 1

1.2 RIS and SIM Architectures 2

1.2.1 Conventional RIS Deployment 2

1.2.2 SIM-Integrated Base Station 3

1.3 Open Challenges 5

1.3.1 Coupling-Aware Modeling 5

1.3.2 Wideband Operation 6

1.3.3 Optimization Complexity 7

1.3.4 Research Questions 8

1.4 Contributions 9

1.4.1 SIM versus RIS in Wideband MU-MIMO 9

1.4.2 Deep Unfolding for Multi-band SIM 10

1.5 Thesis Outline 10

1.6 Publications 11
1.7 Notations 12

2 Background **14**
2.1 Circuit Theory for Communication 14
2.2 Antenna Model 16
2.2.1 Canonical Minimum Scattering Antennas 16
2.2.2 Chu's Antenna Model 17
2.3 Mutual Coupling 19
2.3.1 Two-Antenna Circuit Model 19
2.3.2 Mutual Impedance Between Two Chu Antennas 20
2.3.3 From Two Antennas to Arrays 22
2.4 Wideband and Multi-band Transmission 23
2.4.1 Multi-carrier Wideband Model 23
2.4.2 Carrier Aggregation 23
2.5 Deep Unfolding 25
2.5.1 Iterative Optimization with Projected Gradient Ascent 25
2.5.2 Unfolding Iterations into a Trainable Network 25
2.5.3 Momentum Acceleration 26
2.6 Summary 26

3 SIM versus RIS for Wideband Multi-User MIMO **28**
3.1 Introduction 28
3.2 System Model 29
3.3 Circuit-Theoretic Model 31
3.3.1 Multiport Network Model 31
3.3.2 Mutual Coupling Among Array Elements 34

3.3.3 SIM Model 36
3.3.4 Far-Field Transimpedance Matrix 40
3.3.5 Achievable Sum Rate 41
3.4 Problem Formulation and Solution 42
3.4.1 Problem Formulation 42
3.4.2 Alternating Optimization 43
3.4.3 Convergence and Complexity 47
3.4.4 Partially Reconfigurable SIM 47
3.5 Simulation Results and Discussion 49
3.5.1 Simulation Setup 49
3.5.2 Effect of the Number of Subbands 50
3.5.3 Effect of the Number of Users 51
3.5.4 Impact of Mutual Coupling and Power Allocation 53
3.5.5 Effect of SNR and Bandwidth 54
3.5.6 Frequency Response 55
3.5.7 Goodput of the Partially Reconfigurable SIM 57
3.5.8 Convergence 58
3.6 Summary 59

4 Deep Unfolding for Multi-band SIM Systems 61
4.1 Introduction 61
4.2 System Model 63
4.2.1 Multi-band Setup 63
4.2.2 SIM Model and Effective Channel 63
4.2.3 Signal and Channel Model 65
4.2.4 Achievable Sum Rate 66
4.3 Problem Formulation 67

4.4 Proposed Solution 67
4.4.1 Alternating Optimization Framework 67
4.4.2 From Projected Gradient to Deep Unfolding 68
4.4.3 Multi-band Deep-Unfolding Network 68
4.4.4 Training and Overall Algorithm 69
4.4.5 Complexity Discussion 70
4.5 Simulation Results and Discussion 71
4.5.1 Simulation Setup 71
4.5.2 Convergence and Ablation Study 73
4.5.3 Generalization to Different Numbers of Subcarriers 73
4.6 Summary 74

5 Conclusion and Future Work **77**
5.1 Conclusion 77
5.2 Future Work 78
5.2.1 Imperfect CSI and Hardware Impairments 78
5.2.2 Frequency-Dependent Phase Shifters 79
5.2.3 More General Models 79
5.2.4 Energy Efficiency with RF-Chain Power Models 80
5.2.5 Learning-Based Joint Design 80

Appendices **81**

A Gradient Derivation for Chapter 3 **81**
A.1 Gradient of the Local Impedance Matrix 82
A.2 Gradient of the SIM Transfer Matrix 82
A.3 Gradient of the Objective Function 83

Bibliography **84**

List of Figures

1.1 Conventional RIS deployment versus BS-integrated SIM architecture . . 4
1.2 Structure of the thesis. 12

2.1 Two-port circuit-theoretic model of a SISO communication link. 15
2.2 Equivalent circuit of a CMS antenna in the TM_1 mode 18
2.3 Two-antenna circuit model with mutual coupling 20
2.4 Two TM_1 antennas arbitrarily oriented and separated by a distance d. . . 21
2.5 Types of carrier aggregation 24
2.6 Deep unfolding of projected gradient ascent 26

3.1 SIM-assisted downlink multi-user MIMO system. 30
3.2 Equivalent circuit-theoretic model for the SIM-assisted MIMO communication system. 32
3.3 Rectangular planar array with inter-element spacings d_x and d_y. 35
3.4 Multiport network model of the SIM 37
3.5 Variation of the optimal spectral efficiency with the number of subbands for $K = 5$ and $M_{\text{total}} = 100$. 52
3.6 Spectral efficiency versus the number of users 53
3.7 Impact of mutual coupling and power allocation 55

3.8 Variation of the spectral efficiency with the noise variance for $K = 6$ and $M_{\text{total}} = 100$. 56
3.9 Variation of the spectral efficiency with the bandwidth B for $K = 6$. . . . 57
3.10 Variation of the spectral efficiency with frequency for $K = 5$ and $M_{\text{total}} = 144$. 58
3.11 Variation of the goodput with the number of antenna elements for $K = 2$. 59
3.12 Convergence of the gradient ascent algorithm for $M_{\text{total}} = 100$ and $K = 4$. 60

4.1 SIM-assisted multi-band multi-user MISO downlink system. 64
4.2 One stage of the proposed MBDU-Net 70
4.3 Sum rate versus the number of iterations 74
4.4 Sum rate versus the number of subcarriers per band for $T = 6$ and $I_{\max} = 5$. 75

List of Tables

1.1 Publications resulting from this thesis. 13

3.1 Number of mathematical operations required for different steps of the optimization. 48

3.2 Simulation parameters. 50

4.1 System and training parameters. 72

List of Abbreviations

5G	Fifth Generation
6G	Sixth Generation
AO	Alternating Optimization
AWGN	Additive White Gaussian Noise
BS	Base Station
CA	Carrier Aggregation
CMS	Canonical Minimum Scattering
CSI	Channel State Information
DU	Deep Unfolding
EM	Electromagnetic
EMF	Electromotive Force
GD	Gradient Ascent
LEO	Low Earth Orbit
MBDU	Multi-Band Deep Unfolding
MBDU-Net	Multi-Band Deep-Unfolding Network
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
mmWave	Millimeter Wave
MU-MIMO	Multi-User Multiple-Input Multiple-Output

MU-MISO	Multi-User Multiple-Input Single-Output
OFDM	Orthogonal Frequency-Division Multiplexing
PA	Power Allocation
RF	Radio Frequency
RIS	Reconfigurable Intelligent Surface
SIM	Stacked Intelligent Metasurface
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
sub-THz	Sub-Terahertz
TM	Transverse Magnetic
ULA	Uniform Linear Array

Chapter 1

Introduction

1.1 Overview

The rapid evolution of 6G networks demands high spectral efficiency, wider bandwidth operation, and energy-efficient hardware architectures [1]. While massive multiple-input multiple-output (MIMO) offers significant spatial multiplexing gains, its hardware cost and radio-frequency (RF) chain complexity make it less scalable at millimeter-wave (mmWave) and sub-terahertz (sub-THz) frequencies [2, 3]. In a fully digital massive MIMO transmitter, every antenna element requires a dedicated RF chain, including digital-to-analog converters, mixers, and power amplifiers, so that the power consumption and cost grow with the array size. This motivates the exploration of energy-efficient and less complex architectures, in which part of the beamforming is moved from the digital domain to the electromagnetic (EM) wave domain.

Programmable metasurfaces enable dynamic control of EM waves by electronically tuning the responses of densely packed sub-wavelength meta-atoms distributed over a planar aperture [4]. This capability has stimulated substantial interest in metasurface-assisted wireless communications, where the propagation channel and the transceiver

front-end can be reconfigured to improve spectral efficiency with low power consumption [5]. In parallel, wideband and multi-band transmission through carrier aggregation (CA) has become a key mechanism for improving spectral efficiency by jointly exploiting subcarriers across multiple frequency bands [6]. In this thesis, we combine these two directions and study stacked intelligent metasurfaces (SIMs) integrated into the base station (BS) for wideband and multi-band multi-user downlink transmission, with a focus on physically consistent modeling, joint optimization, and low-complexity design.

1.2 RIS and SIM Architectures

1.2.1 Conventional RIS Deployment

Reconfigurable intelligent surfaces (RISs) have emerged as a promising solution for energy-efficient wireless communication. By intelligently tuning the EM response to the impinging signals, RISs can enhance coverage and improve spectral efficiency with minimal power consumption [4, 5, 7–10]. Both diagonal RIS architectures with phase-only control and beyond-diagonal and non-local RIS architectures [11, 12] have been extensively studied in MIMO and orthogonal frequency-division multiplexing (OFDM) systems, investigating aspects such as joint active and passive beamforming [8], channel estimation [9], and hardware constraints.

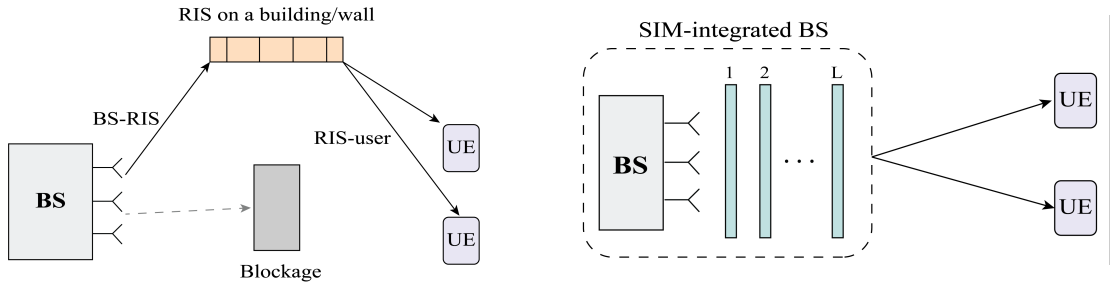
In the conventional RIS paradigm, illustrated in Fig. 1.1(a), the RIS is deployed in the propagation environment, typically on a building facade or wall between the BS and the users, and acts as a passive reflector that creates an additional, controllable propagation path. This deployment has several practical implications. First, the signal reflected by the RIS experiences the path loss of both the BS–RIS and RIS–user links, so that a large surface is often required to obtain a meaningful gain over the direct link. Second, the RIS is physically separated from the BS, which requires a dedicated control

link and the acquisition of the cascaded BS–RIS–user channel. Third, a single reflecting layer provides only a limited number of degrees of freedom, so that conventional single-layer RIS architectures remain limited in their ability to support complex signal manipulation and system performance enhancement. Transmissive RIS architectures, where the surface is placed in front of the transmitter and the signal passes through it, have also been considered [13,14], which paved the way toward integrating metasurfaces directly into the transceiver.

1.2.2 SIM-Integrated Base Station

A new technology has recently been introduced in the form of cascaded layers of transmissive RIS, also known as SIMs [15, 16]. Inspired by diffractive neural networks implemented on programmable metasurfaces [17], a SIM is composed of multiple programmable metasurface layers, where each layer consists of passive meta-atoms that apply an adjustable phase shift to the impinging wavefront. The key architectural novelty of SIM, proposed in [15, 16], is its location: instead of being deployed in the environment between the BS and the users, the SIM is integrated into the BS architecture and placed in close proximity to the BS antenna array, as illustrated in Fig. 1.1(b). The metasurface therefore becomes part of the transmitter itself, forming an intelligent transceiver architecture in which the BS antennas excite the first layer and the signal is progressively shaped as it propagates through the cascaded layers.

Unlike conventional digital or hybrid beamforming systems, SIMs process EM waves directly in the analog domain, enabling more energy-efficient control of signal propagation. By embedding passive, phase-programmable layers into the BS architecture, SIM implements wave-domain beamforming and can reduce the reliance on high-dimensional digital precoding and many RF chains [16,18]. According to [15,16], this BS-integrated architecture offers the following main advantages:



(a) Conventional RIS-assisted downlink (b) BS-integrated SIM architecture

Figure 1.1: Comparison between (a) the conventional deployment of a single-layer RIS in the propagation environment between the BS and the users, and (b) the SIM architecture, in which multiple transmissive metasurface layers are integrated into the BS in front of its antenna array to perform beamforming directly in the wave domain.

- **Fewer RF chains and lower energy consumption:** Since transmit precoding is accomplished automatically as the EM waves propagate through the SIM, holographic MIMO transmission can be implemented without requiring excessive RF chains, which effectively reduces the RF-related energy consumption and hardware cost compared with conventional fully digital MIMO transceivers [15, 16].

- **Additional degrees of freedom:** The cascaded structure provides additional spatial degrees of freedom through cascaded EM processing across multiple transmissive layers, so that a SIM may outperform its single-layer metasurface counterparts, such as RISs and metasurface lenses [15].

- **Reduced processing delay:** Signal processing tasks such as MIMO precoding and multi-user interference mitigation are carried out in the wave domain as the signal propagates through the layers, rather than by high-dimensional digital baseband processing, which reduces the processing delay [16, 17].

However, whether the additional layers are beneficial depends on the operating condi-

tions, which we examine in Chapter 3 using a model that includes mutual coupling and inter-layer reflections. Note that the SIM architecture itself was proposed in [15, 16]; in this thesis, we study how it should be modeled and optimized under wideband conditions.

Several recent studies have explored SIMs to enhance spectral efficiency in various communication systems. In [19], a SIM-aided integrated sensing and communication system was proposed, jointly optimizing the SIM phase shifts and BS power to mitigate inter-user interference. The work in [20] extended SIM applications to holographic MIMO systems, formulating joint optimization of phase shifts and transmit covariance matrices. SIM-assisted transceiver design [16], multi-user downlink beamforming in the wave domain [18], holographic MIMO-aided cell-free networks [21], and low earth orbit (LEO) satellite communications [22] have also been investigated. Other studies have investigated channel estimation and direction-of-arrival estimation [23, 24].

1.3 Open Challenges

However, several aspects of SIM design are still open when SIMs are used for wideband and multi-band multi-user transmission. We discuss them below together with the related work.

1.3.1 Coupling-Aware Modeling

Since the layers of a SIM are placed in close proximity and each layer consists of densely packed meta-atoms, the elements interact electromagnetically through mutual coupling, and part of the wave is reflected back and forth between adjacent layers. Many existing SIM works adopt idealized cascaded models that overlook mutual coupling and inter-layer reflections [16, 20]. To accurately capture these effects, multiport network

theory has been widely adopted as a physically consistent modeling framework, building on the circuit theory of communication [25, 26] and on coupling-aware models of RIS-assisted systems [12, 27–30]. For instance, [31] employed cascaded scattering (S)-parameter representations to explicitly model wave propagation, inter-element coupling, and inter-layer reflections. While physically accurate, cascaded S-parameter models are often analytically intractable and computationally demanding. To address this issue, [32] introduced a transmission (T)-parameter-based formulation that is mathematically equivalent to the cascaded S-parameter model but offers improved tractability. In parallel, [33] developed a general impedance (Z)-parameter-based SIM model that exploits structured impedance matrices to efficiently capture coupling effects. Nevertheless, these existing multiport SIM models are predominantly developed under narrowband assumptions and do not explicitly account for frequency-dependent impedance variations and coupling effects that arise in wideband systems.

1.3.2 Wideband Operation

Although [34] introduced a fully analog wideband SIM beamforming method, it did not incorporate coupling or reflection effects. Moreover, prior SIM research has primarily focused on spatial-domain optimization, while efficient exploitation of frequency-domain characteristics remains largely unexplored.

Carrier aggregation, a key technique in 5G networks [6, 35, 36], enables wideband transmission by combining multiple subband carriers to enhance throughput and spectral efficiency. Compared with single-band operation, multi-band CA can better leverage frequency-selective fading and band diversity, where lower bands typically offer more favorable propagation and reliability, while higher bands provide wider bandwidth but may suffer from higher noise power and less robust scattering [37]. In downlink multi-user systems, these heterogeneous band characteristics provide additional degrees of

freedom for subcarrier selection and power allocation, leading to sum-rate gains. Extending CA to metasurface-assisted systems introduces a joint spatial and frequency-domain design problem, since a common metasurface configuration must operate over multiple subbands. Under wideband operation, frequency-dependent impedance variations and mutual coupling make the effective phase response subband dependent, which can lead to beam squint and phase misalignment across the aggregated carriers. While metasurface-assisted system modeling and carrier-aggregation-aware transmission design have been studied under simplified assumptions, their joint design under physically consistent, coupling-aware wideband metasurface models remains limited.

A related open question is whether stacking multiple layers is always beneficial. For a fixed total number of meta-atoms, a single-layer RIS has a simpler structure, whereas a SIM distributes the elements over several layers and introduces inter-layer coupling and reflections. A fair comparison between the two structures under a physically consistent wideband model is needed to identify the operating regimes in which each structure is preferable.

1.3.3 Optimization Complexity

Integrating SIM-assisted transmission with multi-band CA is challenging because the SIM phase configuration is effectively shared across the aggregated spectrum, whereas the per-subcarrier operating conditions vary substantially across bands. In particular, band-dependent bandwidths and noise levels, together with subcarrier-dependent interference shaped by multi-user scheduling and power allocation, induce gradients of markedly different scales across the low- and high-band components. Consequently, SIM phase design in a multi-band multi-user downlink leads to a tightly coupled optimization problem that is non-convex and strongly frequency coupled. Classical optimization methods for SIM-assisted multi-user transmission typically adopt alternating optimiza-

tion (AO), where the digital precoders are updated for a fixed SIM configuration and the SIM phases are then refined for fixed precoders. However, this approach can be computationally demanding in practical multi-user, multi-subcarrier systems, since the SIM phase update often requires many iterations.

Recently, data-driven methods have been introduced for physical-layer optimization, where a neural network maps channel realizations to transceiver parameters [38], such as beamforming weights or metasurface configurations. Deep unfolding improves interpretability by unrolling iterative optimization algorithms into trainable architectures [39]. However, most unfolding-based beamforming and phase-optimization methods are developed for conventional MIMO/RIS systems and do not explicitly address SIM-assisted multi-band transmission.

1.3.4 Research Questions

The challenges discussed above lead to the following research questions, which are addressed in this thesis:

RQ1: How can a SIM integrated into the BS be modeled in a physically consistent and tractable way, such that mutual coupling, inter-layer reflections, and frequency selectivity over a wide bandwidth are captured, and how can its phase shifts be jointly optimized with the BS precoding and power allocation across subbands?

RQ2: For the same total number of elements, under which operating conditions (bandwidth, signal-to-noise ratio (SNR), number of users, and BS processing) does a multilayer SIM outperform a single-layer RIS, and can the control overhead of the SIM be reduced without a significant performance loss?

RQ3: How can the SIM phase configuration be optimized with low and fixed complexity when a single configuration must serve heterogeneous low and high band

subcarriers under inter-band carrier aggregation?

1.4 Contributions

To answer these questions, we develop a coupling-aware wideband model for SIMs integrated into the BS and use it for both model-based and learning-based optimization. The contributions are summarized below, and the detailed contributions of each technical chapter are listed at the beginning of that chapter.

1.4.1 SIM versus RIS in Wideband MU-MIMO

Chapter 3 addresses RQ1 and RQ2:

- We develop a circuit-theoretic model of a SIM-assisted downlink multi-user wideband MIMO system that captures mutual coupling, inter-layer reflections, and frequency selectivity across subbands, based on the multiport and Chu-based mutual coupling models of Chapter 2. We show that the conventional cascaded SIM model is a special case obtained when these effects are neglected.

- We jointly optimize the SIM phase shifts, the BS precoders, and the power allocation across users and subbands by alternating optimization, where iterative water-filling performs carrier aggregation over favorable subbands.

- We propose updating only the last SIM layer, which reduces the control signaling and part of the computational complexity by a factor equal to the number of layers.

- We show that, for the same total number of elements, the SIM outperforms the single-layer RIS in wideband, high-SNR, and high-user-density scenarios, whereas the RIS is favorable in narrowband and low-SNR regimes, with few users, and

when active beamforming with power allocation is performed at the BS. We also show that modeling mutual coupling improves the spectral efficiency, and that the partially reconfigurable SIM achieves near-optimal performance with the highest goodput for a small number of users.

1.4.2 Deep Unfolding for Multi-band SIM

Chapter 4 addresses RQ3:

- We formulate the joint precoding and SIM phase design for a multi-band MU-MISO downlink with inter-band carrier aggregation, where a common SIM configuration serves a low band and a high band with different bandwidths and noise levels.
- We propose a physically consistent multi-band deep-unfolding network that unrolls projected-gradient phase updates into a fixed-depth architecture with analytic gradients from the SIM model. It separates the low- and high-band gradient contributions and learns band-aware step sizes together with Nesterov-inspired momentum parameters, with only a small number of trainable parameters.
- We show that, for the same number of iterations, MBDU-Net achieves a higher sum rate than classical projected gradient ascent and basic deep unfolding, that its components must be learned jointly, and that it generalizes to numbers of subcarriers not seen during training.

1.5 Thesis Outline

The remainder of this thesis is structured as shown in Fig. 1.2:

- Chapter 2 reviews the background used throughout the thesis: the circuit theory for communication, the Chu antenna model, the mutual coupling model between two antennas and its extension to antenna arrays, wideband and multi-band transmission with carrier aggregation, and deep unfolding.

- Chapter 3 presents the coupling-aware wideband SIM model, the joint precoding and phase-shift optimization with carrier aggregation, the partially reconfigurable SIM, and the comparison between single-layer RISs and SIMs.

- Chapter 4 presents the AO framework and MBDU-Net for SIM-assisted multi-band MU-MISO downlink systems.

- Chapter 5 summarizes the main findings of the thesis and outlines directions for future work.

- Appendix A provides the derivation of the gradients used for the SIM phase optimization in Chapters 3 and 4.

1.6 Publications

The technical chapters of this thesis are based on the journal publications listed in Table 1.1, both of which have been published in IEEE journals.
Other publication during the MSc program:

[C1] **M. Ibrahim**, W. H. Raza, M. Moiz, S. A. Hassan, H. Jung, and M. Gidlund, "On the performance of multi-IRS-assisted networks across real urban, suburban, and rural environments," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2024.

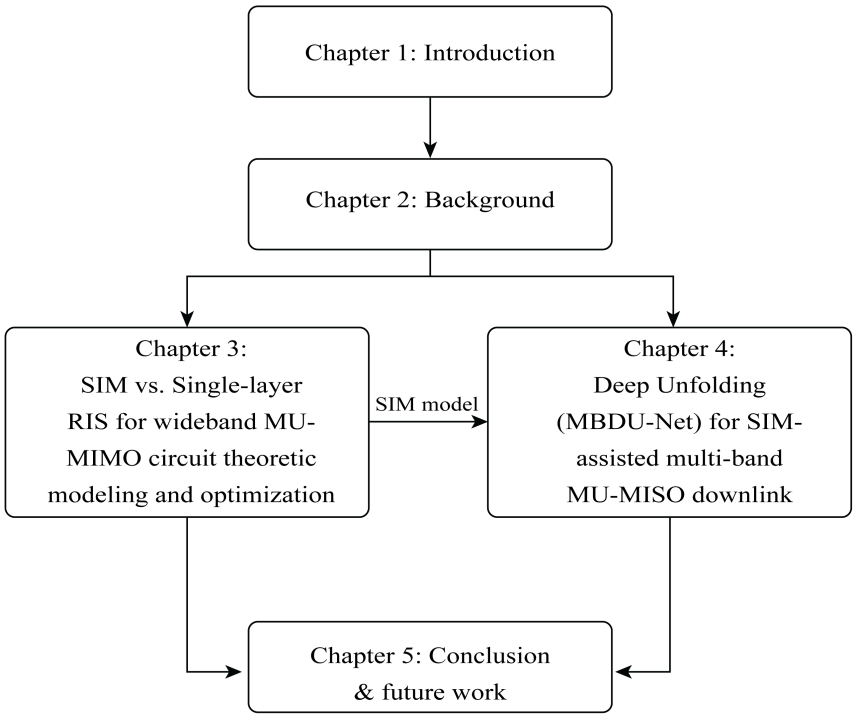


Figure 1.2: Structure of the thesis.

