Reconfigurable Intelligent Surface-Assisted Multi-User Wireless Communications Systems

by

Haseeb Ur Rehman

A Thesis submitted to The Faculty of Graduate 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

August 2021

When you look at yourself from a universal standpoint, something inside always reminds or informs you that there are bigger and better things to worry about.

ALBERT EINSTEIN

Abstract

Intelligent reflective surfaces (IRSs) have recently emerged as a promising technology for 6G wireless communications as it can improve both the spectral and energy efficiencies of wireless systems at a low cost. In this thesis, we first tackle the problem of joint active and passive beamforming optimization for an intelligent reflective surface (IRS)-assisted multi-user downlink multiple-input multiple-output (MIMO) communication system under both ideal and practical IRS phase shifts. We aim to maximize the spectral efficiency of the users by minimizing the sum mean square error (MSE) of the users' received symbols. For this, a joint non-convex optimization problem is formulated under the sum minimum mean square error (MMSE) criterion. An alternating optimization and vector approximate message passing (VAMP)-based approach is presented to solve the joint problem under both ideal and practical constraints on the IRS phase shifts. Then, we introduce a novel approach of utilizing the IRS, now called modulating intelligent surface (MIS), for joint data modulation and signal beamforming in a multi-user downlink cellular network by leveraging the idea of backscatter communication. We present a general framework in which the MIS is used *i*) to beamform the signals for a set of users whose data is modulated by the BS, and at the same time *ii*) to embed the data of a different set of users by passively modulating the deliberately sent signals from the BS. By following the same VAMP-based optimization approach developed earlier, we optimize *i*) the MIS phase-shifts for passive beamforming and data embedding for the BS- and MIS-served users, respectively, *ii*) the active precoder and the receive scaling factors for the BS- and MIS-served users, respectively.

Simulation results are presented to illustrate the performance of both the proposed IRS-based and MIS-based schemes under both perfect and imperfect channel state information (CSI). The results validate the superiority of the proposed method over the

state-of-the-art techniques both in terms of throughput and computational complexity. The results also reveal that the proposed MIS-based approach outperforms the existing IRS-based schemes in terms of throughput while supporting a much higher number of users.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMP	Approximate message passing
BS	Base station
CSI	Channel state information
IRS	Intelligent reflective surface
LMAP	Linear maximum a posteriori
LMMSE	Linear minimum mean square error
MAP	Maximum a posteriori
MIMO	Multiple-input multiple-output
MIS	Modulating intelligent surface
MISO	Multiple-input single-output
MMSE	Minimum mean square error
MSE	Mean square error
NRMSE	Normalized root mean square error
OOVAMP	Optimization oriented vector approximate message passing
SINR	Signal to interference plus noise ratio
VAMP	Vector approximate message passing

Chapter 1

Introduction

1.1 Overview

The need for higher data rates in wireless communication is soaring. This calls for innovative and economically viable communication technologies that can keep up with the increasing network capacity requirement. Massive multiple-input multiple-output (MIMO) technology can fulfill the network capacity requirement for beyond fifth-generation (B5G) wireless networks [1–3]. The basic idea of massive MIMO is to equip the base stations (BSs) with tens (if not hundreds) of antenna elements so as to simultaneously serve multiple mobile devices using the same time/frequency resources. Despite the many advantages of massive MIMO, its practical large-scale deployment is hindered by the associated high hardware cost and energy consumption [4, 5]. Moreover, although millimeter wave (mmWave) communication benefits from massive MIMO due to a symbiotic convergence of technologies, its practical use is still limited by the less penetrative propagation characteristic of mmWave signals in presence of blockages between the BS and the mobile device. [6].

One promising technology that has been introduced recently is intelligent reflec-

tive surfaces (IRSs), also called reconfigurable intelligent surfaces (RISs) [7, 8]. IRS is composed of a planar metasurface consisting of a large number of passive reflective elements. This allows the IRS to passively alter the wireless propagation environment by reconfiguring the phases of its reflective elements through a controller attached to the surface [9]. The key advantages of the IRSs are listed as follows [10]:

- **Easy Installation:** IRSs are nearly-passive devices, composed of electromagnetic (EM) material. Fig. 1.1 shows various possible structures where IRS can be deployed. Unlike traditional BSs which can only be installed at designated locations and high-rise towers, IRSs can be installed on building exteriors, billboards, aerial and road vehicles and even on clothes given its low cost.

- **Spectral efficiency improvement:** Since IRSs are able to modify the wireless propagation environment, they can mitigate the power loss over large distances. IRSs can be utilized to perform passive beamforming. Passive beamforming refers to changing the IRS phases without actively powering the IRS antenna elements as opposed to active beamforming at the BS so as to improve the received power while reducing the interference for unintended users, thereby enhancing the overall throughput of the network [11]. IRSs are especially useful in the scenarios where Line of Sight (LOS) links between the BS and the mobile users are blocked by obstacles, e.g., buildings, or indoor walls. IRS allow to form virtual LOS links between BSs and mobile users via passively reflecting the incident radio signals.

- **Environment friendliness:** IRS does not require a power amplifier for transmission which makes it an energy-efficient technology. Practically, IRS deployment requires a large number of cost-effective phase shifters (PSs) on a surface that can be easily integrated into a traditional wireless network [12].

- **Compatibility with existing networks:** Since the IRSs only reflect radio fre-

quency (RF) signals, they support full duplex full-band transmission. Moreover, IRSs can be easily integrated in the existing wireless networks following the current hardware standards.

Due to the aforementioned reasons, IRS-assisted communication has gained substantial research interest in the wireless research community over the recent few years [11–19].

1.1.1 Applications of IRS

As illustrated by Fig. 1.1, the vast number of use cases for the IRS can be covered under four key scenarios listed as follows [10]:

- **IRS-assisted B5G/6G cellular networks:** In Fig. 1.1(a), IRS-assisted cellular networks are shown where it can different aspects of a wireless systems including but not limited to spectral efficiency, quality of service (QoS) constraints, physical layer security. It has also found its use in enhancing device to device (D2D) networks. Moreover, since IRSs can be made with large number of antenna elements, they can be utilized to harvest enough energy to sustain themselves in simultaneous wireless power and information transfer (SWIPT) networks [20].

- **IRS-assisted indoor communications:** IRSs can mitigate RF power losses due to the unfavorable propagation characteristics of mmWaves by acting as virtual wave guides between the BS and the mobile user. Other indoor applications include but not restricted to enhanced IRS-assisted wireless fidelity (WiFi) [21] and light fidelity (LiFi) networks that offer higher range and data rates than existing networks.

- **Applications in unmanned systems:** As illustrated by Fig. 1.1(c), unmanned aerial vehicle (UAV) enabled wireless networks, UAV connected cellular networks, autonomous vehicular and robotic networks can utilize the IRS for improving the

system performance [22]. For example, one can periodically update the phase shifts of an IRS deployed on a UAV so as to have a portable virtual LOS link between a BS and a mobile device.

- **IRS-enhanced Internet of Things (IoT) networks:** The aforementioned benefits of the IRS can be extended to IoT networks to enhance the usefulness of existing IoT networks at a low additional cost such as in smart agriculture and smart factory [23].

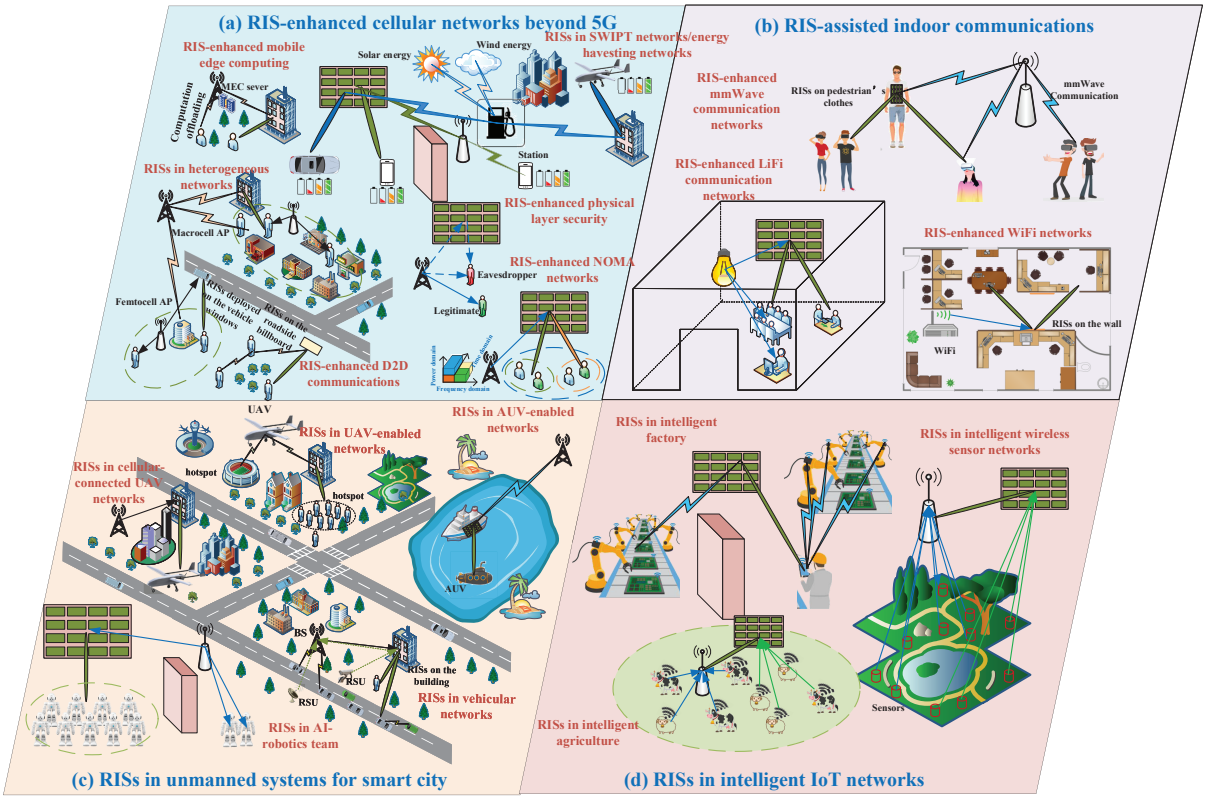

Figure 1.1: RISs/IRSs in wireless communication networks [10].

1.1.2 Challenges in IRS-Assisted Networks

The two major challenges in the IRS-assisted networks are described as follows:

- **CSI (channel state information) acquisition:** CSI knowledge is critical for many of the applications shown in Fig. 1.1. Especially in MIMO-RIS and MISO-IRS wireless networks, CSI knowledge is essential. Perfect CSI is assumed to be available at the BS, the IRS controller and the users in the majority of previous works. However, CSI acquisition in IRS-assisted networks is a difficult task which requires a considerable training overhead.

- **Optimization of phase shaifts to achieve various objectives:** Each application of the IRS-enhanced wireless networks necessitate the optimization of the IRS phase shifts to achieve a certain goal. The challenge is to optimize the IRS phase-shifts efficiently while also taking hardware inconsistencies in consideration. In this thesis, we focus on IRS-assisted cellular wireless networks and design efficient algorithms to optimize its phase shifts to increase the throughput of multi-user cellular wireless systems.

1.2 Related Work and Contributions

In [15], a multi-user multiple-input single-output (MISO) wireless system assisted by a single IRS in the downlink configuration is studied. The authors present a deep reinforcement learning (DRL)-based solution to jointly optimize the IRS phase shifts and the BS precoding under different quality of service constraints. In [19], authors tackle the problem of estimating the cascaded BS-IRS-user channels for an IRS-assisted multi-user MISO system. The author propose a pilot-based solution and improve its efficiency by exploiting the fact that all users share the same BS-IRS channel. The same problem is solved by utilizing the deep residual learning framework in [24]. In [25], an

IRS-assisted multi-cluster MISO system serving multiple users is considered wherein the authors seek to minimize the transmit power under a minimum signal-to-interference-plus-noise ratio (SINR) constraint by jointly optimizing the IRS phase shifts and the transmit precoder. They tackle the underlying problem through alternating direction method of multipliers (ADMM). An IRS-aided MISO and MIMO system with discrete phase shifts for IRS elements is also discussed in [13]. The authors formulate the problem of minimizing the transmit power under minimum SINR constraint and jointly optimize the transmit precoding and the IRS phase shifts in a mixed-integer non-linear programming framework. In [18], a relatively more practical model for IRS reflection coefficients is considered, and then a penalty-based algorithm is used to optimize the phase matrix. Moreover, IRS can also be utilized to simultaneously perform passive beamforming and physical information transfer [20] (e.g., synchronization data or the CSI estimated at the IRS). In [21], the authors present a massive backscatter wireless communication (MBWC) scheme to encode information on Wi-Fi signals reflected by the IRS.

1.3 Motivation

The vast majority of the existing work considers a MISO wireless system assisted by a single or multiple IRSs serving a single user [14, 17, 18]. So far, limited research has been conducted on IRS-aided multi-user MIMO systems. Moreover, IRS reflection coefficients are often modeled as ideal phase shifters and a realistic approach towards modeling reflection coefficients has rarely been investigated. In fact, most of the existing methods are limited to a single-phase shifter model, unimodular phase shifts being the most common one, and hence they are not resilient to the various hardware impairments of the IRS reflection elements [11, 13, 14, 17, 18]. In Chapter 2, we tackle the problem of jointly optimizing the active BS precoding and the IRS phase shifts under both ideal

and practical constraints on the phase shifts for a multi-user MIMO wireless system. Moreover, the idea of using the IRS to embed information besides performing passive beamforming in a multi-user cellular network has not been explored yet. Therefore, in Chapter 3, by leveraging the idea of backscatter communication [21, 26] and building upon the work in Chapter 2 we propose a general framework in which the IRS is used for data embedding on the reflected or re-emitted signals by changing the IRS elements' impedance and making use of the signals deliberately transmitted by a BS. We call the smart surface with such capability as "modulating intelligent surface (MIS)." In the proposed framework, the MIS can be used to either: *i*) perform passive beamforming for users served by a BS, or *ii*) embed information through backscatter communication, or *iii*) do both simultaneously. In this setting, the MIS phase shifts vary with the inevitable changes in the propagation medium to perform passive beamforming for one set of users, and also with every transmit symbol vector in order to modulate the data for another set of users.

Table 1.1: Summary of Scholastic Outputs

Publications	**Appearance**
1. **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*, doi: 10.1109/TCOMM.2021.3094509.	Chapter 2
2. **H. U. Rehman**, F. Bellili, A. Mezghani and E. Hossain, "Modulating Intelligent Surfaces for Multi-User MIMO Systems: Beamforming and Modulation Design," submitted to the *IEEE Transactions on Communications*.	Chapter 3

1.4 Scholastic Outputs and Achievements

This thesis includes material previously published/submitted in peer-reviewed journals as summarized in Table 1.1. I wish to acknowledge Dr. Faouzi Bellili, Dr. Amine Mezghani and Dr. Ekram Hossain for their help and constructive suggestions during the planning and development of this research work.

1.5 Thesis Organization and Notations

We organize the major contents of thesis into two chapters. The brief organization of the thesis is given as follows:

- In Chapter 2, we solve the problem of jointly optimizing the active BS precoding and the IRS phase shifts under the both ideal and practical constraints. Moreover, we also develop an extended version of vector approximate message passing (VAMP) algorithm which we later use for solving the underlying joint problem.
- A modulating intelligent surface (MIS) is proposed in Chapter 3. Specifically, we build upon our work in Chapter 2 and solve the problem of jointly optimizing the BS precoding, the MIS phase shifts for passive beamforming and data modulation, and the receive scaling factors.
- Chapter 4 concludes the thesis while pointing out future research directions.

Notations: Lowercase letters (e.g., r) denote scalar variables. The uppercase italic letters (e.g., N) represent scalar constants. Vectors are denoted by small boldface letters (e.g., $\mathbf{z}$) and the k-th element of $\mathbf{z}$ is denoted as z_k. Exponent on a vector (e.g., $\mathbf{z}^n$) denotes component-wise exponentiation on every element of the vector. Capital boldface letters (e.g., $\mathbf{A}$) are used to denote matrices, while a_{ik} and $\boldsymbol{a}_i$ stand, respectively,

for the (i,k)-th entry and the i-th column of $\mathbf{A}$. The zero matrix of size $M \times N$ is denoted as $\mathbf{0}_{M\times N}$. $\mathbb{C}^{M\times N}$ stands for the set of matrices of size $M \times N$ with complex elements and $\mathbf{A}^{-k}$ means $(\mathbf{A}^{-1})^k$. Rank($\mathbf{A}$) and Tr($\mathbf{A}$), return, respectively, the rank and the trace of any matrix $\mathbf{A}$. We also use $\|.\|_2$, $\|.\|_{\mathrm{F}}$, $(.)^*$, $(.)^{\mathsf{T}}$, $(.)^{\mathsf{H}}$ to denote the $\mathcal{L}_2$ norm, Frobenius norm, the conjugate, the transpose, and the conjugate transpose operators, respectively. The operator $< . >$ returns the empirical average of all the elements/entries of any vector or matrix. Moreover, vec(.) and unvec(.) denote vectorization of a matrix and unvectorization of a vector back to its original matrix form, respectively. Diag(.) operates on a vector and generates a diagonal matrix by placing that vector in the diagonal whereas diag(.) operates on a matrix and returns its main diagonal in a vector. The statistical expectation is denoted as $\mathbb{E}\{.\}$. A random vector with complex normal distribution is represented by $\mathbf{x} \sim \mathcal{CN}(\mathbf{x}; \mathbf{u}, \mathbf{R})$, where $\mathbf{u}$ and $\mathbf{R}$ denote its mean and covariance matrix, respectively. Similarly, a random matrix with complex normal distribution is represented by $\mathbf{X} \sim \mathcal{CMN}(\mathbf{X}; \mathbf{M}, \mathbf{U}, \mathbf{V})$, where $\mathbf{M}$, $\mathbf{U}$ and $\mathbf{V}$ denote its mean and covariance matrices among its rows and columns, respectively. The imaginary unit is represented by $\mathrm{j} = \sqrt{-1}$ and the $\angle(.)$ operator returns the angle of any complex number. The proportional relationship between any two entities (functions or variables) is denoted by $\propto$. Lastly, the operators $\otimes$, $\odot$ and $*$ denote the Kronecker, the Hadamard, and the column-wise Khatri-Rao products, respectively.

Chapter 2

Joint Active and Passive Beamforming Design for IRS-Assisted Multi-User MIMO Systems

IRSs have recently emerged as a promising technology for beyond-5G/6G wireless communications as it can improve both the spectral and energy efficiencies of wireless systems. IRS is an energy-efficient technology since it allows passively to beamform the incoming signal without the need for a power amplifier as in traditional MIMO BSs. It does so by suitably optimizing the phase shifts applied by each reflective element to constructively combine the incoming signals so as to achieve improved received power at the end users. In this chapter, we consider a multi-user IRS-assisted single-cell downlink MIMO system with a single IRS. The IRS is equipped with a large number of passive phase shifters that aid the BS to serve a small number of users. We propose a robust solution for the problem of jointly optimizing the active and passive beamforming

tasks under different models for the IRS reflection coefficients. The main contributions embodied by this chapter are as follows:

- We solve the problem of maximizing the spectral efficiency of the users by jointly optimizing the transmit precoding matrix at the BS and the reflection coefficients at the IRS. To that end, we first formulate the joint optimization problem under the sum MMSE criterion in order to minimize the MSE of the received symbols for all users at the same time.

- To solve the underlying joint optimization problem, we first split it using alternate optimization [27] into two easier sub-optimization tasks of the active precoder at the BS and the reflection coefficients at the IRS. The precoding sub-optimization problem is similar to the MMSE transmit precoder optimization for a traditional MIMO system, which can be solved in closed-form through Lagrange optimization.

- We modify and extend the existing VAMP algorithm [28] and propose a flexible technique to find locally optimal reflective coefficients for the IRS under multiple constraints. Precisely, we find a sub-optimal but good solution for the phase matrix under two different models for the reflection coefficients: *i*) Under the unimodular constraint on the IRS reflection coefficients and *ii*) under a practical constraint, where each IRS element is terminated by a tunable simple reactive load.

- We discuss the convergence and provide the order of complexity of the proposed solution. We present various numerical results to compare the proposed solution with the semi-definite relaxation (SDR) plus MMSE-based IRS beamforming and precoding optimization approach [11, 29], an ADMM-based solution, and a standalone massive MIMO system using MMSE precoder. The results show that, the proposed solution: *i*) outperforms both the SDR-based and the ADMM-based

solutions in terms of throughput while using the same resources and being less computationally demanding, and *ii*) achieves higher throughput than a traditional massive MIMO system while using a significantly smaller number of transmit antennas in typical propagation scenarios. We illustrate the effect of practical phase shifts on the system throughout. We also show the robustness of the proposed solution by assessing its performance under imperfect CSI.

The rest of the chapter is organized as follows: the system model along with the problem formulation for jointly optimizing the active precoder and the reflection coefficients are discussed in Section 2.1. Section 2.2 briefly introduces the VAMP algorithm and then extends it to solve optimization problems. In Section 2.3, we solve the optimization problem at hand using the proposed extended version of VAMP. In Section 2.4, we further solve the underlying optimization problem under the "simple reactive loading" constraint on the IRS reflection elements. Exhaustive numerical results are shown in Section 2.5. Finally, Section 2.6 provides an analysis on the convergence and computational complexity of the proposed solution.

2.1 System Model, Assumptions, and Problem Formulation

Consider a BS that is equipped with N antenna elements serving M ($M < N$) single-antenna users in the downlink. The BS is assisted by an IRS which has K ($K > M$) reflective elements. For each m-th user, we have a direct link to the BS expressed by a channel vector $\mathbf{h}_{\text{b-u},m} \in \mathbb{C}^N$. The channel of the surface-user m link is denoted by $\mathbf{h}_{\text{s-u},m} \in \mathbb{C}^K$. As shown in Fig. 2.1, $\mathbf{H}_{\text{b-s}} \in \mathbb{C}^{K\times N}$ denotes the channel matrix of the MIMO IRS-BS link with $\text{Rank}(\mathbf{H}_{\text{b-s}}) \geq M$. The signal received at the IRS is phase-shifted by a diagonal matrix $\text{Diag}(\boldsymbol{\upsilon}) \in \mathbb{C}^{K\times K}$, where $\boldsymbol{\upsilon} \in \mathbb{C}^K$ is the phase-shift vector

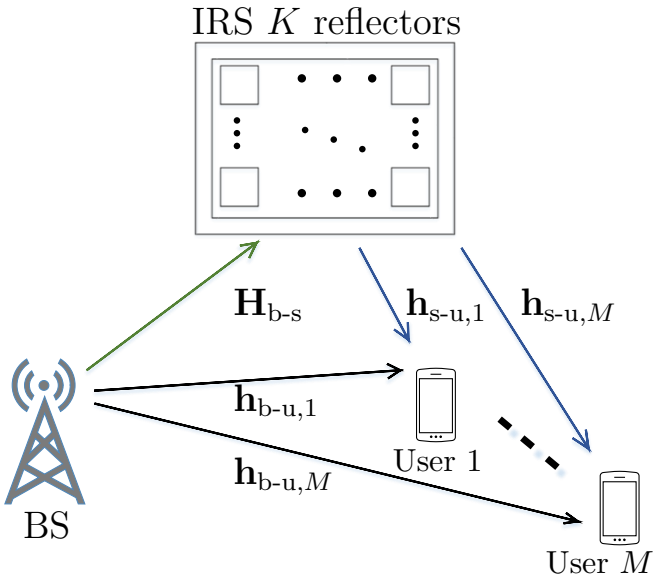

Figure 2.1: IRS-assisted multi-user MIMO system.

having unimodular elements, i.e., $|\upsilon_k| = 1$ for $k = 1, \cdots, K$. In other words, for each reflection element, we have $\upsilon_k = e^{\mathrm{j}\theta_k}$ for some phase shift $\theta_k \in [0, 2\pi]$. The received signal for user m can be expressed as follows:

$$\mathsf{y}_m = \alpha\left(\mathbf{h}_{\text{s-u},m}^{\mathsf{H}}\text{Diag}(\boldsymbol{\upsilon})\mathbf{H}_{\text{b-s}}\sum_{m'=1}^{M}\mathbf{f}_{m'}\mathsf{s}_{m'} + \mathbf{h}_{\text{b-u},m}^{\mathsf{H}}\sum_{m'=1}^{M}\mathbf{f}_{m'}\mathsf{s}_{m'} + \mathsf{w}\right), \quad m = 1, \cdots, M \tag{2.1}$$

where $\mathsf{s}_m \sim \mathcal{CN}(s; 0, 1)$ is the unknown transmit symbol, $\mathsf{w} \sim \mathcal{CN}(w; 0, \sigma_{\text{w}}^2)$ denotes additive white Gaussian noise (AWGN), and $\alpha \in \mathbb{R}$ refers to the receiver scaling which is a common practice in precoding optimization literature [30,31]. Here, $\mathbf{f}_m \in \mathbb{C}^N$ for $m = 1, \cdots, M$ are the precoding vectors that are used for power allocation and beamforming purposes. Let $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_M]$ be the precoding matrix and let P denote the total transmit power. By denoting $\mathbf{s} = [\mathsf{s}_1, \mathsf{s}_2, \cdots, \mathsf{s}_M]^{\mathsf{T}}$, it follows that $\mathbb{E}\left\{\|\mathbf{F}\mathbf{s}\|^2\right\} = P$. Let $\mathbf{H}_{\text{b-u}} = [\mathbf{h}_{\text{b-u},1}, \mathbf{h}_{\text{b-u},2}, \cdots, \mathbf{h}_{\text{b-u},M}]$ and $\mathbf{H}_{\text{s-u}} = [\mathbf{h}_{\text{s-u},1}, \mathbf{h}_{\text{s-u},2}, \cdots, \mathbf{h}_{\text{s-u},M}]$. Then, by stacking all the users' signals in one vector $\mathbf{y} = [\mathsf{y}_1, \mathsf{y}_2, \cdots, \mathsf{y}_M]^{\mathsf{T}}$, we can express the

input-output relationship of the multi-user MIMO system as:

$$\mathbf{y} = \alpha \left(\underbrace{\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} \mathbf{F}\mathbf{s}}_{\text{Users-IRS-BS}} + \underbrace{\mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}\mathbf{s}}_{\text{Users-BS}} + \mathbf{w} \right). \tag{2.2}$$

The overall effective channel matrix for all users is thus given by:

$$\mathbf{H}^{\mathsf{H}} = \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}}. \tag{2.3}$$

We aim to minimize the received symbol error of each user under the MMSE criterion, which consequently maximizes the user SINR. A lower bound on the spectral efficiency for user m can be expressed in terms of the MMSE of its received symbol [32] as follows:

$$C_m^{\text{MMSE}} = \log_2 \left(\frac{1}{\text{MMSE}_m} \right). \tag{2.4}$$

The MSE of the received symbol for user m is given by $\mathbb{E}_{\mathsf{y}_m, \mathsf{s}_m} \left\{ |\mathsf{y}_m - \mathsf{s}_m|^2 \right\}$, and for M users, the sum symbol MSE can be written as:

$$\sum_{m=1}^{M} \mathbb{E}_{\mathsf{y}_m, \mathsf{s}_m} \left\{ |\mathsf{y}_m - \mathsf{s}_m|^2 \right\} \;=\; \mathbb{E}_{\mathbf{y},\mathbf{s}} \left\{ \|\mathbf{y} - \mathbf{s}\|_2^2 \right\}. \tag{2.5}$$

Thus, our problem under the MMSE criterion can be formulated as follows:

$$\underset{\alpha, \mathbf{F}, \boldsymbol{\upsilon}}{\arg\min} \quad \mathbb{E}_{\mathbf{y},\mathbf{s}} \left\{ \|\mathbf{y} - \mathbf{s}\|_2^2 \right\}, \tag{2.6a}$$

$$\text{subject to} \quad \mathbb{E}_{\mathbf{s}} \left\{ \|\mathbf{F}\mathbf{s}\|_2^2 \right\} = P, \tag{2.6b}$$

$$|\upsilon_i| = 1, \quad i = 1, 2, \cdots, K. \tag{2.6c}$$

Remark. The objective function in (2.6a) leads to some fairness among the users by ensuring that the MSE is minimized for each user. The lower bound on sum-spectral-efficiency of M users can be expressed in terms of the MMSE of the users' received symbols [32] as follows:

$$\widehat{C} \;=\; \sum_{m=1}^{M} \log_2 \left(\frac{1}{\text{MMSE}_m} \right) \;=\; \log_2 \left(\prod_{m=1}^{M} \frac{1}{\text{MMSE}_m} \right). \tag{2.7}$$

In other words, maximizing the sum-spectral-efficiency is equivalent to minimizing the product MSE of all users. This can be achieved by minimizing the MSE of the user with the strongest channel, thereby leading to a very unfair solution. On the other hand, aiming for complete fairness results in a very inefficient allocation of resources when it comes to the overall system throughput. In this respect, the sum MMSE criterion is a good balance between the two extremes. Since our aim is to maximize the spectral efficiency of each user rather than the sum-spectral-efficiency, the MMSE criterion is a good fit for our problem formulation.