1.7 Notations

For any matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$, and $\mathbf{A}^{-1}$ represent its transpose, conjugate, conjugate transpose, and inverse, respectively. Bold lowercase letters (e.g., $\mathbf{a}$) denote vectors, bold uppercase letters (e.g., $\mathbf{A}$) denote matrices, and $[\mathbf{A}]_{i,j}$ is the (i,j)-th entry of $\mathbf{A}$. The trace of $\mathbf{A}$ is denoted by $\mathrm{Tr}(\mathbf{A})$. The Kronecker product of two vectors or matrices is represented by $\otimes$. $\mathrm{Diag}(\mathbf{a})$ denotes a diagonal matrix with the elements of vector $\mathbf{a}$ along its diagonal, while $\mathrm{diag}(\mathbf{A})$ returns a vector containing the diagonal elements of the matrix $\mathbf{A}$. $\mathbf{I}_N$ is the $N \times N$ identity matrix and $\mathbb{C}^{N\times M}$ is the set of $N \times M$ complex matrices. $\|\cdot\|$ and $\|\cdot\|_F$ denote the Euclidean norm of a vector and the Frobenius norm of a matrix, and $|\cdot|$ is the absolute value. $\mathbb{E}[\cdot]$ is the expectation operator, and $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ is the circularly symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and

Table 1.1: Publications resulting from this thesis.

Publication	**Chapter**
[**J1**] **M. Ibrahim**, A. Mezghani, and E. Hossain, "Are stacked intelligent metasurfaces (SIMs) better than single-layer reconfigurable intelligent surfaces (RISs) for wideband multi-user MIMO communication systems?" *IEEE Transactions on Wireless Communications*, vol. 25, pp. 18215–18228, 2026, doi: 10.1109/TWC.2026.3696550.	Chapter 3
[**J2**] **M. Ibrahim**, A. Mezghani, and E. Hossain, "Deep unfolding for SIM-assisted multiband MU-MISO downlink systems," *IEEE Wireless Communications Letters*, vol. 15, pp. 3591–3595, 2026, doi: 10.1109/LWC.2026.3701564.	Chapter 4

covariance $\mathbf{R}$. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are the real and imaginary parts, and $j = \sqrt{-1}$. $\nabla_{\mathbf{x}} f$ denotes the gradient of f with respect to $\mathbf{x}$, $[x]^+ = \max(x, 0)$, $\log_2(\cdot)$ and $\ln(\cdot)$ are the binary and natural logarithms, and $\mathcal{U}[a, b]$ denotes the uniform distribution over $[a, b]$. Finally, f, c, $\lambda = c/f$, and $k = 2\pi/\lambda$ denote the frequency, the speed of light, the wavelength, and the wavenumber, respectively. Symbols that are specific to a chapter are defined where they are first used.

Chapter 2

Background

This chapter introduces the fundamental concepts used throughout the thesis. We first review the circuit-theoretic description of wireless links, in which antennas are represented as ports of a multiport network. We then present the antenna model adopted in this thesis and the model for the mutual coupling between two antennas, illustrate it with a two-antenna circuit, and extend it to antenna arrays. Finally, we review wideband and multi-band transmission with carrier aggregation, and the principle of deep unfolding, which are used in Chapters 3 and 4.

2.1 Circuit Theory for Communication

Conventional channel models describe a wireless link through an abstract transfer matrix that maps transmitted symbols to received symbols. Such models do not explicitly contain physical quantities such as the power radiated by the antennas, the interaction between closely spaced antennas, or the physical origin of noise. The circuit theory of communication [25,26,40] addresses these limitations and provides a clear and physically consistent foundation for analyzing MIMO wireless communication systems. In this ap-

proach, antennas are represented as ports in a multiport network, where transmitted and received signals are described by the corresponding port voltages and currents. Capturing the relationship among these port variables is essential for accurately characterizing the communication model.

From a circuit-theoretic perspective, a single-input single-output (SISO) link can be viewed as a two-port network, as shown in Fig. 2.1. The port voltages $(v_{\mathrm{T}}(f), v_{\mathrm{R}}(f))$ and port currents $(i_{\mathrm{T}}(f), i_{\mathrm{R}}(f))$ at the transmit (T) and receive (R) antennas are related by Ohm's law through the impedance matrix $\mathbf{Z}_{\mathrm{SISO}}(f)$ [41]:

$$\begin{bmatrix} v_{\mathrm{T}}(f) \\ v_{\mathrm{R}}(f) \end{bmatrix} = \underbrace{\begin{bmatrix} Z_{\mathrm{T}}(f) & Z_{\mathrm{TR}}(f) \\ Z_{\mathrm{RT}}(f) & Z_{\mathrm{R}}(f) \end{bmatrix}}_{\mathbf{Z}_{\mathrm{SISO}}(f)} \begin{bmatrix} i_{\mathrm{T}}(f) \\ i_{\mathrm{R}}(f) \end{bmatrix}. \tag{2.1}$$

The diagonal entries $Z_{\mathrm{T}}(f)$ and $Z_{\mathrm{R}}(f)$ are the self-impedances of the transmit and receive antennas, i.e., their input impedances when each antenna is considered in isolation. The off-diagonal entries $Z_{\mathrm{RT}}(f)$ and $Z_{\mathrm{TR}}(f)$ are the mutual impedances, which describe how the current at one port induces a voltage at the other port. Since antennas are reciprocal devices, $Z_{\mathrm{RT}}(f) = Z_{\mathrm{TR}}(f)$.

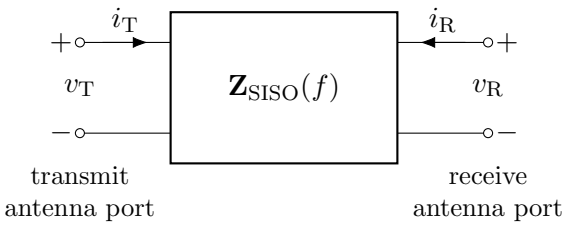


Figure 2.1: Two-port circuit-theoretic model of a SISO communication link.

The same description applies to multi-antenna systems. A MIMO link with N_{T} transmit and N_{R} receive antennas is modeled as a multiport network with $N_{\mathrm{T}} + N_{\mathrm{R}}$

ports, whose voltage and current vectors satisfy

$$\begin{bmatrix} \mathbf{v}_{\mathrm{T}}(f) \\ \mathbf{v}_{\mathrm{R}}(f) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{Z}_{\mathrm{TT}}(f) & \mathbf{Z}_{\mathrm{TR}}(f) \\ \mathbf{Z}_{\mathrm{RT}}(f) & \mathbf{Z}_{\mathrm{RR}}(f) \end{bmatrix}}_{\mathbf{Z}_{\mathrm{MIMO}}(f)} \begin{bmatrix} \mathbf{i}_{\mathrm{T}}(f) \\ \mathbf{i}_{\mathrm{R}}(f) \end{bmatrix}, \tag{2.2}$$

where $\mathbf{Z}_{\mathrm{TT}}(f) \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{T}}}$ and $\mathbf{Z}_{\mathrm{RR}}(f) \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ describe the transmit and receive arrays, and $\mathbf{Z}_{\mathrm{RT}}(f)$ and $\mathbf{Z}_{\mathrm{TR}}(f)$ are the transimpedance matrices between them [26]. The diagonal entries of $\mathbf{Z}_{\mathrm{TT}}$ and $\mathbf{Z}_{\mathrm{RR}}$ are the self-impedances of the ports, while the off-diagonal entries describe the coupling between the elements of the same array. By reciprocity, $\mathbf{Z}_{\mathrm{RT}} = \mathbf{Z}_{\mathrm{TR}}^{T}$. When the transmitter and the receiver are in the far field of each other, the attenuation between them is very large, i.e., $\|\mathbf{Z}_{\mathrm{TR}}(f)\| \ll \min\big(\|\mathbf{Z}_{\mathrm{TT}}(f)\|, \|\mathbf{Z}_{\mathrm{RR}}(f)\|\big)$. This justifies the *unilateral approximation* $\mathbf{Z}_{\mathrm{TR}}(f) \approx \mathbf{0}$ [25], under which the receive antennas do not influence the transmit antennas, while the transmit antennas still drive the receive antennas through $\mathbf{Z}_{\mathrm{RT}}(f)$.

2.2 Antenna Model

The entries of $\mathbf{Z}_{\mathrm{SISO}}$ and $\mathbf{Z}_{\mathrm{MIMO}}$ depend on the antennas used at the transmitter and the receiver. In this thesis, all antenna elements, including the meta-atoms of the SIM, are modeled as canonical minimum scattering (CMS) antennas based on Chu's theory. This choice decouples the analysis from a specific antenna design, while capturing the frequency-dependent behavior of the broadest-band antenna that fits within a given physical size.

2.2.1 Canonical Minimum Scattering Antennas

Minimum scattering antennas are a class of antennas that become electromagnetically invisible, i.e., do not scatter an incident field, under a particular termination of their

port [41]. For a CMS antenna, this termination is the open circuit. As a consequence, the scattering behavior of a CMS antenna, and therefore its coupling with other antennas, is completely determined by its radiation pattern, which makes CMS antennas particularly convenient for modeling mutual coupling in antenna arrays. A short dipole is a classical example of a minimum scattering antenna.

2.2.2 Chu's Antenna Model

Chu showed that the field radiated by an arbitrary, linearly polarized antenna enclosed in a sphere of radius a can be expressed outside the sphere as a superposition of spherical transverse-magnetic (TM_n) modes, and derived an equivalent circuit for each mode [42]. It is well established that maximum bandwidth (lowest Q factor) is obtained when an antenna operates in the lowest order radiation mode, since the higher the radiation mode, the larger the amount of energy stored (non-radiated) in the sphere enclosing the antenna. Therefore, we restrict our analysis to CMS antennas operating in the transverse-magnetic-1 (TM_1) mode only, as this configuration ensures the broadest possible bandwidth [41, 43]. Notably, many common and simple antenna structures, such as dipoles, can be effectively approximated as TM_1 antennas.

For the TM_1 mode, the voltage and current at the antenna port are given by [41, 42]

$$v(f) = \sqrt{\frac{8\pi\eta}{3}}\,\frac{\sqrt{R}\,A_1}{k}\left(1+\frac{1}{jka}-\frac{1}{j(ka)^2}\right)e^{-jka}, \tag{2.3a}$$

$$i(f) = -\sqrt{\frac{8\pi\eta}{3}}\,\frac{A_1}{\sqrt{R}\,k}\left(1+\frac{1}{jka}\right)e^{-jka}, \tag{2.3b}$$

where A_1 denotes the complex coefficient of the TM_1 mode, $k = \frac{2\pi f}{c}$ is the wavenumber, and $\eta = \sqrt{\mu/\epsilon}$ is the intrinsic impedance of the medium, with μ and ϵ representing its permeability and permittivity, respectively. Furthermore, R is the radiation resistance of the antenna and a is the radius of the sphere enclosing the antenna. The corresponding

equivalent circuit, often referred to as Chu's electric antenna, is shown in Fig. 2.2, where $L = \frac{aR}{c}$ and $C = \frac{a}{cR}$ denote the equivalent inductance and capacitance.

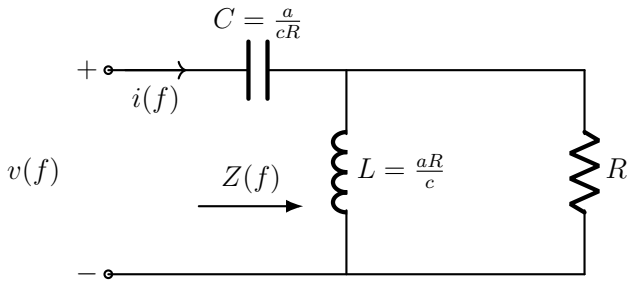


Figure 2.2: Equivalent circuit of a CMS antenna operating in the TM_1 mode (Chu's electric antenna) [41, 42].

By applying basic circuit analysis, the corresponding self-impedance can be obtained as

$$Z(f) = \frac{v(f)}{i(f)} = \left[\frac{c^2 + j2\pi fca - (2\pi fa)^2}{j2\pi fca - (2\pi fa)^2}\right] R. \tag{2.4}$$

Two properties of (2.4) are important for this thesis. First, the self-impedance is frequency dependent, so that an antenna element presents a different impedance on each subband of a wideband system. Second, it depends on the electrical size of the antenna. Defining $x = 2\pi fa/c$, (2.4) can be rewritten as

$$Z(f) = R\,\frac{x^2}{1+x^2} - j\,R\,\frac{1}{x(1+x^2)}. \tag{2.5}$$

For an electrically small antenna ($x \ll 1$), the real part vanishes while the capacitive reactance becomes large, whereas for $x \gg 1$ the impedance approaches the resistance R. This reflects the fundamental size–bandwidth trade-off of electrically small antennas [42, 43].

2.3 Mutual Coupling

When antennas are placed in close proximity, the current flowing on one antenna induces a field, and hence a voltage, on the neighboring antennas. As a result, the voltage at the port of an antenna depends not only on its own current, but also on the currents of all nearby antennas. This electromagnetic interaction is referred to as *mutual coupling*. In the circuit-theoretic description, mutual coupling is captured by the off-diagonal entries of the impedance matrix, i.e., the mutual impedances. For two antennas indexed by n and m, the mutual impedance

$$Z_{n,m}(f) = \left.\frac{v_n(f)}{i_m(f)}\right|_{i_{n'}=0,\ \forall n' \neq m} \tag{2.6}$$

is the open-circuit voltage induced at port n per unit current at port m when all other ports are open-circuited. Mutual coupling alters the input impedance and the radiation pattern of each element and correlates the noise at different ports. It is therefore particularly significant in densely packed arrays and in structures with closely spaced layers, such as SIMs, where the inter-element spacing and the inter-layer distance are typically a fraction of the wavelength.

2.3.1 Two-Antenna Circuit Model

To see how mutual coupling enters a communication link, consider the circuit in Fig. 2.3, which connects two Chu antennas [41]. A generator with voltage v_{T} and internal resistance R drives the transmit antenna, and the receive antenna is terminated by a load R_{L}. Each antenna is represented by the circuit of Fig. 2.2 with radiation resistances R_1 and R_2 and sizes a_{T} and a_{R}, i.e., $C_{\mathrm{T}} = \frac{a_{\mathrm{T}}}{cR_1}$, $L_{\mathrm{T}} = \frac{a_{\mathrm{T}}R_1}{c}$, $C_{\mathrm{R}} = \frac{a_{\mathrm{R}}}{cR_2}$, and $L_{\mathrm{R}} = \frac{a_{\mathrm{R}}R_2}{c}$. The two current-controlled current sources model the mutual coupling: i_{s_2} represents the effect of the transmit antenna on the receive antenna, and i_{s_1} represents the reaction of the receive antenna on the transmit antenna. In terms of (2.1), the self-impedances

$Z_{\mathrm{T}}(f)$ and $Z_{\mathrm{R}}(f)$ are given by (2.4), while the mutual impedances $Z_{\mathrm{RT}}(f) = Z_{\mathrm{TR}}(f)$, which set the strength of the controlled sources, are derived in Section 2.3.2.

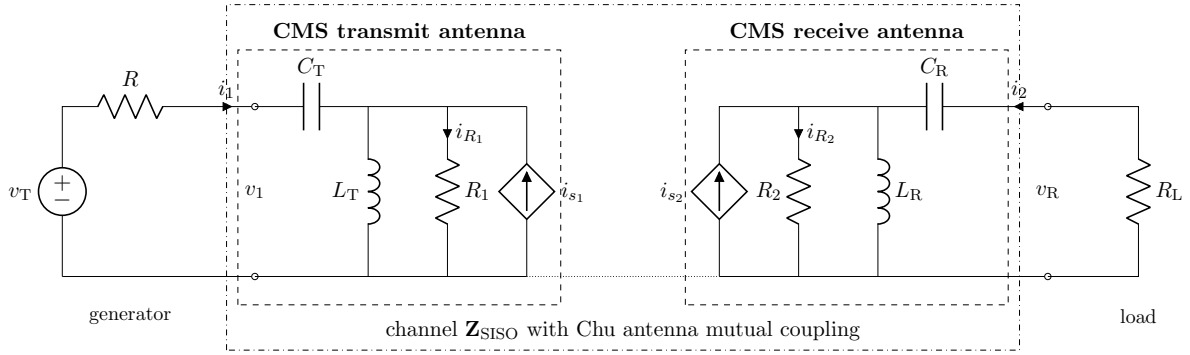


Figure 2.3: Two-antenna circuit model with transmit and receive Chu antennas and mutual coupling [41].

Solving the circuit, the voltage across the load is $v_{\mathrm{R}}(f) = h(f)\, v_{\mathrm{T}}(f)$ with

$$h(f) = \frac{R_{\mathrm{L}}\, Z_{\mathrm{RT}}(f)}{\big(R_{\mathrm{L}} + Z_{\mathrm{R}}(f)\big)\big(R + Z_{\mathrm{T}}(f)\big) - Z_{\mathrm{TR}}(f) Z_{\mathrm{RT}}(f)}. \tag{2.7}$$

The mutual impedance thus appears in two places: in the numerator, where $Z_{\mathrm{RT}}(f)$ carries the signal from the transmitter to the receiver, and in the denominator, where the product $Z_{\mathrm{TR}}(f)Z_{\mathrm{RT}}(f)$ describes how the receive antenna loads the transmit antenna. The channel therefore depends not only on the propagation, but also on the antenna size, the matching, and the mutual coupling. The remaining question is how to compute $Z_{\mathrm{RT}}(f)$ for two Chu antennas, which is addressed next.

2.3.2 Mutual Impedance Between Two Chu Antennas

The mutual impedance between two antennas in the near field is commonly computed with the induced electromotive force (EMF) method. However, this method cannot be directly applied to Chu's antennas, since their fields are only known outside the enclosing sphere. In [41], Akrout *et al.* overcame this limitation in three steps: (i)

they showed that, with an appropriate choice of the mode coefficient A_1, the TM_1 field radiated by a Chu antenna is identical to the field of a Hertz (infinitesimal) dipole and both have the same radiation resistance; (ii) they applied the induced EMF method to two Hertz dipoles to obtain their mutual impedance in closed form; and (iii) they showed that the mutual impedance normalized by $\sqrt{\Re\{Z_{\mathrm{T}}\}\Re\{Z_{\mathrm{R}}\}}$ is identical for the Hertz and Chu pairs, which yields the mutual impedance between two Chu antennas.

Accordingly, the mutual impedance between two TM_1 antennas indexed by n and m, separated by a distance d and oriented with angles α and β with respect to the line joining both antennas, as illustrated in Fig. 2.4, is expressed as [41, 44]

$$\begin{aligned} Z_{n,m}(f) = &- 3\sqrt{\Re\{Z_n(f)\}\,\Re\{Z_m(f)\}}\left[\frac{1}{2}\sin\alpha\sin\beta\left(\frac{1}{jkd}+\frac{1}{(jkd)^2}+\frac{1}{(jkd)^3}\right)\right. \\ &\left. +\cos\alpha\cos\beta\left(\frac{1}{(jkd)^2}+\frac{1}{(jkd)^3}\right)\right]e^{-jkd}, \end{aligned} \tag{2.8}$$

where $\Re\{Z_n(f)\}$ and $\Re\{Z_m(f)\}$ are the real parts of the self-impedances of the two antennas obtained from (2.4). Due to reciprocity, $Z_{n,m}(f) = Z_{m,n}(f)$.

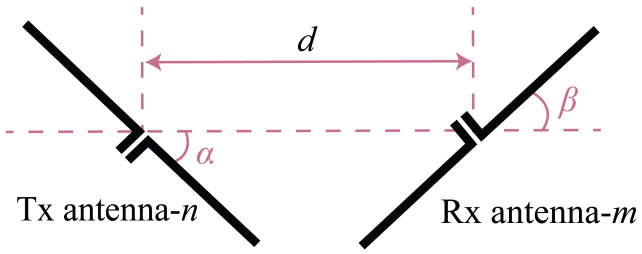


Figure 2.4: Two TM_1 antennas arbitrarily oriented and separated by a distance d.

The mutual coupling model in (2.8) provides several insights. The term proportional to $1/(jkd)$ decays slowly with distance and corresponds to the radiative field, whereas the terms proportional to $1/(jkd)^2$ and $1/(jkd)^3$ decay rapidly and correspond

to the reactive near field. Hence, for closely spaced elements with d comparable to or smaller than the wavelength, all terms contribute and the coupling is strong. Moreover, since $k = 2\pi f/c$, the mutual impedance depends on the electrical distance d/λ and therefore varies with frequency. In a wideband system, the same physical array thus exhibits a different coupling matrix on each subband. Finally, the model depends on the relative orientation of the antennas through α and β, and on their size through the self-impedances.

When the two antennas are in the far field of each other, $Z_{\mathrm{RT}}(f)$ can instead be obtained from the Friis transmission equation. For two Chu antennas with gains G_{T} and G_{R} at distance d, the far-field mutual impedance is [41]

$$Z_{\mathrm{RT}}(f) = \left(\frac{j2\pi f a_{\mathrm{T}}}{c + j2\pi f a_{\mathrm{T}}}\right) \frac{c}{2\pi f d} \sqrt{\frac{G_{\mathrm{T}} G_{\mathrm{R}} R_1}{R_2}} \, \frac{j2\pi f R_2}{j2\pi f + \frac{c}{a_{\mathrm{R}}}}, \tag{2.9}$$

where $G_{\mathrm{T}} = G_{\mathrm{R}} = 3/2$ in the equatorial plane for Chu's CMS antennas. In this case, $|Z_{\mathrm{RT}}(f)|$ is much smaller than the self-impedances, so that the unilateral approximation $Z_{\mathrm{TR}}(f) \approx 0$ applies, the source i_{s_1} vanishes, and (2.7) reduces to $h(f) = R_{\mathrm{L}} Z_{\mathrm{RT}}(f) / \big((R_{\mathrm{L}} + Z_{\mathrm{R}}(f))(R + Z_{\mathrm{T}}(f))\big)$.

2.3.3 From Two Antennas to Arrays

The two-antenna model extends directly to antenna arrays. For an array of N antennas, the impedance matrix $\mathbf{Z}(f) \in \mathbb{C}^{N \times N}$ has the self-impedances (2.4) on its diagonal and the mutual impedances (2.8) between every pair of elements off the diagonal, i.e., $[\mathbf{Z}(f)]_{n,n} = Z_n(f)$ and $[\mathbf{Z}(f)]_{n,m} = Z_{n,m}(f)$ for $n \neq m$, where the distance d and the angles α and β follow from the positions and orientations of elements n and m. The resulting matrix is symmetric due to reciprocity and changes with frequency through $k = 2\pi f/c$. In this way, the blocks $\mathbf{Z}_{\mathrm{TT}}(f)$ and $\mathbf{Z}_{\mathrm{RR}}(f)$ in (2.2), as well as the transimpedance matrices between closely spaced arrays, can be computed. Links to users in the far

field are instead modeled with (2.9), extended to multiple antennas and multipath propagation. This provides the impedance matrices used in the SIM model in Chapter 3, where the geometry of planar arrays and SIM layers is specified.

2.4 Wideband and Multi-band Transmission

2.4.1 Multi-carrier Wideband Model

As shown in Sections 2.2 and 2.3, both the self-impedance and the mutual impedance of the antenna elements vary with frequency. Consequently, a model evaluated at a single carrier frequency cannot capture the behavior of a wideband system. In this thesis, the wideband channel is modeled using a multi-carrier framework, where the total bandwidth B around the center frequency f_c is divided into N_f subbands with center frequencies f_n, $n = 1, \ldots, N_f$, and the channel response within each subband is assumed to be approximately frequency-flat. Each subband is then characterized by its own impedance matrices, evaluated at f_n. The achievable rate of the wideband system is obtained by summing the rates of the individual subbands weighted by their bandwidths.

2.4.2 Carrier Aggregation

Carrier aggregation (CA), a key technique in 5G networks, enables wideband transmission by combining multiple component carriers to enhance throughput and spectral efficiency [6, 35, 36]. Depending on the location of the aggregated component carriers, CA can be classified into three types, as illustrated in Fig. 2.5: (i) *intra-band contiguous* CA, where adjacent carriers within the same band are aggregated; (ii) *intra-band non-contiguous* CA, where carriers within the same band are aggregated although they are separated by a gap; and (iii) *inter-band* CA, where carriers located in different frequency

bands, e.g., a low band and a high band, are aggregated.

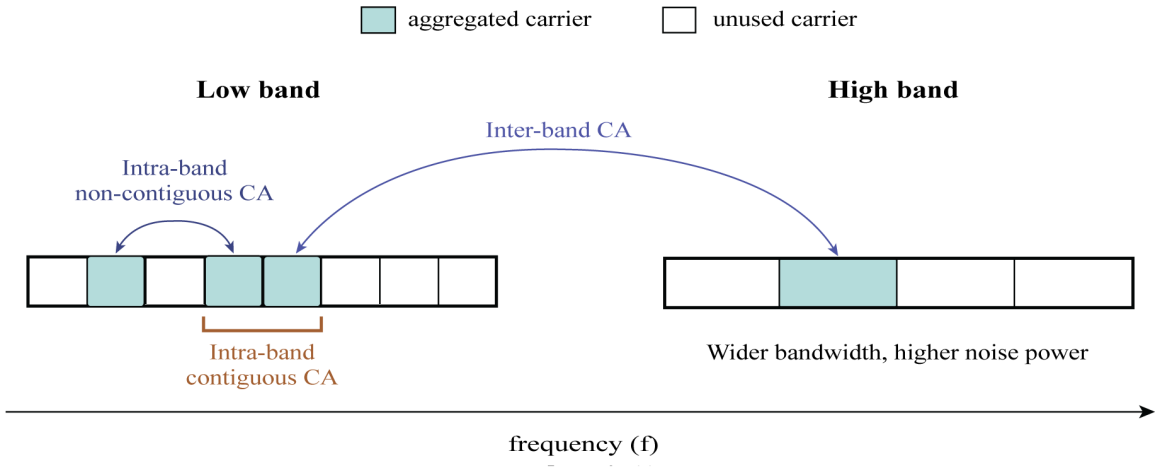


Figure 2.5: Illustration of the three types of carrier aggregation on a common spectrum consisting of a low band and a high band. The high band offers wider carriers but typically suffers from higher noise power.

Compared with single-band operation, multi-band CA can better leverage frequency-selective fading and band diversity, where lower bands typically offer more favorable propagation and reliability, while higher bands provide wider bandwidth but may suffer from higher noise power and less robust scattering [37]. In downlink multi-user systems, these heterogeneous band characteristics provide additional degrees of freedom for subcarrier selection and power allocation. In Chapter 3, CA is modeled in an abstract manner as wideband power allocation across subbands, while Chapter 4 considers inter-band CA with a low band and a high band that differ in bandwidth and noise power.

2.5 Deep Unfolding

2.5.1 Iterative Optimization with Projected Gradient Ascent

Many physical-layer design problems, including the SIM phase-shift design in this thesis, are non-convex problems of the form $\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$, where $\mathcal{X}$ is a feasible set. A standard approach is projected gradient ascent:

$$\mathbf{x}^{(t+1)} = \Pi_{\mathcal{X}}\Big(\mathbf{x}^{(t)} + \eta_t \nabla_{\mathbf{x}} f\big(\mathbf{x}^{(t)}\big)\Big), \tag{2.10}$$

where $\eta_t > 0$ is the step size and $\Pi_{\mathcal{X}}(\cdot)$ denotes the projection onto $\mathcal{X}$. For the SIM, $\mathbf{x}$ collects the phase shifts of all meta-atoms and $\Pi_{\mathcal{X}}$ maps each phase onto $[0, 2\pi)$. Although (2.10) is interpretable and directly tied to the underlying objective, its convergence speed is highly sensitive to the step size and it typically requires many iterations.

2.5.2 Unfolding Iterations into a Trainable Network

Data-driven methods have been introduced for physical-layer optimization, where a neural network maps channel realizations to transceiver parameters, such as beamforming weights or metasurface configurations [38]. Such black-box networks, however, may require large training sets and offer limited interpretability. Deep unfolding combines the two approaches by unrolling a fixed number T of iterations of an optimization algorithm into a network with T stages [39], as illustrated in Fig. 2.6. Each stage preserves the structure of one iteration, e.g., (2.10), while the hand-tuned parameters, such as the step sizes $\{\eta_t\}_{t=0}^{T-1}$, are replaced by trainable parameters. The parameters are learned offline by minimizing a loss function, such as the negative objective $\mathcal{L} = -f(\mathbf{x}^{(T)})$ averaged over a set of channel realizations, using backpropagation through all stages. Online, the trained network is executed with a fixed and small number of stages, which leads to a fixed and low computational complexity.

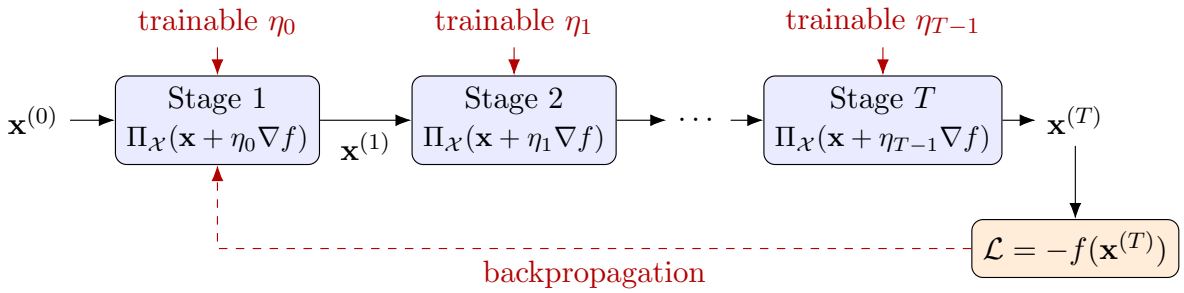


Figure 2.6: Deep unfolding of T projected gradient ascent iterations into a T-stage network with trainable step sizes.

2.5.3 Momentum Acceleration

The convergence of gradient methods can be accelerated by momentum. In Nesterov's accelerated gradient method [45], a velocity (momentum) term $\mathbf{v}^{(t)}$ aggregates past ascent directions, and the gradient is evaluated at a look-ahead point:

$$\mathbf{v}^{(t+1)} = \theta_t\,\mathbf{v}^{(t)} + \eta_t\,\nabla_{\mathbf{x}} f\big(\mathbf{x}^{(t)} + \theta_t \mathbf{v}^{(t)}\big), \tag{2.11a}$$

$$\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)} + \mathbf{v}^{(t+1)}, \tag{2.11b}$$

where $\theta_t \in [0, 1)$ is the momentum factor. In an unfolded network, both the step sizes η_t and the momentum factors θ_t can be learned. Chapter 4 builds on these ideas to design a multi-band deep-unfolding network for the SIM phase optimization, in which separate step sizes are learned for the low- and high-band gradient contributions.

2.6 Summary

In this chapter, we reviewed the multiport circuit model of communication links. We then presented the Chu antenna model and the mutual coupling model between two Chu antennas, illustrated it with a two-antenna circuit, and extended it to the impedance matrices of antenna arrays. We also reviewed the multi-carrier wideband model, carrier aggregation, and deep unfolding. The array impedance model is used in Chapter 3

to build the SIM model, and deep unfolding is used in Chapter 4 for the SIM phase optimization.

Chapter 3

SIM versus RIS for Wideband Multi-User MIMO

3.1 Introduction

This chapter addresses research questions RQ1 and RQ2 of Section 1.3.4, which arise from the narrowband and coupling-unaware modeling of existing SIM works discussed in Section 1.3. In this chapter, we develop a circuit-theory-based wideband SIM framework that incorporates mutual coupling and inter-layer reflections while enabling joint optimization of phase shifts and resource allocation across subbands. Using this framework, we examine whether, and under which conditions, a multilayer SIM outperforms a single-layer RIS with the same total number of elements. The main contributions of this chapter are summarized as follows:

- Development of a circuit-theoretic model for a SIM-assisted downlink multi-user wideband MIMO system that captures mutual coupling, inter-layer reflections, and frequency selectivity across subbands, building on the multiport and mutual coupling models reviewed in Chapter 2.

- Joint optimization of SIM phase shifts and beamforming across users and sub-bands, improving the achievable sum rate relative to narrowband baselines.
- A partial reconfiguration strategy for SIM that reduces computational and control overhead while maintaining high goodput under practical constraints.
- A comparative analysis between single-layer RIS and SIM across different user, bandwidth, and SNR scenarios, identifying operational regimes where SIM provides clear gains.

The material of this chapter has been published in [46]. The rest of the chapter is organized as follows. Section 3.2 presents the system model. Section 3.3 develops the circuit-theoretic model of the SIM-assisted system, including the mutual coupling among the elements, the SIM model, and the far-field transimpedance matrix. Section 3.4 formulates the sum-rate maximization problem and presents the proposed solution, including the partially reconfigurable SIM. Section 3.5 presents the simulation results, and Section 3.6 summarizes the chapter.

3.2 System Model

We consider a SIM-assisted multi-user downlink MIMO system, as shown in Fig. 3.1, where a SIM is deployed near the BS to assist communication with multiple users. The BS is equipped with N_t antennas arranged in a uniform linear array (ULA). The SIM consists of L cascaded transmissive metasurface layers with M meta-atoms (or antenna elements) each, which can be reconfigured by an intelligent controller. The system serves K single-antenna users. Throughout this chapter, the direct BS–user link is assumed to be blocked, and communication occurs exclusively through the metasurface-assisted path.

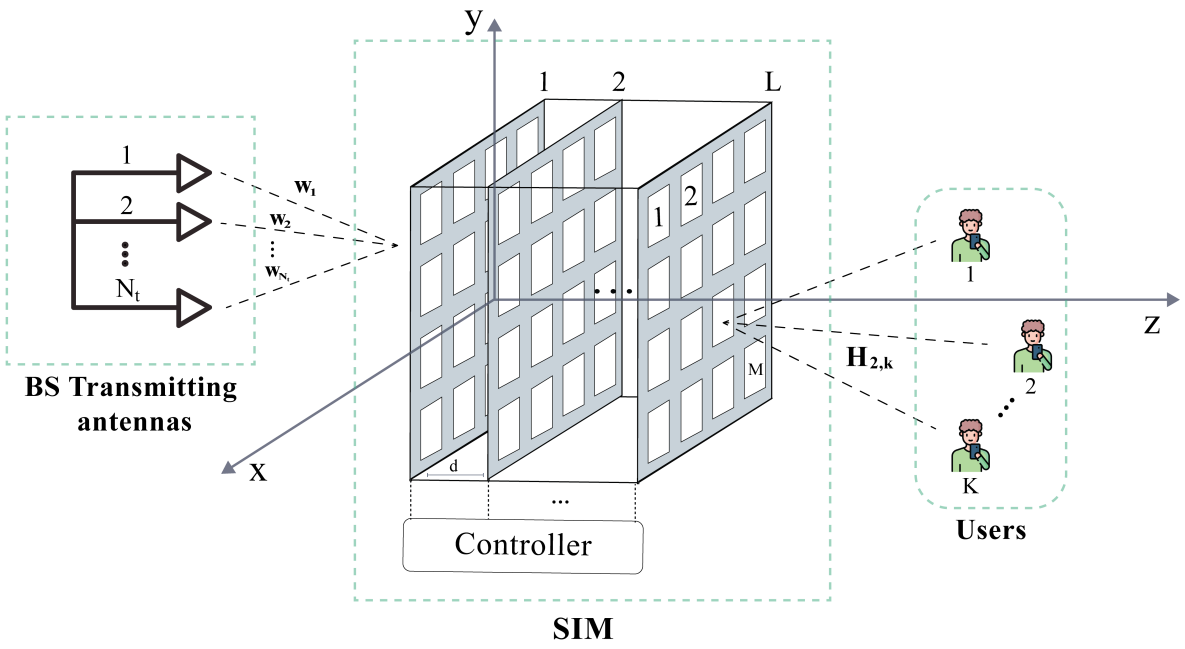


Figure 3.1: SIM-assisted downlink multi-user MIMO system.

To enable a fair comparison between single-layer (RIS) and multilayer (SIM) architectures, the total number of metasurface elements is fixed and denoted by M_{total}. In the single-layer RIS case, all $M_{\text{total}} = M_x \times M_y$ elements are placed on a single planar surface. In the multilayer case, the same total number of elements is evenly distributed across L cascaded layers, such that each layer contains $M = M_{\text{total}}/L = M_x \times (M_y/L)$ meta-atoms arranged on a planar grid with fixed inter-element spacing. Interactions are assumed to occur only between adjacent metasurface layers, consistent with the cascaded SIM structure. With this setup, any performance difference between the two architectures is due to how the elements are arranged, and not to the number of elements.

The system operates over a wideband channel with center frequency f_c and total bandwidth B. Following the multi-carrier model of Section 2.4.1, the total bandwidth B is divided into N_f subbands with center frequencies f_i, $i = 1, \ldots, N_f$, and bandwidths

B_i, and the channel response within each subband is assumed to be approximately frequency-flat. Carrier aggregation is modeled in an abstract manner as wideband power allocation. Digital baseband processing is done at the BS across subbands, whereas the SIM is modeled as a passive analog multiport network whose frequency-dependent response is characterized by impedance parameters. The metasurface phase shifts and the transmit power across subbands are jointly optimized to maximize the achievable sum rate of all users. The metasurface phase shifts are assumed to be common across all subbands, while the effective electromagnetic response varies with frequency due to frequency-dependent self-impedance and mutual coupling. These frequency-dependent effects are captured by the circuit-theoretic model presented next.

3.3 Circuit-Theoretic Model

3.3.1 Multiport Network Model

Following the circuit theory of communication reviewed in Section 2.1, the SIM-assisted MIMO communication channel is modeled using multiport network theory, as shown in Fig. 3.2, with a total of $(N_t + 2M_{\text{total}} + K)$ ports. These correspond to N_t transmit antenna ports, $2M_{\text{total}}$ SIM ports arising from the two facing sides of each transmissive metasurface element, and K single-antenna user ports. As in (2.2), Ohm's law relates all port voltages and currents through a global impedance matrix $\mathbf{Z}(f)$, which fully characterizes both the antenna subsystems and the wireless propagation channel.

To define the port variables explicitly, we denote the port voltage and current vectors at the transmitter (T) by $\mathbf{v}_T(f)$ and $\mathbf{i}_T(f)$, both in $\mathbb{C}^{N_t \times 1}$, those at the SIM (S) by $\mathbf{v}_S(f)$ and $\mathbf{i}_S(f)$, both in $\mathbb{C}^{2M_{\text{total}} \times 1}$, and those at the receiver (R) by $\mathbf{v}_R(f)$ and $\mathbf{i}_R(f)$, both in $\mathbb{C}^{K \times 1}$. For the SIM, $\mathbf{v}_S(f)$ and $\mathbf{i}_S(f)$ include the port voltages and currents of all metasurface elements across all cascaded layers. We consider K independent generator

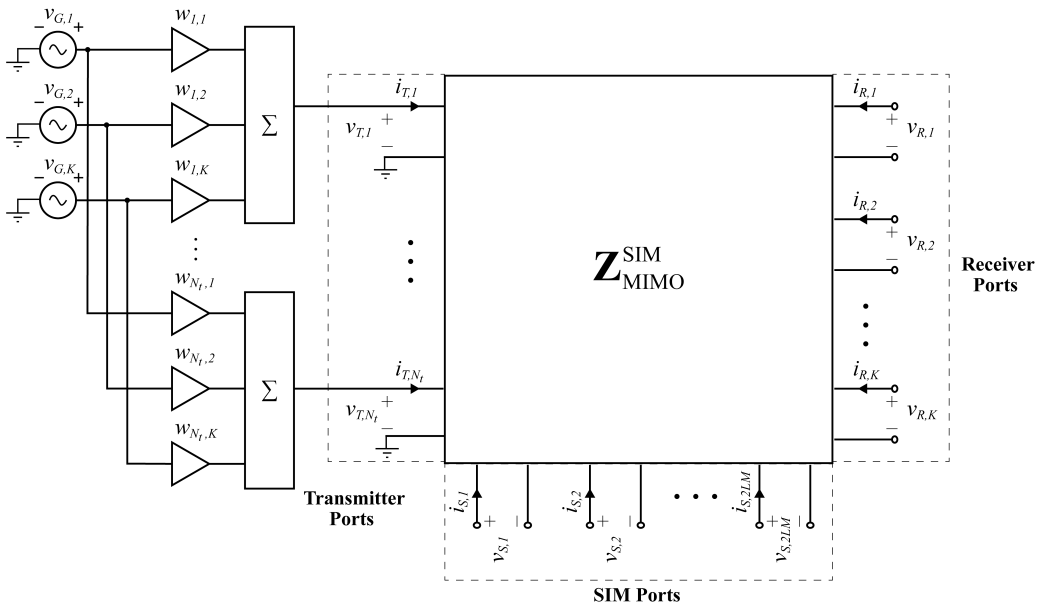


Figure 3.2: Equivalent circuit-theoretic model for the SIM-assisted MIMO communication system.

voltages collected in $\mathbf{v}_G \in \mathbb{C}^{K\times 1}$. These generator voltages are amplified through a gain matrix $\mathbf{W}$, similar to digital beamforming, producing the effective excitation voltages $\mathbf{v}_T$ which serve as inputs to the transmit antenna ports. At the receiver, the load voltages are given by $\mathbf{v}_R$ for the K users, with matched load impedances collected in $\mathbf{R} = \text{Diag}(R_1, \ldots, R_K)$.

In this framework, the SIM is treated as a general group of $2M_{\text{total}}$ ports. The corresponding SIM-related submatrix of $\mathbf{Z}(f)$ captures the self-impedance of each element as well as its coupling with neighboring elements, the transmitter, and the receiver. Hence, the full block structure of the impedance matrix, which extends (2.2) with the

SIM ports, can be expressed as (the frequency argument is dropped for brevity)

$$\begin{bmatrix} \mathbf{v}_T \\ \mathbf{v}_S \\ \mathbf{v}_R \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{TT} & \mathbf{Z}_{TS} & \mathbf{Z}_{TR} \\ \mathbf{Z}_{ST} & \mathbf{Z}_{SS} & \mathbf{Z}_{SR} \\ \mathbf{Z}_{RT} & \mathbf{Z}_{RS} & \mathbf{Z}_{RR} \end{bmatrix} \begin{bmatrix} \mathbf{i}_T \\ \mathbf{i}_S \\ \mathbf{i}_R \end{bmatrix}. \tag{3.1}$$

The voltages and currents at the SIM are related through $\mathbf{v}_S = -\mathbf{Z}_S \mathbf{i}_S$, where $\mathbf{Z}_S$ denotes the impedance matrix of the network to which the SIM ports are connected, i.e., the programmable load network. As in Section 2.1, the diagonal blocks $\mathbf{Z}_{TT}$, $\mathbf{Z}_{SS}$, and $\mathbf{Z}_{RR}$ contain the self-impedances and the intra-array coupling of the transmit array, the SIM, and the receive array, respectively, while the off-diagonal (transimpedance) blocks $\mathbf{Z}_{XX'} \in \mathbb{C}^{N_X \times N_{X'}}$, $X \neq X' \in \{T, S, R\}$, describe the coupling between different subsystems and satisfy $\mathbf{Z}_{XX'} = \mathbf{Z}_{X'X}^T$ by reciprocity.

The input–output relationship between the transmit and receive voltages can be expressed as $\mathbf{v}_R = \mathbf{H}_Z \mathbf{v}_T$, where $\mathbf{H}_Z$ denotes the transfer function. To study the effects of the SIM, we assume that both the transmit and receive antennas are terminated at the reference impedance $Z_0 = 50\ \Omega$ and that there is no mutual coupling among them. Furthermore, we neglect mutual reflections between the transmitter and the SIM, as well as between the SIM and the receiver. Under these assumptions, and following [30], the transfer function is given by

$$\mathbf{H}_Z = \frac{1}{4Z_0}\Big[\mathbf{Z}_{RT} - \mathbf{Z}_{RS}\big(\mathbf{Z}_{SS} + \mathbf{Z}_S\big)^{-1}\mathbf{Z}_{ST}\Big], \tag{3.2}$$

where $\mathbf{Z}_S$ defines the controllable load network that is to be optimized. The first term, $\mathbf{Z}_{RT}$, corresponds to the direct link between the BS and the users, while the second term describes the link through the SIM. The matrix inverse $\big(\mathbf{Z}_{SS} + \mathbf{Z}_S\big)^{-1}$ accounts for all interactions among the SIM ports, including mutual coupling and multiple reflections between the layers.

3.3.2 Mutual Coupling Among Array Elements

We model all antenna elements, i.e., the BS antennas, the SIM meta-atoms, and the user antennas, as CMS antennas operating in the TM_1 mode, as described in Section 2.2. Their self-impedance is therefore given by (2.4), which directly provides the diagonal entries of $\mathbf{Z}_{TT}$, $\mathbf{Z}_{SS}$, and $\mathbf{Z}_{RR}$. The mutual impedance between any two elements is computed with the Chu-based mutual coupling model in (2.8). In the following, we describe how the distance d and the orientation angles α and β in (2.8) are obtained for planar arrays.

Consider a rectangular planar array consisting of $N_x \times N_y$ antenna elements, as shown in Fig. 3.3, with inter-element spacings d_x along the x-axis and d_y along the y-axis. The array is assumed to lie in the xy-plane and to be centered at the origin, such that the position vector of the element with index (n_x, n_y), where $n_x \in \{1, \ldots, N_x\}$ and $n_y \in \{1, \ldots, N_y\}$, is given by

$$\mathbf{r}_{n_x,n_y} = \left[\left(n_x - \tfrac{N_x+1}{2}\right) d_x \quad \left(n_y - \tfrac{N_y+1}{2}\right) d_y \quad 0\right]^T. \tag{3.3}$$

For the mutual coupling between two distinct elements (n_x, n_y) and (n'_x, n'_y), we employ the mutual impedance expression (2.8), where the inter-element distance is

$$d_{(n_x,n_y),(n'_x,n'_y)} = \left[\left((n'_x - n_x)d_x\right)^2 + \left((n'_y - n_y)d_y\right)^2 + z^2\right]^{1/2}. \tag{3.4}$$

Here, z represents the axial separation along the z-axis between the two arrays. For $\mathbf{Z}_{TT}$ and $\mathbf{Z}_{RR}$, i.e., the coupling within the same planar array, we set $z = 0$, while for $\mathbf{Z}_{ST}$, i.e., the coupling between two parallel planar arrays, we set $z = d_z$, where d_z denotes the distance along the z-axis between the two arrays. The orientation angles are computed from the relative geometry of the two elements along the line joining them, and are given by

$$\cos\alpha_{(n_x,n_y),(n'_x,n'_y)} = \frac{(n'_x - n_x)d_x}{d_{(n_x,n_y),(n'_x,n'_y)}}, \qquad \beta_{(n_x,n_y),(n'_x,n'_y)} = \pi - \alpha_{(n_x,n_y),(n'_x,n'_y)}. \tag{3.5}$$

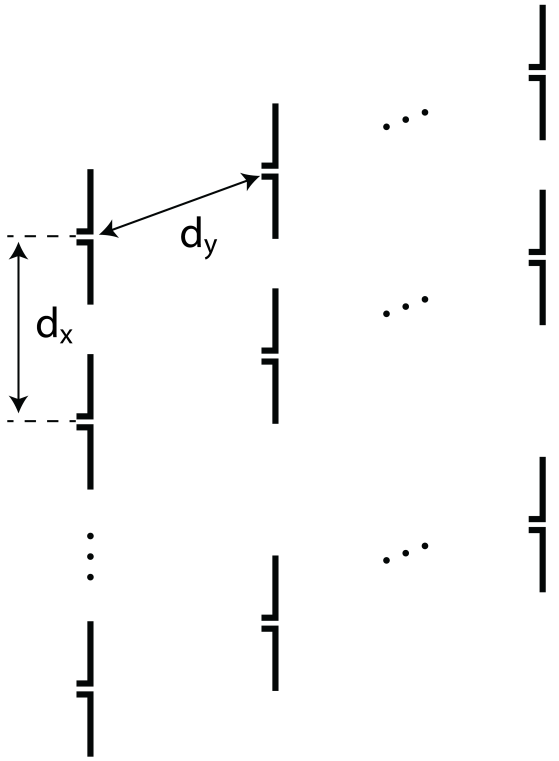


Figure 3.3: Rectangular planar array with inter-element spacings d_x and d_y.