The expectation involved in (2.6a) and (2.6b) is taken with respect to (w.r.t.) the random vectors $\mathbf{s}$ and $\mathbf{w}$. Explicitly writing the objective function in (2.6a) leads to:

$$\begin{aligned}&\mathbb{E}_{\mathbf{w},\mathbf{s}}\Big\{\mathrm{Tr}\big(\alpha^2\mathbf{s}^{\mathsf{H}}\mathbf{F}^{\mathsf{H}}\mathbf{H}\mathbf{H}^{\mathsf{H}}\mathbf{F}\mathbf{s}-\alpha\mathbf{s}^{\mathsf{H}}\mathbf{F}^{\mathsf{H}}\mathbf{H}\mathbf{s}-\alpha\mathbf{s}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{F}\mathbf{s}+\mathbf{s}\mathbf{s}^{\mathsf{H}}+\alpha^2\mathbf{s}^{\mathsf{H}}\mathbf{F}^{\mathsf{H}}\mathbf{H}\mathbf{w}\\&+\alpha^2\mathbf{w}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{F}\mathbf{s}-\alpha\mathbf{w}^{\mathsf{H}}\mathbf{s}-\alpha\mathbf{s}^{\mathsf{H}}\mathbf{w}+\alpha^2\mathbf{w}^{\mathsf{H}}\mathbf{w}\big)\Big\},\end{aligned}\tag{2.8}$$

thereby resulting in the following optimization problem:

$$\underset{\alpha,\mathbf{F},\boldsymbol{\upsilon}}{\arg\min}\quad \left\|\alpha\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon})\mathbf{H}_{\text{b-s}}\mathbf{F}-(\mathbf{I}_M-\alpha\mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F})\right\|_{\text{F}}^2+M\alpha^2\sigma_{\text{w}}^2,\tag{2.9a}$$

$$\text{s.t.}\quad \|\mathbf{F}\|_{\text{F}}^2=P,\tag{2.9b}$$

$$|\upsilon_i|=1,\quad i=1,2,\cdots,K.\tag{2.9c}$$

The optimization problem in (2.9) is a non-convex optimization problem due to the unimodular constraint[1] on the IRS phase shifts in (2.9c). VAMP is a low-complexity algorithm which is designed to solve optimization problems with a linear objective function and non-linear constraints [28]. VAMP has a modular structure that makes it possible to decouple the constraints from the objective function. Therefore, the same objective function can be minimized under different constraints by modifying the VAMP module that satisfies the constraint (simple scalar functions). VAMP automatically

[1]Later, we will solve the same problem under another constraint on the reflection coefficients.

updates the stepsize at a per-iteration basis that leads to a faster convergence compared to other iterative algorithms (e.g., ADMM) [33]. This favorable property makes VAMP tuning-free. The performance of VAMP can be theoretically predicted to establish optimality through the statistical state evolution framework [28, 34].

2.2 Modified VAMP Algorithm for Constrained Optimization

Recently, message passing algorithms [28, 35, 36] have gained attention in estimation theory because of their high performance and fast convergence. Vector approximate message passing (VAMP) [28], in particular, is a low-complexity algorithm that solves quadratic loss optimization of recovering a vector from noisy linear measurements. In this section, we briefly discuss the standard max-sum VAMP algorithm and we further modify it to solve the constrained optimization problem at hand.

2.2.1 Background on Max-Sum VAMP

Approximate message passing (AMP)-based computational techniques have gained a lot of attention since their introduction within the compressed sensing framework [35]. To be precise, AMP solves the standard linear regression problem of recovering a vector $\mathbf{x} \in \mathbb{C}^N$ from noisy linear observations:

$$\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{w}, \tag{2.10}$$

where $\mathbf{A} \in \mathbb{C}^{M \times N}$ (with $M \ll N$) is called sensing matrix and $\mathbf{w} \sim \mathcal{CN}(\mathbf{w}; 0, \gamma_{\mathrm{w}}^{-1}\mathbf{I}_M)$, with $\gamma_{\mathrm{w}} > 0$, so that $p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = \mathcal{CN}(\mathbf{z}; \mathbf{A}\mathbf{x}, \gamma_{\mathrm{w}}^{-1}\mathbf{I}_M)$. Interestingly, the performance of AMP under independent and identically distributed (i.i.d.) Gaussian sensing matrices, $\mathbf{A}$, can be rigorously tracked through scalar state evolution (SE) equations [37]. One

major drawback of AMP, however, is that it often diverges if the sensing matrix, $\mathbf{A}$, is ill-conditioned or has a non-zero mean. To circumvent this problem, vector AMP (VAMP) algorithm was proposed and rigorously analyzed through SE equations in [28]. Although there is no theoretical guarantee that VAMP will always converge, strong empirical evidence suggests that VAMP is more resilient to badly conditioned sensing matrices given that they are right-orthogonally invariant [28]. Consider the joint probability distribution function (pdf) of $\mathbf{x}$ and $\mathbf{z}$, $p_{\mathbf{x},\mathbf{z}}(\mathbf{x},\mathbf{z})$

$$p_{\mathbf{x},\mathbf{z}}(\mathbf{x},\mathbf{z}) = p_{\mathbf{x}}(\mathbf{x})\,\mathcal{C}\mathcal{N}(\mathbf{z};\mathbf{A}\mathbf{x},\gamma_{\mathrm{w}}^{-1}\mathbf{I}_M). \tag{2.11}$$

Here $p_{\mathbf{x}}(\mathbf{x})$ is some prior distribution on the vector $\mathbf{x}$ whose elements are assumed to be i.i.d. with a common prior distribution, $p_{\mathrm{x}}(x)$, i.e.,

$$p_{\mathbf{x}}(\mathbf{x}) = \prod_{i=1}^{N} p_{\mathrm{x}}(x_i). \tag{2.12}$$

Max-sum VAMP can solve the following optimization problem:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{z} - \mathbf{A}\mathbf{x}\|^2, \tag{2.13}$$

by finding the *maximum a posteriori* (MAP) estimate of $\mathbf{x}$ as follows:

$$\widehat{\mathbf{x}} = \arg\max_{\mathbf{x}} p_{\mathbf{x}|\mathbf{z}}(\mathbf{x}|\mathbf{z}). \tag{2.14}$$

The algorithm consists of the following two modules.

Linear MAP/MMSE Estimator

At iteration t, the linear MAP estimator receives extrinsic information (message) from the separable (i.e., element-wise) MAP denoiser of $\mathbf{x}$ in the form of a mean vector, $\mathbf{r}_{t-1}$, and a common scalar precision, γ_{t-1}. Then, under the Gaussian prior, $\mathcal{C}\mathcal{N}(\mathbf{x};\mathbf{r}_{t-1},\gamma_{t-1}^{-1}\mathbf{I}_N)$, it computes the linear MAP estimate, $\bar{\mathbf{x}}_t$, along with the associated posterior precision, $\bar{\gamma}_t$, from the linear observations, $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{w}$ on $\mathbf{x}$. Because

we are dealing with Gaussian densities, the linear MAP estimate is equal to the linear MMSE (LMMSE) and given as follows:

$$\bar{\mathbf{x}}_t = \left(\gamma_{\mathrm{w}}\mathbf{A}^{\mathsf{H}}\mathbf{A} + \gamma_{t-1}\mathbf{I}_N\right)^{-1}\left(\gamma_{\mathrm{w}}\mathbf{A}^{\mathsf{H}}\mathbf{z} + \gamma_{t-1}\mathbf{r}_{t-1}\right), \tag{2.15}$$

$$\bar{\gamma}_t = N\mathrm{Tr}\left(\left[\gamma_{\mathrm{w}}\mathbf{A}^{\mathsf{H}}\mathbf{A} + \gamma_{t-1}\mathbf{I}_N\right]^{-1}\right)^{-1}. \tag{2.16}$$

The extrinsic information on $\mathbf{x}$ is updated as $\mathcal{C}\mathcal{N}(\mathbf{x}; \bar{\mathbf{x}}_t, \bar{\gamma}_t^{-1}\mathbf{I}_N)/\mathcal{C}\mathcal{N}(\mathbf{x}; \mathbf{r}_{t-1}, \gamma_{t-1}^{-1}\mathbf{I}_N)$, and then sent back in the form of a mean vector, $\widetilde{\mathbf{r}}_t = \left(\bar{\mathbf{x}}_t\bar{\gamma}_t - \mathbf{r}_{t-1}\gamma_{t-1}\right)/\left(\bar{\gamma}_t - \gamma_{t-1}\right)$, and a scalar precision, $\widetilde{\gamma}_t = \bar{\gamma}_t - \gamma_{t-1}$, to the separable MAP denoiser of $\mathbf{x}$. The SVD (singular value decomposition) form of VAMP directly computes extrinsic mean vector $\widetilde{\mathbf{r}}_t$ and scalar precision $\widetilde{\gamma}_t$, and can be readily obtained by substituting $\mathbf{A} = \mathbf{U}\mathrm{Diag}(\boldsymbol{\omega})\mathbf{V}^{\mathsf{H}}$ in (2.15) and (2.16).

Separable MAP Denoiser of x

This module computes the MAP estimate, $\widehat{\mathbf{x}}_t$, of $\mathbf{x}$ from the joint distribution $p_{\mathsf{x}}(\mathbf{x})\,\mathcal{C}\mathcal{N}(\mathbf{x}; \widetilde{\mathbf{r}}_t, \widetilde{\gamma}_t^{-1}\mathbf{I}_N)$. Because $\mathbf{x}$ is i.i.d., the MAP estimate can be computed through a component-wise denoising function as follows:

$$\widehat{x}_{i,t} = g_{1,i}(\widetilde{r}_{i,t}, \widetilde{\gamma}_t) \triangleq \arg\max_{x_i}\left[-\widetilde{\gamma}_t|x_i - \widetilde{r}_{i,t}|^2 + \ln p_{\mathsf{x}}(x_i)\right], \tag{2.17}$$

or equivalently,

$$g_{1,i}(\widetilde{r}_{i,t}, \widetilde{\gamma}_t) = \arg\min_{x_i}\left[\widetilde{\gamma}|x_i - \widetilde{r}_{i,t}|^2 - \ln p_{\mathsf{x}}(x_i)\right]. \tag{2.18}$$

The derivative of the scalar MAP denoiser w.r.t. $\widetilde{r}_{i,t}$ is given by [28]:

$$g'_{1,i}(\widetilde{r}_{i,t}, \widetilde{\gamma}_t) \triangleq \frac{\partial g_{1,i}(\widetilde{r}_{i,t}, \widetilde{\gamma}_t)}{\partial \widetilde{r}_{i,t}} = \frac{1}{2}\left(\frac{\partial g_{1,i}\left(\widetilde{r}_{i,t}, \widetilde{\gamma}_t\right)}{\partial \Re\left\{\widetilde{r}_{i,t}\right\}} - \mathrm{j}\frac{\partial g_{1,i}\left(\widetilde{r}_{i,t}, \widetilde{\gamma}_t\right)}{\partial \Im\left\{\widetilde{r}_{i,t}\right\}}\right) = \widetilde{\gamma}_t\widehat{\gamma}_t, \tag{2.19}$$

where $\widehat{\gamma}_t$ is the posterior precision. The vector valued denoiser function and its derivative are defined as follows:

$$\mathbf{g}_1(\widetilde{\mathbf{r}}_t, \widetilde{\gamma}_t) \triangleq \left[g_{1,1}(\widetilde{r}_{1,t}, \widetilde{\gamma}_t), \cdots, g_{1,N}(\widetilde{r}_{N,t}, \widetilde{\gamma}_t)\right]^{\mathsf{T}}, \tag{2.20}$$

$$\mathbf{g}'_1(\widetilde{\mathbf{r}}_t, \widetilde{\gamma}_t) \triangleq \left[g'_{1,1}(\widetilde{r}_{1,t}, \widetilde{\gamma}_t), \cdots, g'_{1,N}(\widetilde{r}_{N,t}, \widetilde{\gamma}_t)\right]^{\mathsf{T}}. \tag{2.21}$$

Similar to the LMMSE module, the MAP denoiser module computes an extrinsic mean vector, $\mathbf{r}_t = (\widehat{\mathbf{x}}_t\widehat{\gamma}_t - \widetilde{\mathbf{r}}_t\widetilde{\gamma}_t)\,/\,(\widehat{\gamma}_t - \widetilde{\gamma}_t)$, and a scalar precision, $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$, and sends them back to the LMMSE module for the next iteration. The process is repeated until convergence. It is worth mentioning that the extrinsic parameters, i.e., the extrinsic mean vector and the scalar precision, calculated by each module act as a Gaussian prior on the succeeding estimate of the adjacent module, thus making VAMP parameter-free. Another key advantage of VAMP is that it decouples the prior information, $p_{\mathbf{x}}(\mathbf{x})$, and the observations, $p_{\mathbf{x}|\mathbf{z}}(\mathbf{z}|\mathbf{x})$, into two separate modules. Moreover, it also enables the denoising function to be separable even if the elements of $\mathbf{x}$ are correlated in which case the LMMSE module can easily incorporate such correlation information. The steps of the standard max-sum VAMP algorithm are shown in **Algorithm 1**.

2.2.2 Optimization Oriented VAMP

In this section, we explain how max-sum VAMP can be applied to constrained optimization problems. Given the knowledge of three matrices $\mathbf{A} \in \mathbb{C}^{M\times N}$, $\mathbf{B} \in \mathbb{C}^{Q\times N}$ and $\mathbf{Z} \in \mathbb{C}^{M\times Q}$, the goal is to solve an optimization problem of the form:

$$\underset{\mathbf{x}\,\in\,\mathbb{C}^N}{\arg\min} \quad \left\|\mathbf{A}\mathrm{Diag}(\mathbf{x})\mathbf{B}^{\mathsf{T}} - \mathbf{Z}\right\|_{\mathrm{F}}^2 \tag{2.22a}$$

$$\text{s.t.} \quad f_i(x_i) = 0 \quad i = 1, \cdots, N. \tag{2.22b}$$

In the context of optimization, the observation matrix, $\mathbf{Z}$, is considered as the desired output matrix and it is also assumed to be known. Unlike the estimation problem in (2.13), we do not have a prior distribution on $\mathbf{x}$. Yet, the optimization problem in (2.22) can be solved by modifying the modules of standard max-sum VAMP.

Algorithm 1 Max-sum VAMP SVD

Given $\mathbf{A} \in \mathbb{C}^{M\times N}$, $\mathbf{z} \in \mathbb{C}^{M}$, a precision tolerance (ϵ) and a maximum number of iterations (T_{MAX})

Initialize $\mathbf{r}_0$, $\gamma_0 \geq 0$ and $t \leftarrow 1$
Compute economy-size SVD $\mathbf{A} = \mathbf{U}\text{Diag}(\boldsymbol{\omega})\mathbf{V}^{\mathsf{H}}$
$R_{\text{A}} = \text{Rank}(\mathbf{A}) = \text{length}(\boldsymbol{\omega})$
Compute $\widetilde{\mathbf{z}} = \text{Diag}(\boldsymbol{\omega})^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{z}$
repeat
    // LMMSE SVD Form.
    $\mathbf{d}_t = \gamma_{\text{w}}\text{Diag}(\gamma_{\text{w}}\boldsymbol{\omega}^2 + \gamma_{t-1}\mathbf{1}_{R_{\text{A}}})^{-1}\boldsymbol{\omega}^2$
    $\widetilde{\mathbf{r}}_t = \mathbf{r}_{t-1} + \frac{N}{R_{\text{A}}}\mathbf{V}\text{Diag}\left(\mathbf{d}_t/\langle\mathbf{d}_t\rangle\right)\left(\widetilde{\mathbf{z}} - \mathbf{V}^{\mathsf{H}}\mathbf{r}_{t-1}\right)$
    $\widetilde{\gamma}_t = \gamma_{t-1}\langle\mathbf{d}_t\rangle / \left(\frac{N}{R_{\text{A}}} - \langle\mathbf{d}_t\rangle\right)$
    // MAP Denoiser
    $\widehat{\mathbf{x}}_t = \mathbf{g}_1(\widetilde{\mathbf{r}}_t, \widetilde{\gamma}_t)$
    $\widehat{\gamma}_t = \langle\mathbf{g}_1'(\widetilde{\mathbf{r}}_t, \widetilde{\gamma}_t)\rangle / \widetilde{\gamma}_t$
    $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$
    $\mathbf{r}_t = (\widehat{\gamma}_t\widehat{\mathbf{x}}_t - \widetilde{\gamma}_t\widetilde{\mathbf{r}}_t)/\gamma_t$
    $t \leftarrow t + 1$
until $\|\widehat{\mathbf{x}}_t - \widehat{\mathbf{x}}_{t-1}\|_2^2 \leq \epsilon\,\|\widehat{\mathbf{x}}_{t-1}\|_2^2$ or $t > T_{\text{MAX}}$
return $\widehat{\mathbf{x}}_t$

Extended LMMSE/LMAP

Through vectorization, the objective function in (2.22a) can be written in the same form as the quadratic objective function in (2.13) in the following way:

$$\text{vec}(\mathbf{Z}) = (\mathbf{B} \otimes \mathbf{A})\text{vec}(\text{Diag}(\mathbf{x})). \tag{2.23}$$

We then define a matrix, $\mathbf{D} \in \mathbb{C}^{MQ \times N}$, as follows:

$$\mathbf{D} \triangleq \mathbf{B} * \mathbf{A} = [\mathbf{b}_1 \otimes \mathbf{a}_1, \cdots, \mathbf{b}_K \otimes \mathbf{a}_K]. \tag{2.24}$$

Then, the objective function in (2.22a) is equivalently expressed in a standard form that is amenable to VAMP as follows:

$$\underset{\mathbf{x} \in \mathbb{C}^N}{\arg\min} \quad \|\mathbf{D}\mathbf{x} - \mathrm{vec}(\mathbf{Z})\|_2^2 \tag{2.25a}$$

$$\text{s.t.} \quad f_i(x_i) = 0 \quad i = 1, \cdots, N. \tag{2.25b}$$

The column-wise Khatri-Rao structure can be exploited to avoid taking SVD of the large matrix $\mathbf{D}$ as explained in the sequel. Let $\mathbf{A} = \mathbf{U}_{\mathrm{A}}\mathrm{Diag}(\boldsymbol{\omega}_{\mathrm{A}})\mathbf{V}_{\mathrm{A}}^{\mathsf{H}}$, $\mathbf{B} = \mathbf{U}_{\mathrm{B}}\mathrm{Diag}(\boldsymbol{\omega}_{\mathrm{B}})\mathbf{V}_{\mathrm{B}}^{\mathsf{H}}$, $\mathbf{D} = \mathbf{U}\mathrm{Diag}(\boldsymbol{\omega})\mathbf{V}^{\mathsf{H}}$ and $\mathbf{V}_{\mathrm{BA}} = \left(\mathbf{V}_{\mathrm{B}}^{\mathsf{H}} * \mathbf{V}_{\mathrm{A}}^{\mathsf{H}}\right)^{\mathsf{H}}$. By defining the normalization vector:

$$\mathbf{v}_{\mathrm{n}} = \left[\, \|\mathbf{v}_{\mathrm{BA},1}\|_2, \ \|\mathbf{v}_{\mathrm{BA},2}\|_2, \cdots, \|\mathbf{v}_{\mathrm{BA},MQ}\|_2 \,\right]^{\mathsf{T}}, \tag{2.26}$$

it can be shown that the SVD of the matrix $\mathbf{D}$ is given by:

$$\mathbf{D} = \underbrace{(\mathbf{U}_{\mathrm{B}} \otimes \mathbf{U}_{\mathrm{A}})}_{\mathbf{U}} \underbrace{\mathrm{Diag}\big((\boldsymbol{\omega}_{\mathrm{B}} \otimes \boldsymbol{\omega}_{\mathrm{A}}) \odot \mathbf{v}_{\mathrm{n}}\big)}_{\mathrm{Diag}(\boldsymbol{\omega})} \underbrace{\left(\mathbf{V}_{\mathrm{B}}^{\mathsf{H}} * \mathbf{V}_{\mathrm{A}}^{\mathsf{H}}\right) \odot \left(\mathbf{v}_{\mathrm{n}}^{-1}\mathbf{1}_N^{\mathsf{T}}\right)}_{\mathbf{V}^{\mathsf{H}}}. \tag{2.27}$$

These steps can be easily incorporated in the **Algorithm** 1 accordingly. Similar to the standard max-sum VAMP, at iteration t, the LMMSE module receives an extrinsic mean vector, $\mathbf{r}_{t-1}$, and a scalar precision, γ_{t-1}, from the separable MAP estimator. The SVD form of VAMP allows for exploiting the Kronecker structure inside the algorithm to avoid any large matrix multiplication. The product of a Kronecker-structured matrix and a vector can be computed in an efficient way through reverse vectorization or *un-vectorization* by computing the product of three smaller matrices, and then vectorizing the result. Therefore, line 4 of **Algorithm** 1 can be modified as follows:

$$\widetilde{\mathbf{z}} = \mathrm{Diag}(\boldsymbol{\omega})^{-1}\mathbf{U}^{\mathsf{H}}\mathrm{vec}(\mathbf{Z}) \tag{2.28}$$

$$= \mathrm{Diag}(\boldsymbol{\omega})^{-1}\mathrm{vec}\left(\mathbf{U}_{\mathrm{A}}^{\mathsf{H}}\mathbf{Z}\mathbf{U}_{\mathrm{B}}^{*}\right). \tag{2.29}$$

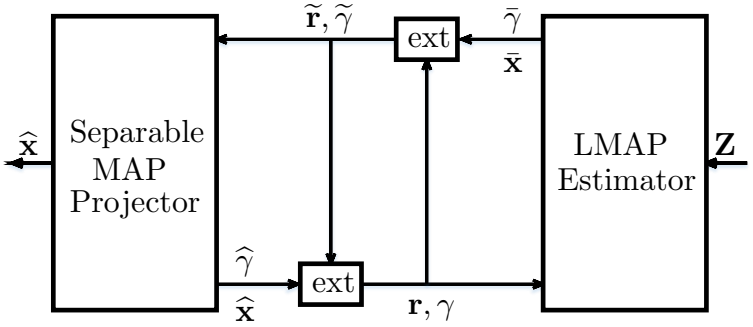

Figure 2.2: Block diagram of VAMP for optimization. The calculation of extrinsic information is performed by the "ext" blocks.