By substituting these parameters into (2.8), we obtain the off-diagonal entries of $\mathbf{Z}_{TT}$. Due to reciprocity, $\mathbf{Z}_{TT}$ is symmetric. Similarly, since the BS transmit array and the SIM array are located on two parallel planes separated by d_z along the z-axis, the entries of $\mathbf{Z}_{ST}$ can be obtained using the same formulation by setting $z = d_z$ in the distance computation. The same procedure is used for the coupling blocks between adjacent SIM layers, as described next. Since (2.4) and (2.8) are evaluated at the center frequency f_i of each subband, all impedance matrices, and therefore the effective channel, differ from one subband to another.

3.3.3 SIM Model

We consider a SIM implemented as an L-layer cascaded structure of transmissive RISs. Each transmissive RIS is modeled as two facing layers with a phase shifter embedded in between. In this configuration, the first layer receives the incident signal, the phase shift is applied, and the second layer subsequently re-transmits the modified signal to the next layer [12]. We assume that the number of antenna elements in each layer is the same, so that each RIS consists of $2M$ ports. Consequently, the total number of ports in the SIM is $2ML$. The effect of the SIM is described by $\mathbf{G} = (\mathbf{Z}_{SS} + \mathbf{Z}_S)^{-1}$ from (3.2), where $\mathbf{G}$ consists of $2L \times 2L$ submatrices $\mathbf{G}_{i,j} \in \mathbb{C}^{M \times M}$:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{1,1} & \mathbf{G}_{1,2} & \cdots & \mathbf{G}_{1,2L} \\ \mathbf{G}_{2,1} & \mathbf{G}_{2,2} & \cdots & \mathbf{G}_{2,2L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_{2L,1} & \mathbf{G}_{2L,2} & \cdots & \mathbf{G}_{2L,2L} \end{bmatrix}. \tag{3.6}$$

Inspired by [33], the SIM model is shown in Fig. 3.4. Both $\mathbf{Z}_{SS}$ and $\mathbf{Z}_S$ exhibit a banded structure and can be decomposed into submatrices of size $M \times M$. The interconnection matrix is

$$\mathbf{Z}_{SS}(f) = \begin{bmatrix} \mathbf{Z}_{2,2}^{(0)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{1,1}^{(1)} & \mathbf{Z}_{1,2}^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{2,1}^{(1)} & \mathbf{Z}_{2,2}^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Z}_{1,1}^{(2)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z}_{1,1}^{(L)} \end{bmatrix}. \tag{3.7}$$

The submatrices $\mathbf{Z}_{i,j}^{(\ell)} \in \mathbb{C}^{M \times M}$, with $\ell = 1, 2, \ldots, L-1$ and $i, j \in \{1, 2\}$, characterize the interconnections between consecutive layers of the SIM. In particular, $\mathbf{Z}_{i,j}^{(\ell)}$ relates the i-th facing of the 2ℓ-th layer to the j-th facing of the $(2\ell+1)$-th layer. We compute

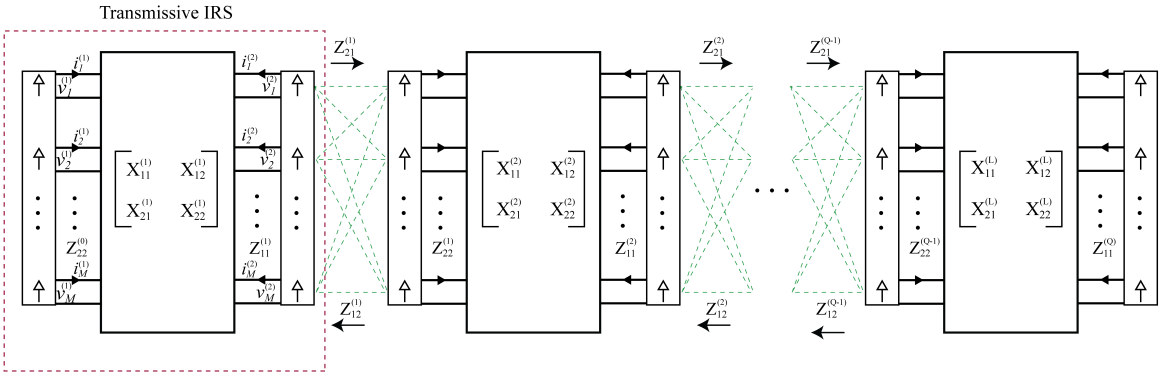


Figure 3.4: Multiport network model of the SIM [33].

these blocks using the method described in Section 3.3.2, where the diagonal blocks contain the self-impedances and the intra-layer coupling, and the off-diagonal blocks contain the coupling between the two facing layers separated by the inter-layer distance. Additionally, $\mathbf{Z}_{2,2}^{(0)} \in \mathbb{C}^{M\times M}$ characterizes the ports associated with the first layer of the SIM, while $\mathbf{Z}_{1,1}^{(L)} \in \mathbb{C}^{M\times M}$ characterizes the ports of the last layer. Since all these blocks are obtained from (2.4) and (2.8), $\mathbf{Z}_{SS}(f)$ is frequency dependent.

The load matrix $\mathbf{Z}_S(\boldsymbol{\phi})$ can be decomposed into submatrices $\mathbf{X}_{i,j}^{(\ell)}(\boldsymbol{\phi}_\ell) \in \mathbb{C}^{M\times M}$, with $\ell = 1, 2, \ldots, L$ and $i, j \in \{1, 2\}$. These $M \times M$ blocks describe the load-network relations within each SIM layer. In particular, $\mathbf{X}_{i,j}^{(\ell)}$ specifies the impedance block associated with the i-th facing of the $(2\ell-1)$-th port set and the j-th facing of the 2ℓ-th port set of the same layer, parameterized by the control vector $\boldsymbol{\phi}_\ell$. For simplicity, we omit the explicit dependence on $\boldsymbol{\phi}_\ell$ in the notation. The global load matrix can be represented by block

submatrices as

$$\mathbf{Z}_S(\boldsymbol{\phi}) = \begin{bmatrix} \mathbf{X}_{1,1}^{(1)} & \mathbf{X}_{1,2}^{(1)} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{X}_{2,1}^{(1)} & \mathbf{X}_{2,2}^{(1)} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_{1,1}^{(2)} & \mathbf{X}_{1,2}^{(2)} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_{2,1}^{(2)} & \mathbf{X}_{2,2}^{(2)} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_{1,1}^{(L)} & \mathbf{X}_{1,2}^{(L)} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_{2,1}^{(L)} & \mathbf{X}_{2,2}^{(L)} \end{bmatrix}. \tag{3.8}$$

We assume that each layer of the SIM applies diagonal phase shifts. Hence, for $i, j \in \{1, 2\}$ and $\ell = 1, 2, \ldots, L$, the submatrices $\mathbf{X}_{i,j}^{(\ell)}(\boldsymbol{\phi}_\ell)$ in (3.8) are diagonal matrices whose entries correspond to the tunable load parameters, i.e.,

$$\mathbf{X}_{i,j}^{(\ell)}(\boldsymbol{\phi}_\ell) = \mathrm{Diag}\Big(x_1^{(\ell)}(i,j),\, x_2^{(\ell)}(i,j),\, \ldots,\, x_M^{(\ell)}(i,j)\Big), \tag{3.9}$$

where $x_m^{(\ell)}(i,j)$ denotes the m-th diagonal entry, parameterized by the phase shift $\phi_{\ell,m}$ associated with the m-th meta-atom of the ℓ-th layer. This diagonal structure corresponds to phase-only control, while any amplitude variations arise implicitly from inter-layer coupling and frequency-dependent effects captured by the circuit-theoretic formulation. By allowing $\mathbf{X}_{i,j}^{(\ell)}$ to be full matrices, the proposed framework can accommodate more general hardware models with joint phase–amplitude control using non-local interactions [11, 29].

To characterize each diagonal entry $x_m^{(\ell)}(i,j)$, we model the m-th transmissive meta-atom of the ℓ-th layer as a two-port network with the scattering parameter matrix

$$\mathbf{S}_m^{(\ell)} = \begin{bmatrix} 0 & e^{j\phi_{\ell,m}} \\ e^{j\phi_{\ell,m}} & 0 \end{bmatrix}, \tag{3.10}$$

which represents full transmission with a tunable phase shift $\phi_{\ell,m}$. Transforming (3.10)

into the Z-parameter domain, we obtain the corresponding two-port impedance matrix

$$\mathbf{Z}_m^{(\ell)} = jZ_0 \begin{bmatrix} \dfrac{\cos\phi_{\ell,m}}{\sin\phi_{\ell,m}} & \dfrac{1}{\sin\phi_{\ell,m}} \\ \dfrac{1}{\sin\phi_{\ell,m}} & \dfrac{\cos\phi_{\ell,m}}{\sin\phi_{\ell,m}} \end{bmatrix}. \tag{3.11}$$

Consequently, each diagonal block $\mathbf{X}_{i,j}^{(\ell)}(\boldsymbol{\phi}_\ell)$ in (3.9) is obtained directly from the Z-parameter representation in (3.11). In this formulation, each block is fully characterized by the set of phase shifts $\{\phi_{\ell,1}, \phi_{\ell,2}, \ldots, \phi_{\ell,M}\}$, which serve as the tunable load parameters of the ℓ-th layer.

Relation to the conventional SIM model. Let us consider the case where inter-layer reflections and mutual coupling are ignored, and all ports are matched at the reference impedance Z_0. In particular, let $\mathbf{Z}_{1,2}^{(\ell)} = \mathbf{0}$ and $\mathbf{Z}_{1,1}^{(\ell)} = \mathbf{Z}_{2,2}^{(\ell-1)} = Z_0\mathbf{I}_M$ for all layer indices ℓ. Under these assumptions, the SIM model reduces to the conventional cascaded model:

$$\mathbf{G}_{2L,1} = \left(-\frac{1}{2Z_0}\right)^L e^{j\boldsymbol{\phi}_L} \prod_{\ell=L-1}^{1} \mathbf{Z}_{2,1}^{(\ell)}\, e^{j\boldsymbol{\phi}_\ell}, \tag{3.12}$$

where $\frac{1}{2Z_0}\mathbf{Z}_{2,1}^{(\ell)}$ is modeled by the Rayleigh–Sommerfeld diffraction equation in conventional models [15, 16], which is questionable as it fails when the antenna element size is not much larger than the wavelength. The proposed model therefore generalizes the conventional model by retaining the coupling and reflection terms that are neglected in (3.12).

End-to-end channel. Moreover, we assume that there is no direct path between the transmitter and the receiver, i.e., $\mathbf{Z}_{RT} = \mathbf{0}$. Under this condition, and absorbing the constant factor $-\frac{1}{4Z_0}$, the transfer function in (3.2) can be written as $\mathbf{H}_Z = \mathbf{Z}_{RS}(\mathbf{Z}_{SS} + \mathbf{Z}_S)^{-1}\mathbf{Z}_{ST}$. Here, $\mathbf{Z}_{ST}$ has only the first M nonzero rows, because only the first facing of the first layer is excited by the transmitter, while $\mathbf{Z}_{RS}$ has only the last M nonzero

columns, since only the last facing of the last layer of the SIM is connected to the receiver. Let $\mathbf{H}_1 \in \mathbb{C}^{M \times N_t}$ denote the matrix formed by the first M rows of $\mathbf{Z}_{ST}$, and $\mathbf{H}_2 \in \mathbb{C}^{K \times M}$ denote the matrix formed by the last M columns of $\mathbf{Z}_{RS}$. Then, the transfer function can be expressed as

$$\mathbf{H}_Z = \mathbf{H}_2 \, \mathbf{G}_{2L,1} \, \mathbf{H}_1. \tag{3.13}$$

Thus, the SIM acts on the signal only through the block $\mathbf{G}_{2L,1}$, which maps the port quantities at the first SIM face to those at the last SIM face through the stacked structure. Note that the single-layer RIS is obtained as the special case $L = 1$ with $M = M_{\text{total}}$.

3.3.4 Far-Field Transimpedance Matrix

The external propagation environment from the last SIM layer to the users is incorporated through the far-field transimpedance matrix $\mathbf{H}_2(f)$, while $\mathbf{H}_1(f)$ characterizes the near-field coupling between the BS transmit array and the first SIM layer and is computed as described in Section 3.3.2. As the users are in the far field, using circuit-theoretic principles together with the Friis transmission equation (cf. Section 2.3.2), the multipath transimpedance can be written as [28, 44]

$$\begin{aligned} \mathbf{H}_2(f) = \frac{c\sqrt{G_S G_R}}{2\pi f \, d_{RS}^{\alpha_{\text{PL}}/2}} \operatorname{Diag}\big(\Re\{\mathbf{Z}_{RR}(f)\}\big)^{1/2} \left(\sum_{e=1}^{E} \mathbf{a}_R(\theta_{R,e}) \, \mathbf{a}_S^T(\varphi_{S,e}, \psi_{S,e}) \right) \\ \times \operatorname{Diag}\big(\Re\{\mathbf{Z}_{1,1}^{(L)}(f)\}\big)^{1/2} e^{-j\phi_0}, \end{aligned} \tag{3.14}$$

where

$$\phi_0 = \pi - \arctan\left(\frac{2\pi f \, a_S}{c}\right) - \arctan\left(\frac{2\pi f \, a_R}{c}\right). \tag{3.15}$$

Here, α_{PL} is the path-loss exponent and d_{RS} denotes the distance between the SIM and the receiver. The SIM and the receiver have gains G_S and G_R, respectively. For the e-th propagation path, the angle of arrival at the receiver is $\theta_{R,e}$, while $(\varphi_{S,e}, \psi_{S,e})$

denote the elevation and azimuth angles measured from the last layer of the $M_x \times M_y$ planar SIM array. These angles are defined with respect to the broadside axes of the receive and SIM arrays, respectively. The diagonal factors $\mathrm{Diag}(\Re\{\mathbf{Z}_{RR}(f)\})^{1/2}$ and $\mathrm{Diag}(\Re\{\mathbf{Z}_{1,1}^{(L)}(f)\})^{1/2}$ embed the port-wise self-impedances at the receiver and at the last layer of the SIM. The array response vectors are

$$\mathbf{a}_R(\theta_R) = \left[1,\ e^{j\frac{2\pi d_R}{\lambda}\cos\theta_R},\ \ldots,\ e^{j\frac{2\pi d_R}{\lambda}(N_R-1)\cos\theta_R}\right]^T, \tag{3.16}$$

$$\mathbf{a}_S(\varphi,\psi) = \begin{bmatrix} 1 \\ e^{j\frac{2\pi d_S}{\lambda}\sin\varphi\,\sin\psi} \\ \vdots \\ e^{j\frac{2\pi d_S}{\lambda}(M_x-1)\sin\varphi\,\sin\psi} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ e^{j\frac{2\pi d_S}{\lambda}\sin\varphi\,\cos\psi} \\ \vdots \\ e^{j\frac{2\pi d_S}{\lambda}(M_y-1)\sin\varphi\,\cos\psi} \end{bmatrix}, \tag{3.17}$$

where d_S denotes the inter-element spacing of the SIM uniform planar array and d_R denotes the inter-element spacing of the receive array. The k-th row of $\mathbf{H}_2(f)$ corresponds to user k, which is at distance $d_{RS,k}$ and whose propagation paths have their own angles. Note that the wavelength $\lambda = c/f$ in the array responses changes across subbands, which captures the beam squint of the wideband system.

3.3.5 Achievable Sum Rate

We consider $\mathbf{h}_{k,i} \in \mathbb{C}^{1\times N_t}$, the k-th row of $\mathbf{H}_Z$ evaluated at f_i, as the effective downlink channel between the BS and user k on subband i, and $\mathbf{w}_{k,i} \in \mathbb{C}^{1\times N_t}$ as the corresponding transmit beamforming vector. The BS employs linear precoding across all users and subbands to mitigate multi-user interference. Inter-user interference is modeled as colored noise. Accordingly, the interference-plus-noise covariance matrix in the transmit domain for user k on subband i is given by

$$\mathbf{C}_{k,i} = \sigma_i^2\mathbf{I}_{N_t} + \sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{h}_{j,i}^H\mathbf{w}_{j,i}\mathbf{w}_{j,i}^H\mathbf{h}_{j,i}, \qquad \mathbf{C}_{k,i} \in \mathbb{C}^{N_t\times N_t}. \tag{3.18}$$

The achievable sum rate of the SIM-assisted system can then be written in determinant form as

$$R = \sum_{i=1}^{N_f} \sum_{k=1}^{K} B_i \Big[\log_2 \Big| \mathbf{C}_{k,i} + \mathbf{h}_{k,i}^H \mathbf{w}_{k,i} \mathbf{w}_{k,i}^H \mathbf{h}_{k,i} \Big| - \log_2 \big| \mathbf{C}_{k,i} \big| \Big]. \tag{3.19}$$

Since $\mathbf{C}_{k,i}$ is positive semidefinite, using the eigendecomposition $\mathbf{C}_{k,i} = \mathbf{U}_{k,i} \mathbf{\Lambda}_{k,i} \mathbf{U}_{k,i}^H$, we obtain the whitened channel $\widetilde{\mathbf{h}}_{k,i} = \mathbf{h}_{k,i} \mathbf{U}_{k,i} \mathbf{\Lambda}_{k,i}^{-1/2}$. Applying the matrix determinant lemma, the sum rate can be rewritten as

$$R = \sum_{i=1}^{N_f} \sum_{k=1}^{K} B_i \log_2 \Big| \mathbf{I} + \widetilde{\mathbf{h}}_{k,i}^H \mathbf{w}_{k,i} \mathbf{w}_{k,i}^H \widetilde{\mathbf{h}}_{k,i} \Big|. \tag{3.20}$$

3.4 Problem Formulation and Solution

3.4.1 Problem Formulation

For each subband $i \in \{1, \ldots, N_f\}$, let $\mathbf{H}_{1,i} \in \mathbb{C}^{M \times N_t}$ denote the impedance matrix from the BS with N_t antennas to the first facing of the SIM with $M = M_x M_y$ antenna elements per layer, and let $\mathbf{H}_{2,i} \in \mathbb{C}^{K \times M}$ denote the impedance matrix from the last facing of the SIM to the K single-antenna users. The n-th column of $\mathbf{H}_{1,i}$ is $\mathbf{h}_{1,i}^{(n)} \in \mathbb{C}^{M \times 1}$, so that $\mathbf{H}_{1,i} = [\, \mathbf{h}_{1,i}^{(1)}, \ldots, \mathbf{h}_{1,i}^{(N_t)} \,]$. The k-th row of $\mathbf{H}_{2,i}$ is $(\mathbf{h}_{2,i}^{(k)})^T$ with $\mathbf{h}_{2,i}^{(k)} \in \mathbb{C}^{M \times 1}$, hence $\mathbf{H}_{2,i} = [\, \mathbf{h}_{2,i}^{(1)}, \ldots, \mathbf{h}_{2,i}^{(K)} \,]^T$. The SIM phase vector is $\boldsymbol{\phi} = [\phi_{1,1}, \ldots, \phi_{L,M}]^T$ for L layers. Let $\mathbf{w}_{k,i} \in \mathbb{C}^{1 \times N_t}$ be the transmit beamformer for user k on subband i. The effective BS–user channel on subband i is

$$\mathbf{h}_{k,i} = \big(\mathbf{h}_{2,i}^{(k)}\big)^T \mathbf{G}_{2L,1}(\boldsymbol{\phi}, f_i)\, \mathbf{H}_{1,i} \in \mathbb{C}^{1 \times N_t}. \tag{3.21}$$

The optimization problem to maximize the wideband sum rate is formulated as

$$R^\star = \max_{\{\mathbf{w}_{k,i}\},\{p_{k,i}\},\boldsymbol{\phi}} \quad \sum_{i=1}^{N_f}\sum_{k=1}^{K} B_i \, \log_2\left|\mathbf{I}+\boldsymbol{\gamma}_{k,i}\right| \tag{3.22a}$$

$$\text{s.t.} \quad p_{k,i} = \mathrm{Tr}\big(\mathbf{w}_{k,i}\mathbf{w}_{k,i}^H\big), \quad \forall i,k, \tag{3.22b}$$

$$\sum_{i=1}^{N_f}\sum_{k=1}^{K} p_{k,i} \le P_{\mathrm{tot}}, \qquad p_{k,i}\ge 0, \tag{3.22c}$$

$$0 \le \phi_{\ell,m} < 2\pi, \quad \forall \ell,m, \tag{3.22d}$$

where $\boldsymbol{\gamma}_{k,i} = \widetilde{\mathbf{h}}_{k,i}^H \mathbf{w}_{k,i}\mathbf{w}_{k,i}^H \widetilde{\mathbf{h}}_{k,i}$, $p_{k,i}$ is the power allocated to user k on subband i, and P_{tot} is the total transmit power budget. Constraint (3.22d) reflects the phase-only control of the meta-atoms.

3.4.2 Alternating Optimization

Problem (3.22) is non-convex, since the beamformers and the SIM phase shifts are coupled through the effective channels, and the phase shifts enter the channel through the matrix inverse in $\mathbf{G}$. To obtain a solution, we adopt an alternating optimization (AO) algorithm that decomposes the problem into the following two subproblems:

$$\textbf{P1:} \quad \max_{\{p_{k,i}\},\{\mathbf{w}_{k,i}\}} \quad \sum_{i=1}^{N_f}\sum_{k=1}^{K} B_i \log_2\left|\mathbf{I}+\widetilde{\mathbf{h}}_{k,i}^H \mathbf{w}_{k,i}\mathbf{w}_{k,i}^H \widetilde{\mathbf{h}}_{k,i}\right| \quad \text{s.t.} \quad (3.22\text{b}),\ (3.22\text{c}), \tag{3.23}$$

$$\textbf{P2:} \quad \max_{\boldsymbol{\phi}} \quad \sum_{i=1}^{N_f}\sum_{k=1}^{K} B_i \log_2\left|\mathbf{I}+\widetilde{\mathbf{h}}_{k,i}^H \mathbf{w}_{k,i}\mathbf{w}_{k,i}^H \widetilde{\mathbf{h}}_{k,i}\right| \quad \text{s.t.} \quad (3.22\text{d}). \tag{3.24}$$

In this AO approach, the beamforming vectors and the phase shifts are optimized alternately: **P1** optimizes the beamformers and the power allocation for fixed SIM phases, and **P2** optimizes the SIM phases for fixed beamformers. This is a widely adopted

strategy, since it decouples a non-convex problem into subproblems that can be solved iteratively with reduced computational overhead. The overall procedure is summarized in Algorithm 1.

Power Allocation and Beamforming Optimization

Subproblem **P1** maximizes the wideband sum rate with fixed SIM phases $\boldsymbol{\phi}$. We use the iterative water-filling approach of [47] and adapt it to the multi-band setting in [37]. For user k on subband i, all other users are treated as interference, captured by $\mathbf{C}_{k,i}$ in (3.18). The corresponding scalar gain is

$$\lambda_{k,i} = \|\widetilde{\mathbf{h}}_{k,i}\|^2 = \mathbf{h}_{k,i}\mathbf{C}_{k,i}^{-1}\mathbf{h}_{k,i}^H. \tag{3.25}$$

Let $\mathbf{Q}_{k,i}\,\boldsymbol{\Delta}_{k,i}^{1/2}\,\mathbf{R}_{k,i}^H$ and $\mathbf{S}_{k,i}\,\boldsymbol{\nabla}_{k,i}^{1/2}\,\mathbf{T}_{k,i}^H$ denote the singular value decompositions of $\widetilde{\mathbf{h}}_{k,i}^H$ and $\mathbf{w}_{k,i}^H$, respectively, where $\mathbf{Q}_{k,i}, \mathbf{R}_{k,i}, \mathbf{S}_{k,i}, \mathbf{T}_{k,i}$ are unitary matrices. The objective is maximized by aligning the beamformer with the whitened channel, i.e., by choosing $\mathbf{S}_{k,i} = \mathbf{Q}_{k,i}$. Hence, the maximization reduces to the power allocation problem

$$\begin{aligned} \max_{\{p_{k,i}\}} \; & \sum_{i=1}^{N_f}\sum_{k=1}^{K} B_i \,\log_2\left(1+\lambda_{k,i}\,p_{k,i}\right) \\ \text{s.t.} \; & \sum_{i=1}^{N_f}\sum_{k=1}^{K} p_{k,i} \le P_{\text{tot}}, \quad p_{k,i} \ge 0. \end{aligned} \tag{3.26}$$

We initialize with any feasible $\{p_{k,i}^{(0)}\}$. At iteration t, for each user–subband pair we construct $\mathbf{C}_{k,i}^{(t)}$ from the current powers, compute $\widetilde{\mathbf{h}}_{k,i}^{(t)}$, and set $\lambda_{k,i}^{(t)} = \|\widetilde{\mathbf{h}}_{k,i}^{(t)}\|^2$. The power is then updated by iterative water-filling as

$$p_{k,i}^{(t+1)} = \Big[\mu^{(t)} - \frac{1}{\lambda_{k,i}^{(t)}}\Big]^+, \tag{3.27}$$

where the water level $\mu^{(t)}$ enforces $\sum_{i=1}^{N_f}\sum_{k=1}^{K} p_{k,i}^{(t+1)} = P_{\text{tot}}$, and the operator $[x]^+ = \max(x, 0)$ returns x if $x \ge 0$ and 0 otherwise. Thus, user–subband pairs with stronger

effective gains receive more power, while pairs with $\lambda_{k,i}^{(t)} < 1/\mu^{(t)}$ are switched off. After the power update, the optimal beamformer is obtained as

$$\left(\mathbf{w}_{k,i}^{(t+1)}\right)^T = \mathbf{Q}_{k,i} \begin{bmatrix} \sqrt{p_{k,i}^{(t+1)}} \\ \mathbf{0} \end{bmatrix}. \tag{3.28}$$

This optimization process inherently performs carrier aggregation over the wideband spectrum. Through the water-filling mechanism, the transmit power is selectively distributed across subbands according to their instantaneous channel gains and interference conditions. Consequently, only subbands that contribute meaningfully to the overall spectral efficiency are allocated nonzero power, while those with unfavorable propagation conditions remain inactive. This adaptive subband activation effectively emulates the carrier aggregation behavior of practical multi-band systems, thereby ensuring efficient utilization of the spectral resources in the SIM-assisted wideband system.

Phase Shift Optimization

With the beamformers fixed, we maximize the sum rate with respect to the SIM phases $\boldsymbol{\phi}$. The objective of **P2** can be written as

$$\mathcal{J}(\boldsymbol{\phi}) = \sum_{i=1}^{N_f} \sum_{k=1}^{K} B_i \, \log_2\left(1 + s_{k,i}(\boldsymbol{\phi})\right), \qquad s_{k,i}(\boldsymbol{\phi}) = p_{k,i} \, \mathbf{h}_{k,i} \, \mathbf{C}_{k,i}^{-1} \, \mathbf{h}_{k,i}^H, \tag{3.29}$$

where $s_{k,i}$ is the effective signal-to-interference-plus-noise ratio (SINR) of user k on subband i, and the effective channel $\mathbf{h}_{k,i}$ depends on $\boldsymbol{\phi}$ through (3.21). To optimize the phase shifts, we use gradient ascent with step size $\eta_t > 0$:

$$\phi_{\ell,m}^{(t+1)} = \phi_{\ell,m}^{(t)} + \eta_t \, \frac{\partial \mathcal{J}}{\partial \phi_{\ell,m}}, \tag{3.30}$$

followed by mapping each phase onto $[0, 2\pi)$. To obtain the gradient, let $\mathbf{F}_{\ell,m}^i(\boldsymbol{\phi})$ denote the derivative of the SIM block $\mathbf{G}_{2L,1}$ with respect to the phase of the m-th element of

the ℓ-th layer on subband i (derived in Appendix A), and define

$$\boldsymbol{\Delta}_{k,i} = -\Big[\mathbf{H}_{2,i}\,\mathbf{F}^{i}_{\ell,m}(\boldsymbol{\phi})\,\mathbf{H}_{1,i}\Big]_{k,:}, \qquad \mathbf{v}_{k,i} = \mathbf{C}^{-1}_{k,i}\mathbf{h}^{H}_{k,i}, \tag{3.31}$$

where $[\cdot]_{k,:}$ denotes the k-th row, so that $\boldsymbol{\Delta}_{k,i} = \partial \mathbf{h}_{k,i}/\partial\phi_{\ell,m}$. The derivative of the SINR $s_{k,i}$ with respect to $\phi_{\ell,m}$ is given by (see Appendix A)

$$\frac{\partial s_{k,i}}{\partial \phi_{\ell,m}} = p_{k,i}\left[2\,\Re\big\{\boldsymbol{\Delta}_{k,i}\,\mathbf{v}_{k,i}\big\} - 2\sum_{k'\neq k} p_{k',i}\,\Re\Big\{\big(\mathbf{h}_{k',i}\mathbf{v}_{k,i}\big)^{*}\big(\boldsymbol{\Delta}_{k',i}\mathbf{v}_{k,i}\big)\Big\}\right]. \tag{3.32}$$

Combining this with the derivative of (3.29) gives the gradient of the objective function:

$$\frac{\partial \mathcal{J}}{\partial \phi_{\ell,m}} = \sum_{i=1}^{N_f}\sum_{k=1}^{K} \frac{B_i}{\ln 2}\,\frac{1}{1+s_{k,i}}\,\frac{\partial s_{k,i}}{\partial \phi_{\ell,m}}. \tag{3.33}$$

The phase shifts are then updated using (3.30). Note that the gradient accumulates the contributions of all subbands, so that the common phase configuration is optimized for the whole wideband spectrum rather than for a single frequency.

Algorithm 1 Alternating optimization for the SIM-assisted wideband system

Require: P_{tot}, $\{\mathbf{H}_{1,i}, \mathbf{H}_{2,i}, \mathbf{Z}_{SS}(f_i)\}_{i=1}^{N_f}$, tolerance ϵ

1: Initialize $\boldsymbol{\phi}$ and a feasible power allocation $\{p_{k,i}\}$

2: **repeat**

3: Compute $\mathbf{G}_{2L,1}(\boldsymbol{\phi}, f_i)$ and $\mathbf{h}_{k,i}$ for all i,k using (3.21)

4: **P1:** update $\{p_{k,i}\}$ by iterative water-filling (3.27) and $\{\mathbf{w}_{k,i}\}$ by (3.28) until convergence

5: **P2:** update $\boldsymbol{\phi}$ by gradient ascent (3.30) using (3.33)

6: **until** $|\mathcal{J}^{(t+1)} - \mathcal{J}^{(t)}|/|\mathcal{J}^{(t)}| < \epsilon$

Ensure: $\{\mathbf{w}_{k,i}\}$, $\{p_{k,i}\}$, $\boldsymbol{\phi}$

3.4.3 Convergence and Complexity

The proposed AO algorithm exhibits a monotonic behavior, where each update of the beamforming and phase-shift variables ensures a non-decreasing objective function value. The iterations proceed until the change in the objective value between two consecutive iterations becomes sufficiently small. Specifically, the convergence criterion is defined as $|\mathcal{J}^{(t+1)} - \mathcal{J}^{(t)}|/|\mathcal{J}^{(t)}| < \epsilon$, where ϵ is a small positive threshold (10^{-7}).

The number of mathematical operations required for the different steps of the algorithm is summarized in Table 3.1. During the beamforming optimization stage, the steps corresponding to B_1 through B_8 are computed only once, as the equivalent channel remains fixed in this stage. Subsequently, for each of the I_{P1} beamforming iterations, the updates in B_9 and B_{10} are executed. Once the beamforming is optimized, it is kept constant, and the phase shifts are updated. In this stage, the equivalent channel $\widetilde{\mathbf{h}}_{k,i}$ is recomputed at each iteration, and the operations corresponding to B_{11}, B_{12}, and B_{13} are used. I_{P2} iterations are used to perform this stage. Therefore, the total computational complexity of one full iteration of the AO algorithm can be expressed as

$$(B_1 + B_2 + \cdots + B_8 + B_{11} + B_{12} + B_{13})\, I_{P2} + (B_9 + B_{10})\, I_{P1}. \tag{3.34}$$

3.4.4 Partially Reconfigurable SIM

As the number of antenna elements increases, the overall system complexity and control overhead of fully reconfigurable SIM and RIS architectures grow significantly. Each additional element increases the number of control bits that must be transmitted whenever the metasurface configuration is updated, which makes real-time optimization challenging in large-scale deployments. To address this issue, we consider a partially reconfigurable SIM design that leverages the flexibility of the SIM while reducing the complexity of reconfiguration. In this configuration, the first layer is optimized once for a given

Table 3.1: Number of mathematical operations required for different steps of the optimization.

Step	**Quantity**	**Number of mathematical operations**
B_1	$\mathbf{G}_{:,1}$ and $\mathbf{G}_{2L,:}$	$(18L-6)\,N_f\,M^3 + (2N_f-1)\,N\,K$
B_2	$\mathbf{H}_Z$	$2N_t\,(M^2+KM)\,N_f$
B_3	$\mathbf{F}$	$(3M^2+6M)\,L\,M\,N_f$
B_4	$\partial\mathbf{H}_Z/\partial\boldsymbol{\phi}$	$2N_t\,(M^2+KM)\,L\,M\,N_f$
B_5	$\mathbf{C}_{k,i}$	$\big(\big((2N_t-1)+(2N_t^2+1)\big)(K-1)+N_t^2\big)\,K\,N_f$
B_6	$\mathbf{U}_{k,i}$ and $\boldsymbol{\Lambda}_{k,i}$	$\mathcal{O}(N_t^4)\,K\,N_f$
B_7	$\widetilde{\mathbf{h}}_{k,i}$	$K\,(2N_t^2+N_t)\,N_f$
B_8	$s_{k,i}$	$2K\,N_t\,N_f$
B_9	Power allocation	$3N_fK-1$
B_{10}	$\mathbf{w}_{k,i}^{(t+1)}$ update	$2K\,N_t\,N_f$
B_{11}	$\partial s_{k,i}/\partial\phi_{\ell,m}$	$N_f\,K\,\big(2N_t+(K-1)(2N_t^2+2N_t)\big)$
B_{12}	$\nabla_{\boldsymbol{\phi}}\mathcal{J}(\boldsymbol{\phi})$	$2L\,M\,K\,N_f$
B_{13}	Phase update $\boldsymbol{\phi}^{(t+1)}$	$2L\,M$

channel realization and then kept fixed, so that it provides a stable wavefront transformation, while only the last layer is adaptively reconfigured to compensate for residual multi-user interference and wideband phase variations. This approach substantially reduces the number of control parameters and the corresponding signaling updates, thereby improving the operational efficiency of the system.

To quantify the impact of control overhead, we evaluate the system performance in terms of goodput, which accounts for both the achievable rate and the signaling cost.

The goodput is expressed as [11]

$$\mathcal{G} = \left(1 - \frac{b\, M_{\text{upd}}}{\eta_c\, N_s}\right) R, \tag{3.35}$$

where the penalty term reflects the fraction of each slot used for metasurface control signaling. Here, M_{upd} denotes the number of reconfigurable elements that are updated in a slot, and each element is controlled by b bits. Therefore, updating M_{upd} elements requires $b\, M_{\text{upd}}$ control bits per slot. Moreover, η_c denotes the spectral efficiency of the control channel and N_s denotes the number of transmitted symbols per slot, so that $\frac{b\, M_{\text{upd}}}{\eta_c N_s}$ represents the normalized signaling cost. For a fully reconfigurable SIM, $M_{\text{upd}} = LM$, where L and M denote the number of layers and elements per layer, respectively. In contrast, for the partially reconfigurable configuration where only the last layer is adjustable, $M_{\text{upd}} = M$, which reduces the required control signaling by a factor of L. In addition to the reduction in control bits, the computational complexity of the optimization is also reduced. In particular, at each iteration, the number of mathematical operations required to compute B_3, B_4, B_{12}, and B_{13} in Table 3.1 is reduced by a factor of L.

3.5 Simulation Results and Discussion

3.5.1 Simulation Setup

In this section, Monte Carlo simulation results are presented to numerically evaluate the developed model. The spectral efficiency R denotes the system-wide achievable sum rate in bps/Hz, and the band is partitioned into N_f subbands of equal width. The inter-element spacing is half the wavelength at the center frequency; since the system operates over a wide bandwidth, the electrical spacing becomes frequency dependent across subbands, which naturally captures wideband effects such as beam squint and

frequency-dependent array responses. Single-antenna users are randomly placed in the three-dimensional region in front of the SIM. The main simulation parameters are summarized in Table 3.2, and are used unless mentioned otherwise.

Table 3.2: Simulation parameters.

Parameter	**Value**
Center frequency f_c	27 GHz
Total bandwidth B	15 GHz
Total transmit power P_{tot}	50 W
Antenna element radius a	2.7 mm ($\lambda_c/4$)
Inter-element spacing $d_x = d_y$	5.56 mm ($\lambda_c/2$)
Distance between BS array and SIM	$5\lambda_c$
User distance from the SIM	$\mathcal{U}[10, 20]$ m
Noise power σ^2	-70 dB
Path-loss exponent α_{PL}	2
Number of propagation paths E	1
Number of BS antennas N_t	K
Control bits per element b	2
Control channel spectral efficiency η_c	2
Symbols per slot N_s	700
Convergence threshold ϵ	10^{-7}

3.5.2 Effect of the Number of Subbands

We first investigate the effect of optimizing the phase shifts across multiple subbands over a wideband channel under uniform power allocation. Fig. 3.5 illustrates the influence of the number of subbands on the spectral efficiency for a system employing a total of 100 antenna elements. The single-layer RIS configuration is designed as a 10×10 array. For the multilayer case ($L = 2$), each layer consists of a 10×5 metasur-

face, ensuring that the total number of antenna elements remains constant across both configurations.

The results show that increasing the number of subbands significantly enhances the spectral efficiency, as the optimization captures the frequency-selective channel behavior more accurately. At lower subband counts, the multilayer metasurface ($L = 2$) exhibits superior performance compared to the single-layer configuration due to its cascaded structure. However, as the number of subbands increases, the performance gap between the single-layer and multilayer configurations narrows. Beyond approximately 15 subbands, the achievable spectral efficiency tends to saturate, indicating that the wideband optimization has effectively converged and that further subband divisions yield negligible additional gain. Therefore, $N_f = 15$ is used in the remaining multi-subband results.

3.5.3 Effect of the Number of Users

Next, we investigate the effect of the number of users on the performance of the single-layer RIS and the multilayer SIM. Fig. 3.6(a) illustrates the variation of the spectral efficiency with the number of users for a total of 100 antenna elements. For a small number of users, where inter-user interference is relatively low, the single-layer RIS exhibits better performance due to its larger effective aperture area. However, as the number of users increases, inter-user interference becomes more dominant. In this regime, the single-layer structure struggles to generate sufficiently complex beam patterns capable of effectively separating users. In contrast, the multilayer SIM, due to its additional spatial degrees of freedom, begins to outperform the single-layer RIS while maintaining the same total number of elements. It is observed that beyond approximately $K = 5$ users, the SIM achieves a higher spectral efficiency than the RIS under wideband operation with the same total number of antenna elements. Moreover, the performance gap

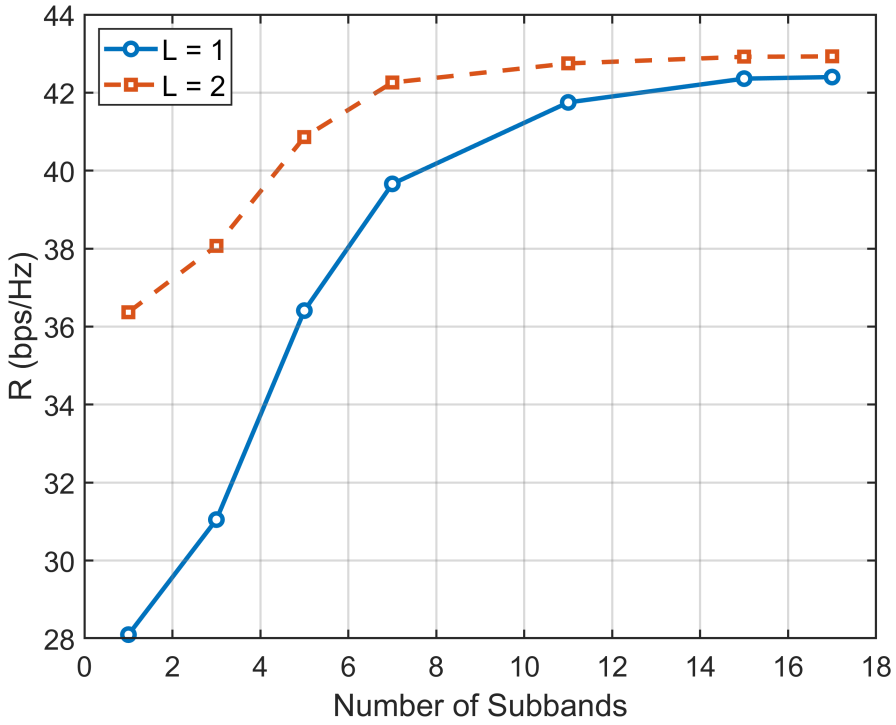
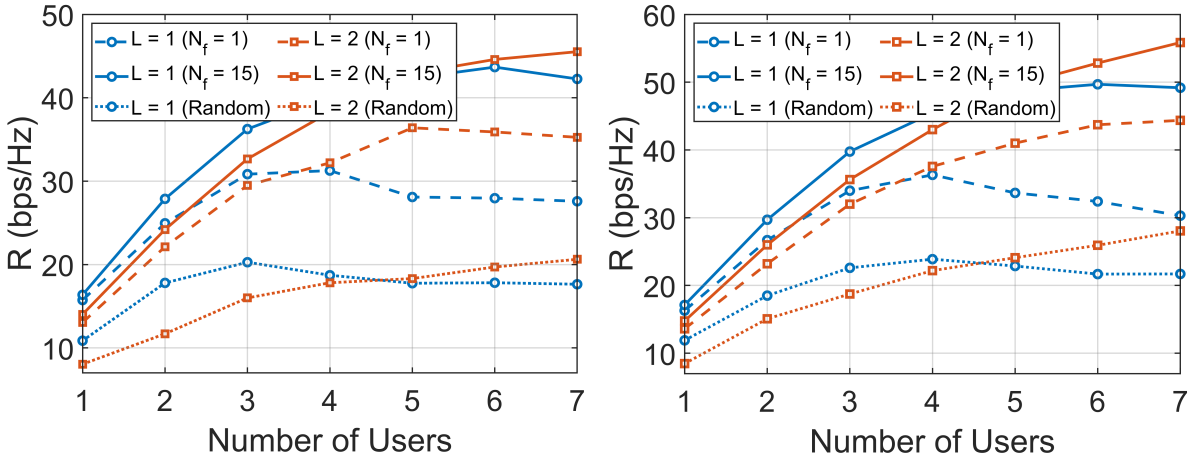


Figure 3.5: Variation of the optimal spectral efficiency with the number of subbands for $K = 5$ and $M_{\text{total}} = 100$.

between single-subband ($N_f = 1$) and multiple-subband ($N_f = 15$) optimization widens as the number of users increases, highlighting the increasing frequency selectivity of the channel in multi-user wideband scenarios. This demonstrates the necessity of optimizing the metasurface phase shifts across multiple subbands to fully exploit the wideband potential of metasurface-assisted systems.

Similarly, Fig. 3.6(b) presents the results for a larger metasurface consisting of 144 antenna elements. As the size of the metasurface increases, the difference in spectral efficiency between the single-layer and multilayer configurations becomes more pronounced, especially at higher user counts. This behavior arises because larger RISs exhibit stronger frequency-selective responses. Therefore, employing a cascaded (multilayer) structure proves more beneficial than merely increasing the number of elements

within a single layer, as the additional depth provides a more effective way to enhance spectral efficiency.



(a) $M_{\text{total}} = 100$ (b) $M_{\text{total}} = 144$

Figure 3.6: Spectral efficiency versus the number of users for the single-layer RIS and the SIM, with single-subband ($N_f = 1$) and multi-subband ($N_f = 15$) optimization.

3.5.4 Impact of Mutual Coupling and Power Allocation

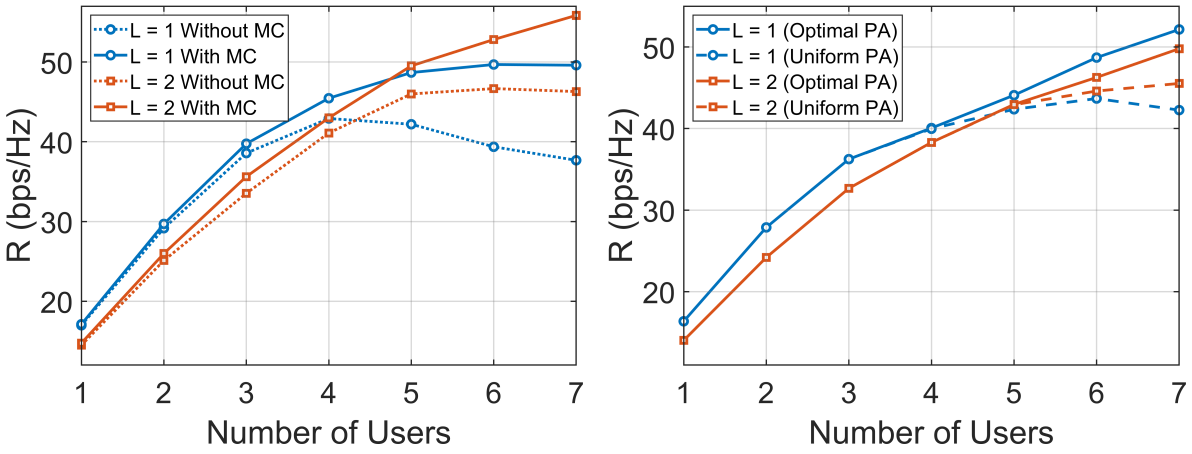
In Fig. 3.7(a), for a total of 144 antenna elements, we illustrate the effect of mutual coupling and mutual reflections in the proposed model, in comparison with the conventional model, where these electromagnetic effects are ignored. We observe that accounting for these electromagnetic effects improves the spectral efficiency relative to the conventional model. For a small number of users, the effect of ignoring mutual coupling is relatively insignificant, as inter-user interference and inter-element interactions remain limited. However, as the number of users increases, the impact of mutual coupling becomes more pronounced, resulting in a noticeable performance gap between the proposed and conventional models. The single-layer RIS, having a larger effective aperture area, is more susceptible to distortions introduced by coupling among closely spaced elements.

This highlights the importance of modeling mutual coupling and mutual reflections, especially for higher user densities and closely spaced elements.

In Fig. 3.7(b), we illustrate the impact of employing optimal power allocation (PA) at the BS on the spectral efficiency of both the single-layer RIS and the multilayer SIM. For the RIS, the effect of power allocation becomes significant beyond $K = 4$ users, where the combination of active beamforming at the transmitter and passive beamforming at the metasurface enables it to outperform the SIM with the same total number of antenna elements. For the SIM, power allocation provides a noticeable improvement after approximately $K = 5$ users, as the system benefits from additional degrees of freedom for reducing interference. Nevertheless, when both active and passive beamforming are jointly optimized, the single-layer RIS consistently achieves a higher spectral efficiency than the multilayer SIM under the same total number of antenna elements. This is because the RIS has a higher raw aperture gain due to its larger surface, while the SIM has an advantage in inter-user interference separation. When active beamforming is performed at the BS, the digital precoder already mitigates inter-user interference, making the advantage of the SIM less significant. As a result, the overall performance is dominated by the raw aperture gain.