The steps for computing the extrinsic mean vector, $\widetilde{\mathbf{r}}_t$, and the scalar precision, $\widetilde{\gamma}_t$, remain unchanged and they are computed directly without the need for computing the LMMSE estimate, $\bar{\mathbf{x}}_t$, and the posterior precision, $\bar{\gamma}_t$. Hence, the only Kronecker product required for the LMMSE is of the two vectors $\boldsymbol{\omega}_{\mathrm{B}}$ and $\boldsymbol{\omega}_{\mathrm{A}}$.

Scalar MAP Projector

Because the constraint on $\mathbf{x}$ is component-wise, we model the constraint on its entries, x_i, as a prior with some precision, γ_{p}, i.e., $p_{\mathrm{x}}(x_i) \propto \exp\left(-\gamma_{\mathrm{p}}|f_i(x_i)|^2\right)$ with $\gamma_{\mathrm{p}} \to \infty$. We then define the scalar denoising function (now called projector function in the context of optimization) as follows:

$$\widehat{x}_{i,t} \;=\; g_i(\widetilde{r}_{i,t}, \widetilde{\gamma}_t) \;\triangleq\; \arg\min_{x_i} \left[\widetilde{\gamma}_t |x_i - \widetilde{r}_{i,t}|^2 - \ln p_{\mathrm{x}}(x_i)\right], \tag{2.30}$$

or equivalently:

$$g_i(\widetilde{r}_{i,t}, \widetilde{\gamma}_t) = \arg\min_{x_i} \left[\widetilde{\gamma}_t |x_i - \widetilde{r}_{i,t}|^2 + \gamma_{\mathrm{p}} |f_i(x_i)|^2\right]. \tag{2.31}$$

The parameter γ_{p} in (2.31) accounts for the weight given to the prior on x_i inside the scalar MAP optimization. Therefore, taking $\gamma_p \to \infty$ enforces the constraint. Taking

the derivative of the scalar projector function w.r.t. $\widetilde{r}_{i,t}$ as defined in equation (2.19) yields:

$$g_i'(\widetilde{r}_{i,t}, \widetilde{\gamma}_t) = \widetilde{\gamma}_t \widehat{\gamma}_t, \tag{2.32}$$

where $\widehat{\gamma}_t$ is the posterior precision. The vector valued projector function, $\mathbf{g}(\widetilde{\mathbf{r}}_t, \widetilde{\gamma})$, and its derivative, $\mathbf{g}'(\widetilde{\mathbf{r}}_t, \widetilde{\gamma})$, are defined in the same way as (2.20) and (2.21) respectively. Similar to the denoiser module, extrinsic information from the projector module is calculated in the form of the mean vector, $\mathbf{r}_t = \left(\widehat{\mathbf{x}}_t \widehat{\gamma}_t - \widetilde{\mathbf{r}}_t \widetilde{\gamma}_t\right) / (\widehat{\gamma}_t - \widetilde{\gamma}_t)$, and scalar precision, $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$, which are then fed to the LMMSE module. In an analogous way to sum-product VAMP, the max-sum VAMP (for optimization) decouples the constraint from the objective function and also enables the projector function to be separable. While the LMMSE module optimizes the objective function with no constraints, the latter are enforced by the projector function. This modular property makes VAMP a robust algorithm for solving optimization problems in the presence of linear mixing and under various component-wise constraints. The block diagram and the algorithmic steps for the optimization-oriented VAMP are presented in Fig. 2.2 and **Algorithm 2**, respectively.

2.3 VAMP-Based Solution for the Joint Beamforming Problem

In this section, we apply the optimization-oriented VAMP algorithm, described in Section 2.2.2, to simultaneously optimize the vector of phase shifters, $\boldsymbol{\upsilon}$, as well as the optimal precoding matrix $\mathbf{F}$. We decouple the joint optimization problem into two sub-problems through alternate optimization. In one side we optimize $\boldsymbol{\upsilon}$ by utilizing the modified max-sum VAMP and, on the other side, we find the optimal transmit precoding $\mathbf{F}$.

Algorithm 2 Max-sum VAMP SVD for optimization

Given $\mathbf{A} \in \mathbb{C}^{M \times N}$, $\mathbf{B} \in \mathbb{C}^{Q \times N}$, $\mathbf{Z} \in \mathbb{C}^{M \times Q}$, a precision tolerance (ϵ) and a maximum number of iterations (T_{MAX})

Select initial $\mathbf{r}_0$, $\gamma_0 \geq 0$ and $t \leftarrow 1$

Compute economy-size SVD $\mathbf{A} = \mathbf{U}_{\mathrm{A}} \mathrm{Diag}(\boldsymbol{\omega}_{\mathrm{A}}) \mathbf{V}_{\mathrm{A}}^{\mathsf{H}}$

Compute economy-size SVD $\mathbf{B} = \mathbf{U}_{\mathrm{B}} \mathrm{Diag}(\boldsymbol{\omega}_{\mathrm{B}}) \mathbf{V}_{\mathrm{B}}^{\mathsf{H}}$

Compute $\mathbf{V}_{\mathrm{BA}} = \left(\mathbf{V}_{\mathrm{B}}^{\mathsf{H}} * \mathbf{V}_{\mathrm{A}}^{\mathsf{H}} \right)^{\mathsf{H}}$

Compute vector $\mathbf{v}_{\mathrm{n}} = \left[\left\| \mathbf{v}_{\mathrm{BA},1} \right\|_2, \cdots, \left\| \mathbf{v}_{\mathrm{BA},MQ} \right\|_2 \right]^{\mathsf{T}}$

Compute $\mathbf{V}^{\mathsf{H}} = \mathbf{V}_{\mathrm{BA}}^{\mathsf{H}} \odot \left(\mathbf{v}_{\mathrm{n}}^{-1} \mathbf{1}_N^{\mathsf{T}} \right)$

Compute $\boldsymbol{\omega} = (\boldsymbol{\omega}_{\mathrm{B}} \otimes \boldsymbol{\omega}_{\mathrm{A}}) \odot \mathbf{v}_{\mathrm{n}}$

Compute $\widetilde{\mathbf{z}} = \mathrm{Diag}(\boldsymbol{\omega})^{-1} \mathrm{vec} \left(\mathbf{U}_{\mathrm{A}}^{\mathsf{H}} \mathbf{Z} \mathbf{U}_{\mathrm{B}}^{*} \right)$

$R_{\mathrm{BA}} = \mathrm{Rank}(\mathbf{B} * \mathbf{A}) = \mathrm{length}(\boldsymbol{\omega})$

repeat

  // LMMSE SVD Form.

  $\mathbf{d}_t = \gamma_{\mathrm{w}} \mathrm{Diag}(\gamma_{\mathrm{w}} \boldsymbol{\omega}^2 + \gamma_{t-1} \mathbf{1}_{R_{\mathrm{BA}}})^{-1} \boldsymbol{\omega}^2$

  $\widetilde{\mathbf{r}}_t = \mathbf{r}_{t-1} + \frac{N}{R_{\mathrm{BA}}} \mathbf{V} \mathrm{Diag} \left(\mathbf{d}_t / \langle \mathbf{d}_t \rangle \right) \left(\widetilde{\mathbf{z}} - \mathbf{V}^{\mathsf{H}} \mathbf{r}_{t-1} \right)$

  $\widetilde{\gamma}_t = \gamma_{t-1} \langle \mathbf{d}_t \rangle \, / \left(\frac{N}{R_{\mathrm{BA}}} - \langle \mathbf{d}_t \rangle \right)$

  // Separable MAP Projector.

  $\widehat{\mathbf{x}}_t = \mathbf{g}(\widetilde{\mathbf{r}}_t, \widetilde{\gamma})$

  $\widehat{\gamma}_t = \widetilde{\gamma}_t^{-1} \langle \mathbf{g}'(\widetilde{\mathbf{r}}_t, \widetilde{\gamma}) \rangle$

  $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$

  $\mathbf{r}_t = \gamma_t^{-1} (\widehat{\gamma}_t \widehat{\mathbf{x}}_t - \widetilde{\gamma}_t \widetilde{\mathbf{r}}_t)$

  $t \leftarrow t + 1$

until $\left\| \widehat{\mathbf{x}}_t - \widehat{\mathbf{x}}_{t-1} \right\|_2^2 \leq \epsilon \left\| \widehat{\mathbf{x}}_{t-1} \right\|_2^2$ or $t > T_{\mathrm{MAX}}$

return $\widehat{\mathbf{x}}_t$

2.3.1 Alternating Optimization

We use alternating minimization which is the two-block version of the block coordinate descent (BCD) algorithm. It is a simple iterative approach that optimizes one variable at a time[2] (while fixing the others) and the process is repeated for every variable. Although it is hard to analytically establish the optimality of the alternating minimization technique for non-convex optimization problems, it is known that it performs really well for various non-convex optimization problems especially for large system sizes [38–41]. More specifically, we divide the optimization problem in (2.6) into the following two sub-optimization problems:

1.

$$\arg\min_{\boldsymbol{\upsilon}} \quad \mathbb{E}_{\mathbf{y},\mathbf{s}} \left\{ \|\mathbf{y}-\mathbf{s}\|_2^2 \right\} \tag{2.33a}$$

$$|\upsilon_i| = 1, \quad i = 1, 2, \cdots, K. \tag{2.33b}$$

2.

$$\arg\min_{\alpha,\mathbf{F}} \quad \mathbb{E}_{\mathbf{y},\mathbf{s}} \left\{ \|\mathbf{y}-\mathbf{s}\|_2^2 \right\} \tag{2.34a}$$

$$\text{s.t.} \quad \mathbb{E}_{\mathbf{s}} \|\mathbf{F}\mathbf{s}\|_2^2 = P. \tag{2.34b}$$

Let us define the error at iteration t as follows:

$$E_t \triangleq \left\| \widehat{\alpha}_t \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\widehat{\boldsymbol{\upsilon}}_t) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \right) \widehat{\mathbf{F}}_t - \mathbf{I}_M \right\|_{\text{F}}^2 + M \widehat{\alpha}_t^2 \sigma_{\text{w}}^2. \tag{2.35}$$

The algorithm stops iterating when $|E_t - E_{t-1}| < \epsilon E_{t-1}$, where $\epsilon \in \mathbb{R}_+$ is some precision tolerance. The algorithmic steps for alternating minimization (after evaluating the expectation) are shown in **Algorithm 3**.

[2]Note here that a variable can be a scalar, a vector, or a whole matrix.

Algorithm 3 Alternating minimization

Given $\mathbf{H}_{\text{s-u}}$, $\mathbf{H}_{\text{b-u}}$, $\mathbf{H}_{\text{b-s}}$, a precision tolerance (ϵ), and a maximum number of iterations (T_{MAX})

1: Initialize $\widehat{\boldsymbol{\upsilon}}_0$ and $t \leftarrow 1$.

2: **repeat**

3:
$$[\widehat{\alpha}_t, \widehat{\mathbf{F}}_t] = \arg\min_{\alpha, \mathbf{F}} \left\| \alpha \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\widehat{\boldsymbol{\upsilon}}_{t-1}) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \right) \mathbf{F} - \mathbf{I}_M \right\|_{\text{F}}^2 + M\alpha^2\sigma_{\text{w}}^2$$
$$\text{s.t.} \quad \|\mathbf{F}\|_{\text{F}}^2 = P$$

4:
$$\widehat{\boldsymbol{\upsilon}}_t = \arg\min_{\boldsymbol{\upsilon}} \quad \left\| \widehat{\alpha}_t \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \right) \widehat{\mathbf{F}}_t - \mathbf{I}_M \right\|_{\text{F}}^2$$
$$\text{s.t.} \quad \upsilon_{ik} = 0 \quad i \neq k,$$
$$|\upsilon_{ii}| = 1 \quad i = 1, 2, \cdots, K$$

5: $t \leftarrow t + 1$

6: **until** $|E_t - E_{t-1}| < \epsilon E_{t-1}$ or $t > T_{\text{MAX}}$

7: **return** $\widehat{\boldsymbol{\upsilon}}_t$, $\widehat{\mathbf{F}}_t$, $\widehat{\alpha}_t$.

2.3.2 Optimization of the Phase Vector

Here, we specialize optimization-oriented VAMP algorithm introduced in Section 2.2.2 in order to optimize the phase vector, $\boldsymbol{\upsilon}$. Let us restate the associated optimization

after explicitly finding the expectation in (2.33) as follows:

$$\arg\min_{\boldsymbol{\upsilon}} \quad \left\|\alpha\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\text{Diag}(\boldsymbol{\upsilon})\mathbf{H}_{\text{b-s}}\mathbf{F} - (\mathbf{I}_M - \alpha\mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F})\right\|_{\text{F}}^2 \tag{2.36a}$$

$$\text{s.t.} \quad |\upsilon_i| = 1 \quad i = 1, 2, \cdots, K. \tag{2.36b}$$

The solution is obtained by setting $\mathbf{A} = \alpha\mathbf{H}_{\text{s-u}}^{\mathsf{H}}$, $\mathbf{B} = (\mathbf{H}_{\text{b-s}}\mathbf{F})^{\mathsf{T}}$ and $\mathbf{Z} = \mathbf{I}_M - \alpha\mathbf{H}_{\text{b-u}}\mathbf{F}$ in **Algorithm 2** and then choosing a suitable projector function to satisfy the constraints on the reflection coefficients. The unconstrained minimization of the objective function in (2.36a) is performed by the LMMSE module. We define the projector function that enforces the constraint on the reflection coefficients as:

$$g_{2,i}\left(\widetilde{r}_i, \widetilde{\gamma}\right) \triangleq \arg\min_{\upsilon_i} \left[\widetilde{\gamma}|\upsilon_i - \widetilde{r}_i|^2 + \gamma_{\text{p}}\big||\upsilon_i| - 1\big|^2\right]. \tag{2.37}$$

Solving the optimization problem in (2.37) results in the following closed-form expression for the underlying projector function:

$$g_{2,i}\left(\widetilde{r}_i, \widetilde{\gamma}\right) = \frac{\widetilde{\gamma}}{\widetilde{\gamma} + \gamma_{\text{p}}}\widetilde{r}_i + \frac{\gamma_{\text{p}}}{\widetilde{\gamma} + \gamma_{\text{p}}}\widetilde{r}_i|\widetilde{r}_i|^{-1}. \tag{2.38}$$

As $\gamma_{\text{p}} \to \infty$, we have, $\frac{\widetilde{\gamma}}{\widetilde{\gamma}+\gamma_{\text{p}}} \to 0$ and $\frac{\gamma_{\text{p}}}{\widetilde{\gamma}+\gamma_{\text{p}}} \to 1$. Therefore, the projector function simplifies to:

$$g_{2,i}\left(\widetilde{r}_i\right) = \widetilde{r}_i|\widetilde{r}_i|^{-1}. \tag{2.39}$$

The derivative of the projector function (2.39) w.r.t. $\widetilde{r}_i$ is obtained according to equation (2.19) as follows:

$$g'_{2,i}\left(\widetilde{r}_i\right) = \frac{1}{2}|\widetilde{r}_i|^{-1}. \tag{2.40}$$

Finally, the projector function, $\mathbf{g}_2(\widetilde{\mathbf{r}}_t)$, and its derivative $\mathbf{g}'_2(\widetilde{\mathbf{r}}_t)$ are obtained by following (2.20) and (2.21), respectively.

2.3.3 Optimal Precoding

The sub-optimization problem in (2.34) is a constrained MMSE transmit precoding optimization for traditional MIMO systems. It can be solved by jointly optimizing $\mathbf{F}$

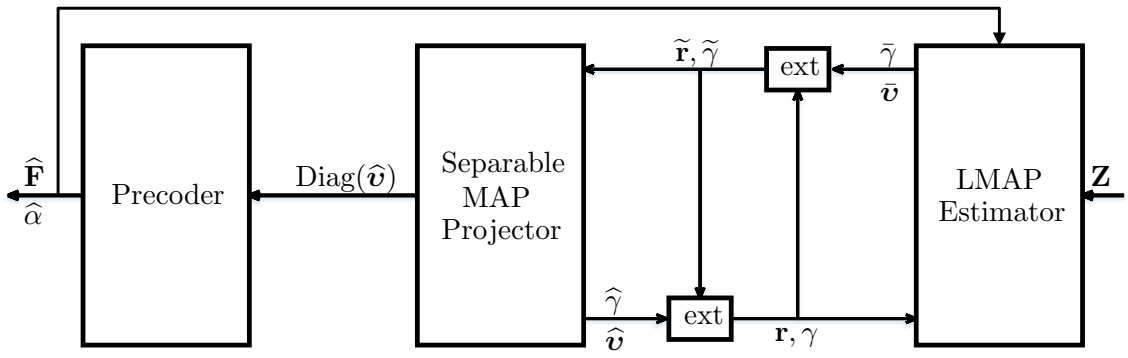

Figure 2.3: Block diagram of the proposed algorithm.

and α using Lagrange optimization. After finding the expectation, we construct the Lagrangian function associated to the problem in (2.34) as follows:

$$\mathcal{L}(\mathbf{F}, \alpha, \lambda) = \left\|\alpha \mathbf{H}^{\mathsf{H}}\mathbf{F} - \mathbf{I}_M\right\|_{\mathrm{F}}^2 + M\alpha^2\sigma_{\mathrm{w}}^2 + \lambda(\mathrm{Tr}(\mathbf{F}\mathbf{F}^{\mathsf{H}}) - P), \tag{2.41}$$

with $\lambda \in \mathbb{R}$ being the Lagrange multiplier. The closed-form solutions for optimal α and $\mathbf{F}$ are given below and we refer the reader to [30] for more details:

$$\alpha^{\mathrm{opt}} = g_3(\mathbf{H}) \triangleq \sqrt{\frac{1}{P}}\sqrt{\mathrm{Tr}\left(\left[\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{M\sigma_{\mathrm{w}}^2\mathbf{I}_N}{P}\right]^{-2}\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)}. \tag{2.42}$$

$$\begin{aligned}\mathbf{F}^{\mathrm{opt}} = g_4(\mathbf{H}) &\triangleq \frac{\sqrt{P}\left[\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{M\sigma_{\mathrm{w}}^2\mathbf{I}_N}{P}\right]^{-1}\mathbf{H}}{\sqrt{\mathrm{Tr}\left(\left[\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{M\sigma_{\mathrm{w}}^2\mathbf{I}_N}{P}\right]^{-2}\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)}} \\ &= \alpha^{\mathrm{opt}^{-1}}\left[\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{M\sigma_{\mathrm{w}}^2\mathbf{I}_N}{P}\right]^{-1}\mathbf{H}.\end{aligned} \tag{2.43}$$

Note that, the scalar, α, merely represents a scaling factor at the receiver that is used to scale the incident signal as so to obtain the transmitted constellation symbols and

this a common practice in MMSE precoding optimization [30, 31]. Choosing a common α for all users results in better tractability and makes it possible to derive a closed-form solution for the optimal α. By inspecting the closed-form solution of the precoding matrix, we observe that it is scaled by $\alpha^{\text{opt}^{-1}}$. This allows the transmitter to optimally scale all the transmit symbols based on the available transmit power whereas the receiver upscales the received signal plus noise to get back the original transmitted symbols while keeping the SNR unaffected. It is also worth mentioning that the optimal scaling factor, α^{opt}, is only utilized in optimizing the precoding matrix since the receivers can blindly estimate this scalar based on the received symbol sequence [30, 31]. Now that we have solved both sub-optimization problems in (2.33) and (2.34), separately, we substitute their solutions into **Algorithm 3**. The overall block diagram and algorithmic steps are respectively shown in Fig. 2.3 and **Algorithm 4**.

Remark. It is possible to include per-user data requirement by formulating the problem under the weighted MMSE (WMMSE) criterion where we scale the MSEs of the users with weights according to each user's data requirement and then minimize the sum MSE. To that end, we define a positive semi-definite real diagonal matrix, $\mathbf{Q}$, containing user weights, $\{q_m\}_{m=1}^{M}$, in its diagonal, i.e., $\mathbf{Q} = \text{Diag}(q_1, \cdots, q_M)$. Then, the optimization problem under the WMMSE criterion is given by:

$$\underset{\alpha,\mathbf{F},\boldsymbol{\upsilon}}{\arg\min} \quad \mathbb{E}_{\mathbf{y},\mathbf{s}} \left\{ \left\| \mathbf{Q}^{1/2}(\mathbf{y} - \mathbf{s}) \right\|_2^2 \right\}, \tag{2.44a}$$

$$\text{subject to} \quad \mathbb{E}_{\mathbf{s}} \left\{ \|\mathbf{F}\mathbf{s}\|_2^2 \right\} = P, \tag{2.44b}$$

$$|\upsilon_i| = 1, \quad i = 1, 2, \cdots, K. \tag{2.44c}$$

In this formulation, WMMSE precoding optimization is performed instead of the ordinary MMSE precoding optimization, wherein the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{Z}$ are adjusted accordingly inside the VAMP part of the **Algorithm 4**.

2.4 Joint Beamforming Under Reactive Loading at the IRS

We consider a reflective element that is combined with a tunable reactive load [3] instead of an ideal phase shifter, i.e., [4] $\upsilon_i = -(1+\mathrm{j}\chi_i)^{-1}$, where $\chi_i \in \mathbb{R}$ is a scalar reactance value that has to be optimized for each reflection coefficient. Under the unimodular constraint, the idealistic IRS has a full field of view (FOV) and the reflection coefficients correspond to ideal phase-shifters and are of the form $\upsilon_i = e^{\mathrm{j}\theta_i}$, where $\theta_i \in [0, 2\pi]$[5], whereas under the practical constraint we have a restriction on the possible values of the IRS phase shifts i.e., $\angle -(1+\mathrm{j}\chi)^{-1} \in \left[\text{-}\frac{\pi}{2}, \frac{\pi}{2}\right]$. Moreover, the magnitude of each phase shift under this constraint is always less than 1 for any $\chi \neq 0$. Practically, this introduces the phase-dependent amplitude attenuation in the incident wave. We rewrite the objective function under the new constraint on phases as follows:

$$\arg\min_{\boldsymbol{\upsilon}} \quad \left\|\alpha \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} \mathbf{F} - (\mathbf{I}_M - \alpha \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F})\right\|_{\text{F}}^2 \tag{2.45a}$$

$$\text{s.t.} \quad \upsilon_i = \frac{-1}{1+\mathrm{j}\chi_i}, \quad i = 1, 2, \cdots, K. \tag{2.45b}$$

To find the sub-optimal phase vector under the new constraint, we change the projector function accordingly as follows:

$$g_{5,i}\left(\widetilde{r}_i, \widetilde{\gamma}\right) \triangleq \arg\min_{\upsilon_i} \left[\widetilde{\gamma}|\upsilon_i - \widetilde{r}_i|^2 + \gamma_{\text{p}}\left|\upsilon_i + \frac{1}{1+\mathrm{j}\chi_i^{\text{opt}}}\right|^2\right], \tag{2.46}$$

[3] This can be implemented for instance by an antenna array composed of omni-directional dipole elements loaded with the reactive elements in the absence of a ground plane to allow for bidirectional beamforming and not just hemispherical coverage.

[4] The value 1 is the normalized resistive part of the element impedance whereas χ_i is the normalized reactive part of the antenna plus reactive termination. Accordingly υ_i represents the induced current flowing across the antenna. We assume the antenna elements to be uncoupled which holds approximately for half-wavelength element spacing.

[5] In practice, this assumption is difficult from a practical standpoint. With the reactive-loading constraint, the assumption of an IRS with full FOV becomes more acceptable.

Algorithm 4 VAMP-based joint optimization algorithm

Given $\mathbf{H}_{\text{s-u}}$, $\mathbf{H}_{\text{b-u}}$, $\mathbf{H}_{\text{b-s}}$, a precision tolerance (ϵ), and a maximum number of iterations (T_{MAX})

Initialize $\widehat{\boldsymbol{\upsilon}}_0,\ \mathbf{r}_0,\ \gamma_0 \geq 0$ and $t \leftarrow 1$

$\widehat{\mathbf{H}}_0 = \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\widehat{\boldsymbol{\upsilon}}_0)\mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}}\right)^{\mathsf{H}}$

$\widehat{\alpha}_0 = g_3\left(\widehat{\mathbf{H}}_0\right)$

$\widehat{\mathbf{F}}_0 = g_4\left(\widehat{\mathbf{H}}_0\right)$

repeat

  // LMMSE SVD Form.

  Set $\mathbf{A} = \widehat{\alpha}_{t-1}\mathbf{H}_{\text{s-u}}^{\mathsf{H}},\ \mathbf{B} = \left(\mathbf{H}_{\text{b-s}}\widehat{\mathbf{F}}_{t-1}\right)^{\mathsf{T}}$ and $\mathbf{Z} = \mathbf{I}_M - \widehat{\alpha}_{t-1}\mathbf{H}_{\text{b-u}}^{\mathsf{H}}\widehat{\mathbf{F}}_{t-1}$.

  Compute economy-size SVD $\mathbf{A} = \mathbf{U}_{\text{A}}\text{Diag}(\boldsymbol{\omega}_{\text{A}})\mathbf{V}_{\text{A}}^{\mathsf{H}}$

  Compute economy-size SVD $\mathbf{B} = \mathbf{U}_{\text{B}}\text{Diag}(\boldsymbol{\omega}_{\text{B}})\mathbf{V}_{\text{B}}^{\mathsf{H}}$

  Compute $\mathbf{V}_{\text{BA}} = \left(\mathbf{V}_{\text{B}}^{\mathsf{H}} * \mathbf{V}_{\text{A}}^{\mathsf{H}}\right)^{\mathsf{H}}$

  Compute vector $\mathbf{v}_{\text{n}} = \left[\, \|\mathbf{v}_{\text{BA},1}\|_2, \cdots, \|\mathbf{v}_{\text{BA},M^2}\|_2 \,\right]^{\mathsf{T}}$

  Compute $\mathbf{V}^{\mathsf{H}} = \mathbf{V}_{\text{BA}}^{\mathsf{H}} \odot \left(\mathbf{v}_{\text{n}}^{-1}\mathbf{1}_K^{\mathsf{T}}\right)$

  Compute $\boldsymbol{\omega} = (\boldsymbol{\omega}_{\text{B}} \otimes \boldsymbol{\omega}_{\text{A}}) \odot \mathbf{v}_{\text{n}}$

  Compute $\widetilde{\mathbf{z}} = \boldsymbol{\omega}^{-1} \odot \text{vec}\left(\mathbf{U}_{\text{A}}^{\mathsf{H}}\mathbf{Z}\mathbf{U}_{\text{B}}^{*}\right)$

  $R_{\text{BA}} = \text{Rank}(\mathbf{B} * \mathbf{A}) = \text{length}(\boldsymbol{\omega})$

  $\mathbf{d}_t = \gamma_{\text{w}}(\gamma_{\text{w}}\boldsymbol{\omega}^2 + \gamma_{t-1}\mathbf{1}_{R_{\text{BA}}})^{-1} \odot \boldsymbol{\omega}^2$

  $\widetilde{\mathbf{r}}_t = \mathbf{r}_{t-1} + \frac{K}{R_{\text{BA}}}\mathbf{V}\left(\frac{\mathbf{d}_t}{\langle \mathbf{d}_t \rangle} \odot \left(\widetilde{\mathbf{z}} - \mathbf{V}^{\mathsf{H}}\mathbf{r}_{t-1}\right)\right)$

  $\widetilde{\gamma}_t = \gamma_{t-1}\langle \mathbf{d}_t \rangle / \left(\frac{K}{R_{\text{BA}}} - \langle \mathbf{d}_t \rangle\right)$

  // Separable MAP Projector

  $\widehat{\boldsymbol{\upsilon}}_t = \mathbf{g}_2\left(\widetilde{\mathbf{r}}_t\right)$

  $\widehat{\gamma}_t = \widetilde{\gamma}_t^{-1}\left\langle \mathbf{g}_2'\left(\widetilde{\mathbf{r}}_t\right)\right\rangle$.

  $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$

  $\mathbf{r}_t = \gamma_t^{-1}\left(\widehat{\gamma}_t\widehat{\boldsymbol{\upsilon}}_t - \widetilde{\gamma}_t\widetilde{\mathbf{r}}_t\right)$

  //Find α and $\mathbf{F}$ through their closed-form solutions.

  $\widehat{\mathbf{H}}_t = \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\widehat{\boldsymbol{\upsilon}}_t)\mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}}\right)^{\mathsf{H}}$

  $\widehat{\alpha}_t = g_3\left(\widehat{\mathbf{H}}_t\right)$

  $\widehat{\mathbf{F}}_t = g_4\left(\widehat{\mathbf{H}}_t\right)$

  $t \leftarrow t + 1$

until $|E_t - E_{t-1}| < \epsilon E_{t-1}$ or $t > T_{\text{MAX}}$

return $\widehat{\boldsymbol{\upsilon}}_t,\ \widehat{\mathbf{F}}_t,\ \widehat{\alpha}_t$.

where

$$\chi_i^{\text{opt}} \;=\; g_6\left(\widetilde{r}_i\right) \;\triangleq\; \arg\min_{\chi_i}\left|\widetilde{r}_i + \frac{1}{1+\mathrm{j}\chi_i}\right|^2. \tag{2.47}$$

The optimization problem in (2.46) is a bi-level one [42]. The solution to (2.47) is substituted in (2.46) which is then solved as ordinary MAP optimization. We show in **Appendix A** that the solution to (2.47) is given by:

$$g_6\left(\widetilde{r}_i\right) \;=\; \frac{1}{2\Im\left\{\widetilde{r}_i\right\}}\left(1+2\Re\left\{\widetilde{r}_i\right\}+\sqrt{\left(1+2\Re\left\{\widetilde{r}_i\right\}\right)^2+4\Im\left\{\widetilde{r}_i\right\}^2}\right). \tag{2.48}$$

Substituting (2.48) back into (2.46) and solving the minimization leads to the following result:

$$g_{5,i}(\widetilde{r}_i,\widetilde{\gamma}) = \frac{\widetilde{\gamma}}{\widetilde{\gamma}+\gamma_{\mathrm{p}}}\widetilde{r}_i - \frac{\gamma_{\mathrm{p}}}{\widetilde{\gamma}+\gamma_{\mathrm{p}}}\left(1+\mathrm{j}g_6\left(\widetilde{r}_i\right)\right)^{-1}, \tag{2.49}$$

where $\gamma_{\mathrm{p}} \to \infty$. Thus, the projector function can be expressed as:

$$g_{5,i}\left(\widetilde{r}_i\right) = -\left(1+\mathrm{j}g_6\left(\widetilde{r}_i\right)\right)^{-1}, \tag{2.50}$$

whose derivative is obtained as defined in equation (2.19) as follows:

$$g'_{5,i}\left(\widetilde{r}_i\right) = \mathrm{j}g'_6\left(\widetilde{r}_i\right)\left(1+\mathrm{j}g_6\left(\widetilde{r}_i\right)\right)^{-2}, \tag{2.51}$$

where

$$g'_6\left(\widetilde{r}_i\right) = \frac{1}{2}\left(\frac{\partial g_6\left(\widetilde{r}_i\right)}{\partial\Re\left\{\widetilde{r}_i\right\}} - \mathrm{j}\frac{\partial g_6\left(\widetilde{r}_i\right)}{\partial\Im\left\{\widetilde{r}_i\right\}}\right). \tag{2.52}$$

The partial derivatives involved in (2.52) are given by:

$$\frac{\partial g_6\left(\widetilde{r}_i\right)}{\partial\Re\left\{\widetilde{r}_i\right\}} = \Im\left\{\widetilde{r}_i\right\}^{-1} + \left(1+2\Re\left\{\widetilde{r}_i\right\}\right)\left(\Im\left\{\widetilde{r}_i\right\}\sqrt{\left(1+2\Re\left\{\widetilde{r}_i\right\}\right)^2+4\Im\left\{\widetilde{r}_i\right\}^2}\right)^{-1}, \tag{2.53}$$

and

$$\frac{\partial g_6\left(\widetilde{r}_i\right)}{\partial\Im\left\{\widetilde{r}_i\right\}} = -\left(1+2\Re\left\{\widetilde{r}_i\right\}\right)\left(2\Im\left\{\widetilde{r}_i\right\}^2\right)^{-1} - \left(1+2\Re\left\{\widetilde{r}_i\right\}\right)^2\left(2\Im\left\{\widetilde{r}_i\right\}^2\sqrt{\left(1+2\Re\left\{\widetilde{r}_i\right\}\right)^2+4\Im\left\{\widetilde{r}_i\right\}^2}\right)^{-1}. \tag{2.54}$$

Since the derivative is required to be a real scalar, we take the absolute value of the complex derivative and, therefore, we modify the derivative of the projector function (2.51) as follows:

$$g'_{5,i}\left(\widetilde{r}_i\right) = \left|\mathrm{j}g'_6\left(\widetilde{r}_i\right)\left(1+\mathrm{j}g_6\left(\widetilde{r}_i\right)\right)^{-2}\right|. \tag{2.55}$$

Lastly, we obtain the vector valued projector function, $\mathbf{g}_5(\widetilde{\mathbf{r}}_t)$, and its derivative $\mathbf{g}'_5(\widetilde{\mathbf{r}}_t)$ according to (2.20) and (2.21), respectively, and replace $\mathbf{g}_2(\widetilde{\mathbf{r}}_t)$ and $\mathbf{g}'_2(\widetilde{\mathbf{r}}_t)$ in lines 19 and 20 of **Algorithm 4**.

2.5 Numerical Results: Performance Analysis

2.5.1 Simulation Model and Parameters

We present exhaustive Monte-Carlo simulation results to assess the performance of the proposed algorithm. We assume that the IRS is located at a fixed distance of 500 m from the BS and the users are spread uniformly at a radial distance of 10 m to 50 m from the IRS. A path-based propagation channel model, also known as parametric channel model [32], is used. Such a model is more appropriate for systems with large antenna arrays. One key parameter of such a channel model is the number of multi-path components of the BS-IRS channel which governs the effect of channel correlation. The channel between the IRS and the BS is generated according to:

$$\mathbf{H}_{\text{b-s}} \;=\; \sqrt{L(d_{\text{IRS}})}\sum_{q=1}^{Q_{\text{IRS}}} \mathsf{c}_q \mathbf{a}_{\text{IRS}}(\varphi_q,\psi_q)\mathbf{a}_{\text{BS}}(\phi_q)^{\mathsf{T}}. \tag{2.56}$$

Here, Q_{IRS} and $L(d_{\text{IRS}})$ denote the number of channel paths and the distance-dependent path-loss factor, respectively. The vectors $\mathbf{a}_{\text{BS}}(\phi)$ and $\mathbf{a}_{\text{IRS}}(\varphi,\psi)$ are the array response vectors for the BS and the IRS, respectively. The coefficients c_q in (2.56) denote the path gains which are modeled by a complex normal distribution, i.e., $\mathsf{c}_q \sim \mathcal{CN}(c_q;0,1)$. Assuming that a uniform linear array (ULA) with N antennas is used at the BS, we

Table 2.1: Simulation parameters, their notations, and values.

Parameter	**Notation, Value**	**Parameter**	**Notation, Value**
Number of channel paths IRS-BS link	$Q_{\text{IRS}} = 10$	IRS-BS distance	$d_{\text{IRS}} = 500$ m
Number of channel paths BS-user link	$Q_{\text{b-u}} = 2$	User-BS distance	$d = 500$ m
Number of channel paths IRS-user link	$Q_{\text{s-u}} = 2$	User-IRS distance	$d' \in [10, 50]$ m
Path-loss exponent IRS-BS, IRS-user link	$\eta = 2.5$	Noise variance	$\sigma_{\text{w}}^2 = -100$ dBm
Path-loss exponent BS-user link	$\eta = 3.7$	Channel path gain	$c_q \sim \mathcal{CN}(0, 1)$
Reference distance	$d_0 = 1$ m	Path-loss at reference distance	$C_0 = -30$ dB

have $\mathbf{a}_{\text{BS}}(\phi) = [1, e^{2\pi \text{j} \frac{d_{\text{b}}}{\lambda} \cos\phi}, \cdots, e^{2\pi \text{j} \frac{d_{\text{b}}}{\lambda}(N-1)\cos\phi}]^{\mathsf{T}}$ wherein λ, ϕ, and d_{b} represent the wavelength, the angle of departure (AOD), and the inter-antenna spacing at the BS, respectively. The IRS is equipped with a (square) uniform planar array (UPA) with K antenna elements which are assumed to have a cosine embedded element pattern. By defining the z-axis as the normal vector to the array, the array response vector for the

IRS is expressed as follows [43]:

$$\mathbf{a}_{\text{IRS}}(\varphi,\psi)=\sqrt{|\cos\varphi|}\begin{bmatrix}1\\ e^{2\pi\mathrm{j}\frac{d_{\text{s}}}{\lambda}\sin\varphi\sin\psi}\\ \vdots\\ e^{2\pi\mathrm{j}\frac{d_{\text{s}}}{\lambda}(\sqrt{K}-1)\sin\varphi\sin\psi}\end{bmatrix}\otimes\begin{bmatrix}1\\ e^{2\pi\mathrm{j}\frac{d_{\text{s}}}{\lambda}\sin\varphi\cos\psi}\\ \vdots\\ e^{2\pi\mathrm{j}\frac{d_{\text{s}}}{\lambda}(\sqrt{K}-1)\sin\varphi\cos\psi}\end{bmatrix}. \tag{2.57}$$

Here d_{s} represents the inter-antenna spacing at the IRS whereas φ and ψ are the angles of elevation and azimuth, respectively. In simulations we set $d_{\text{b}}=d_{\text{s}}=\lambda/2$. The angles ψ_q and ϕ_q are uniformly distributed in the interval $[0,2\pi)$ and the φ_q's are uniformly distributed in $[0,\pi)$. The channel of the direct link between the BS and each m-th single-antenna user, with $Q_{\text{b-u}}$ paths, is modeled as follows:

$$\mathbf{h}_{\text{b-u},m} \;=\; \sqrt{L(d_m)}\sum_{q=1}^{Q_{\text{b-u}}}\mathsf{c}_{m,q}\mathbf{a}_{\text{BS}}(\phi_{m,q}),\quad m=1,\cdots,M. \tag{2.58}$$

Similar to the IRS-BS channel, $\mathsf{c}_{m,q}\sim\mathcal{CN}(c_{m,q};0,1)$ and each angle $\phi_{m,q}$ is uniformly distributed in $[0,2\pi)$. The channel vectors in (2.58) are assumed to be independent across all users. Finally, the channel vector for the link between each m-th user, and the IRS with $Q_{\text{s-u}}$ channel paths, is modeled as follows:

$$\mathbf{h}_{\text{s-u},m} \;=\; \sqrt{L(d'_m)}\sum_{q=1}^{Q_{\text{s-u}}}\mathsf{c}_{m,q}\mathbf{a}_{\text{IRS}}(\varphi_{m,q},\psi_{m,q}),\quad m=1,\cdots,M. \tag{2.59}$$

The term $L(d)=C_0(d/d_0)^{\eta}$ in (2.56), (2.58), (2.59) is the distance-dependent path-loss factor, where C_0 denotes the path-loss at a reference distance $d_0=1$ m, and η is the path-loss exponent. Moreover, to account for the line-of-sight (LOS) component, the gain of one channel path is set to at least of 5 dB higher than the other path gains. To account for channel correlation effects, we have set the number of multi-path components lower than the number of BS antennas and the IRS antenna elements for the BS-IRS channel thereby making the channel matrix rank-deficient. Therefore, in simulations we have set the number of users lower than the rank of BS-IRS channel

matrix $\mathbf{H}_{\text{b-s}}$. In the simulations, we fix $d_{\text{IRS}} = 500$ m for the IRS-BS channel whereas the user-BS distance, d, and the user-IRS distance, d', vary for each user according to its location from the BS and the IRS, respectively. In all simulations, we also set $C_0 = -30$ dB, $\eta = 3.7$ (NLOS BS-user channel), $\eta = 2.5$ (NLOS IRS-BS and IRS-user channels), $Q_{\text{b-u}} = 2,\ Q_{\text{s-u}} = 2,\ \epsilon = 10^{-3}$ and $\sigma_{\text{w}}^2 = -100$ dBm. The results are averaged over 1000 independent Monte Carlo simulations.

The following two scenarios are studied. First, we consider the case where only the BS-IRS channel contains a LOS component. Then we consider the scenario where both the BS-IRS and the IRS-user channels have a LOS component but all the direct BS-user channels do not have a LOS component. The proposed VAMP-based algorithm is compared against the following four different configurations:

i. A MIMO system assisted by one IRS where the SDR technique is used to optimize the IRS reflection coefficients in combination with MMSE precoding.

ii. A MIMO system assisted by one IRS where the joint optimization of the phase matrix the and the precoding is solved through alternate optimization and penalty-based ADMM.

iii. A massive MIMO system with a large number of BS antennas with MMSE precoding.

iv. An IRS-assisted MIMO system with unoptimized IRS phases and MMSE transmit precoding.

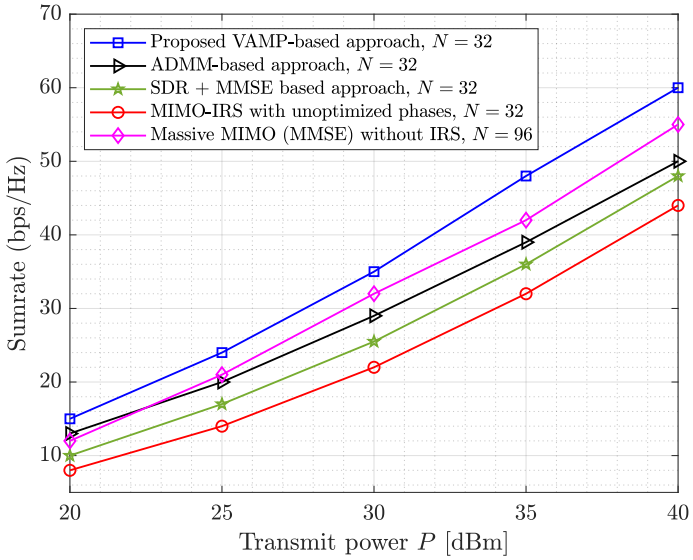

Figure 2.4: LOS IRS-BS channel: Sum-rate versus transmit power with $M = 8$ and $K = 256$.

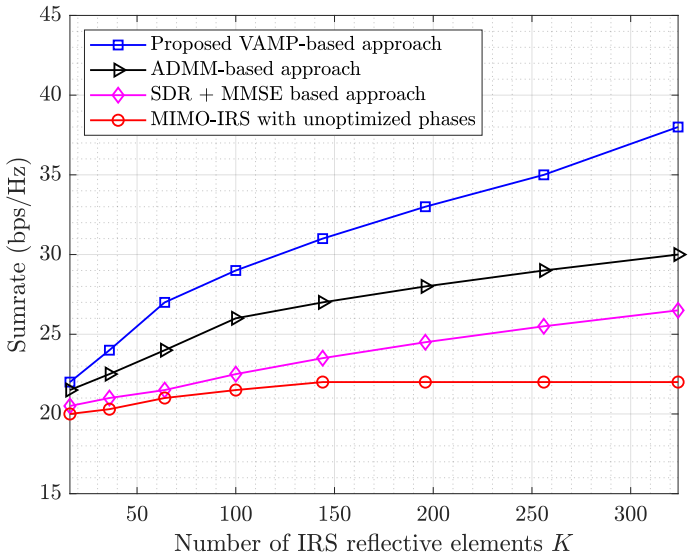

Figure 2.5: LOS IRS-BS channel: Sum-rate versus the number of IRS reflective elements with $M = 8,\ N = 32$ and $P = 30$ dBm.

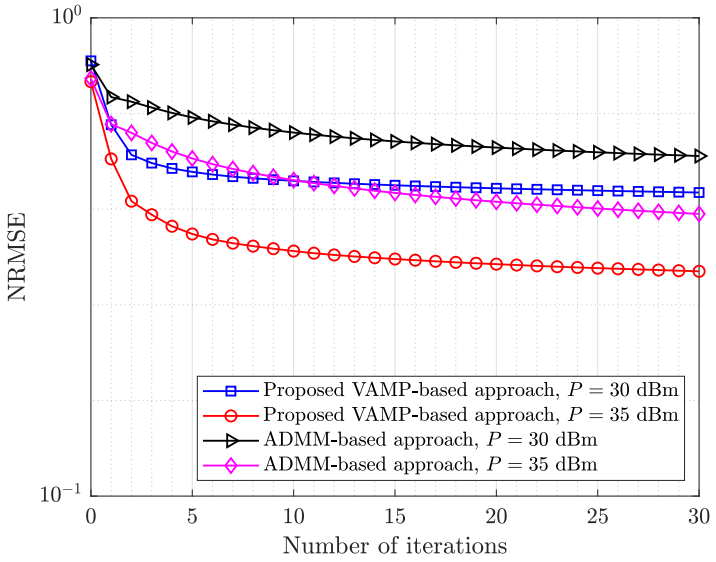

Figure 2.6: LOS IRS-BS channel: NRMSE versus the number of iterations (BS-user link excluded) with $M = 8,\ N = 32$, and $K = 256$.

2.5.2 Benchmarking Metrics

We use two metrics for performance evaluation, namely the *sum-rate*, $\widehat{C}$, and the *normalized root mean square error* (NRMSE) which are defined as follows:

$$\widehat{C} \;=\; \sum_{m=1}^{M} \log_2 \left(1 + \frac{\left|\mathbf{h}_m^{\mathsf{H}} \mathbf{f}_m\right|^2}{\sigma_{\mathrm{w}}^2 + \sum\limits_{i \neq m} \left|\mathbf{h}_m^{\mathsf{H}} \mathbf{f}_i\right|^2} \right), \tag{2.60}$$

where $\mathbf{h}_m^H = \mathbf{h}_{\text{s-u},m}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} + \mathbf{h}_{\text{b-u},m}^{\mathsf{H}}$.

$$\begin{aligned} &\mathrm{NRMSE}(\alpha, \boldsymbol{\upsilon}, \mathbf{F}) \;\triangleq \\ &\frac{1}{\sqrt{M}} \sqrt{\left\| \alpha \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \right) \mathbf{F} - \mathbf{I}_M \right\|_{\mathrm{F}}^2 + M \alpha^2 \sigma_{\mathrm{w}}^2}. \end{aligned} \tag{2.61}$$

2.5.3 Performance Results With Perfect CSI

BS-IRS channel with LOS component

This situation is encountered in a typical urban or suburban environments where the BS is located far away from the users and has no direct LOS component. However, the IRS is installed at a location where a LOS component is present in the BS-IRS link but not in the user-IRS link. Here we set the number of users to $M = 8$ and the number of BS antennas to $N = 32$ for every configuration except for massive MIMO for which we use $N = 96$. Fig. 2.4, depicts the achievable sum-rate versus the transmit power, P, for the different considered transmission schemes. The proposed algorithm in this scenario outperforms the massive MIMO system even with a significantly smaller number of transmit antennas. VAMP automatically updates the step size at a per-iteration basis that leads to a faster convergence compared to other iterative algorithms Since the proposed algorithm is based on VAMP, it outperforms the ADMM-based solution as it automatically updates the step size at a per-iteration basis (by means of calculating extrinsic information at each step) that leads to a faster convergence compared to ADMM, where the penalty parameter must be manually chosen. As per the IRS-assisted configuration, where one uses the SDR technique to optimize the IRS reflection coefficients, a significant gap is observed between the achieved sum-rates as compared to the proposed algorithm. Fig. 2.5 shows a plot of sum-rate against the number of IRS reflective elements. It is observed that even with a small number of active transmit antennas and merely ten paths between the IRS and the BS, the sum-rate for the proposed solution keeps increasing with the number of reflective elements. In contrast, the sum-rate saturates after a small gain when the IRS reflection coefficients are not optimized. Compared to the ADMM-based solution and the SDR method, the proposed algorithm shows higher throughput at every point.

The convergence of the proposed algorithm is investigated in Fig. 2.6 which depicts

the NRMSE as a function of the number of iterations. Here we exclude the direct BS-user link to highlight the throughput of the BS-IRS-user link after optimizing the IRS phase shifts. Observe that the major portion of the gain is achieved in the first few iterations. The small number of iterations required for convergence in combination with the low per-iteration complexity makes the proposed algorithm very attractive from the practical implementation point of view. The superiority of the proposed VAMP-based algorithm over the ADMM-based approach stems from the feedback mechanism of VAMP. In fact, such feedback controls the weight given to the update of $\boldsymbol{v}$ at each iteration compared to that of the preceding iteration. This is achieved with the help of scalar precision parameters that act as weighting coefficients for the $\boldsymbol{v}$ updates that are computed in the current and the preceding iteration. In addition to the plots shifting downward, the increase in transmit power widens the gap between ADMM and the proposed VAMP-based algorithm. This demonstrates that the latter utilizes the available transmit power in a more efficient way than ADMM.

BS-IRS and IRS-user channels with LOS components

Fig. 2.7 illustrates the sum-rate versus the transmit power for this configuration. As expected, the results show that by adding a LOS component, the use of an IRS together with the proposed joint beamforming optimization solution yields considerably higher sum-rates compared to a massive MIMO system with no IRS. Moreover, although the ADMM-based solution now matches the performance of massive MIMO, the advantage of the proposed VAMP-based solution over the ADMM- and SDR-based solutions is higher when compared to the NLOS configuration.

The results in Fig. 2.8, i.e., sum-rate versus the number of IRS reflective elements, also exhibit the same trends as in the NLOS scenario yet with a broader gap between the curves, thereby, corroborating the superiority of the proposed solution. Intuitively,

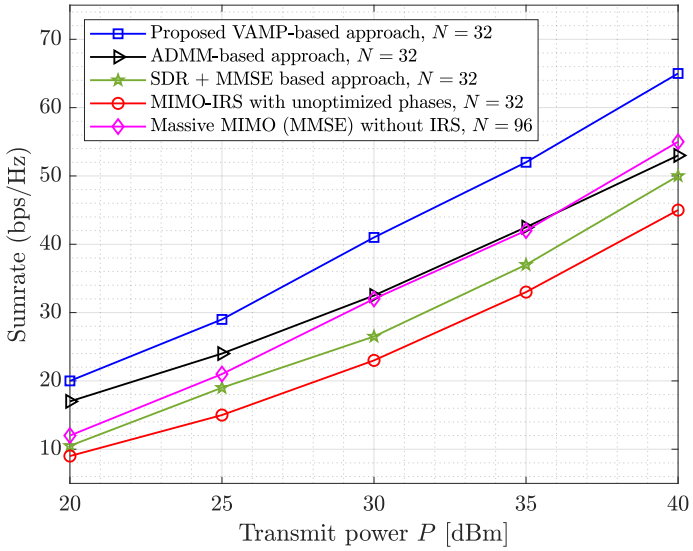

Figure 2.7: LOS IRS-user and BS-IRS channels: (a) Sum-rate versus transmit power with $M = 8$ and $K = 256$.