3.5.5 Effect of SNR and Bandwidth

Next, we analyze the effect of the SNR (varying noise variance) on the spectral efficiency of both the RIS and the SIM in a multi-user system. Fig. 3.8 presents a comparison between the two configurations for a total of 100 antenna elements. At low SNR values (high noise variance), the RIS shows slightly better performance, since noise dominates and the impact of inter-user interference remains limited. However, as the SNR increases, the SIM begins to outperform the RIS. This is because, in high-SNR regimes, inter-user interference becomes the primary performance-limiting factor, and the addi-



(a) With and without mutual coupling

(b) With and without power allocation

Figure 3.7: Spectral efficiency versus the number of users: (a) proposed coupling-aware model versus the conventional model, and (b) with and without optimal power allocation at the BS.

tional spatial degrees of freedom offered by the SIM enable more effective interference suppression and beam control.

In Fig. 3.9, we examine the impact of the bandwidth on the spectral efficiency. It can be observed that in the narrowband regime, the single-layer configuration achieves a slightly higher spectral efficiency. However, as the bandwidth increases, the multilayer SIM begins to outperform the single-layer RIS, demonstrating its superior capability in mitigating frequency-selective effects and maintaining performance over wideband channels.

3.5.6 Frequency Response

In Fig. 3.10, we illustrate the variation of the spectral efficiency with frequency for both the RIS and the SIM when optimized over a single subband and multiple subbands, for $K = 5$ users and 144 antenna elements. When optimized over a single subband

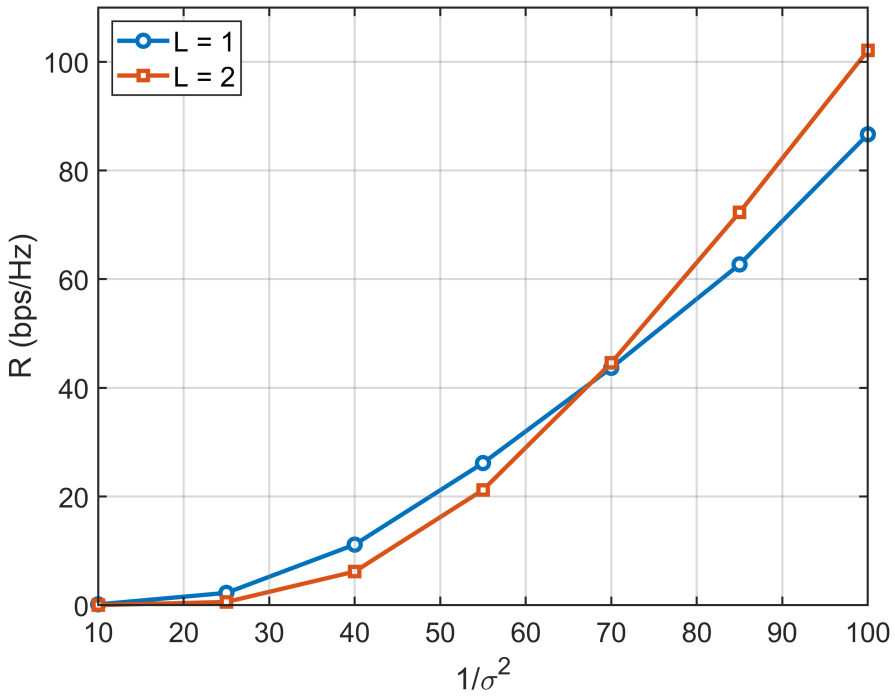


Figure 3.8: Variation of the spectral efficiency with the noise variance for $K = 6$ and $M_{\text{total}} = 100$.

($N_f = 1$) at the center frequency $f_c = 27$ GHz, the system achieves its highest gain near the center frequency but exhibits performance degradation at other frequencies due to frequency selectivity. The SIM shows a slightly lower peak gain at the center frequency compared to the single-layer RIS. However, it maintains a better performance across the wider bandwidth owing to its multilayer structure, which provides additional spatial degrees of freedom. When the optimization is extended across multiple subbands ($N_f = 15$), the overall gain at the center frequency slightly decreases, but the response becomes significantly flatter across the spectrum, resulting in improved robustness and more consistent spectral efficiency over the entire wideband range.

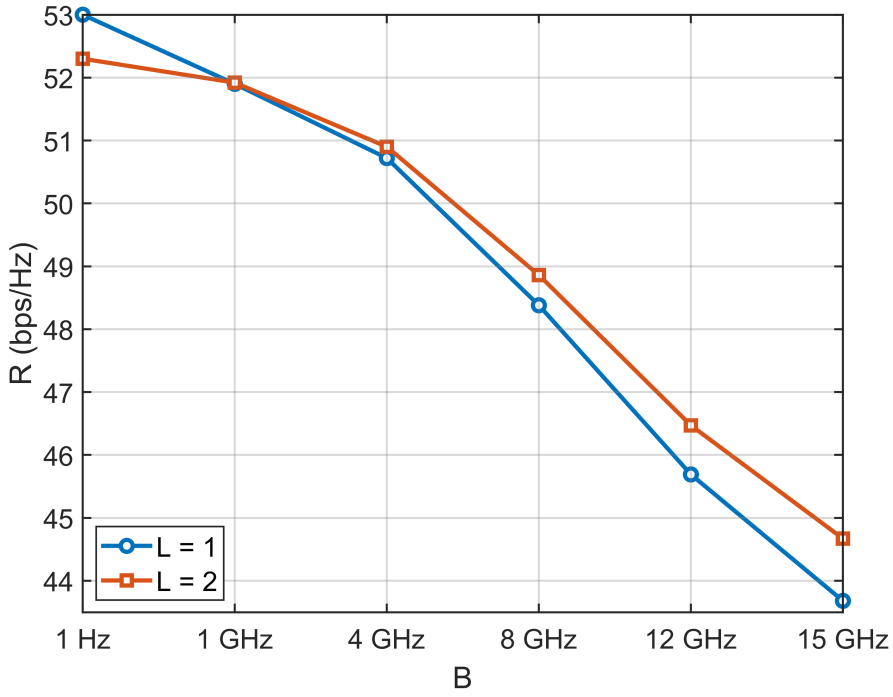


Figure 3.9: Variation of the spectral efficiency with the bandwidth B for $K = 6$.

3.5.7 Goodput of the Partially Reconfigurable SIM

The previous results showed that the single-layer RIS performs better for a small number of users, since inter-user interference is limited, whereas the SIM begins to outperform the RIS in higher SNR regimes and in scenarios with more users, where interference suppression becomes more critical. In Fig. 3.11, we present the variation of the goodput (3.35) with the number of antenna elements for $K = 2$ at SNR $= 110$ dB. The results show that the SIM architecture provides additional flexibility through its multi-layer design. Specifically, when only the second layer of the SIM is reconfigurable while the first layer remains fixed, the system achieves a higher goodput compared to both the single-layer RIS and the fully reconfigurable SIM. This indicates that partial reconfiguration can offer a favorable trade-off between performance and complexity for systems with a small number of users, whereas a fully reconfigurable SIM is more advantageous

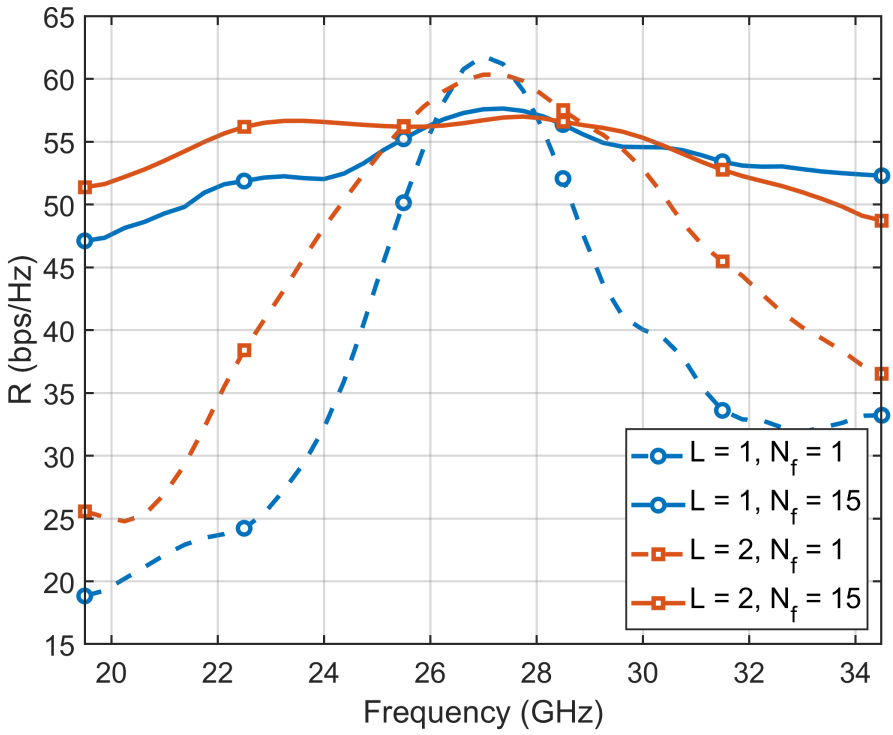


Figure 3.10: Variation of the spectral efficiency with frequency for $K = 5$ and $M_{\text{total}} = 144$.

in multi-user or high-SNR scenarios, where enhanced degrees of freedom are required to manage interference effectively.

3.5.8 Convergence

Finally, Fig. 3.12 shows the convergence of the proposed algorithm for $M_{\text{total}} = 100$ and $K = 4$. All schemes converge stably, and the SIM converges faster than the single-layer RIS. Moreover, the partially reconfigurable SIM not only uses fewer control bits but also converges within a few hundred iterations. In contrast, the single-layer RIS attains a slightly higher final spectral efficiency but requires more iterations, highlighting a trade-off between performance and complexity.

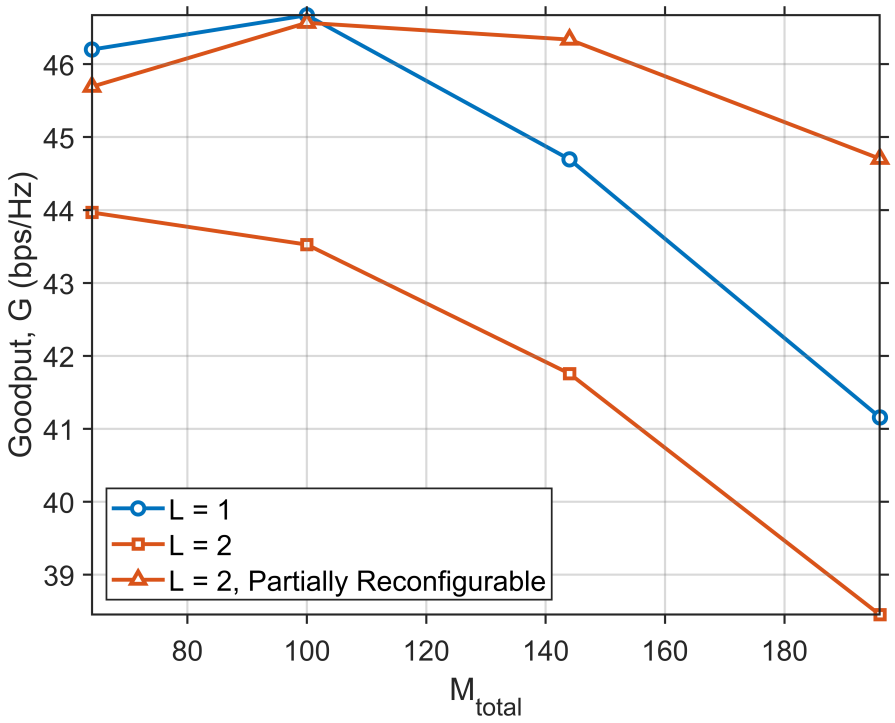


Figure 3.11: Variation of the goodput with the number of antenna elements for $K = 2$.

3.6 Summary

In this chapter, we have presented a circuit-theoretic framework for a SIM-assisted wideband multi-user MIMO system that captures mutual coupling, inter-layer reflections, and frequency selectivity. We have jointly optimized the SIM phases, the BS beamforming, and the power allocation across subbands, where the water-filling solution performs carrier aggregation. For the same total number of elements, the SIM has outperformed the single-layer RIS beyond approximately five users, at high SNR, and over wide bandwidths, while the RIS has performed better with few users, at low SNR, in the narrowband case, and when power allocation is used at the BS. We have also observed that modeling mutual coupling improves the spectral efficiency, that optimizing over approximately 15 subbands gives a much flatter response than single-frequency

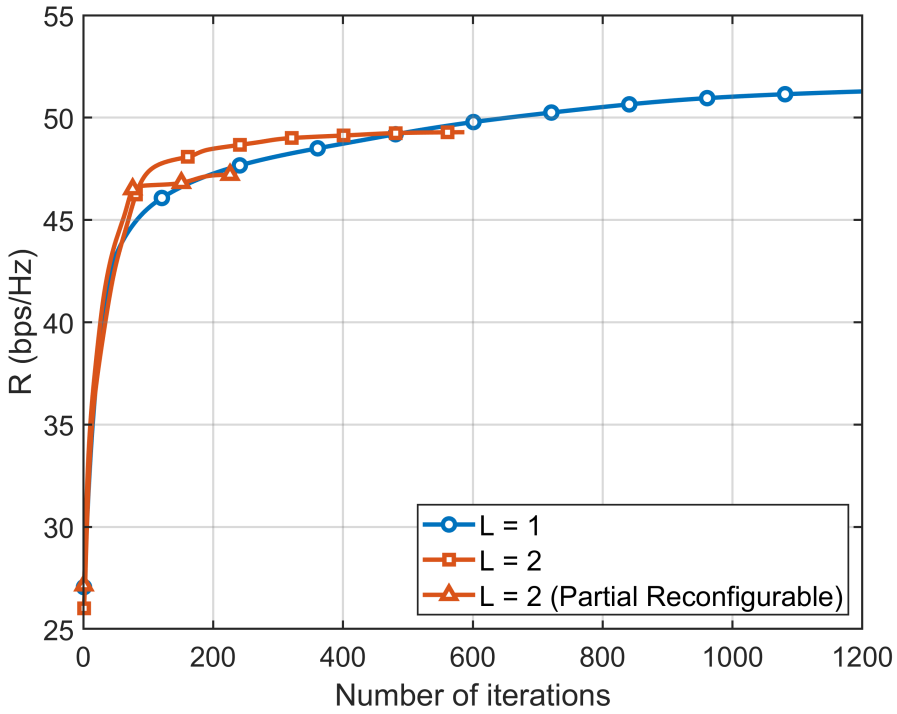


Figure 3.12: Convergence of the gradient ascent algorithm for $M_{\text{total}} = 100$ and $K = 4$.

optimization, and that the partially reconfigurable SIM gives the highest goodput for two users.

The gradient-ascent phase update used in this chapter needs many iterations, and each iteration becomes more expensive as the number of subcarriers, users, and SIM elements grows. In the next chapter, we replace this phase update with a deep-unfolding network for multi-band transmission with inter-band carrier aggregation.

Chapter 4

Deep Unfolding for Multi-band SIM Systems

4.1 Introduction

This chapter addresses research question RQ3 of Section 1.3.4. In Chapter 3, the SIM phase shifts were optimized with gradient ascent inside an alternating optimization (AO) framework, which typically requires many iterations. As discussed in Section 1.3.3, this becomes more costly in multi-band transmission with inter-band carrier aggregation, where a single SIM phase configuration has to serve the subcarriers of a low band and a high band whose gradients have markedly different scales.

In this chapter, we propose an AO framework for SIM-assisted multi-band multi-user multiple-input single-output (MU-MISO) downlink transmission, where we optimize the digital precoding and the SIM phase shifts. For the SIM phase optimization subproblem, we develop a physically consistent multi-band deep-unfolding network (MBDU-Net), building on the deep-unfolding principles reviewed in Section 2.5. The key idea is to unroll a fixed number of projected-gradient phase updates into a finite-depth ar-

chitecture, where each stage preserves the underlying update structure while replacing hand-tuned update parameters with lightweight trainable variables. The main contributions of this chapter are summarized as follows:

- An AO framework that decomposes the joint design into a power-constrained precoder update and a SIM phase update, under the physically consistent impedance-domain SIM model of Chapter 3.
- MBDU-Net, which unrolls projected-gradient phase updates into a compact trainable architecture. Each stage computes an analytic gradient from the cascaded SIM channel model. Unlike standard unfolding designs with a single learned update rule, MBDU-Net separates the low- and high-band gradient contributions and learns band-aware step sizes together with momentum-related parameters.
- Numerical results showing that MBDU-Net achieves faster convergence and consistent sum-rate gains on unseen channel realizations compared with classical projected gradient ascent and basic deep unfolding, and that it maintains strong performance when the number of subcarriers differs from the training configuration.

The material of this chapter has been published in [48]. The rest of the chapter is organized as follows. Section 4.2 presents the system model. Section 4.3 formulates the optimization problem, and Section 4.4 presents the AO framework and MBDU-Net. Section 4.5 presents the simulation results, and Section 4.6 summarizes the chapter.

4.2 System Model

4.2.1 Multi-band Setup

We consider a SIM-assisted multi-band multi-user downlink, as shown in Fig. 4.1, where the SIM is integrated into the BS architecture. The BS employs a uniform linear array (ULA) of N_t active antennas and a transmissive SIM comprising L cascaded layers, each with M passive meta-atoms (elements), and serves K single-antenna users. The system operates over two bands using inter-band carrier aggregation (cf. Section 2.4.2), with M_L and M_H orthogonal subcarriers allocated to the low- and high-frequency bands, respectively. The corresponding per-subcarrier bandwidths are denoted by B_L and B_H. The BS operates under a total downlink transmit power budget P_T. Owing to the larger bandwidth in the high band, we allow different noise variances across bands and model the additive white Gaussian noise (AWGN) power as $\sigma_H^2 > \sigma_L^2$. For notational convenience, we use a unified subcarrier index i across both bands and define the overall frequency set as

$$\mathcal{F} = \{f_{L,1}, \ldots, f_{L,M_L},\ f_{H,1}, \ldots, f_{H,M_H}\}, \tag{4.1}$$

with $N_c \triangleq M_L + M_H$ subcarriers in total. We further denote by $\mathcal{I}_L$ and $\mathcal{I}_H$ the index sets of the low-band and high-band subcarriers, respectively, and by $B_i \in \{B_L, B_H\}$ the bandwidth of subcarrier i.

4.2.2 SIM Model and Effective Channel

We adopt the physically consistent circuit-theoretic SIM model of Section 3.3.3. The phases of the ℓ-th layer are collected in $\boldsymbol{\phi}_\ell = [\phi_{\ell,1}, \ldots, \phi_{\ell,M}]^T$, and all phases in the matrix $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_L]$. The SIM transfer matrix on subcarrier i is $\mathbf{G}(\boldsymbol{\Phi}, f_i) = \big(\mathbf{Z}_{SS}(f_i) + \mathbf{Z}_S(\boldsymbol{\Phi})\big)^{-1}$, where $\mathbf{Z}_{SS}(f_i)$ and $\mathbf{Z}_S(\boldsymbol{\Phi})$ are given in (3.7) and (3.8). Follow-

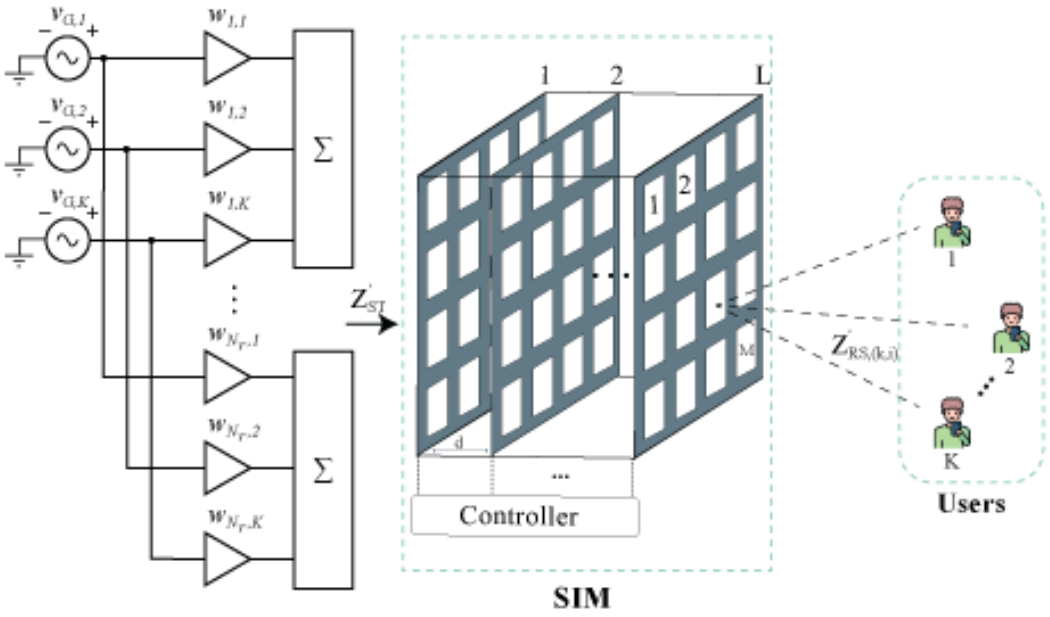


Figure 4.1: SIM-assisted multi-band multi-user MISO downlink system.

ing (3.13), the effective channel from the BS to the users on subcarrier i is

$$\mathbf{H}_i(\mathbf{\Phi}) = \mathbf{H}_{2,i}\,\mathbf{G}_{2L,1}(\mathbf{\Phi}, f_i)\,\mathbf{H}_{1,i}, \tag{4.2}$$

where $\mathbf{H}_{1,i} \in \mathbb{C}^{M\times N_t}$ and $\mathbf{H}_{2,i} \in \mathbb{C}^{K\times M}$ are the BS–SIM and SIM–user transimpedance matrices on subcarrier i. The phase variables enter the channel only through $\mathbf{Z}_S(\mathbf{\Phi})$, and the channel is differentiable with respect to every phase $\phi_{\ell,m}$, as shown in Appendix A.

Common phase configuration across bands. The phase matrix $\mathbf{\Phi}$ represents one programmable SIM configuration that is jointly optimized over all aggregated bands during a scheduling interval, instead of assigning an independent phase profile to each band or subcarrier. It should be interpreted as the design state of the phase-shifting network embedded between the two faces of each transmissive layer, i.e., the phase-shifting elements are modeled to provide a common programmable transmission phase across the aggregated bands. This does not imply that the SIM response is frequency inde-

pendent. Even with the same $\mathbf{\Phi}$, the transfer matrix $\mathbf{G}(\mathbf{\Phi}, f_i)$ varies across subcarriers through $\mathbf{Z}_{SS}(f_i)$, which includes the frequency-dependent self-impedances, the inter-layer coupling, and the propagation between adjacent layers (cf. Sections 2.2 and 2.3), as well as through the frequency-dependent transimpedance matrices $\mathbf{H}_{1,i}$ and $\mathbf{H}_{2,i}$. The objective is thus to find a shared SIM configuration that provides the best aggregate performance across the frequency-dependent low- and high-band channels. A more detailed hardware model with frequency-dependent phase shifters, e.g., based on tunable capacitances, would require an additional circuit-level characterization of the phase-shifting elements and is left for future work.

4.2.3 Signal and Channel Model

On subcarrier i, the BS transmits K independent data streams $\mathbf{d}_i = [d_{1,i}, \ldots, d_{K,i}]^T$, with $\mathbb{E}\{\mathbf{d}_i \mathbf{d}_i^H\} = \mathbf{I}_K$, using the linear precoder $\mathbf{W}_i = [\mathbf{w}_{1,i}, \ldots, \mathbf{w}_{K,i}] \in \mathbb{C}^{N_t \times K}$. The received signal is

$$\mathbf{y}_i = \mathbf{H}_i(\mathbf{\Phi})\, \mathbf{W}_i \mathbf{d}_i + \mathbf{n}_i, \tag{4.3}$$

where $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \sigma_i^2 \mathbf{I}_K)$ with $\sigma_i^2 \in \{\sigma_L^2, \sigma_H^2\}$ depending on the band of subcarrier i. Thus, the model uses one independent stream per scheduled user on each subcarrier, and users may receive independent symbols over subcarriers assigned across both bands. The role of the SIM in this multi-stream setting is to provide an additional passive wave-domain processing stage after the digital beamforming at the BS: the digital precoder first forms the multi-user transmit signal, and the SIM then further shapes the transmitted wavefront before propagation to the users. In this way, the SIM refines the effective multi-user channel by providing additional passive degrees of freedom for user separation and inter-user interference management.