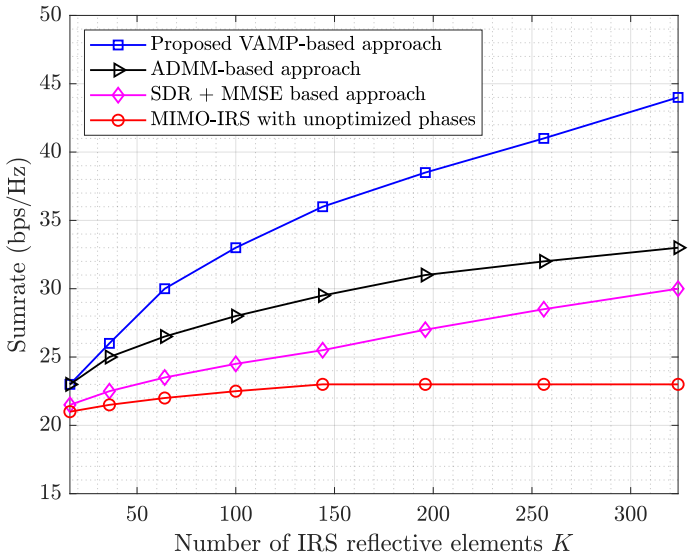

Figure 2.8: LOS IRS-user and BS-IRS channels: Sum-rate versus the number of IRS reflection elements with $M = 8,\ N = 32$ and $P = 30$ dBm.

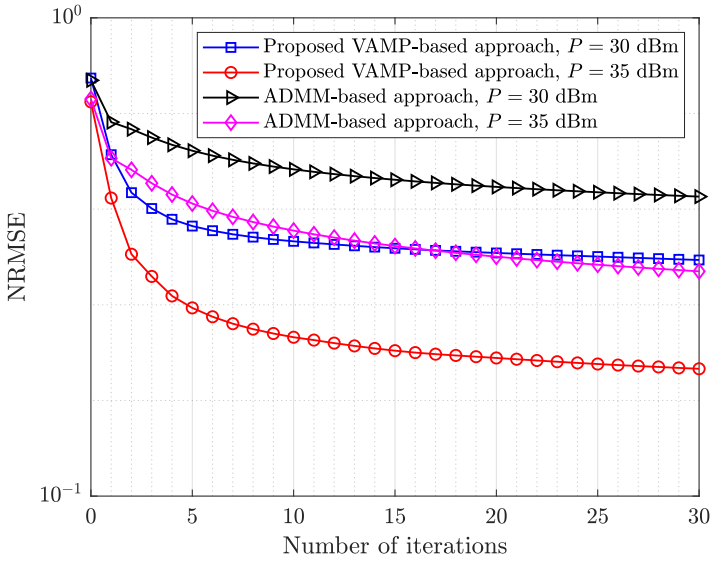

Figure 2.9: LOS IRS-user and BS-IRS channels: NRMSE versus the number of iterations (BS-user link excluded) $M = 8,\ N = 32$, and $K = 256$.

the presence of a LOS component helps the VAMP-based joint beamforming scheme to focus most of the transmit/reflected energy in that direction. This is clearly depicted in Fig. 2.9, where the NRMSE achieved by the proposed algorithm is approaching the NRMSE achieved by the ADMM-based solution but at almost 5 dB lower transmit power.

Practical IRS phase shifts

In this subsection, we assess the effect of replacing the unimodular constraint on the reflection coefficients by reactively loaded omni-directional elements. We use the same channel configuration as in Section 2.5.3. But, we rely on optimizing just the reactive part of the reflection coefficients. Therefore, as portrayed by Fig. 2.11, the new constraint decreases the throughput when compared with the ideal phase shifters setup. However, the resulting sum-rate is still much higher than the one obtained by using

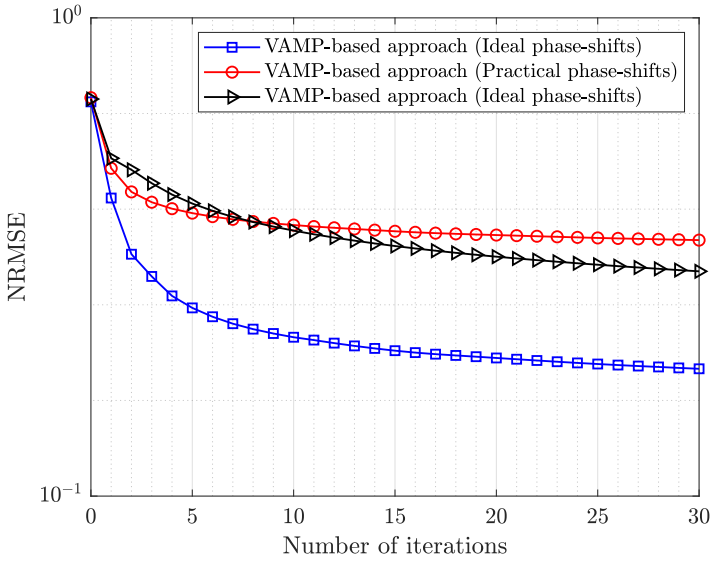

Figure 2.10: LOS IRS-user and BS-IRS channels: NRMSE versus the number of iterations (BS-user link excluded) $M = 8,\ N = 32,\ K = 256$ and $P = 30$ dBm.

unoptimized IRS reflection coefficients. In fact, when the number of IRS elements is higher than a certain value, the proposed approach with practical phase shifts achieves higher throughput than both the SDR- and ADMM-based solutions with ideal phase shifts. Similarly, due to having less room for optimizing the reflection coefficients, Fig. 2.10 shows that the NRMSE saturates sooner and at a higher value as compared to the case of a unimodular constraint (i.e., ideal phase shifts). Nonetheless, even with the more practical reactive load constraint, the resulting VAMP-based NRMSE is close to the NRMSE achieved by ADMM with ideal phase shifters.

2.5.4 Performance Results With Imperfect CSI

In this section, we measure the performance of the proposed solution in the presence of channel estimation errors. Specifically, we consider a scenario where pilot training followed by MMSE estimation algorithms are used to estimate the cascaded BS-IRS-

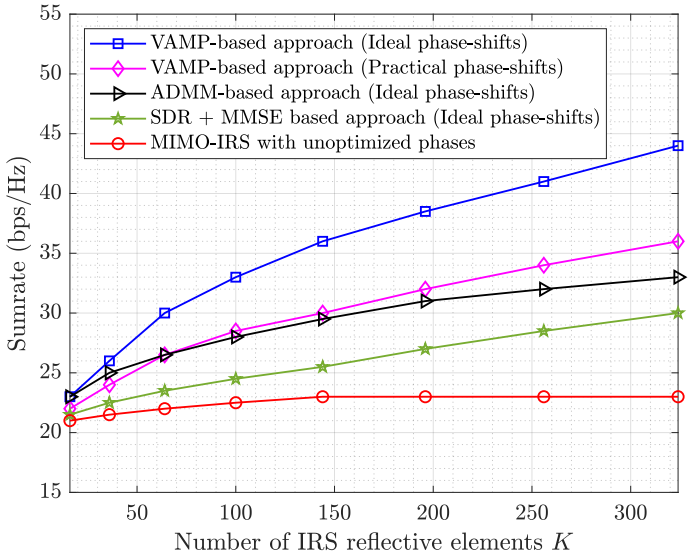

Figure 2.11: LOS IRS-user and BS-IRS channels: Sum-rate versus IRS elements with practical phase shifts with $M = 8,\ N = 32$ and $P = 30$ dBm.

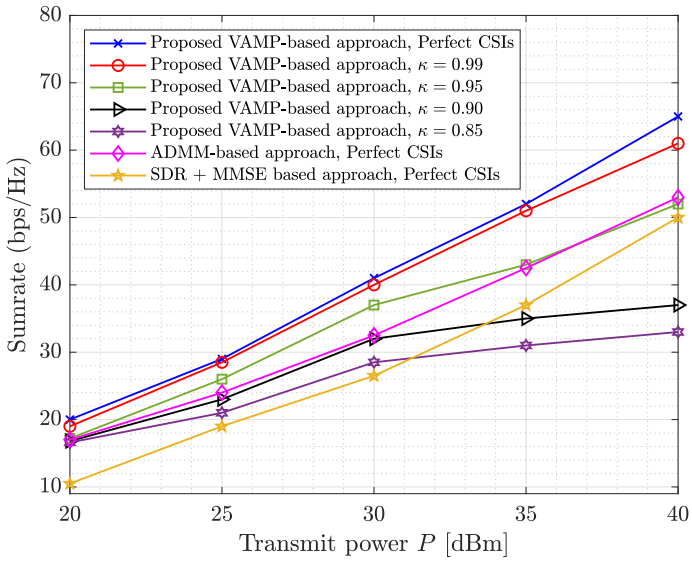

Figure 2.12: LOS IRS-user and BS-IRS channels: Sum-rate versus transmit power under imperfect CSI with $M = 8,\ N = 32$, and $K = 256$.

users and the direct BS-user channels [44, 45]. We model the estimated channel matrix and vectors using the statistical CSI error model proposed in [46–48] as follows:

$$\widehat{\mathbf{H}}_{\text{b-s}} = \kappa \mathbf{H}_{\text{b-s}} + \sqrt{(1-\kappa^2)L(d_{\text{IRS}})}\boldsymbol{\Delta}_{\text{b-s}}, \tag{2.62}$$

$$\widehat{\mathbf{h}}_{\text{b-u},m} = \kappa \mathbf{h}_{\text{b-u},m} + \sqrt{(1-\kappa^2)L(d_m)}\boldsymbol{\delta}_{\text{b-u},m}, \quad m = 1, \cdots, M, \tag{2.63}$$

$$\widehat{\mathbf{h}}_{\text{s-u},m} = \kappa \mathbf{h}_{\text{s-u},m} + \sqrt{(1-\kappa^2)L(d'_m)}\boldsymbol{\delta}_{\text{s-u},m}, \quad m = 1, \cdots, M, \tag{2.64}$$

where $\kappa \in [0, 1]$ denotes the channel estimation accuracy and $\boldsymbol{\Delta}_{\text{b-s}}$, $\boldsymbol{\delta}_{\text{b-u},m}$ and $\boldsymbol{\delta}_{\text{s-u},m}$ follow the circularly symmetric complex Gaussian (CSCG) distribution, i.e., $\text{vec}(\boldsymbol{\Delta}_{\text{b-s}}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{1}_{N\times N} \otimes \mathbf{I}_K)$, $\boldsymbol{\delta}_{\text{b-u},m} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ and $\boldsymbol{\delta}_{\text{s-u},m} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$. We first optimize the matrix $\mathbf{F}$ and vector $\boldsymbol{\upsilon}$ under imperfect CSI and then use the exact CSI matrices to calculate the sum-rate. Fig. 2.12 plots the sum-rate versus transmit power for different values of the channel estimation accuracy parameter κ. We also include plots for the other beamforming schemes under perfect CSI for reference. The results show the resilience of the proposed VAMP-based approach against small channel estimation errors. At low SNR, it is observed that the proposed design with a low channel estimation accuracy of $\kappa = 0.85$ performs better than the SDR based approach and nearly as good as the ADMM-based approach under perfect CSIs. Moreover, the performance loss with a high channel estimation accuracy value of $\kappa = 0.99$ is negligible.

2.6 Convergence, Optimality, and Complexity Analysis

According to the monotone convergence theorem in real analysis [49], a monotonically decreasing sequence with a lower bound is convergent. In our case, since the objective function,

$$\left\|\alpha \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \text{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} \mathbf{F} - (\mathbf{I}_M - \alpha \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F})\right\|_{\text{F}}^2 + M\alpha^2\sigma_{\text{w}}^2 \tag{2.65}$$

has a lower bound of zero, the proposed algorithm will always converge to a solution if the MSE monotonically decreases in both steps of the algorithm, i.e., the step of optimizing the phase-shifts (VAMP part) and the step of optimizing the active precoding. For the latter, we have a closed-form optimal solution. Therefore, it is necessary that the MSE decreases monotonically inside the VAMP step in every iteration to guarantee the convergence of the entire algorithm. In practice, most of the approximate message passing-based algorithms (including VAMP) add damping steps inside the algorithm to avoid any oscillations in the resultant MSE and thus, ensuring convergence [28]. The lines 18 and 20 inside the VAMP part of the **Algorithm 4** are, respectively, replaced by the damped versions:

$$\widetilde{\gamma}_t = \varrho\gamma_{t-1}\left\langle \mathbf{d}_t \right\rangle / \left(\frac{K}{R_{\mathrm{BA}}} - \left\langle \mathbf{d}_t \right\rangle \right) + (1-\varrho)\widetilde{\gamma}_{t-1}. \tag{2.66}$$

$$\widehat{\boldsymbol{\upsilon}}_t = \varrho \mathbf{g}_1(\widetilde{\mathbf{r}}_t) + (1-\varrho)\widehat{\boldsymbol{\upsilon}}_{t-1}, \tag{2.67}$$

for all iterations $t > 1$ where $\varrho \in (0, 1]$ is a suitably chosen damping factor. The optimal-

Table 2.2: Comparison between the CPU execution time of the proposed VAMP-based algorithm, the ADMM-based algorithm and the SDR-based algorithm for different design configurations. The algorithms terminate when $|\mathrm{NRMSE}_t - \mathrm{NRMSE}_{t-1}| < 10^{-3}\mathrm{NRMSE}_{t-1}$ or $t > 100$.

Design Parameters	**VAMP-based algorithm** $\mathcal{O}(MN(K+N))$ (msec)	**ADMM-based algorithm** $\mathcal{O}(MN(K+N))$ (msec)	**SDR-based algorithm** $\mathcal{O}(MN+K^6)$ (msec)
$M = 2,\ N = 16,\ \ K = 64$	14	26	2100
$M = 4,\ N = 32,\ \ K = 256$	104	340	12500

ity of the proposed VAMP-based approach can be investigated through statistical state

evolution analysis of the proposed algorithm which we have left for a future work. Please note that for non convex optimization problems like optimizing the phase-shifts vector under uni-modular constraint, asymptotic (for large matrix sizes) optimality can be claimed for i.i.d. matrices, if the proximal functions (projector function and its derivative inside **Algorithm 4**) are shown to be Lipschitz continuous, and the state evolution analysis reveals that the VAMP-based algorithm has only one fixed point [28, 34]. For implementation purpose, we choose the maximum possible value for precision tolerance, ϵ, for which the MSE approximately saturates before the algorithm is stopped. For the proposed solution we have found out that $\epsilon = 10^{-3}$ does the trick as the MSE achieved by choosing any lower values for ϵ is approximately equal to the MSE achieved by choosing $\epsilon = 10^{-3}$. The maximum number of iterations, $T_{\max}$, is a hardware-dependent parameter and is manually chosen to have a limit on the number of iterations.

Note that, by utilizing the Kronecker structure, we avoid any large matrix multiplication or even taking SVD of Kronecker or Khatri-Rao product of matrices. Let $\mathbf{A} = \alpha \mathbf{H}_{\text{s-u}}^{\mathsf{H}}$ and $\mathbf{B} = (\mathbf{H}_{\text{b-u}}\mathbf{F})^{\mathsf{T}}$. For our system model, the matrices $\mathbf{A}$ and $\mathbf{B}$ are of the same size $M \times K$. Assuming that the matrices $\mathbf{A}$ and $\mathbf{B}$ are of full rank, the complexity of the truncated SVDs of the matrices is of $\mathcal{O}(M^2K)$. The computational complexity of the column-wise Khatri-Rao product in line 10 and the following operations in lines 11 and 12 of **Algorithm 4** has a complexity of $\mathcal{O}(M^2K)$. The Kronecker product of two vectors in line 13 and the component-wise operations of vectors in lines 16 and 17 are of order $\mathcal{O}(M^2)$. The projector function and its derivative has a complexity in the order of $\mathcal{O}(K)$. The functions $g_3(\mathbf{H})$ and $g_4(\mathbf{H})$ can be implemented efficiently by using the matrix inversion lemma, thereby entailing a complexity of $\mathcal{O}(M^3 + MN^2)$. The complexity of all other matrix multiplications elsewhere including the LMMSE part is of order $\mathcal{O}(MNK + M^2K)$. Therefore, the overall per-iteration complexity of the algorithm is of order $\mathcal{O}(M^3 + M^2K + MNK + MN^2)$. Since $M < N$ and $M < K$ in our case, the overall per-iteration complexity simplifies to $\mathcal{O}(MN(K + N))$ or $\mathcal{O}(MNK)$

for $N \leq K$.

Table 2.2 provides a comparison of CPU (central processing unit) run time between the VAMP-based approach, the ADMM-based approach and the SDR-based approach for different design configurations. For comparison purpose, we measure the time until the NRMSE saturates with a tolerance, ϵ. Therefore, we run the algorithms until $|\text{NRMSE}_t - \text{NRMSE}_{t-1}| < \epsilon \text{NRMSE}_{t-1}$ or $t > T_{\text{MAX}}$, while setting $\epsilon = 10^{-3}$ and $T_{\text{MAX}} = 100$. We set the channel simulation parameters as in Section 2.5.3 with $P = 30$ dBm. The algorithms are simulated using MATLAB R2020a on a laptop having a Core i7-4720HQ processor and 16 GB of RAM running Windows 10 operating system. As expected, the simulation results confirm that the proposed approach is significantly faster in terms of convergence time, especially when there is a high number of IRS antenna elements.

2.7 Summary

We investigated the problem of joint active and passive beamforming design for an IRS-assisted downlink multi-user MIMO system under both ideal and practical models for the IRS phase shifts. The associated joint non-convex optimization has been formulated under sum-MMSE criterion. Using alternating minimization, the joint optimization has been decomposed into two sub-optimization tasks, i.e., optimizing the IRS phase shifts and the BS precoding matrix separately. Regarding the phase shifts, we have presented a novel approach that relies on the approximate message passing framework to solve the associated sub-optimization problem. For this, we have first extended the traditional VAMP algorithm, and then used the extended version to find a local optimum, for the phase-shifts matrix under both ideal and practical constraints. The optimal precoder at the BS, however, was found in closed-form using Lagrange optimization. Simulation results illustrate the superiority of the proposed approach over existing beamforming

schemes (e.g., the SDR-and ADMM-based approaches) both in terms of throughput and convergence speed. The results also illustrate that the reduction in the throughput of the system under more restrictive phase shifts is not significant. Moreover, it has been shown that the performance of the proposed approach is largely unaffected by small channel estimation errors.

Chapter 3

Modulating Intelligent Surfaces for Multi-User MIMO Systems: Beamforming and Modulation Design

Modulating intelligent surface (MIS) refers to an intelligent surface that has the capability of doing both *i*) beamforming for a set of users whose data is modulated by a BS, and *ii*) modulating data for another set of users on an unmodulated carrier signal transmitted by a BS by appropriately designing the MIS phase shifts. The benefits of MISs – that can beamform and modulate signals at the same time – are three-fold:

i. In traditional purely reflective IRS-based schemes, the users' received signals are subject to severe attenuation stemming from the product path loss of the BS-IRS and IRS-users links since both links are used for data communication. Using a MIS, however, allows the BS to focus all of the transmit power (of a reference signal) towards the strongest path in the BS-MIS link and the modulated signals

are passively generated at the MIS by appropriately designing its phase shifts. In this way, the information-bearing signals undergo the path loss of the MIS-users link only before reaching the intended users.

ii. The total number of users which can be served is not limited by the number of channel paths available in the BS-IRS link, but rather by the number of MIS elements. Therefore, more users than the number of BS antennas can be served by allocating power towards an MIS having a high number of antenna elements and then serving the users through the MIS.

iii. The MIS can serve users by recycling the incoming signals transmitted by the BS without any RF chains, thus making the entire approach very cost-effective.

A practical implementation of the MIS system will require the BS to transmit the optimized phase shifts to the MIS through a high speed communication link, and also a deliberately transmitted carrier signal by the BS towards the MIS to obtain the RF (radio frequency) power required to serve different users. Ideally, the MIS must be installed at a location where a line-of-sight (LOS) path is present between the MIS and the BS to minimize the power loss. There are two possible ways transmit the optimized phase shifts for modulation (i.e., phase shifts which encode the data) to the MIS: *i*) through an optical fibre link between the BS and the MIS, *ii*) through a high-speed reliable wireless link (e.g., Terahertz wireless links) between the BS and the MIS, if they are in close proximity to each other.

In this chapter, We consider a single-cell downlink MIMO system assisted by a single MIS which is equipped with a large number of passive phase shifter elements. The MIS helps the BS in better serving one portion of the users through passive beamforming and also embeds the information-bearing data for the remaining users on a separate carrier signal that it receives from the BS. In this regard, we build on our work in Chapter 2 and propose a general method to jointly optimize:

- The active BS Precoder.
- The receive scaling factor for the BS- and MIS-served users.
- The MIS phase-shifts of passive beamforming for the BS-served users.
- The embedding of user data on the reflected signal (i.e., modulation of the reflected signal) for the MIS-served users.

The major contributions embodied by this chapter can be summarized as follows:

- To the best of our knowledge, this is the first work that studies the use of the MIS for passive beamforming and data embedding at the same time in a multi-user setup. We solve the problem of maximizing the spectral-efficiency of the users by jointly optimizing the transmit precoding matrix at the BS, the receive scaling factor for the MIS-served users, and the MIS phase shifts. To do so, we follow the optimization-oriented vector approximate message passing (OOVAMP)-based approach developed in Chapter 2 while formulating the joint optimization problem under the sum minimum mean-square error (MMSE) criterion in order to minimize the mean-square error (MSE) of the received symbols for all users at the same time.
- To solve the underlying joint optimization problem, we first split it using alternate optimization [27] into two simpler sub-optimization tasks, one for finding the sub-optimal MIS phase shifts and the other for jointly optimizing the active BS precoder and the receive scaling factor for both the BS- and MIS-served users. The solution to the latter sub-optimization task is provided in closed-form.
- We apply the OOVAMP algorithm to optimize the MIS phase shifts under two different constraints. Specifically, we optimize the phase matrix under two different models for the phase shifts: *i*) under the unimodular constraint on the MIS

phase shifts, and *ii*) under a more practical constraint, where each MIS element is terminated by a tunable reactive load.

- We present various numerical results to compare the proposed scheme against the standard approach in which an IRS is being used for passive beamforming only while active MMSE precoding is used at the BS for all users. In this context, we consider two baseline techniques that rely on *i*) semi-definite relaxation (SDR) [11, 29], and *ii*) OOVAMP-based alternate optimization [50] to find the adequate IRS phase shifts. The latter approach boils down to a special case of the herein proposed scheme when the number of MIS-served users is equal to zero. The simulation results show that using MISs for joint beamforming and information embedding significantly outperforms the classical schemes in which IRSs are solely used for passively beamforming the signals they receive from the BS. We also study the resilience of the proposed scheme under channel state information (CSI) mismatch stemming from imperfect CSI acquisition in practice.

The rest of the chapter is organized as follows: the system model along with the problem formulation for jointly optimizing the active BS precoder and the MIS phase-shifts are discussed in Section 3.1. Section 3.2 discusses the matrix OOVAMP algorithm and its constituent modules. In Section 3.3, we solve the optimization problem at hand under multiple constraints by applying the OOVAMP algorithm. Lastly, numerical results are presented in Section 3.4 before drawing out some concluding remarks in Section 3.5.

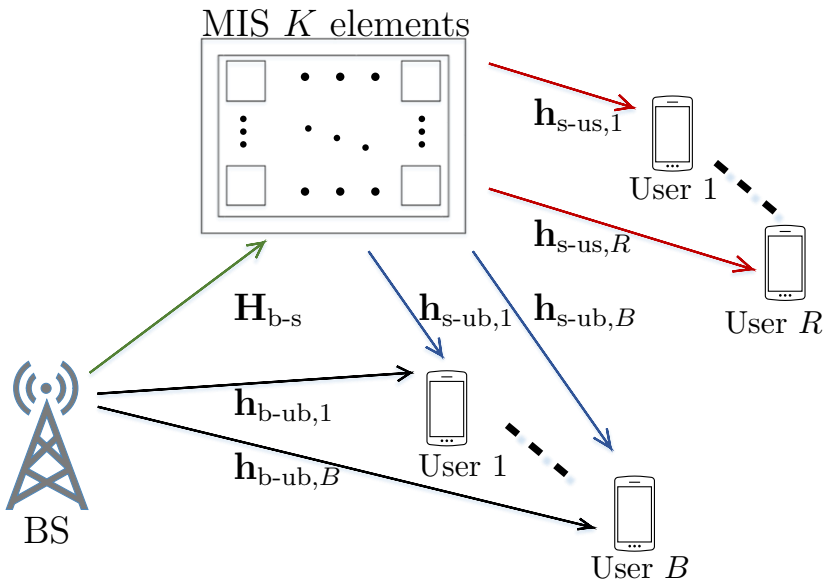

Figure 3.1: MIS-assisted multi-user MIMO system in which the MIS is being concurrently used for beamforming and data embedding.

3.1 System Model, Assumptions, and Problem Formulation

Consider a BS that is equipped with N antenna elements that is serving a total of M single-antenna users (in the downlink) with the help of an MIS that has $K > M$ reflective elements. Also consider a scheme in which the data for B out of the total M users are directly modulated/encoded by the BS (in baseband). Those B users ($B < N$) are referred to as the *BS-served* users. For the remaining $R = M - B$ users, the BS will simply send a known/reference signal which will then be modulated by appropriately phase-shifting it using the MIS reflective elements. For this reason, we call those R users as the *MIS-served* users although, strictly speaking, both types of

users are being served by the BS[1]. The goal is to optimally design the MIS phase shifts not only to modulate the data for the MIS-served users but also to passively beamform the signals for the BS-served users. As illustrated by Fig. 3.1, for each b-th BS-served user, its direct link to the BS is expressed by a channel vector $\mathbf{h}_{\text{b-ub},b} \in \mathbb{C}^N$. The BS-user channel vector for each r-th MIS-served user[2] is denoted by $\mathbf{h}_{\text{b-us},r} \in \mathbb{C}^N$. The channels of the surface-user link for the b-th BS-served user and the r-th MIS-served user are denoted, respectively, by $\mathbf{h}_{\text{s-ub},b} \in \mathbb{C}^K$ and $\mathbf{h}_{\text{s-us},r} \in \mathbb{C}^K$. Let $\mathbf{H}_{\text{b-s}} \in \mathbb{C}^{K\times N}$ denote the channel matrix of the MIMO MIS-BS link with $\text{Rank}(\mathbf{H}_{\text{b-s}}) \geq B$. The signal received at the MIS is phase-shifted by a diagonal matrix $\text{Diag}(\boldsymbol{\upsilon}) \in \mathbb{C}^{K\times K}$, where $\boldsymbol{\upsilon} \in \mathbb{C}^K$ is the phase-shift vector. Under the unimodular constraint we have $|\upsilon_k| = 1$ for $k = 1, \cdots, K$. In other words, for each reflection element, we have $\upsilon_k = e^{\mathrm{j}\theta_k}$ for some phase shift $\theta_k \in [0, 2\pi]$. Let $\mathbf{H}_{\text{b-ub}} = [\mathbf{h}_{\text{b-ub},1}, \mathbf{h}_{\text{b-ub},2}, \cdots, \mathbf{h}_{\text{b-ub},B}] \in \mathbb{C}^{N\times B}$, $\mathbf{H}_{\text{b-us}} = [\mathbf{h}_{\text{b-us},1}, \mathbf{h}_{\text{b-us},2}, \cdots, \mathbf{h}_{\text{b-us},R}] \in \mathbb{C}^{N\times R}$, $\mathbf{H}_{\text{s-ub}} = [\mathbf{h}_{\text{s-ub},1}, \mathbf{h}_{\text{s-ub},2}, \cdots, \mathbf{h}_{\text{s-ub},B}] \in \mathbb{C}^{K\times B}$ and $\mathbf{H}_{\text{s-us}} = [\mathbf{h}_{\text{s-us},1}, \mathbf{h}_{\text{s-us},2}, \cdots, \mathbf{h}_{\text{s-us},R}] \in \mathbb{C}^{K\times R}$. For mathematical convenience, we stack the channel matrices of the BS- and MIS-served users in the following combined matrices:

$$\mathbf{H}_{\text{b-u}} = [\mathbf{H}_{\text{b-ub}}\ \mathbf{H}_{\text{b-us}}] \in C^{N\times M}, \tag{3.1}$$

$$\mathbf{H}_{\text{s-u}} = [\mathbf{H}_{\text{s-ub}}\ \mathbf{H}_{\text{s-us}}]\ \ \in C^{K\times M}. \tag{3.2}$$

Let $\bar{\mathsf{y}}_m$ be the *noiseless* signal received by the m-th user and define $\bar{\mathbf{y}} \triangleq [\bar{\mathsf{y}}_1, \bar{\mathsf{y}}_2, \cdots, \bar{\mathsf{y}}_M]^{\mathsf{T}}$. In the sequel, without loss of generality, we assume that $\{\bar{\mathsf{y}}_1, \bar{\mathsf{y}}_2, \cdots, \bar{\mathsf{y}}_B\}$ and $\{\bar{\mathsf{y}}_{B+1}, \bar{\mathsf{y}}_{B+2}, \cdots, \bar{\mathsf{y}}_M\}$ pertain to the BS-served and IRS-served users, respectively. Then, $\bar{\mathbf{y}}$ can be decomposed as $\bar{\mathbf{y}} = \bar{\mathbf{y}}_{\text{b}} + \bar{\mathbf{y}}_{\text{s}}$ where $\bar{\mathbf{y}}_{\text{b}}$ (resp. $\bar{\mathbf{y}}_{\text{s}}$) is the signal intended to the

[1]Indeed, although being applied at the MIS, the information-bearing phase shifts are designed centrally at the BS as function of the users data.

[2]Although the BS is not transmitting any data to the MIS-served users, they will experience interference from the direct BS-user link of BS-served users.

BS-served (resp. IRS-served) users which are given by:

$$\bar{\mathbf{y}}_{\mathrm{b}} \;=\; \mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon})\mathbf{H}_{\text{b-s}}\mathbf{F}\mathbf{s}_{\mathrm{b}} \;+\; \mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F}\mathbf{s}_{\mathrm{b}}, \tag{3.3}$$

$$\bar{\mathbf{y}}_{\mathrm{s}} \;=\; \mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon})\mathbf{H}_{\text{b-s}}\left(\sqrt{P_{\mathrm{s}}}\mathbf{v}_{\mathrm{b}}\right). \tag{3.4}$$

In (3.3), $\mathbf{s}_{\mathrm{b}} \sim \mathcal{CN}(\mathbf{s}; \mathbf{0}, \mathbf{I}_B)$ is the unknown symbol vector being transmitted by the BS (to the B-served users) and $\boldsymbol{\upsilon} = [\upsilon_1, \upsilon_2, \cdots, \upsilon_K]^{\mathsf{T}} \in \mathbb{C}^K$ is a vector that gathers all the phase shifts used by the IRS. Moreover, $\mathbf{F} \in \mathbb{C}^{N\times B}$ is the active precoding matrix that is used for beamforming purposes at the BS, which satisfies $\|\mathbf{F}\|_{\mathrm{F}}^2 = P_{\mathrm{b}}$ where P_{b} is the fraction of power being allocated to the BS-served users. In (3.4), $\sqrt{P_{\mathrm{s}}}\mathbf{v}_{\mathrm{b}}$ is a separate constant vector being transmitted by the BS towards the IRS with $\|\mathbf{v}_{\mathrm{b}}\|_2^2 = 1$ and P_{s} is the fraction of transmit power being allocated to the IRS-served users. The total transmit power is denoted by $P = P_{\mathrm{b}} + P_{\mathrm{s}}$. Now, we let $\mathbf{w}_{\mathrm{b}} \sim \mathcal{CN}(\mathbf{w};\ \mathbf{0},\ \sigma_{\mathrm{w}}^2\mathbf{I}_B)$ and $\mathbf{w}_{\mathrm{s}} \sim \mathcal{CN}(\mathbf{w};\ \mathbf{0},\ \sigma_{\mathrm{w}}^2\mathbf{I}_R)$ denote the additive white Gaussian noise (AWGN) vectors pertaining to the BS-served and IRS-served users, respectively. Therefore, the *noisy* received signal at all the users, $\mathbf{y}$, is given by:

$$\mathbf{y} = \alpha_{\mathrm{b}}\bar{\mathbf{y}}_{\mathrm{b}} + \alpha_{\mathrm{s}}\bar{\mathbf{y}}_{\mathrm{s}} + \left[\alpha_{\mathrm{b}}\mathbf{w}_{\mathrm{b}}^{\mathsf{T}},\ \alpha_{\mathrm{s}}\mathbf{w}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}, \tag{3.5}$$

wherein α_{b} and α_{s} are some real-valued receive scaling factors[3]. They are only utilized to facilitate the optimization of the other variables (i.e., the BS precoder and the IRS phase shifts) since the receivers can blindly estimate these scalars based on the received symbol sequence [30, 31].

Remark. The system model in (3.7) is an approximation of the exact system model which is expressed as follows:

$$\begin{bmatrix} \alpha_{\mathrm{b}}\mathbf{I}_B & \mathbf{0}_{B\times R} \\ \mathbf{0}_{R\times B} & \alpha_{\mathrm{s}}\mathbf{I}_R \end{bmatrix}\left(\bar{\mathbf{y}}_{\mathrm{b}} + \bar{\mathbf{y}}_{\mathrm{s}} + \left[\mathbf{w}_{\mathrm{b}}^{\mathsf{T}},\ \mathbf{w}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}\right). \tag{3.6}$$

[3]This is a common practice in precoding optimization literature [30, 31].

The approximation will allow better tractability for the receive scaling factors, α_{b} and α_{s}, and the precoding matrix, $\mathbf{F}$, by decoupling α_{s} from the other two variables in the joint optimization problem solved in Section 3.3.2. Therefore, we use the system model in (3.7) to optimize the variables and then use those optimized variables together with the exact system model in (3.6) to compute the performance metrics such as sumrate in the numerical results section (Section 3.4).

By using (3.3) and (3.4) in (3.5), it follows that:

$$\mathbf{y} = \alpha_{\mathrm{b}} \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} \mathbf{F} \mathbf{s}_{\mathrm{b}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \mathbf{s}_{\mathrm{b}}\right) + \alpha_{\mathrm{s}} \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}) \mathbf{H}_{\text{b-s}} \sqrt{P_{\mathrm{s}}} \mathbf{v}_{\mathrm{b}}\right) + \left[\alpha_{\mathrm{b}} \mathbf{w}_{\mathrm{b}}^{\mathsf{T}},\ \alpha_{\mathrm{s}} \mathbf{w}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}. \tag{3.7}$$

The sub-optimal MIS phase-shift vector varies in space (with changes in the channel) and in time (with every transmit symbol vector). We now extend the system model in (3.7) for a transmit block of length, L, as follows:

$$\begin{aligned} \mathbf{Y} = {} & \alpha_{\mathrm{b}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{F} \mathbf{s}_{\mathrm{b},1} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \mathbf{s}_{\mathrm{b},1}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{F} \mathbf{s}_{\mathrm{b},L} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \mathbf{s}_{\mathrm{b},L}\Big] \\ & + \alpha_{\mathrm{s}} \sqrt{P_{\mathrm{s}}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}}\Big] \\ & + \Big[\alpha_{\mathrm{b}} \mathbf{W}_{\mathrm{b}}^{\mathsf{T}},\ \alpha_{\mathrm{s}} \mathbf{W}_{\mathrm{s}}^{\mathsf{T}}\Big]^{\mathsf{T}}. \end{aligned} \tag{3.8}$$

Let $\mathbf{S}_{\mathrm{b}} = [\mathbf{s}_{\mathrm{b},1}, \cdots, \mathbf{s}_{\mathrm{b},L}]$ and let $\mathbf{S}_{\mathrm{s}} \in \mathbb{C}^{R \times L}$ denote the exact information symbols for the MIS-served users, then the information symbols for both types of users are gathered in a single matrix $\mathbf{S} = \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{T}},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}$. Lastly, we define the phase shifts matrix, $\boldsymbol{\Upsilon} \in \mathbb{C}^{K \times L}$, containing the phase-shift vectors corresponding to all time indices $l = 1, \cdots, L$ as $\boldsymbol{\Upsilon} = [\boldsymbol{\upsilon}_1,\ \boldsymbol{\upsilon}_2, \cdots, \boldsymbol{\upsilon}_L]$. The goal is to maximize each user's signal-to-interference-plus-noise ratio (SINR) by minimizing the error in its received symbols under the sum MMSE criterion. A lower bound on the spectral efficiency for user m can be expressed in terms of the MMSE of its received symbol error [32] as follows:

$$C_m^{\mathrm{MMSE}} = \log_2 \left(\frac{1}{\mathrm{MMSE}_m}\right). \tag{3.9}$$

The MSE of the received symbol for user m and time index l is given by $\mathbb{E}_{\mathsf{y}_{ml},\mathsf{s}_{ml}}\left\{\left|\mathsf{y}_{ml}-\mathsf{s}_{ml}\right|^2\right\}$, and for M users and a transmit block of length L, the sum symbol MSE is given by:

$$\sum_{m=1}^{M}\sum_{l=1}^{L}\mathbb{E}_{\mathsf{y}_{ml},\mathsf{s}_{ml}}\left\{\left|\mathsf{y}_{ml}-\mathsf{s}_{ml}\right|^2\right\} \;=\; \mathbb{E}_{\mathbf{Y},\mathbf{S}}\left\{\|\mathbf{Y}-\mathbf{S}\|_{\mathrm{F}}^2\right\}. \tag{3.10}$$

Thus, our optimization problem under the sum MMSE criterion can be formulated as follows:

$$\underset{\alpha_{\mathrm{s}},\alpha_{\mathrm{b}},\mathbf{F},\boldsymbol{\Upsilon}}{\arg\min} \quad \mathbb{E}_{\mathbf{Y},\mathbf{S}}\left\{\|\mathbf{Y}-\mathbf{S}\|_{\mathrm{F}}^2\right\}, \tag{3.11a}$$

$$\text{subject to} \quad \|\mathbf{F}\|_{\mathrm{F}}^2 = P_{\mathrm{b}}, \tag{3.11b}$$

$$|\upsilon_{kl}| = 1, \quad k=1,2,\cdots,K, \;\; l=1,2,\cdots,L. \tag{3.11c}$$

We take the expectation involved in (3.11a) with respect to (w.r.t.) the random matrices $\mathbf{S}$, $\mathbf{W}_{\mathrm{b}}$ and $\mathbf{W}_{\mathrm{s}}$ to further simplify the objective function (see **Appendix B**) thereby resulting in the following optimization problem:

$$\begin{aligned}\underset{\alpha_{\mathrm{s}},\alpha_{\mathrm{b}},\mathbf{F},\boldsymbol{\Upsilon}}{\arg\min} \quad & \Big\|\alpha_{\mathrm{b}}\Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_1)\mathbf{H}_{\text{b-s}}\mathbf{F}+\mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F},\cdots,\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_L)\mathbf{H}_{\text{b-s}}\mathbf{F}+\mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F}\Big] \\ & -\Big[[\mathbf{I}_{B,1},\;\mathbf{0}_{B\times R,1}]^{\mathsf{T}},\cdots,[\mathbf{I}_{B,L},\;\mathbf{0}_{B\times R,L}]^{\mathsf{T}}\Big]\Big\|_{\mathrm{F}}^2 \\ & +\Big\|\alpha_{\mathrm{s}}\sqrt{P_{\mathrm{s}}}\Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_1)\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}},\cdots,\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_L)\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}}\Big]-\left[\mathbf{0}_{L\times B}\;\mathbf{S}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}\Big\|_{\mathrm{F}}^2 \\ & +LB\sigma_{\mathrm{w}}^2\alpha_{\mathrm{b}}^2+LR\sigma_{\mathrm{w}}^2\alpha_{\mathrm{s}}^2.\end{aligned} \tag{3.12a}$$

$$\text{s.t.} \quad \|\mathbf{F}\|_{\mathrm{F}}^2 = P_{\mathrm{b}}, \tag{3.12b}$$

$$|\upsilon_{kl}| = 1, \quad k=1,2,\cdots,K, \;\; l=1,2,\cdots,L. \tag{3.12c}$$

We shall denote the objective function in (3.12a) by $f(\alpha_{\mathrm{s}},\alpha_{\mathrm{b}},\mathbf{F},\boldsymbol{\Upsilon})$ throughout the rest of the chapter. The optimization problem in (3.12) is non-convex due to the unimodular

constraint on the MIS phase-shifts in (3.12c). We aim to solve the problem by utilizing OOVAMP algorithm in the same way as done in Chapter 2.

3.2 Optimization Oriented VAMP for Matrices

We developed OOVAMP as an extension of the standard max-sum VAMP algorithm [28] in Chapter 2 which solves constrained optimization problems involving linear objective functions under both linear and non-linear constraints. Moreover, asymptotic optimality can be claimed for the computed solution under certain mild conditions using state evolution arguments. In this section, we present an extended version of the OOVAMP algorithm to optimize matrices involving linear mixing. Given the knowledge of two matrices, $\mathbf{A} \in \mathbb{C}^{M\times N}$ and $\mathbf{Z} \in \mathbb{C}^{M\times K}$, the OOVAMP algorithm solves the following optimization problem:

$$\underset{\mathbf{X} \in \mathbb{C}^{N\times K}}{\arg\min} \quad \|\mathbf{A}\mathbf{X} - \mathbf{Z}\|_{\mathrm{F}}^2 \tag{3.13a}$$

$$\text{s.t.} \quad f_{ik}(x_{ik}) = 0 \quad i = 1, \cdots, N, \quad k = 1, \cdots, K. \tag{3.13b}$$

The algorithm consists of the following two modules.

Linear MAP Estimator

At iteration t, the linear MAP (LMAP) estimator receives extrinsic information (message) from the separable (i.e., entry-wise) MAP projector of $\mathbf{X}$ in the form of a mean matrix, $\mathbf{R}_{t-1}$, and a common scalar precision, γ_{t-1}. Then, under the Gaussian prior, $\mathcal{CMN}\left(\mathbf{X}; \mathbf{R}_{t-1}, \gamma_{t-1}^{-1}\mathbf{I}_N, \mathbf{I}_K\right)$, it computes the LMAP estimate, $\bar{\mathbf{X}}_t$, along with the associated posterior precision, $\bar{\gamma}_t$, as follow:

$$\bar{\mathbf{X}}_t = \left(\mathbf{A}^{\mathsf{H}}\mathbf{A} + \gamma_{t-1}\mathbf{I}_N\right)^{-1}\left(\mathbf{A}^{\mathsf{H}}\mathbf{Z} + \gamma_{t-1}\mathbf{R}_{t-1}\right), \tag{3.14}$$

$$\bar{\gamma}_t = N\mathrm{Tr}\left(\left[\mathbf{A}^{\mathsf{H}}\mathbf{A} + \gamma_{t-1}\mathbf{I}_N\right]^{-1}\right)^{-1}. \tag{3.15}$$

The extrinsic information on $\mathbf{X}$ is updated as:

$$\mathcal{CMN}(\mathbf{X}; \bar{\mathbf{X}}_t, \bar{\gamma}_t^{-1}\mathbf{I}_N, \mathbf{I}_K) / \mathcal{CMN}(\mathbf{X}; \mathbf{R}_{t-1}, \gamma_{t-1}^{-1}\mathbf{I}_N, \mathbf{I}_K),$$

and then sent back in the form of a mean matrix, $\widetilde{\mathbf{R}}_t = \left(\bar{\mathbf{X}}_t\bar{\gamma}_t - \mathbf{R}_{t-1}\gamma_{t-1}\right) / \left(\bar{\gamma}_t - \gamma_{t-1}\right)$, and a scalar precision, $\widetilde{\gamma}_t = \bar{\gamma}_t - \gamma_{t-1}$, to the separable MAP projector of $\mathbf{X}$.

Separable MAP Projector

Because the constraint on $\mathbf{X}$ is component-wise, the constraint on its entries, x_{ik}, is modeled as a prior with some precision, γ_{p}, i.e., $p_{\mathsf{x}}(x_{ik}) \propto \exp\left(-\gamma_{\mathrm{p}}|f_{ik}(x_{ik})|^2\right)$ with $\gamma_{\mathrm{p}} \to \infty$, which results in the following prior distribution on $\mathbf{X}$:

$$p_{\mathbf{X}}(\mathbf{X}) = \prod_{i=1}^{N}\prod_{k=1}^{Q} p_{\mathsf{x}}(x_{ik}). \tag{3.16}$$

This module computes the MAP estimate, $\widehat{\mathbf{X}}_t$, of $\mathbf{X}$ from the joint distribution $p_{\mathbf{X}}(\mathbf{X})\mathcal{CMN}\left(\mathbf{X}; \widetilde{\mathbf{R}}_t, \widetilde{\gamma}_t^{-1}\mathbf{I}_N, \mathbf{I}_K\right)$. The MAP estimate can be computed through a component-wise projector function as follows:

$$\widehat{x}_{ik,t} \;=\; g_{ik}(\widetilde{r}_{ik,t}, \widetilde{\gamma}_t) \;\triangleq\; \arg\min_{x_{ik}} \left[\widetilde{\gamma}_t|x_{ik} - \widetilde{r}_{ik,t}|^2 - \ln p_{\mathsf{x}}(x_{ik})\right], \tag{3.17}$$

or equivalently:

$$g_{ik}(\widetilde{r}_{ik,t}, \widetilde{\gamma}_t) = \arg\min_{x_{ik}} \left[\widetilde{\gamma}_t|x_{ik} - \widetilde{r}_{ik,t}|^2 + \gamma_{\mathrm{p}}|f_{ik}(x_{ik})|^2\right]. \tag{3.18}$$

The parameter γ_{p} in (3.18) accounts for the weight given to the prior on x_i inside the scalar MAP optimization. Therefore, taking $\gamma_p \to \infty$ enforces the constraint. The derivative of the scalar MAP projector w.r.t. $\widetilde{r}_{ik,t}$ is given by [28]:

$$g'_{ik}(\widetilde{r}_{ik,t}, \widetilde{\gamma}_t) \;\triangleq\; \frac{\partial g_{ik}(\widetilde{r}_{ik,t}, \widetilde{\gamma}_t)}{\partial \widetilde{r}_{ik,t}} \;=\; \frac{1}{2}\left(\frac{\partial g_{ik}\left(\widetilde{r}_{ik,t}, \widetilde{\gamma}_t\right)}{\partial \Re\left\{\widetilde{r}_{ik,t}\right\}} - \mathrm{j}\frac{\partial g_{ik}\left(\widetilde{r}_{ik,t}, \widetilde{\gamma}_t\right)}{\partial \Im\left\{\widetilde{r}_{ik,t}\right\}}\right) \;=\; \widetilde{\gamma}_t\widehat{\gamma}_t, \tag{3.19}$$

where $\widehat{\gamma}_t$ is the posterior precision. The matrix-valued projector function and its derivative are defined as follows:

$$\mathbf{G}(\widetilde{\mathbf{R}}_t, \widetilde{\gamma}_t) \triangleq \begin{bmatrix} g_{11}(\widetilde{r}_{11,t}, \widetilde{\gamma}_t) & g_{12}(\widetilde{r}_{12,t}, \widetilde{\gamma}_t) & \cdots & g_{1K}(\widetilde{r}_{1K,t}, \widetilde{\gamma}_t) \\ g_{21}(\widetilde{r}_{21,t}, \widetilde{\gamma}_t) & g_{22}(\widetilde{r}_{22,t}, \widetilde{\gamma}_t) & \cdots & g_{2K}(\widetilde{r}_{2K,t}, \widetilde{\gamma}_t) \\ \vdots & \vdots & \ddots & \vdots \\ g_{N1}(\widetilde{r}_{N1,t}, \widetilde{\gamma}_t) & g_{N2}(\widetilde{r}_{N2,t}, \widetilde{\gamma}_t) & \cdots & g_{NK}(\widetilde{r}_{NK,t}, \widetilde{\gamma}_t) \end{bmatrix}, \tag{3.20}$$

$$\mathbf{G}'(\widetilde{\mathbf{R}}_t, \widetilde{\gamma}_t) \triangleq \begin{bmatrix} g'_{11}(\widetilde{r}_{11,t}, \widetilde{\gamma}_t) & g'_{12}(\widetilde{r}_{12,t}, \widetilde{\gamma}_t) & \cdots & g'_{1K}(\widetilde{r}_{1K,t}, \widetilde{\gamma}_t) \\ g'_{21}(\widetilde{r}_{21,t}, \widetilde{\gamma}_t) & g'_{22}(\widetilde{r}_{22,t}, \widetilde{\gamma}_t) & \cdots & g'_{2K}(\widetilde{r}_{2K,t}, \widetilde{\gamma}_t) \\ \vdots & \vdots & \ddots & \vdots \\ g'_{N1}(\widetilde{r}_{N1,t}, \widetilde{\gamma}_t) & g'_{N2}(\widetilde{r}_{N2,t}, \widetilde{\gamma}_t) & \cdots & g'_{NK}(\widetilde{r}_{NK,t}, \widetilde{\gamma}_t) \end{bmatrix}. \tag{3.21}$$

Similar to the LMAP module, the MAP projector module computes an extrinsic mean matrix, $\mathbf{R}_t = \left(\widehat{\mathbf{X}}_t\widehat{\gamma}_t - \widetilde{\mathbf{R}}_t\widetilde{\gamma}_t\right) / (\widehat{\gamma}_t - \widetilde{\gamma}_t)$, and a scalar precision, $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$, and sends them back to the LMAP module for the next iteration. The process is repeated until convergence.

It is worth mentioning that the extrinsic parameters, i.e., the extrinsic mean matrix and the scalar precision, calculated by each module act as a Gaussian prior on the subsequent estimate of the adjacent module, thus making OOVAMP parameter-free. Another major advantage of OOVAMP is that it decouples the constraint from the objective function and also allows the projector function to be separable. While the LMAP module optimizes the objective function with no constraints, the latter are enforced by the projector function. This modular property makes OOVAMP an attractive algorithm for solving optimization problems involving linear mixing and under various component-wise constraints. The algorithmic steps of OOVAMP are shown in **Algorithm 5**.

Algorithm 5 Optimization-oriented max-sum matrix VAMP

Given $\mathbf{A} \in \mathbb{C}^{M\times N}$, $\mathbf{Z} \in \mathbb{C}^{M\times Q}$, a precision tolerance (ϵ) and a maximum number of iterations (T_{MAX})

Initialize $\mathbf{R}_0$, $\gamma_0 \geq 0$ and $t \leftarrow 1$

repeat

  // LMAP.