The BS–SIM matrix $\mathbf{H}_{1,i}$ and the SIM blocks in $\mathbf{Z}_{SS}(f_i)$ are computed with the coupling-aware method of Section 3.3.2. Unlike the deterministic far-field model (3.14)

of Chapter 3, the SIM–user link is modeled by Rayleigh fading based on the multiport model in [26]:

$$\mathbf{H}_{2,i} = \frac{c}{2\pi f_i \, d^{\alpha_{\mathrm{PL}}/2}} \big(\Re\{\mathbf{Z}_{RR}(f_i)\}\big)^{1/2} \mathbf{F}_i \, \big(\Re\{\mathbf{Z}_{1,1}^{(L)}(f_i)\}\big)^{1/2}, \tag{4.4}$$

where d is the link distance, α_{PL} is the path-loss exponent, and $\mathbf{F}_i \in \mathbb{C}^{K\times M}$ has independent and identically distributed $\mathcal{CN}(0,1)$ entries. Assuming negligible mutual coupling at the user side and frequencies in the GHz range, $\big(\Re\{\mathbf{Z}_{RR}(f_i)\}\big)^{1/2} \approx \sqrt{R_a}\,\mathbf{I}$, where R_a denotes the antenna resistance. This rich-scattering model allows the unfolded network to be trained and tested over many independent channel realizations.

4.2.4 Achievable Sum Rate

The interference-plus-noise covariance $\mathbf{C}_{k,i}$, the whitened channel $\widetilde{\mathbf{h}}_{k,i}$, and the sum rate are defined as in Section 3.3.5, where $\mathbf{h}_{k,i}$ is now the k-th row of $\mathbf{H}_i(\mathbf{\Phi})$, the noise variance σ_i^2 depends on the band, and the summation runs over the N_c subcarriers. With the beamformers aligned to the whitened channels (Section 4.4.1), the sum rate reduces to

$$R = \sum_{i=1}^{N_c} \sum_{k=1}^{K} B_i \, \log_2\big(1 + s_{k,i}\big), \qquad s_{k,i} = p_{k,i}\, \mathbf{h}_{k,i} \mathbf{C}_{k,i}^{-1} \mathbf{h}_{k,i}^{H}, \tag{4.5}$$

where $p_{k,i} = \|\mathbf{w}_{k,i}\|^2$ and $s_{k,i}$ is the effective SINR, as in (3.29). We further define the low-band and high-band sum rates R_L and R_H by restricting the subcarrier summation in (4.5) to $\mathcal{I}_L$ and $\mathcal{I}_H$, respectively, so that $R = R_L + R_H$. This split is the key to the band-aware design of Section 4.4.3.

4.3 Problem Formulation

Our goal is to jointly design the per-subcarrier precoders $\{\mathbf{W}_i\}_{i=1}^{N_c}$ and the SIM phase configuration $\mathbf{\Phi}$ to maximize the multi-band sum rate:

$$\max_{\{\mathbf{W}_i\},\,\mathbf{\Phi}} \quad R(\{\mathbf{W}_i\},\mathbf{\Phi}) \tag{4.6a}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_c} \mathrm{Tr}\big(\mathbf{W}_i\mathbf{W}_i^H\big) \le P_T, \tag{4.6b}$$

$$0 \le \phi_{\ell,m} < 2\pi, \qquad \forall\, \ell,\ m. \tag{4.6c}$$

Similar to (3.22), this problem is non-convex because the precoders and the SIM phases are coupled through the effective channels. In addition, the same $\mathbf{\Phi}$ must serve all N_c subcarriers of two bands with different bandwidths and noise levels, which makes the phase design strongly frequency coupled.

4.4 Proposed Solution

4.4.1 Alternating Optimization Framework

We adopt the AO structure of Section 3.4.2, which alternates between two subproblems for $I_{\max}$ outer iterations. In **P1**, for fixed $\mathbf{\Phi}$, the powers are allocated by the iterative water-filling rule (3.27) over the effective gains $\lambda_{k,i} = \mathbf{h}_{k,i}\mathbf{C}_{k,i}^{-1}\mathbf{h}_{k,i}^H$ of all user–subcarrier pairs of both bands, and each beamformer is matched to the corresponding whitened channel, i.e.,

$$\mathbf{w}_{k,i} = \sqrt{p_{k,i}}\,\frac{\widetilde{\mathbf{h}}_{k,i}^H}{\|\widetilde{\mathbf{h}}_{k,i}\|}. \tag{4.7}$$

Since the two bands have different bandwidths and noise levels, the water-filling solution distributes the power over the subcarriers of both bands according to their effective gains, which realizes the inter-band carrier aggregation. In **P2**, for fixed precoders, the

SIM phases $\mathbf{\Phi}$ are updated. The main contribution of this chapter is an efficient solver for **P2**, described next.

4.4.2 From Projected Gradient to Deep Unfolding

A standard approach to **P2** is projected gradient ascent (2.10), which we refer to as GD in the following, with the feasible set $[0, 2\pi)$ for every phase, where the gradient $\nabla_{\mathbf{\Phi}} R$ is given by (3.33) with the summation over the N_c subcarriers, and is therefore analytic and follows the physical structure of the SIM model (Appendix A). The gradients $\nabla_{\mathbf{\Phi}} R_L$ and $\nabla_{\mathbf{\Phi}} R_H$ are obtained in the same way by restricting the summation to $\mathcal{I}_L$ and $\mathcal{I}_H$. As discussed in Section 2.5.1, the convergence of projected gradient ascent is highly sensitive to the step size and typically requires many iterations, especially here, where a single $\mathbf{\Phi}$ must jointly accommodate heterogeneous propagation and noise levels across $\mathcal{F}$.

Following Section 2.5.2, the basic deep-unfolding (DU) approach unrolls T projected-gradient iterations into a T-stage network with trainable step sizes η_t, $t = 0, \ldots, T-1$. A limitation of basic DU is that a single step size per stage must accommodate heterogeneous subcarriers across $\mathcal{F}$. In particular, the low-band and high-band components can exhibit different gradient magnitudes due to their distinct bandwidths and noise levels, which may lead to conservative updates or oscillatory behavior. To better match this multi-band heterogeneity and to accelerate convergence without increasing the number of unfolding stages, we adopt a momentum-accelerated unfolding inspired by Nesterov's method (Section 2.5.3) [45].

4.4.3 Multi-band Deep-Unfolding Network

The proposed MBDU-Net introduces a velocity state $\mathbf{V}^{(t)}$, which has the same dimension as $\mathbf{\Phi}$ and aggregates past ascent directions. At stage t, we first form a look-ahead

phase point by extrapolating along the current velocity. The gradients of the low-band and high-band objectives are then evaluated at this look-ahead point and combined using separate trainable step sizes. Next, the velocity is updated through a trainable momentum factor, and the phase matrix is projected onto the feasible interval. The resulting T-stage update is given by

$$\widetilde{\mathbf{\Phi}}^{(t)} = \mathbf{\Phi}^{(t)} + \xi_t \, \mathbf{V}^{(t)}, \tag{4.8a}$$

$$\mathbf{D}^{(t)} = \eta_{L,t} \, \nabla_{\mathbf{\Phi}} R_L\big(\widetilde{\mathbf{\Phi}}^{(t)}\big) + \eta_{H,t} \, \nabla_{\mathbf{\Phi}} R_H\big(\widetilde{\mathbf{\Phi}}^{(t)}\big), \tag{4.8b}$$

$$\mathbf{V}^{(t+1)} = \theta_t \, \mathbf{V}^{(t)} + \mathbf{D}^{(t)}, \tag{4.8c}$$

$$\mathbf{\Phi}^{(t+1)} = \Pi_{[0,2\pi)}\Big(\widetilde{\mathbf{\Phi}}^{(t)} + \mathbf{V}^{(t+1)}\Big), \tag{4.8d}$$

with initialization $\mathbf{V}^{(0)} = \mathbf{0}$. In (4.8b), $\nabla_{\mathbf{\Phi}} R_L$ and $\nabla_{\mathbf{\Phi}} R_H$ follow from (3.33) by restricting the subcarrier summation to $\mathcal{I}_L$ and $\mathcal{I}_H$, respectively. One stage of MBDU-Net is illustrated in Fig. 4.2.

The parameters $\eta_{L,t}$ and $\eta_{H,t}$ control the step sizes of the two bands and compensate for their different bandwidth and noise conditions, while θ_t governs the amount of momentum carried across stages. The extrapolation parameter ξ_t determines the look-ahead point used to evaluate the gradients. Collectively, ξ_t, θ_t, $\eta_{L,t}$, and $\eta_{H,t}$ increase the degrees of freedom of the unfolded solver, enabling larger progress per stage and earlier convergence for a fixed depth T, instead of relying on a large number of projected-gradient iterations. Note that the basic DU is recovered as a special case of (4.8) with $\xi_t = \theta_t = 0$ and $\eta_{L,t} = \eta_{H,t} = \eta_t$.

4.4.4 Training and Overall Algorithm

In the overall AO, the above T-stage unfolded block is further unrolled over $I_{\max}$ outer iterations, where each outer iteration contains a depth-T phase-update module. Moreover, each outer iteration is assigned its own trainable parameters. Hence, for outer

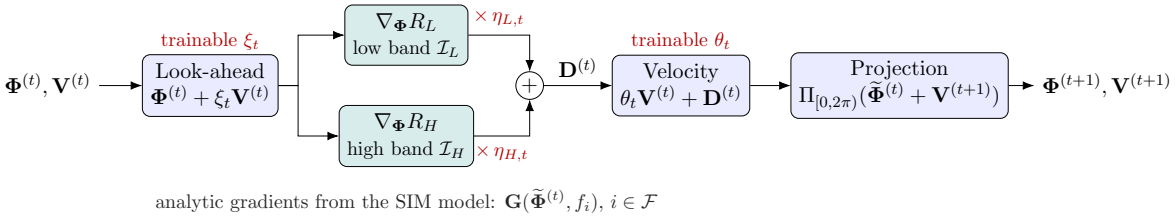


Figure 4.2: One stage of the proposed MBDU-Net. The gradients of the low-band and high-band sum rates are evaluated at a look-ahead point and combined with trainable band-aware step sizes, followed by a trainable momentum update and a projection onto the feasible phase interval.

iteration n, the corresponding parameter set is $\{\xi_{n,t}, \theta_{n,t}, \eta_{L,n,t}, \eta_{H,n,t}\}_{t=0}^{T-1}$, and the full architecture consists of $I_{\max}$ cascaded unfolded blocks, each having depth T. The total number of trainable parameters is therefore only $4\,T\,I_{\max}$, independent of the number of SIM elements, users, and subcarriers.

Let $\boldsymbol{\Phi}^{(I_{\max},T)}$ denote the network output after $I_{\max}$ outer iterations, each containing T unfolded stages. The trainable parameters are learned offline by minimizing the negative sum rate induced by $\boldsymbol{\Phi}^{(I_{\max},T)}$ over a set of channel realizations, i.e.,

$$\mathcal{L} = -R\big(\boldsymbol{\Phi}^{(I_{\max},T)}\big). \tag{4.9}$$

During training, the parameters are updated using the Adam optimizer, with gradients obtained by backpropagation through all T unfolded stages across the $I_{\max}$ outer AO iterations. Online, the trained network is executed for a given channel realization with a fixed number of stages. The overall procedure is summarized in Algorithm 2.

4.4.5 Complexity Discussion

The main advantage of the proposed deep-unfolding framework is not a reduction in the per-iteration computational complexity of the phase update. In fact, GD, DU, and MBDU all rely on the same underlying channel model and gradient computation, so their per-stage computational costs are of the same order, dominated by the evaluation

of the SIM transfer blocks and their derivatives on every subcarrier (cf. Table 3.1). MBDU-Net requires two partial gradients, $\nabla_{\mathbf{\Phi}} R_L$ and $\nabla_{\mathbf{\Phi}} R_H$, but these are obtained by splitting the same subcarrier summation, so no additional gradient computations are needed. The reduction in complexity comes from the number of iterations required to reach a good operating point. In the conventional GD-based approach, a large number of projected-gradient iterations is typically required for convergence, which can be impractical for repeated channel realizations and scheduling intervals. In contrast, DU and MBDU learn the update parameters offline and use only a small fixed number of unfolded stages during online inference.

4.5 Simulation Results and Discussion

4.5.1 Simulation Setup

We consider a two-carrier multi-band system with a low band and a high band centered at $f_L = 3.5$ GHz and $f_H = 17.5$ GHz, respectively. The subcarrier bandwidths are $B_L = 120$ kHz and $B_H = 480$ kHz, and each band contains $M_L = 4$ and $M_H = 4$ subcarriers unless stated otherwise. The BS total downlink transmit power is fixed to $P_T = 10$ W. Since the noise power scales linearly with the occupied bandwidth, the high-band noise variance is set larger than the low-band noise variance by the same factor as the bandwidth ratio. The distance between the BS antenna array and the last layer of the SIM is set to $5\lambda_L$, where λ_L denotes the wavelength at f_L. Unless stated otherwise, we consider $K = 2$ single-antenna users and a BS equipped with $N_t = 2$ active antennas. The SIM has $L = 2$ transmissive layers, each consisting of a 7×7 planar array of meta-atoms.

For the deep-unfolding-based methods, training is performed offline using independently generated channel realizations with the above transmit-power and noise settings.

The trainable step sizes are initialized to 0.1, the momentum and extrapolation parameters are initialized to 0.2 and 0.5, respectively, and the SIM phases are initialized randomly over $[0, 2\pi)$. The Adam optimizer is used with a learning rate of 10^{-3}, and each model is trained for 40 epochs. The classical GD baseline uses a fixed step size of 0.1. The system and training parameters are summarized in Table 4.1.

Table 4.1: System and training parameters.

Parameter	**Value**
Low-band / high-band center frequency f_L / f_H	3.5 GHz / 17.5 GHz
Subcarrier bandwidth B_L / B_H	120 kHz / 480 kHz
Subcarriers per band M_L / M_H	4 / 4
Noise variance ratio σ_H^2/σ_L^2	$B_H/B_L = 4$
Total transmit power P_T	10 W
Number of users K / BS antennas N_t	2 / 2
SIM layers L / elements per layer M	2 / 7×7
Distance between BS array and SIM	$5\lambda_L$
Initial step sizes η	0.1
Initial momentum θ / extrapolation ξ	0.2 / 0.5
Initial SIM phases	$\mathcal{U}[0, 2\pi)$
Optimizer / learning rate	Adam / 10^{-3}
Training epochs	40
GD baseline step size	0.1

We compare three SIM phase update methods within the same AO framework: (i) classical projected gradient ascent (GD) with a fixed step size, (ii) basic deep unfolding (DU) obtained by unrolling T projected-gradient updates with trainable step sizes, and (iii) the proposed multi-band deep unfolding (MBDU), which augments the unfolded updates with a momentum state and separate trainable step sizes for the low-band and high-band gradient contributions. For all three methods, the phase update is per-

formed over $I_{\max}$ outer iterations, and the performance is evaluated on unseen channel realizations.

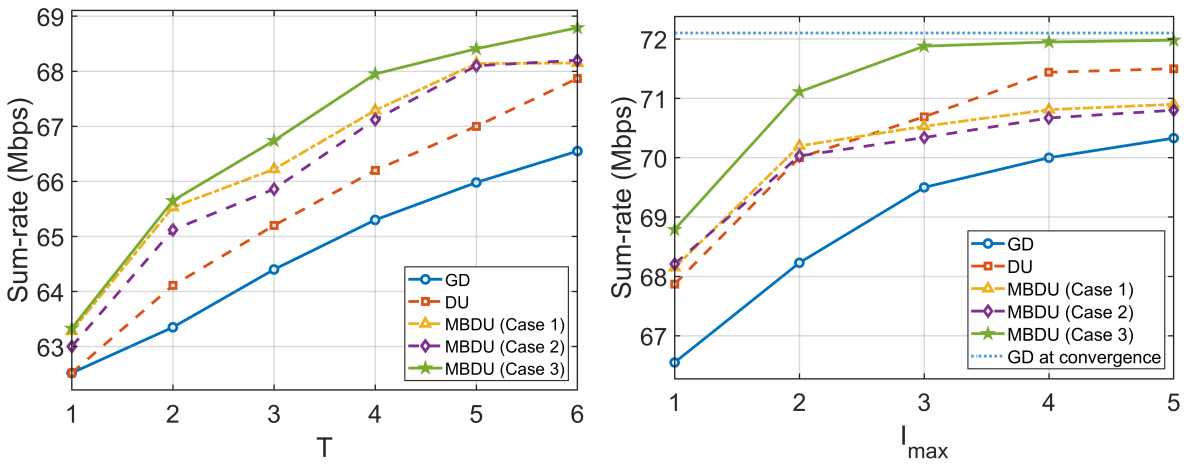
4.5.2 Convergence and Ablation Study

Fig. 4.3 shows the sum-rate performance under variations of the inner and outer iterations. In Fig. 4.3(a), we fix the number of outer iterations to $I_{\max} = 1$ and vary the number of inner iterations T. In Fig. 4.3(b), we instead fix the number of inner iterations to $T = 6$ and vary the maximum number of outer iterations $I_{\max}$. Separate DU and MBDU models are trained for each operating point corresponding to the considered values of T and $I_{\max}$. In both cases, the classical GD baseline improves slowly, indicating that many AO iterations are needed to reach a good operating point. By learning per-stage step sizes, DU achieves a faster convergence than GD.

To further examine the contribution of the different MBDU components, we also include three cases. In Case 1, the momentum and extrapolation parameters are fixed, while the band-aware step sizes are learnable. In Case 2, the step sizes are fixed, while the momentum and extrapolation parameters are learnable. In Case 3, all MBDU parameters are learnable. The results show that separately optimizing only the band-aware step sizes or only the momentum-related parameters yields a lower sum rate than the jointly learned design, which indicates that the two groups of parameters are coupled and should be jointly optimized. Case 3, where all MBDU update parameters are learnable, achieves the fastest convergence and the highest sum rate.

4.5.3 Generalization to Different Numbers of Subcarriers

Fig. 4.4 evaluates the robustness with respect to the number of subcarriers per band for $T = 6$ and $I_{\max} = 5$. The unfolded models are trained using $M_L = M_H = 4$ and then tested on other subcarrier configurations without retraining. The GD baseline remains



(a) Varying T for $I_{\max} = 1$

(b) Varying $I_{\max}$ for $T = 6$

Figure 4.3: Sum rate versus the number of iterations for GD, DU, and MBDU: (a) varying the number of inner iterations for $I_{\max} = 1$, (b) varying the number of outer iterations for $T = 6$.

consistently below the unfolded approaches across the entire range, indicating that the unfolded methods generalize well to unseen configurations. This is expected, since the learned parameters in (4.8) govern only the update dynamics, while the gradients themselves are computed analytically from the SIM channel model for any number of subcarriers.

4.6 Summary

In this chapter, we have investigated SIM-assisted multi-band multi-user downlink transmission under a physically consistent impedance-domain SIM model, where carrier aggregation introduces heterogeneous subcarrier conditions across bands, while the SIM phase shifts remain shared. This coupling makes conventional alternating-optimization solvers costly, because the SIM phase update often requires many iterations. To address

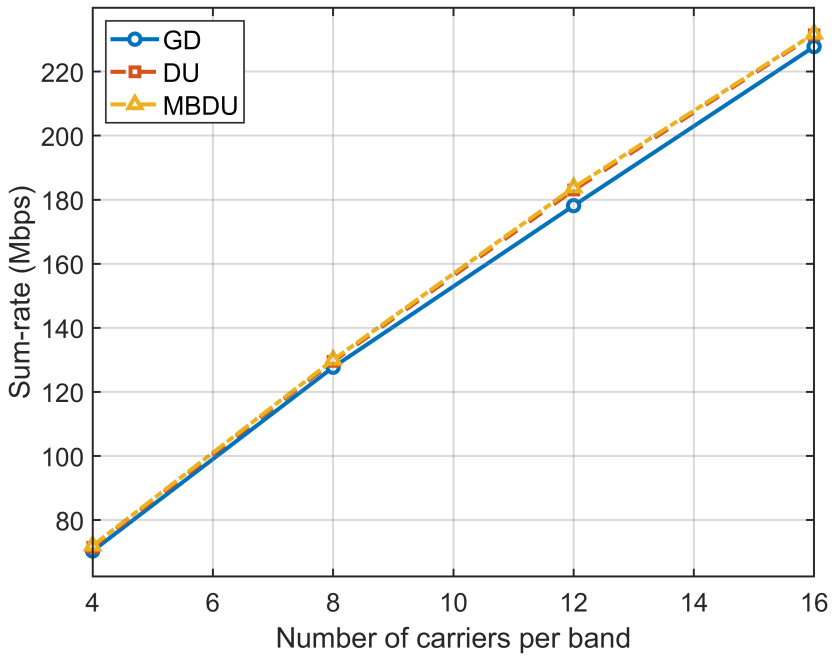


Figure 4.4: Sum rate versus the number of subcarriers per band for $T = 6$ and $I_{\max} = 5$.

this limitation, we have proposed MBDU-Net, a physically consistent multi-band deep-unfolding network for SIM phase optimization. MBDU-Net unrolls a projected-gradient phase update into a fixed-depth architecture that retains the analytic gradient structure of the SIM channel model, while learning lightweight parameters that govern the update dynamics, including band-aware step sizes and momentum. Numerical results have shown that, for the same number of iterations, the proposed unfolding achieves a higher sum rate than classical projected-gradient methods, and that it maintains its performance when the number of subcarriers differs from the training configuration.

Algorithm 2 AO with MBDU-Net for SIM-assisted multi-band MU-MISO downlink

Require: Total power P_T, outer iterations $I_{\max}$, unfolding depth T, trained parameters $\{\xi_{n,t}, \theta_{n,t}, \eta_{L,n,t}, \eta_{H,n,t}\}$

Initialize $\mathbf{\Phi}$ and $\{\mathbf{W}_i\}$
for $n = 1$ to $I_{\max}$ **do**
  Compute $\{\mathbf{H}_i(\mathbf{\Phi})\}_{i=1}^{N_c}$ using (4.2)
  Allocate the powers $\{p_{k,i}\}$ using (3.27)
  Update the beamformers $\{\mathbf{w}_{k,i}\}$ using (4.7)
  Set $\mathbf{V}^{(0)} = \mathbf{0}$ and $\mathbf{\Phi}^{(0)} = \mathbf{\Phi}$
  for $t = 0$ to $T - 1$ **do**
    Form the look-ahead point $\widetilde{\mathbf{\Phi}}^{(t)}$ using (4.8a)
    Compute $\nabla_{\mathbf{\Phi}} R_L$ and $\nabla_{\mathbf{\Phi}} R_H$ using (3.33)
    Compute $\mathbf{D}^{(t)}$ using (4.8b)
    Update $\mathbf{V}^{(t+1)}$ using (4.8c)
    Update $\mathbf{\Phi}^{(t+1)}$ using (4.8d)
  end for
  Set $\mathbf{\Phi} \leftarrow \mathbf{\Phi}^{(T)}$
end for

Ensure: $\{\mathbf{W}_i\}$, $\mathbf{\Phi}$

Chapter 5

Conclusion and Future Work

5.1 Conclusion

In this thesis, we have studied SIMs that are integrated into the BS for wideband and multi-band multi-user downlink transmission. We focused on two problems: how to model and optimize a SIM when mutual coupling, inter-layer reflections, and frequency selectivity are taken into account, and how to optimize the SIM phases with low complexity when one configuration has to serve several bands.