  $\bar{\mathbf{X}}_t = \left(\mathbf{A}^{\mathsf{H}}\mathbf{A} + \gamma_{t-1}\mathbf{I}_N\right)^{-1}\left(\mathbf{A}^{\mathsf{H}}\mathbf{Z} + \gamma_{t-1}\mathbf{R}_{t-1}\right)$

  $\bar{\gamma}_t = N\text{Tr}\left(\left[\mathbf{A}^{\mathsf{H}}\mathbf{A} + \gamma_{t-1}\mathbf{I}_N\right]^{-1}\right)^{-1}$

  $\widetilde{\gamma}_t = \bar{\gamma}_t - \gamma_{t-1}$

  $\widetilde{\mathbf{R}}_t = \widetilde{\gamma}_t^{-1}\left(\bar{\mathbf{X}}_t\bar{\gamma}_t - \mathbf{R}_{t-1}\gamma_{t-1}\right)$

  // Separable MAP Projector

  $\widehat{\mathbf{X}}_t = \mathbf{G}\Big(\widetilde{\mathbf{R}}_t, \widetilde{\gamma}_t\Big)$

  $\widehat{\gamma}_t = \widetilde{\gamma}_t^{-1}\left\langle \mathbf{G}'\Big(\widetilde{\mathbf{R}}_t, \widetilde{\gamma}_t\Big)\right\rangle$

  $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$

  $\mathbf{R}_t = \gamma_t^{-1}\Big(\widehat{\gamma}_t\widehat{\mathbf{X}}_t - \widetilde{\gamma}_t\widetilde{\mathbf{R}}_t\Big)$

  $t \leftarrow t + 1$

until $\left\|\widehat{\mathbf{X}}_t - \widehat{\mathbf{X}}_{t-1}\right\|_2^2 \leq \epsilon\left\|\widehat{\mathbf{X}}_{t-1}\right\|_2^2$ or $t > T_{\text{MAX}}$

return $\widehat{\mathbf{X}}_t$

3.3 OOVAMP-Based Solution for the Optimization Problem in (3.12)

In this section, we apply the OOVAMP algorithm, introduced in Section 3.2, to simultaneously optimize the matrix of phase shifters, $\boldsymbol{\Upsilon}$, the optimal precoding matrix $\mathbf{F}$, and the scaling factors, α_{b} and α_{s}. We follow the optimization procedure presented in Chapter 2, and decouple the joint optimization problem into two sub-problems through alternate optimization. In one side, we optimize $\boldsymbol{\Upsilon}$ by utilizing OOVAMP and, on the

other side, we find the optimal transmit precoding, $\mathbf{F}$, and scalars α_{b} and α_{s}. It is a simple iterative approach that optimizes a subset of all variables at a time while fixing the other set of variables and the process is repeated until convergence. More specifically, we divide the optimization problem in (3.11) into the following two sub-optimization problems:

1.

$$\widehat{\boldsymbol{\Upsilon}} = \arg\min_{\boldsymbol{\Upsilon}} \quad f(\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}, \boldsymbol{\Upsilon}) \tag{3.22a}$$

$$\text{s.t.} \quad |\upsilon_{kl}| = 1, \quad k = 1, 2, \cdots, K, \quad l = 1, 2, \cdots, L. \tag{3.22b}$$

2.

$$\arg\min_{\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}} \quad f(\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}, \widehat{\boldsymbol{\Upsilon}}) \tag{3.23a}$$

$$\text{s.t.} \quad \|\mathbf{F}\|_{\mathrm{F}}^2 = P_{\mathrm{b}}. \tag{3.23b}$$

3.3.1 Optimizing the MIS Phase Shifts

Here, we derive the OOVAMP modules (i.e., LMAP estimator and separable projector function) to solve the sub-optimization problem in (3.22) which is restated explicitly as follows:

$$\begin{aligned}\arg\min_{\boldsymbol{\Upsilon}} \quad & \Big\| \alpha_{\mathrm{b}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \Big] \\ & - \Big[[\mathbf{I}_{B,1},\ \mathbf{0}_{B \times R,1}]^{\mathsf{T}}, \cdots, [\mathbf{I}_{B,L},\ \mathbf{0}_{B \times R,L}]^{\mathsf{T}} \Big] \Big\|_{\mathrm{F}}^2 \\ & + \Big\| \alpha_{\mathrm{s}} \sqrt{P_{\mathrm{s}}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}} \Big] - \big[\mathbf{0}_{L \times B}\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}} \big]^{\mathsf{T}} \Big\|_{\mathrm{F}}^2 \\ & + LB\sigma_{\mathrm{w}}^2 \alpha_{\mathrm{b}}^2 + LR\sigma_{\mathrm{w}}^2 \alpha_{\mathrm{s}}^2. \end{aligned} \tag{3.24a}$$

$$\text{s.t.} \quad |\upsilon_{kl}| = 1, \quad k = 1, 2, \cdots, K, \quad l = 1, 2, \cdots, L. \tag{3.24b}$$

Next, we re-express the objective function in (3.24a) in a form that is similar to the general OOVAMP objective function in (3.13a). In fact, by introducing the following matrices:

$$\mathbf{A} = \alpha_{\mathrm{b}} \mathbf{H}_{\text{s-u}}^{\mathsf{H}}, \qquad (M \times K), \tag{3.25}$$

$$\mathbf{B} = (\mathbf{H}_{\text{b-s}} \mathbf{F})^{\mathsf{T}}, \qquad (B \times K), \tag{3.26}$$

$$\mathbf{D} = \left[\mathbf{b}_1 \otimes \mathbf{a}_1, \cdots, \mathbf{b}_K \otimes \mathbf{a}_K\right], \qquad (MB \times K), \tag{3.27}$$

$$\mathbf{M} = \sqrt{P_{\mathrm{s}}} \alpha_{\mathrm{s}} \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}}), \qquad (M \times K), \tag{3.28}$$

$$\mathbf{X} = \left[\mathrm{vec}\left([\mathbf{I}_B,\ \mathbf{0}_{B\times R}]^{\mathsf{T}} - \alpha_{\mathrm{b}} \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}\right), \cdots, \mathrm{vec}\left([\mathbf{I}_B,\ \mathbf{0}_{B\times R}]^{\mathsf{T}} - \alpha_{\mathrm{b}} \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}\right)\right], \quad (MB \times L), \tag{3.29}$$

$$\mathbf{Z} = \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}, \qquad (M \times L), \tag{3.30}$$

we show in **Appendix C** that the optimization problem in (3.24) can be rewritten as follows:

$$\arg\min_{\boldsymbol{\Upsilon}} \quad \|\mathbf{D}\boldsymbol{\Upsilon} - \mathbf{X}\|_{\mathrm{F}}^2 \ + \ \|\mathbf{M}\boldsymbol{\Upsilon} - \mathbf{Z}\|_{\mathrm{F}}^2 \ + \ LB\sigma_{\mathrm{w}}^2\alpha_{\mathrm{b}}^2 \ + \ LR\sigma_{\mathrm{w}}^2\alpha_{\mathrm{s}}^2 \tag{3.31a}$$

$$\text{s.t.} \quad |\upsilon_{kl}| = 1, \quad k = 1, 2, \cdots, K, \quad l = 1, 2, \cdots, L. \tag{3.31b}$$

The steps to derive both OOVAMP modules are detailed in the sequel.

LMAP Estimator

The LMAP module performs the minimization of the objection function in (3.31a) under the Gaussian prior, $\mathcal{CMN}\left(\boldsymbol{\Upsilon}; \mathbf{R}_{t-1}, \gamma_{t-1}^{-1}\mathbf{I}_K, \mathbf{I}_L\right)$, by solving the following un-

constrained optimization problem:

$$\arg\min_{\boldsymbol{\Upsilon}} \quad \frac{1}{2}\left\|\mathbf{D}\boldsymbol{\Upsilon}-\mathbf{X}\right\|_{\mathrm{F}}^{2}+\frac{1}{2}\left\|\mathbf{M}\boldsymbol{\Upsilon}-\mathbf{Z}\right\|_{\mathrm{F}}^{2}+\frac{\gamma_{t-1}}{2}\left\|\boldsymbol{\Upsilon}-\mathbf{R}_{t-1}\right\|_{\mathrm{F}}^{2}. \tag{3.32}$$

The solution (i.e., the LMAP estimate and the associated posterior precision) to the optimization problem in (3.32) is given as follows:

$$\bar{\boldsymbol{\Upsilon}}_{t} = \left(\mathbf{D}^{\mathsf{H}}\mathbf{D}+\mathbf{M}^{\mathsf{H}}\mathbf{M}+\gamma_{t-1}\mathbf{I}_{K}\right)^{-1}\left(\mathbf{D}^{\mathsf{H}}\mathbf{X}+\mathbf{M}^{\mathsf{H}}\mathbf{Z}+\gamma_{t-1}\mathbf{R}_{t-1}\right), \tag{3.33}$$

$$\bar{\gamma}_{t} = K\mathrm{Tr}\left(\left[\mathbf{D}^{\mathsf{H}}\mathbf{D}+\mathbf{M}^{\mathsf{H}}\mathbf{M}+\gamma_{t-1}\mathbf{I}_{K}\right]^{-1}\right)^{-1}. \tag{3.34}$$

Scalar MAP Projector

In this section, we reuse the two projector functions defined in Chapter 2 to satisfy the two types of constraint on the MIS reflection coefficients, i.e., *i*) the unimodular constraint, and *ii*) a practical constraint on the MIS phase shifts in which each antenna element is terminated by a variable reactive load. The projector function and its derivative for the unimodular constraint is given as follows:

$$g_{1,kl}\left(\widetilde{r}_{kl}\right) = \widetilde{r}_{kl}|\widetilde{r}_{kl}|^{-1}, \tag{3.35}$$

$$g'_{1,kl}\left(\widetilde{r}_{kl}\right) = \frac{1}{2}|\widetilde{r}_{kl}|^{-1}, \tag{3.36}$$

where the derivative is taken according to (3.19).

To optimize the MIS phase shifts under a practical constraint, we consider a reflective element that is combined with a tunable reactive load instead of an ideal phase-shifter, i.e., $\upsilon_{kl}=-(1+\mathrm{j}\chi_{kl})^{-1}$, where $\chi_{kl}\in\mathbb{R}$ is a scalar reactance value that must be optimized for each reflection coefficient. Under the unimodular constraint, the idealistic MIS has a full field of view (FOV) and the reflection coefficients correspond to ideal phase-shifters and are of the form $\upsilon_{kl}=e^{\mathrm{j}\theta_{kl}}$, where $\theta_{kl}\in[0,2\pi]$, whereas under the practical constraint we have a restriction on the possible values of the MIS phase-shifts, i.e., $\angle-(1+\mathrm{j}\chi)^{-1}\in\left[\text{-}\frac{\pi}{2},\frac{\pi}{2}\right]$. Moreover, the magnitude of each phase-shift under this

constraint is always less than 1 for any $\chi \neq 0$. Practically, this introduces the phase-dependent amplitude attenuation in the incident wave. The projector function under this new constraint is defined as:

$$g_{2,kl}\left(\widetilde{r}_{kl},\widetilde{\gamma}\right) \triangleq \arg\min_{\upsilon_{kl}} \left[\widetilde{\gamma}|\upsilon_{kl}-\widetilde{r}_{kl}|^2+\gamma_{\mathrm{p}}\left|\upsilon_{kl}+\frac{1}{1+\mathrm{j}\chi_{kl}^{\mathrm{opt}}}\right|^2\right], \tag{3.37}$$

where

$$\chi_{kl}^{\mathrm{opt}} = g_3\left(\widetilde{r}_{kl}\right) \triangleq \arg\min_{\chi_{kl}}\left|\widetilde{r}_{kl}+\frac{1}{1+\mathrm{j}\chi_{kl}}\right|^2. \tag{3.38}$$

The solution to the optimization problem in (3.38), the projector function, and its derivative are given by:

$$g_3\left(\widetilde{r}_{kl}\right) = \frac{1}{2\Im\left\{\widetilde{r}_{kl}\right\}}\left(1+2\Re\left\{\widetilde{r}_{kl}\right\}+\sqrt{\left(1+2\Re\left\{\widetilde{r}_{kl}\right\}\right)^2+4\Im\left\{\widetilde{r}_{kl}\right\}^2}\right), \tag{3.39}$$

$$g_{2,kl}\left(\widetilde{r}_{kl}\right) = -\left(1+\mathrm{j}g_3\left(\widetilde{r}_{kl}\right)\right)^{-1}, \tag{3.40}$$

$$g'_{2,kl}\left(\widetilde{r}_{kl}\right) = \left|\mathrm{j}g'_3\left(\widetilde{r}_{kl}\right)\left(1+\mathrm{j}g_3\left(\widetilde{r}_{kl}\right)\right)^{-2}\right|. \tag{3.41}$$

The matrix valued projector functions, $\mathbf{G}_1(\widetilde{\mathbf{R}}_t,\widetilde{\gamma}_t)$ and $\mathbf{G}_2(\widetilde{\mathbf{R}}_t,\widetilde{\gamma}_t)$, and their derivatives, $\mathbf{G}'_1(\widetilde{\mathbf{R}}_t,\widetilde{\gamma}_t)$ and $\mathbf{G}'_2(\widetilde{\mathbf{R}}_t,\widetilde{\gamma}_t)$, are obtained according to (3.20) and (3.21). Lastly, the constant transmitted vector by the BS, $\mathbf{v}_{\mathrm{b}}$, is set to the right singular vector of the matrix $\mathbf{H}_{\text{b-s}}$ that corresponds to the largest eigenvalue.

3.3.2 Optimal Precoding and Scaling Factors

The receive scaling factor, α_{s}, is decoupled from the other optimization variables, i.e., $\mathbf{F}$ and α_{b}, in the objective function (3.23a). Therefore, it can be optimized independently of the other two variables as follows:

$$\begin{aligned}\arg\min_{\alpha_{\mathrm{s}}} \quad & \left\|\alpha_{\mathrm{s}}\sqrt{P_{\mathrm{s}}}\left[\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_1)\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}},\cdots,\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_L)\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}}\right]-\left[\mathbf{0}_{L\times B}\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}\right\|_{\mathrm{F}}^2 \\ & + LR\sigma_{\mathrm{w}}^2\alpha_{\mathrm{s}}^2.\end{aligned} \tag{3.42}$$

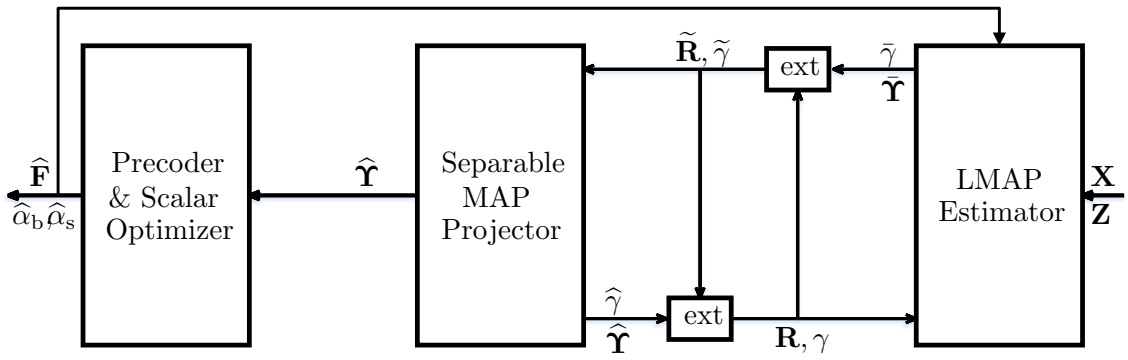

Figure 3.2: Block diagram of the proposed algorithm. The calculation of extrinsic information is performed by the "ext" blocks.

By defining the matrix:

$$\mathbf{C} \triangleq \sqrt{P_{\mathrm{s}}}\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}})\boldsymbol{\Upsilon}, \tag{3.43}$$

we rewrite (3.42) as:

$$\underset{\alpha_{\mathrm{s}}}{\arg\min} \quad \left\|\alpha_{\mathrm{s}}\mathbf{C} - \left[\mathbf{0}_{L\times B}\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}\right\|_{\mathrm{F}}^{2} + LR\sigma_{\mathrm{w}}^{2}\alpha_{\mathrm{s}}^{2}. \tag{3.44}$$

From (3.44), we establish the closed-form solution to the optimization problem in (3.42) as follows:

$$\alpha_{\mathrm{s}}^{\mathrm{opt}} = g_4(\mathbf{C}) \ \triangleq \ \frac{\mathrm{Tr}\left(\mathbf{C}^{\mathsf{H}}\left[\mathbf{0}_{L\times B}\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}} + \left[\mathbf{0}_{L\times B}\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}}\right]\mathbf{C}\right)}{2\left(\|\mathbf{C}\|_{\mathrm{F}}^{2} + LR\sigma_{\mathrm{w}}^{2}\right)}. \tag{3.45}$$

Now, the optimal precoding matrix, $\mathbf{F}$, and receive scaling factor, α_{b}, are obtained as a solution to the following optimization problem:

$$\begin{aligned}\underset{\alpha_{\mathrm{b}},\mathbf{F}}{\arg\min} \quad & \Big\| \alpha_{\mathrm{b}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_1) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_L) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \Big] \\ & - \Big[[\mathbf{I}_{B,1},\ \mathbf{0}_{B\times R,1}]^{\mathsf{T}}, \cdots, [\mathbf{I}_{B,L},\ \mathbf{0}_{B\times R,L}]^{\mathsf{T}} \Big] \Big\|_{\mathrm{F}}^2 + LB\sigma_{\mathrm{w}}^2 \alpha_{\mathrm{b}}^2. \end{aligned} \tag{3.46a}$$

$$\text{s.t.} \quad \|\mathbf{F}\|_{\mathrm{F}}^2 = P_{\mathrm{b}}. \tag{3.46b}$$

By defining the matrices:

$$\mathbf{K} \triangleq \sum_{l=1}^{L} \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_l) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \right)^{\mathsf{H}} \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_l) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \right) \tag{3.47}$$

$$\mathbf{E} \triangleq [\mathbf{I}_B\ \mathbf{0}_{B\times R}] \sum_{l=1}^{L} \left(\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_l) \mathbf{H}_{\text{b-s}} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \right), \tag{3.48}$$

the optimization problem in (3.46) becomes a constrained MMSE transmit precoding optimization for MIMO systems. The problem can be solved jointly by Lagrange optimization. We construct the Lagrangian function for the optimization problem in (3.46) as follows:

$$\mathcal{L}(\mathbf{F}, \alpha_{\mathrm{b}}, \lambda) = \mathrm{Tr}\left(\alpha_{\mathrm{b}}^2 \mathbf{K}\mathbf{F}\mathbf{F}^{\mathsf{H}} - \alpha_{\mathrm{b}} \mathbf{E}\mathbf{F} - \alpha_{\mathrm{b}} \mathbf{F}^{\mathsf{H}} \mathbf{E}^{\mathsf{H}} \right) + LB + LB\alpha_{\mathrm{b}}^2 \sigma_{\mathrm{w}}^2 + \lambda \left(\mathrm{Tr}\left(\mathbf{F}\mathbf{F}^{\mathsf{H}} \right) - P_{\mathrm{b}} \right), \tag{3.49}$$

with $\lambda \in \mathbb{R}$ being the Lagrange multiplier. The closed-form solutions for optimal α_{b} and $\mathbf{F}$ are given below and we refer the reader to [30] for more details:

$$\alpha_{\mathrm{b}}^{\mathrm{opt}} = g_5(\mathbf{K},\ \mathbf{E}) \triangleq \sqrt{\frac{1}{P_{\mathrm{b}}}} \sqrt{\mathrm{Tr}\left(\left[\mathbf{K} + \frac{LB\sigma_{\mathrm{w}}^2 \mathbf{I}_N}{P_{\mathrm{b}}} \right]^{-2} \mathbf{E}^{\mathsf{H}} \mathbf{E} \right)}, \tag{3.50}$$

$$\mathbf{F}^{\mathrm{opt}} = g_6(\mathbf{K},\ \mathbf{E}) \triangleq \frac{\sqrt{P_{\mathrm{b}}} \left[\mathbf{K} + \frac{LB\sigma_{\mathrm{w}}^2 \mathbf{I}_N}{P_{\mathrm{b}}} \right]^{-1} \mathbf{E}^{\mathsf{H}}}{\sqrt{\mathrm{Tr}\left(\left[\mathbf{K} + \frac{LB\sigma_{\mathrm{w}}^2 \mathbf{I}_N}{P_{\mathrm{b}}} \right]^{-2} \mathbf{E}^{\mathsf{H}} \mathbf{E} \right)}}. \tag{3.51}$$

We have found the solution to both optimization problems in (3.22) and (3.23). Therefore, we can combine their solutions together into one algorithm. We define the MSE at iteration t as follows:

$$E_t \triangleq f\big(\widehat{\alpha}_{\mathrm{s},t},\ \widehat{\alpha}_{\mathrm{b},t},\ \widehat{\mathbf{F}}_t,\ \widehat{\boldsymbol{\Upsilon}}_t\big). \tag{3.52}$$

The algorithm stops when $|E_t - E_{t-1}| < \epsilon E_{t-1}$, where $\epsilon \in \mathbb{R}_+$ is some precision tolerance. The overall block diagram and the algorithmic steps are shown, respectively, in Fig. 3.2 and **Algorithm 6**. The convergence and complexity of the OOVAMP-based approach are discussed in Chapter 2. Because of the monotone convergence theorem in real analysis [49], **Algorithm 6** is guaranteed to converge since the MSE is minimized in every step and the objective function, $f(\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}, \boldsymbol{\Upsilon})$, is lower bounded by zero. The algorithm can be efficiently implemented by exploiting matrix structures, and by using the singular value decomposition (SVD) form of OOVAMP so that the computational complexity is of $\mathcal{O}(MNL(K+N))$.

Algorithm 6 OOVAMP-based joint optimization algorithm

Given $\mathbf{H}_{\text{s-u}}$, $\mathbf{H}_{\text{b-u}}$, $\mathbf{H}_{\text{b-s}}$, $\mathbf{S}_{\text{s}}$, a precision tolerance (ϵ), and a maximum number of iterations (T_{MAX})

Initialize $\widehat{\boldsymbol{\Upsilon}}_0,\ \mathbf{R}_0$, $\gamma_0 \geq 0$ and $t \leftarrow 1$, and obtain $\mathbf{v}_{\text{b}}$ from $\mathbf{H}_{\text{b-s}}$

Compute $\widehat{\mathbf{C}}_0,\ \widehat{\mathbf{K}}_0$ and $\widehat{\mathbf{E}}_0$ by substituting $\widehat{\boldsymbol{\Upsilon}}_0$ into (3.43), (3.47) and (3.48).

$\widehat{\alpha}_{\text{s},0} = g_4\left(\widehat{\mathbf{C}}_0\right)$

$\widehat{\alpha}_{\text{b},0} = g_5\left(\widehat{\mathbf{K}}_0,\ \widehat{\mathbf{E}}_0\right)$

$\widehat{\mathbf{F}}_0 = g_6\left(\widehat{\mathbf{K}}_0,\ \widehat{\mathbf{E}}_0\right)$

repeat

  // LMAP Estimator

  Compute matrices $\mathbf{A},\ \mathbf{B},\ \mathbf{D},\ \mathbf{M},\ \mathbf{X}$ and $\mathbf{Z}$ by substituting $\widehat{\alpha}_{\text{s},t-1},\ \widehat{\alpha}_{\text{b},t-1}$ and $\widehat{\mathbf{F}}_{t-1}$ into (3.25) to (3.30).

  $\bar{\boldsymbol{\Upsilon}}_t = \left(\mathbf{D}^{\mathsf{H}}\mathbf{D} + \mathbf{M}^{\mathsf{H}}\mathbf{M} + \gamma_{t-1}\mathbf{I}_K\right)^{-1}\left(\mathbf{D}^{\mathsf{H}}\mathbf{X} + \mathbf{M}^{\mathsf{H}}\mathbf{Z} + \gamma_{t-1}\mathbf{R}_{t-1}\right)$

  $\bar{\gamma}_t = K\text{Tr}\left(\left[\mathbf{D}^{\mathsf{H}}\mathbf{D} + \mathbf{M}^{\mathsf{H}}\mathbf{M} + \gamma_{t-1}\mathbf{I}_K\right]^{-1}\right)^{-1}$

  $\widetilde{\gamma}_t = \bar{\gamma}_t - \gamma_{t-1}$

  $\widetilde{\mathbf{R}}_t = \widetilde{\gamma}_t^{-1}\left(\bar{\boldsymbol{\Upsilon}}_t\bar{\gamma}_t - \mathbf{R}_{t-1}\gamma_{t-1}\right)$

  // Separable MAP Projector

  $\widehat{\boldsymbol{\Upsilon}}_t = \mathbf{G}_1\left(\widetilde{\mathbf{R}}_t\right)$

  $\widehat{\gamma}_t = \widetilde{\gamma}_t^{-1}\left\langle \mathbf{G}_1'\left(\widetilde{\mathbf{R}}_t\right)\right\rangle$.

  $\gamma_t = \widehat{\gamma}_t - \widetilde{\gamma}_t$

  $\mathbf{R}_t = \gamma_t^{-1}\left(\widehat{\gamma}_t\widehat{\boldsymbol{\Upsilon}}_t - \widetilde{\gamma}_t\widetilde{\mathbf{R}}_t\right)$

  //Find $\alpha_{\text{s}},\ \alpha_{\text{b}}$ and $\mathbf{F}$ through their closed-form solutions.

  Compute $\widehat{\mathbf{C}}_t,\ \widehat{\mathbf{K}}_t$ and $\widehat{\mathbf{E}}_t$ by substituting $\widehat{\boldsymbol{\Upsilon}}_{\text{t}}$ into (3.43), (3.47) and (3.48).