In Chapter 3, we have developed a circuit-theoretic model of a SIM-assisted wideband multi-user MIMO system based on the multiport and Chu-based mutual coupling models of Chapter 2. The phase shifts, the BS precoder, and the power allocation over subbands were optimized jointly by alternating optimization, where the water-filling step performs carrier aggregation. We compared single-layer and multilayer structures with the same total number of elements and observed that the answer depends on the operating conditions. With uniform power allocation, the SIM achieved a higher spectral efficiency beyond approximately five users, at high SNR, and over wide bandwidths, because the additional layers help to separate the users and keep the response more con-

sistent across frequency. With few users, at low SNR, and in the narrowband case, the single-layer RIS performed better due to its larger aperture. When power allocation and active beamforming were performed at the BS, the digital precoder already handled most of the inter-user interference and the single-layer RIS performed better. We have also shown that including mutual coupling in the model improves the spectral efficiency, that optimizing the phases over approximately 15 subbands gives a much flatter response than optimizing at the center frequency only, and that updating only the last layer of the SIM gives near-optimal performance and the highest goodput for two users.

In Chapter 4, we have considered a multi-band MU-MISO downlink with a low band at 3.5 GHz and a high band at 17.5 GHz that share one SIM configuration. Since the classical phase update needs many iterations, we have proposed MBDU-Net, which unrolls projected-gradient updates into a fixed number of stages, keeps the analytic gradient of the SIM model, and learns separate step sizes for the two bands together with momentum parameters. For the same number of iterations, MBDU-Net achieved a higher sum rate than projected gradient ascent and basic deep unfolding. The ablation cases showed that the band-aware step sizes and the momentum parameters need to be learned together, and the network kept its advantage when the number of subcarriers was different from the training setup.

5.2 Future Work

Several extensions of this work are possible.

5.2.1 Imperfect CSI and Hardware Impairments

Throughout this thesis, we assumed that the BS has perfect channel state information (CSI). In practice, the channels have to be estimated, which is difficult because the

SIM is passive and the effective channel depends on all layers. Coupling-aware channel estimation for SIMs and the effect of CSI errors on the proposed designs should therefore be studied. The effect of hardware impairments, such as phase noise and finite phase resolution, on SIM-assisted wideband systems is also left for future work.

5.2.2 Frequency-Dependent Phase Shifters

We modeled the phase shifters inside each transmissive layer with a common programmable phase over the aggregated bands, while the frequency dependence of the SIM entered through the impedance matrices. A hardware model in which the phase response itself changes with frequency, for example, based on tunable capacitances, would be closer to practical metasurfaces. This requires a circuit-level model of the phase-shifting elements and its inclusion in the optimization.

5.2.3 More General Models

All antenna elements were modeled as CMS antennas operating in the TM_1 mode. The framework can be extended to non-CMS antennas and higher-order radiation modes. Moreover, the load matrices of the SIM model do not have to be diagonal, so the same framework can be used for joint phase–amplitude control and non-local (beyond-diagonal) SIM architectures. Finally, the users were assumed to be in the far field of the SIM. With large apertures and short distances at mmWave frequencies, the users can be in the radiating near field, where the wavefront is spherical and the SIM can focus the beam on a location rather than only on a direction. Since the Chu-based mutual coupling model already contains the near-field terms, the same circuit-theoretic framework can be used to model the SIM–user links in the near field and to study near-field beamfocusing with SIMs.

5.2.4 Energy Efficiency with RF-Chain Power Models

The main motivation for integrating a SIM into the BS is to reduce the number of RF chains and the power consumption. In this thesis, however, the architectures were compared in terms of spectral efficiency under a total transmit power budget. A more complete comparison should include the power consumed by the RF chains, i.e., the digital-to-analog converters, mixers, local oscillators, filters, and power amplifiers, together with the power needed to control the SIM layers. The circuit-theoretic model also distinguishes between the available power at the transmitter and the power actually radiated by the antennas, which allows the power-amplifier efficiency and the matching losses to be included. Using such models, the energy efficiency of a SIM-assisted BS can be compared with that of a fully digital array of the same aperture, which would show when the reduction in the number of RF chains outweighs the loss in spectral efficiency.

5.2.5 Learning-Based Joint Design

In this thesis, deep unfolding was applied only to the SIM phase update, while the precoders were obtained by water-filling. Unfolding the complete alternating optimization, including the power allocation and beamforming, and adapting the network to time-varying channels and user scheduling are natural next steps toward real-time operation.

Appendix A

Gradient Derivation for Chapter 3

In this appendix, we present the derivation of the gradient of the sum rate with respect to the SIM phase parameters $\boldsymbol{\phi}$, which is used in Section 3.4.2. The frequency and subband indices are omitted where they are clear from the context. Let

$$\mathbf{G}(\boldsymbol{\phi}) = \big(\mathbf{Z}_{SS} + \mathbf{Z}_S(\boldsymbol{\phi})\big)^{-1}, \qquad \mathbf{G} = \big[\mathbf{G}_{r,c}\big]_{r,c=1}^{2L}, \tag{A.1}$$

where $\mathbf{G}$ is partitioned into $2L \times 2L$ blocks $\mathbf{G}_{r,c} \in \mathbb{C}^{M \times M}$. Then, the end-to-end transfer matrix can be expressed as

$$\mathbf{H}_Z = \mathbf{Z}_{RS}\,\big(\mathbf{Z}_{SS} + \mathbf{Z}_S(\boldsymbol{\phi})\big)^{-1}\mathbf{Z}_{ST} = \mathbf{Z}_{RS}\,\mathbf{G}(\boldsymbol{\phi})\,\mathbf{Z}_{ST}. \tag{A.2}$$

The transmitter is connected to the first-facing ports of the first SIM layer, and the receiver is connected to the last-facing ports of the last layer. Let $\mathbf{S}_{\text{in}} \in \mathbb{R}^{2LM \times M}$ and $\mathbf{S}_{\text{out}} \in \mathbb{R}^{2LM \times M}$ denote the input and output selection matrices, respectively. Then,

$$\mathbf{G}_{2L,1} = \mathbf{S}_{\text{out}}^T\,\mathbf{G}\,\mathbf{S}_{\text{in}}. \tag{A.3}$$

With $\mathbf{H}_1 \in \mathbb{C}^{M \times N_t}$ formed by the first M rows of $\mathbf{Z}_{ST}$ and $\mathbf{H}_2 \in \mathbb{C}^{K \times M}$ formed by the last M columns of $\mathbf{Z}_{RS}$, the effective channel becomes

$$\mathbf{H}_Z = \mathbf{H}_2\,\mathbf{G}_{2L,1}\,\mathbf{H}_1. \tag{A.4}$$

A.1 Gradient of the Local Impedance Matrix

The derivative of the local two-port impedance matrix (3.11) of the m-th element of the ℓ-th layer with respect to its phase $\phi_{\ell,m}$ is

$$\frac{\partial \mathbf{Z}_m^{(\ell)}}{\partial \phi_{\ell,m}} = -\,jZ_0 \begin{bmatrix} 1+\left(\dfrac{\cos\phi_{\ell,m}}{\sin\phi_{\ell,m}}\right)^2 & \dfrac{\cos\phi_{\ell,m}}{\sin^2\phi_{\ell,m}} \\ \dfrac{\cos\phi_{\ell,m}}{\sin^2\phi_{\ell,m}} & 1+\left(\dfrac{\cos\phi_{\ell,m}}{\sin\phi_{\ell,m}}\right)^2 \end{bmatrix} \triangleq \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}. \tag{A.5}$$

Embedding (A.5) within the global load matrix $\mathbf{Z}_S$ yields the sparse matrix $\mathbf{E}_{\ell,m} = \partial \mathbf{Z}_S / \partial \phi_{\ell,m}$, which is nonzero only at the two ports corresponding to element (ℓ, m).

A.2 Gradient of the SIM Transfer Matrix

The derivative of $\mathbf{G}(\boldsymbol{\phi})$ with respect to $\phi_{\ell,m}$ follows from the derivative of the matrix inverse:

$$\frac{\partial \mathbf{G}}{\partial \phi_{\ell,m}} = -\,\mathbf{G}\,\mathbf{E}_{\ell,m}\,\mathbf{G}. \tag{A.6}$$

Substituting (A.6) into the derivative of the end-to-end transfer matrix gives

$$\frac{\partial \mathbf{H}_Z}{\partial \phi_{\ell,m}} = -\,\mathbf{H}_2\,\mathbf{S}_{\text{out}}^T \mathbf{G}\,\mathbf{E}_{\ell,m}\,\mathbf{G}\,\mathbf{S}_{\text{in}}\,\mathbf{H}_1. \tag{A.7}$$

Since $\mathbf{E}_{\ell,m}$ is nonzero only at the ports of element (ℓ, m), only a few columns and rows of $\mathbf{G}$ are needed. Let $\mathbf{e}_m \in \mathbb{R}^M$ be a vector whose m-th entry is one and all other entries are zero, and define

$$\mathbf{c}_1 = \mathbf{G}_{2L,\,2\ell-1}\mathbf{e}_m, \quad \mathbf{c}_2 = \mathbf{G}_{2L,\,2\ell}\mathbf{e}_m, \quad \mathbf{r}_1^T = \mathbf{e}_m^T\mathbf{G}_{2\ell-1,\,1}, \quad \mathbf{r}_2^T = \mathbf{e}_m^T\mathbf{G}_{2\ell,\,1}. \tag{A.8}$$

Following the formulation in [33], $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{r}_1$, and $\mathbf{r}_2$ can be obtained with an iterative solution, and the matrix product simplifies to

$$\mathbf{F}_{\ell,m} = \big(d_{11}\mathbf{c}_1 + d_{21}\mathbf{c}_2\big)\mathbf{r}_1^T + \big(d_{12}\mathbf{c}_1 + d_{22}\mathbf{c}_2\big)\mathbf{r}_2^T, \tag{A.9}$$

so that

$$\frac{\partial \mathbf{H}_Z}{\partial \phi_{\ell,m}} = -\,\mathbf{H}_2\,\mathbf{F}_{\ell,m}\,\mathbf{H}_1. \tag{A.10}$$

A.3 Gradient of the Objective Function

For user k on subband i, define

$$\begin{aligned} \mathbf{h}_{k,i} &= (\mathbf{e}_k^T\mathbf{H}_2)\,\mathbf{G}_{2L,1}\,\mathbf{H}_1, & p_{k,i} &= \mathbf{w}_{k,i}\mathbf{w}_{k,i}^H, \\ \mathbf{C}_{k,i} &= \sigma_i^2\mathbf{I}_{N_t} + \sum_{k'\neq k} p_{k',i}\,\mathbf{h}_{k',i}^H\mathbf{h}_{k',i}, & \mathbf{v}_{k,i} &= \mathbf{C}_{k,i}^{-1}\,\mathbf{h}_{k,i}^H, \end{aligned} \tag{A.11}$$

where $\mathbf{e}_k$ is the k-th unit vector. The whitened gain term is

$$s_{k,i} = p_{k,i}\,\mathbf{h}_{k,i}\,\mathbf{C}_{k,i}^{-1}\,\mathbf{h}_{k,i}^H = p_{k,i}\,\mathbf{h}_{k,i}\,\mathbf{v}_{k,i}. \tag{A.12}$$

Since $\mathbf{C}_{k,i}$ depends on $\boldsymbol{\phi}$ via $\mathbf{h}_{k',i}$, we use $\partial\mathbf{C}_{k,i}^{-1} = -\,\mathbf{C}_{k,i}^{-1}\,(\partial\mathbf{C}_{k,i})\,\mathbf{C}_{k,i}^{-1}$. Using the product rule and $\dot{\mathbf{h}}_{j,i} \triangleq \partial\mathbf{h}_{j,i}/\partial\phi_{\ell,m} = -(\mathbf{e}_j^T\mathbf{H}_2)\,\mathbf{F}_{\ell,m}\,\mathbf{H}_1$, we obtain

$$\frac{\partial s_{k,i}}{\partial \phi_{\ell,m}} = 2\,p_{k,i}\,\Re\{\dot{\mathbf{h}}_{k,i}\,\mathbf{v}_{k,i}\} - \,p_{k,i}\,\mathbf{h}_{k,i}\,\mathbf{C}_{k,i}^{-1}\big(\partial\mathbf{C}_{k,i}\big)\mathbf{C}_{k,i}^{-1}\,\mathbf{h}_{k,i}^H. \tag{A.13}$$

Because

$$\partial\mathbf{C}_{k,i} = \sum_{k'\neq k} p_{k',i}\Big[\dot{\mathbf{h}}_{k',i}^H\mathbf{h}_{k',i} + \mathbf{h}_{k',i}^H\dot{\mathbf{h}}_{k',i}\Big], \tag{A.14}$$

and $\mathbf{h}_{k,i}\mathbf{C}_{k,i}^{-1} = \mathbf{v}_{k,i}^H$ since $\mathbf{C}_{k,i}$ is Hermitian, the second term in (A.13) simplifies to

$$\mathbf{h}_{k,i}\,\mathbf{C}_{k,i}^{-1}\big(\partial\mathbf{C}_{k,i}\big)\mathbf{C}_{k,i}^{-1}\,\mathbf{h}_{k,i}^H = 2\sum_{k'\neq k} p_{k',i}\,\Re\Big\{(\mathbf{h}_{k',i}\mathbf{v}_{k,i})^*\,(\dot{\mathbf{h}}_{k',i}\,\mathbf{v}_{k,i})\Big\}. \tag{A.15}$$

Therefore,

$$\frac{\partial s_{k,i}}{\partial \phi_{\ell,m}} = 2\,p_{k,i}\,\Re\{\dot{\mathbf{h}}_{k,i}\mathbf{v}_{k,i}\} - \,2\,p_{k,i}\sum_{k'\neq k} p_{k',i}\,\Re\Big\{(\mathbf{h}_{k',i}\mathbf{v}_{k,i})^*\,(\dot{\mathbf{h}}_{k',i}\mathbf{v}_{k,i})\Big\}, \tag{A.16}$$

which corresponds to (3.32). Substituting (A.16) into (3.33) yields the gradient of the objective function, which is used in the phase update (3.30).

Bibliography

[1] Z. Ning, T. Li, Y. Wu, X. Wang, Q. Wu, F. R. Yu, and S. Guo, "6G communication new paradigm: The integration of autonomous aerial vehicles and intelligent reflecting surfaces," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 6, pp. 3382–3416, 2025.

[2] Z. Mokhtari, M. Sabbaghian, and R. Dinis, "Massive MIMO downlink based on single carrier frequency domain processing," *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1164–1175, 2016.

[3] A. Judice, S. Malhotra, and M. Sharma, "Recent developments in MIMO antenna architectures and their future directions," in *2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)*, vol. 1, pp. 1798–1803, IEEE, 2024.

[4] 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.

[5] M. Ibrahim, W. H. Raza, M. Moiz, S. A. Hassan, H. Jung, and M. Gidlund, "On the performance of multi-IRS-assisted networks across real urban, suburban,

and rural environments," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 01–06, IEEE, 2024.

[6] B. Khan, A. Mihovska, R. Prasad, F. J. Velez, *et al.*, "A study on cross-carrier scheduler for carrier aggregation in beyond 5G networks," in *2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting (AT-AP-RASC)*, pp. 1–4, IEEE, 2022.

[7] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2019.

[8] H. U. Rehman, F. Bellili, A. Mezghani, and E. Hossain, "Joint active and passive beamforming design for IRS-assisted multi-user MIMO systems: A VAMP-based approach," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6734–6749, 2021.

[9] Q. Tao, T. Xie, X. Hu, S. Zhang, and D. Ding, "Channel estimation and detection for intelligent reflecting surface-assisted orthogonal time frequency space systems," *IEEE Transactions on Wireless Communications*, vol. 23, no. 8, pp. 8419–8431, 2024.

[10] I. Nimnadini, A. Mezghani, and E. Hossain, "Low-overhead Kronecker-based intelligent reflective surfaces for 5G and beyond," in *2024 IEEE 10th World Forum on Internet of Things (WF-IoT)*, pp. 382–387, IEEE, 2024.

[11] A. Mezghani, F. Bellili, and E. Hossain, "Nonlocal reconfigurable intelligent surfaces for wireless communication: Modeling and physical layer aspects," *arXiv:2210.05928*, 2022.

[12] D. Wijekoon, A. Mezghani, G. C. Alexandropoulos, and E. Hossain, "Physically-consistent modeling and optimization of non-local RIS-assisted multi-user MIMO communication systems," *arXiv:2406.05617*, 2024.

[13] Z. Li, W. Chen, Q. Wu, X. Zhu, H. Qin, K. Wang, and J. Li, "Toward transmissive RIS transceiver enabled uplink communication systems: Design and optimization," *IEEE Internet of Things Journal*, vol. 11, no. 4, pp. 6788–6801, 2023.

[14] Y. Sun, K. An, C. Li, Z. Lin, H. Niu, D. W. K. Ng, J. Wang, and N. Al-Dhahir, "Joint transmissive and reflective RIS-aided secure MIMO systems design under spatially-correlated angular uncertainty and coupled PSEs," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 3606–3621, 2023.

[15] J. An, C. Xu, D. W. K. Ng, G. C. Alexandropoulos, C. Huang, C. Yuen, and L. Hanzo, "Stacked intelligent metasurfaces for efficient holographic MIMO communications in 6G," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 8, pp. 2380–2396, 2023.

[16] J. An, C. Yuen, C. Xu, H. Li, D. W. K. Ng, M. Di Renzo, M. Debbah, and L. Hanzo, "Stacked intelligent metasurface-aided MIMO transceiver design," *IEEE Wireless Communications*, vol. 31, no. 4, pp. 123–131, 2024.

[17] C. Liu, Q. Ma, Z. J. Luo, Q. R. Hong, Q. Xiao, H. C. Zhang, L. Miao, W. M. Yu, Q. Cheng, L. Li, *et al.*, "A programmable diffractive deep neural network based on a digital-coding metasurface array," *Nature Electronics*, vol. 5, no. 2, pp. 113–122, 2022.

[18] J. An, M. Di Renzo, M. Debbah, H. V. Poor, and C. Yuen, "Stacked intelligent metasurfaces for multiuser downlink beamforming in the wave domain," *IEEE Transactions on Wireless Communications*, 2025.

[19] H. Niu, J. An, A. Papazafeiropoulos, L. Gan, S. Chatzinotas, and M. Debbah, "Stacked intelligent metasurfaces for integrated sensing and communications," *IEEE Wireless Communications Letters*, 2024.

[20] A. Papazafeiropoulos, J. An, P. Kourtessis, T. Ratnarajah, and S. Chatzinotas, "Achievable rate optimization for stacked intelligent metasurface-assisted holographic MIMO communications," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 13173–13186, 2024.

[21] Q. Li, M. El-Hajjar, C. Xu, J. An, C. Yuen, and L. Hanzo, "Stacked intelligent metasurfaces for holographic MIMO-aided cell-free networks," *IEEE Transactions on Communications*, vol. 72, no. 11, pp. 7139–7151, 2024.

[22] S. Lin, J. An, L. Gan, M. Debbah, and C. Yuen, "Stacked intelligent metasurface enabled LEO satellite communications relying on statistical CSI," *IEEE Wireless Communications Letters*, vol. 13, no. 5, pp. 1295–1299, 2024.

[23] X. Yao, J. An, L. Gan, M. Di Renzo, and C. Yuen, "Channel estimation for stacked intelligent metasurface-assisted wireless networks," *IEEE Wireless Communications Letters*, vol. 13, no. 5, pp. 1349–1353, 2024.

[24] J. An, C. Yuen, Y. L. Guan, M. Di Renzo, M. Debbah, H. V. Poor, and L. Hanzo, "Two-dimensional direction-of-arrival estimation using stacked intelligent metasurfaces," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 10, pp. 2786–2802, 2024.

[25] M. T. Ivrlač and J. A. Nossek, "Toward a circuit theory of communication," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 7, pp. 1663–1683, 2010.

[26] M. T. Ivrlač and J. A. Nossek, "The multiport communication theory," *IEEE Circuits and Systems Magazine*, vol. 14, no. 3, pp. 27–44, 2014.

[27] 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.

[28] M. Akrout, F. Bellili, A. Mezghani, and J. A. Nossek, "Physically consistent models for intelligent reflective surface-assisted communications under mutual coupling and element size constraint," in *2023 57th Asilomar Conference on Signals, Systems, and Computers*, pp. 1589–1594, IEEE, 2023.

[29] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 1229–1243, 2021.

[30] A. Abrardo, A. Toccafondi, and M. Di Renzo, "Design of reconfigurable intelligent surfaces by using S-parameter multiport network theory–optimization and full-wave validation," *IEEE Transactions on Wireless Communications*, 2024.

[31] M. Nerini and B. Clerckx, "Physically consistent modeling of stacked intelligent metasurfaces implemented with beyond diagonal RIS," *IEEE Communications Letters*, vol. 28, no. 7, pp. 1693–1697, 2024.

[32] H. Yahya, M. Nerini, B. Clerckx, and M. Debbah, "T-parameters based modeling for stacked intelligent metasurfaces: Tractable and physically consistent model," *IEEE Wireless Communications Letters*, 2025.

[33] A. Abrardo, G. Bartoli, and A. Toccafondi, "A novel comprehensive multiport network model for stacked intelligent metasurfaces (SIM) characterization and optimization," *IEEE Transactions on Communications*, 2025.

[34] Z. Li, J. An, and C. Yuen, "Stacked intelligent metasurfaces for fully-analog wideband beamforming design," in *2024 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*, pp. 1–5, IEEE, 2024.

[35] R. Joda, M. Elsayed, H. Abou-Zeid, R. Atawia, A. B. Sediq, G. Boudreau, and M. Erol-Kantarci, "Carrier aggregation with optimized UE power consumption in 5G," *IEEE Networking Letters*, vol. 3, no. 2, pp. 61–65, 2021.

[36] A. Hussein, *Operating multi-user transmission for 5G and beyond cellular systems*. PhD thesis, University of Waterloo, 2023.

[37] N. Balasuriya, A. Mezghani, and E. Hossain, "Physically-consistent multi-band massive MIMO systems: A radio resource management model," *IEEE Transactions on Wireless Communications*, 2025.

[38] W. Xia, G. Zheng, Y. Zhu, J. Zhang, J. Wang, and A. P. Petropulu, "A deep learning framework for optimization of MISO downlink beamforming," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1866–1880, 2019.

[39] L. Pellaco, M. Bengtsson, and J. Jaldén, "Matrix-inverse-free deep unfolding of the weighted MMSE beamforming algorithm," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 65–81, 2021.

[40] A. Mezghani, M. Akrout, M. R. Castellanos, S. Saab, B. Hochwald, R. W. Heath, and J. A. Nossek, "Reincorporating circuit theory into information theory," *IEEE BITS the Information Theory Magazine*, vol. 4, no. 1, pp. 40–58, 2023.

[41] M. Akrout, V. Shyianov, F. Bellili, A. Mezghani, and R. W. Heath, "Achievable rate of near-field communications based on physically consistent models," *IEEE Transactions on Wireless Communications*, vol. 22, no. 2, pp. 1266–1280, 2022.

[42] L. J. Chu, "Physical limitations of omni-directional antennas," *Journal of Applied Physics*, vol. 19, no. 12, pp. 1163–1175, 1948.

[43] V. Shyianov, M. Akrout, F. Bellili, A. Mezghani, and R. W. Heath, "Achievable rate with antenna size constraint: Shannon meets Chu and Bode," *IEEE Transactions on Communications*, vol. 70, no. 3, pp. 2010–2024, 2022.

[44] M. Akrout, V. Shyianov, F. Bellili, A. Mezghani, and R. W. Heath, "Super-wideband massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 8, pp. 2414–2430, 2023.

[45] Y. Nesterov, "A method for solving the convex programming problem with convergence rate $O(1/k^2)$," in *Doklady Akademii Nauk SSSR*, vol. 269, p. 543, 1983.

[46] M. Ibrahim, A. Mezghani, and E. Hossain, "Are stacked intelligent metasurfaces (SIMs) better than single-layer reconfigurable intelligent surfaces (RISs) for wideband multi-user MIMO communication systems?," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 18215–18228, 2026.

[47] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 145–152, 2004.

[48] M. Ibrahim, A. Mezghani, and E. Hossain, "Deep unfolding for SIM-assisted multiband MU-MISO downlink systems," *IEEE Wireless Communications Letters*, 2026.