  $\widehat{\alpha}_{\text{s},t} = g_4\left(\widehat{\mathbf{C}}_t\right)$

  $\widehat{\alpha}_{\text{b},t} = g_5\left(\widehat{\mathbf{K}}_t,\ \widehat{\mathbf{E}}_t\right)$

  $\widehat{\mathbf{F}}_{\text{t}} = g_6\left(\widehat{\mathbf{K}}_t,\ \widehat{\mathbf{E}}_t\right)$

  $t \leftarrow t + 1$

until $|E_t - E_{t-1}| < \epsilon E_{t-1}$ or $t > T_{\text{MAX}}$

return $\widehat{\boldsymbol{\upsilon}}_t,\ \widehat{\mathbf{F}}_t,\ \widehat{\alpha}_t$.

3.4 Numerical Results and Performance Analysis

3.4.1 Simulation Model and Parameters

We present Monte-Carlo simulation results to assess the performance of the proposed algorithm. We use the channel models introduced in 2.5 besides considering the same setup for the location of users, the MIS and the BS. Assuming a uniform linear array with N antennas at the BS and a square uniform planar array with K antenna elements at the MIS, the channel between the MIS and the BS is generated according to:

$$\mathbf{H}_{\text{b-s}} \;=\; \sqrt{L(d_{\text{MIS}})}\sum_{q=1}^{Q_{\text{MIS}}} \mathsf{c}_q \mathbf{a}_{\text{MIS}}(\varphi_q,\psi_q)\mathbf{a}_{\text{BS}}(\phi_q)^{\mathsf{T}}. \tag{3.53}$$

The channel vectors for the link between each single antenna m-th user and the MIS, and each m-th user and the BS are modeled, respectively, as follows:

$$\mathbf{h}_{\text{s-u},m} \;=\; \sqrt{L(d'_m)}\sum_{q=1}^{Q_{\text{s-u}}} \mathsf{c}_{m,q}\mathbf{a}_{\text{MIS}}(\varphi_{m,q},\psi_{m,q}), \qquad m=1,\cdots,M, \tag{3.54}$$

$$\mathbf{h}_{\text{b-u},m} \;=\; \sqrt{L(d_m)}\sum_{q=1}^{Q_{\text{b-u}}} \mathsf{c}_{m,q}\mathbf{a}_{\text{BS}}(\phi_{m,q}), \qquad m=1,\cdots,M. \tag{3.55}$$

To account for channel correlation effects, we set the number of multi-path components lower than the number of BS antennas and the MIS antenna elements for the BS-MIS channel, and set the number of BS-served users lower than the rank of the BS-MIS channel matrix $\mathbf{H}_{\text{b-s}}$. Moreover, we split the total transmit power, P, between the BS-served and MIS-served users according to the share of each type of users, i.e., $P_{\text{b}} = \frac{B}{M}P$ and $P_{\text{s}} = \frac{R}{M}P$. The results are averaged over 1000 independent Monte Carlo trials.

We use the *sum-rate*, $\widehat{C}$, for performance evaluation which is defined as follows:

$$\widehat{C} \;=\; \sum_{m=1}^{M}\log_2\left(\frac{1}{\text{MMSE}_m}\right), \tag{3.56}$$

where MMSE_m refers to the MSE of each m-th user's received symbol. Since, the approach of using the MIS as a modulating surface is novel, we benchmark the proposed

solution against traditional MIS-assisted systems where it is merely used for beamforming purposes. The proposed OOVAMP-based approach is compared against the following two schemes:

i. SCHEME 1: a multi-user MIMO system assisted by one IRS where the joint optimization of the phase matrix is solved through alternate optimization and OOVAMP along with MMSE precoding.

ii. SCHEME 2: a multi-user MIMO system assisted by one IRS where the SDR technique is used to optimize the IRS reflection coefficients for beamforming in combination with MMSE precoding.

3.4.2 Performance Results With Perfect CSI

We consider a typical urban or suburban environment where the BS is located faraway from the users and has no LOS to them. However, the MIS is installed at a location where a LOS component is present in the BS-MIS link but not in the user-MIS link. We also set the number of BS antennas to $N = 32$ and the number of MIS antennas to $K = 256$. Fig. 3.3, depicts the achievable sum-rate versus the transmit power, P, for the different considered transmission schemes. The total number of users is equal to $M = 8$. Here we consider two configurations for the proposed approach: i) when all the users are served by just the MIS, and ii) a hybrid case where the MIS and the BS both serve 4 users each. The configuration in which all the users are served by the MIS significantly outperforms the scheme in which an IRS is only used for beamforming (i.e., SCHEME 1). This is because the MIS-served users do not suffer from the path-loss between the BS and the MIS. Moreover, at low transmit power (e.g., $P = 20$ dBm), the sum-rate for the users solely served by the MIS is, respectively, two and four times the sum-rate of the BS-served users for the beamforming-only IRS-assisted OOVAMP-based and SDR-based approaches. For the hybrid configuration, the resulting sum-rate

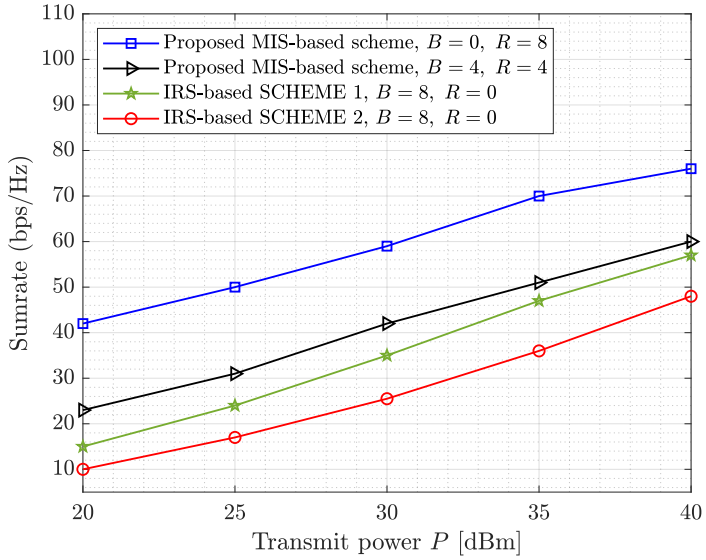

Figure 3.3: Sum-rate versus transmit power with $M = 8,\ N = 32$, and $K = 256$.

edges SCHEME 1 but, it is lower than the case when all users are served by the MIS. This confirms that it is more beneficial to serve the users by the MIS.

Fig. 3.4 shows a plot of the sum-rate against the share of MIS-served users among the total number of users. It is observed that the combined sum-rate first decreases and then monotonically increases with the number of MIS-served users. This is because of the presence of cross-user interference among the MIS-served and the BS-served users since the MIS is performing both tasks, i.e., beamforming to assist the BS and also data embedding to serve another set of users. This implies that there is more loss than gain when the ratio of the MIS-served users to the total number of users becomes small.

Fig. 3.5 illustrates the sum-rate versus the total number of users being served. The users are solely served by the BS for one plot and by the MIS for the other. Here we show the benefits of the approach of using the MIS as a modulating surface to directly serve users. Although the sum-rate of the BS-served users is higher when the number of users becomes small, the sum-rate of the MIS-served users keeps increasing with the

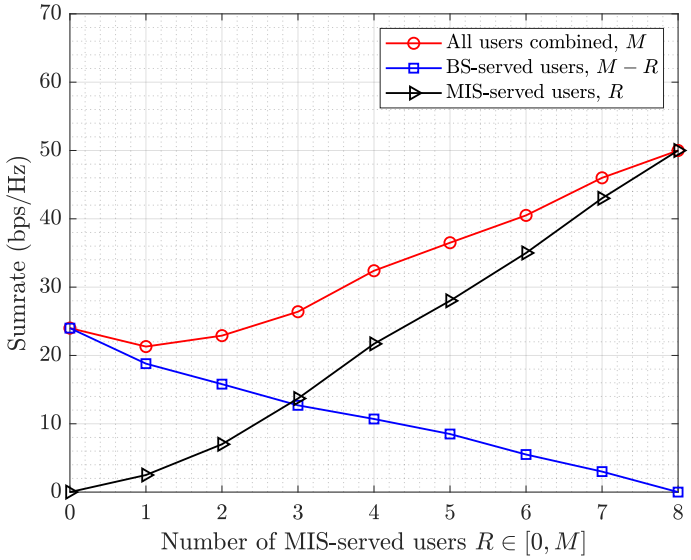

Figure 3.4: Sum-rate versus the number of MIS-served users with $M = 8,\ N = 32$, and $P = 25$ dBm.

number of users while the sum-rate of the BS-served users only increases up to the number of available channel paths which is set to 10. This is because the number of users that can be served by the MIS is independent of the number of BS antennas and the correlation in the MIS-BS channel. The upper limit for the number of MIS-served users is equal to the number of MIS antenna elements K.

3.4.3 Performance Results With Imperfect CSI

In this section, we asses the performance of the proposed scheme in the presence of channel estimation errors. Specifically, we consider a scenario where pilot training followed by MMSE estimation algorithms are used to estimate the cascaded BS-MIS-users and the direct BS-users channels [44, 45]. We model the estimated channel matrix

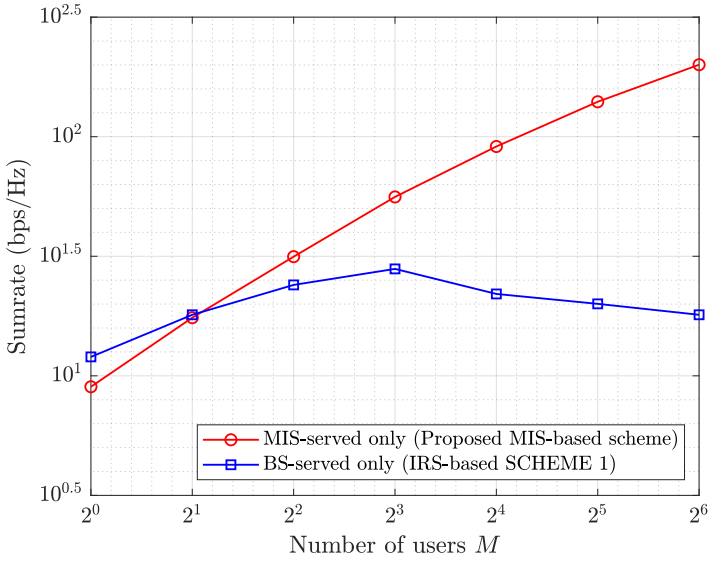

Figure 3.5: Sum-rate versus the number of users with $N = 32,\ K = 256$ and $P = 30$ dBm.

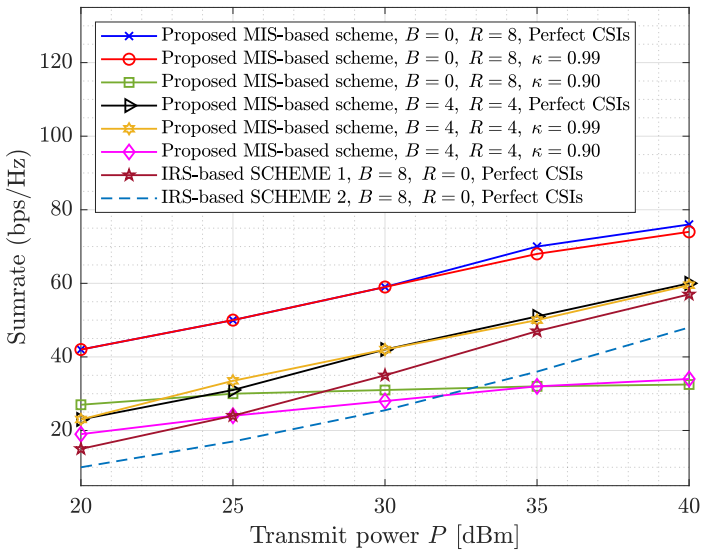

Figure 3.6: Sum-rate versus transmit power under imperfect CSI with $M = 8,\ N = 32$, and $K = 256$.

and vectors using the statistical CSI error model proposed in [46–48] as follows:

$$\widehat{\mathbf{H}}_{\text{b-s}} = \kappa \mathbf{H}_{\text{b-s}} + \sqrt{(1-\kappa^2)L(d_{\text{MIS}})}\boldsymbol{\Delta}_{\text{b-s}}, \tag{3.57}$$

$$\widehat{\mathbf{h}}_{\text{b-u},m} = \kappa \mathbf{h}_{\text{b-u},m} + \sqrt{(1-\kappa^2)L(d_m)}\boldsymbol{\delta}_{\text{b-u},m}, \quad m = 1, \cdots, M, \tag{3.58}$$

$$\widehat{\mathbf{h}}_{\text{s-u},m} = \kappa \mathbf{h}_{\text{s-u},m} + \sqrt{(1-\kappa^2)L(d'_m)}\boldsymbol{\delta}_{\text{s-u},m}, \quad m = 1, \cdots, M, \tag{3.59}$$

where $\kappa \in [0, 1]$ denotes the channel estimation accuracy and $\boldsymbol{\Delta}_{\text{b-s}}$, $\boldsymbol{\delta}_{\text{b-u},m}$ and $\boldsymbol{\delta}_{\text{s-u},m}$ follow the circularly symmetric complex Gaussian (CSCG) distribution, i.e., $\text{vec}(\boldsymbol{\Delta}_{\text{b-s}}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{1}_{N\times N} \otimes \mathbf{I}_K)$, $\boldsymbol{\delta}_{\text{b-u},m} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ and $\boldsymbol{\delta}_{\text{s-u},m} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$. We first optimize the variables α_{s}, α_{b}, $\mathbf{F}$ and $\boldsymbol{\Upsilon}$ under imperfect CSI and then use the exact CSI matrices to calculate the sum-rate.

Fig. 3.6 plots the sum-rate versus transmit power for different values of the channel estimation accuracy parameter κ. We also include plots for the other schemes (i.e., SCHEMES 1 AND 2) which use an IRS as a purely reflective surface) under perfect CSI for reference. The results show the resilience of the proposed MIS-based approach against small channel estimation errors. At low SNR, it is observed that the proposed design with a low channel estimation accuracy of $\kappa = 0.90$ performs better than the baseline schemes even under perfect CSIs. Moreover, the performance loss with a high channel estimation accuracy value of $\kappa = 0.99$ is negligible.

3.5 Summary

We have presented a novel approach of employing the MIS for performing passive beamforming and data embedding for the BS-served and the MIS-served users, respectively, in a multi-user MIS-assisted downlink MIMO system. The associated joint convex optimization problem has been formulated under the sum MMSE criterion in order to maximize the users' spectral efficiency. Alternating minimization has been used to split

the original optimization problem into two tasks, i.e., separately optimizing the MIS phase shifts and jointly optimizing the BS precoding and the receive scaling factors for the BS- and MIS-served users. The optimal solution to the joint optimization problem for precoding is found in closed form. We have optimized the MIS phase shifts using OOVAMP by deriving the problem-specific modules of the OOVAMP algorithm. Moreover, the original joint problem has been solved under both the ideal and a practical constraint on the MIS phase shifts. Simulation results illustrate highly superior system throughput performance of the proposed MIS-based scheme over two baseline schemes in which an IRS is used for beamforming purposes only. Moreover, the proposed approach can support more number of users simultaneously than existing beamforming approaches wherein the users are served by the BS only. Finally, the results under imperfect CSI confirm that the performance remains nearly unchanged in the presence of small channel estimation errors.

Chapter 4

Conclusion and Future Directions

4.1 Concluding Remarks

IRS/MIS-assisted wireless networks provide an effective solution for many next generation (e.g., 6G) cellular wireless applications, especially when the link between the BS and the end user is not favorable. We considered a multi-user IRS/MIS-assisted wireless cellular system, and provided a general framework to utilize the surface for passive beamforming as well as modulation. In Chapter 2, we investigated the problem of joint active and passive beamforming design for an IRS-assisted downlink multi-user MIMO system under both ideal and practical models for the IRS phase shifts and provided an alternating optimization OOVAMP-based solution. Chapter 3 extended the work done in Chapter 2 by introducing MIS that can simultaneously perform passive beamforming for and data modulation for the BS- and MS-served users, respectively. From the numerical results we conclude that *i*) IRS-based beamforming approach performs better than existing IRS-based beamforming schemes in terms of throughput and computational complexity, *ii*) MIS-based schemes greatly outperforms traditional IRS-based schemes in which IRS is only used for beamforming and, *iii*) MIS-based schemes can support

much more number of users than IRS-based beamforming schemes.

4.2 Future Directions

This work can be extended in a few directions to provide a comprehensive framework for multi-user IRS/MIS-assisted cellular wireless systems.

4.2.1 Digital Intelligent Surface

So far, the IRS/MIS does not have the capability to perform any kind of digital signal processing on the incident signals. Although the MIS is able to modulate user data on the carrier signal, it does so by only changing the phase of the reflected carrier signals. Therefore, MIS can only be utilized in the downlink. The problem of designing an intelligent surface that can also receive user data in the uplink besides doing passive beamforming and data modulation in the downlink must be investigated to fully utilize the potential of intelligent surfaces in cellular systems.

4.2.2 IRS-Based Passive Beamforming Under LOS MIMO

The performance of IRS-based beamforming scheme can be studied under LOS MIMO links between the BS and the IRS. LOS MIMO links are possible when the IRS is deployed close to the BS which thereby makes the BS-IRS channel matrix full rank. The maximum number of users that can be supported in such scenario is equal to the number of BS antennas instead of the number of channel paths in the BS-IRS channel.

4.2.3 More Physically Consistent Phase Shift Models and Establishing Optimality

Since the proposed solution provides flexibility in terms of choosing the constraint on the IRS/MIS phase shifts, it opens up the possibility of solving the joint optimization problem using more physically-consistent models for the IRS elements. The performance of the OOVAMP-based approach can be theoretically predicted to establish optimality through the statistical state evolution framework [28, 34].

Appendix A

We solve the following optimization problem:

$$\arg\min_{\chi} \; f(\chi), \text{ where} \tag{A.1}$$

$$f(\chi) \triangleq \left| \widetilde{r} + \frac{1}{1+\mathrm{j}\chi} \right|^2, \tag{A.2}$$

in which $\chi \in \mathbb{R}$ and $\widetilde{r} \in \mathbb{C}$. Expanding the objective function, we re-express it as follows:

$$\arg\min_{\chi} \; \widetilde{r}^*\widetilde{r} + \frac{\widetilde{r}^*}{1+\mathrm{j}\chi} + \frac{\widetilde{r}}{1-\mathrm{j}\chi} + \frac{1}{(1-\mathrm{j}\chi)(1+\mathrm{j}\chi)}. \tag{A.3}$$

Let $a \triangleq \Re\{\widetilde{r}\}$ and $b \triangleq \Im\{\widetilde{r}\}$. We substitute a and b into (A.3) and simplify the objective function as follows:

$$\arg\min_{\chi} \; a^2 + b^2 + \frac{1+2a}{1+\chi^2} - \frac{2b\chi}{1+\chi^2}. \tag{A.4}$$

By defining $c \triangleq (1+2a)$, we take the derivative w.r.t. χ and set it to zero to obtain:

$$f'(\chi) \;=\; -\frac{2b(1-\chi^2)}{(1+\chi^2)^2} - \frac{2c\chi}{(1+\chi^2)^2} \;=\; 0. \tag{A.5}$$

Simplifying (A.5) leads to:

$$b\chi^2 - c\chi - b \;=\; 0. \tag{A.6}$$

The roots of the quadratic equation in (A.6) are real and distinct and are given by:

$$\chi_1 \;=\; \frac{c + \sqrt{c^2+4b^2}}{2b}, \tag{A.7}$$

and

$$\chi_2 = \frac{c - \sqrt{c^2 + 4b^2}}{2b}, \tag{A.8}$$

where $b \neq 0$. By taking the second derivative of the objective function in (A.4) w.r.t. χ and resorting to some straightforward algebraic manipulations, we also obtain:

$$f''(\chi) \;=\; \frac{2}{(1+\chi^2)^3}(6b\chi - 2b\chi^3 + 3c\chi^2 - c). \tag{A.9}$$

Substituting $\chi = \chi_1$ in (A.9) and simplifying the result yields:

$$\begin{aligned} f''(\chi_1) = \frac{1}{(1+\chi_1^2)^3}\Bigg(\frac{1}{b^2}\left(c^3 + c^2\sqrt{c^2+4b^2}\right) \\ + 4\left(c + \sqrt{c^2+4b^2}\right)\Bigg). \end{aligned} \tag{A.10}$$

Since $b \neq 0$, we have $c^2\sqrt{c^2+4b^2} > |c^3|$ and $\sqrt{c^2+4b^2} > |c|$ which implies that $f''(\chi_1) > 0$. Similarly, we have:

$$\begin{aligned} f''(\chi_2) = \frac{1}{(1+\chi_2^2)^3}\Bigg(\frac{1}{b^2}\left(c^3 - c^2\sqrt{c^2+4b^2}\right) \\ + 4\left(c - \sqrt{c^2+4b^2}\right)\Bigg) < 0, \quad b \neq 0. \end{aligned} \tag{A.11}$$

Thus, we choose:

$$\chi^{\text{opt}} \;=\; \chi_1 \;=\; \frac{1 + 2a + \sqrt{(1+2a)^2 + 4b^2}}{2b}. \tag{A.12}$$

Interestingly, the solution χ_1 results in the same sign for both $\Im\{-(1+\mathrm{j}\chi_1)^{-1}\}$ and $\Im\{\widetilde{r}\}$.

Appendix B

We have the following function:

$$f(\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}, \boldsymbol{\Upsilon}) = \mathbb{E}_{\mathbf{Y},\mathbf{S}} \left\{ \|\mathbf{Y} - \mathbf{S}\|_{\mathrm{F}}^2 \right\}, \tag{B.1}$$

where

$$\begin{aligned}\mathbf{Y} = {} & \alpha_{\mathrm{b}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{F} \mathbf{s}_{\mathrm{b},1} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \mathbf{s}_{\mathrm{b},1}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{F} \mathbf{s}_{\mathrm{b},L} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \mathbf{s}_{\mathrm{b},L} \Big] \\ & + \alpha_{\mathrm{s}} \sqrt{P_{\mathrm{s}}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}} \Big] \\ & + \Big[\alpha_{\mathrm{b}} \mathbf{W}_{\mathrm{b}}^{\mathsf{T}},\ \alpha_{\mathrm{s}} \mathbf{W}_{\mathrm{s}}^{\mathsf{T}} \Big]^{\mathsf{T}},\end{aligned} \tag{B.2}$$

$$\mathbf{S} = \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{T}},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}} \right]^{\mathsf{T}}. \tag{B.3}$$

We explicitly write the function in (B.1) and then expand it as follows:

$$f(\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}, \boldsymbol{\Upsilon}) = \mathbb{E}_{\mathbf{Y},\mathbf{S}} \left\{ \mathrm{Tr} \left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y} - \mathbf{Y}^{\mathsf{H}}\mathbf{S} - \mathbf{S}^{\mathsf{H}}\mathbf{Y} + \mathbf{S}^{\mathsf{H}}\mathbf{S} \right) \right\} \tag{B.4}$$

$$\begin{aligned}f(\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}, \boldsymbol{\Upsilon}) = {} & \mathbb{E}_{\mathbf{Y},\mathbf{S}_{\mathrm{b}}} \Big\{ \mathrm{Tr} \Big(\mathbf{Y}^{\mathsf{H}}\mathbf{Y} - \mathbf{Y}^{\mathsf{H}} \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{T}},\ \mathbf{0}_{L\times R} \right]^{\mathsf{T}} - \mathbf{Y}^{\mathsf{H}} \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}} \right]^{\mathsf{T}} - \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R} \right] \mathbf{Y} \\ & - \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}} \right] \mathbf{Y} + \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R} \right] \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{T}},\ \mathbf{0}_{L\times R} \right]^{\mathsf{T}} + \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R} \right] \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}} \right]^{\mathsf{T}} \\ & + \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}} \right] \left[\mathbf{S}_{\mathrm{b}}^{\mathsf{T}},\ \mathbf{0}_{L\times R} \right]^{\mathsf{T}} + \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}} \right] \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}} \right]^{\mathsf{T}} \Big) \Big\}\end{aligned} \tag{B.5}$$

By defining the matrices:

$$\mathbf{Y}_{\mathrm{e}} = \alpha_{\mathrm{b}}\Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_1)\mathbf{H}_{\text{b-s}}\mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_L)\mathbf{H}_{\text{b-s}}\mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F}\Big], \tag{B.6}$$

$$\mathbf{Y}_{\mathrm{s}} = \alpha_{\mathrm{s}}\sqrt{P_{\mathrm{s}}}\Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_1)\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_L)\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}}\Big], \tag{B.7}$$

$$\mathbf{W} = \Big[\alpha_{\mathrm{b}}\mathbf{W}_{\mathrm{b}}^{\mathsf{T}},\ \alpha_{\mathrm{s}}\mathbf{W}_{\mathrm{s}}^{\mathsf{T}}\Big]^{\mathsf{T}}, \tag{B.8}$$

$$\mathbf{Y}_{\mathrm{b}} = \alpha_{\mathrm{b}}\Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_1)\mathbf{H}_{\text{b-s}}\mathbf{F}\mathbf{s}_{\mathrm{b},1} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F}\mathbf{s}_{\mathrm{b},1}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\upsilon}_L)\mathbf{H}_{\text{b-s}}\mathbf{F}\mathbf{s}_{\mathrm{b},L} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}}\mathbf{F}\mathbf{s}_{\mathrm{b},L}\Big], \tag{B.9}$$

we expand the terms in (B.5) and take expectation w.r.t. the random matrices $\mathbf{S}_{\mathrm{b}}$ and $\mathbf{W}$ thereby leading to:

$$\begin{aligned}\mathbb{E}_{\mathbf{Y},\mathbf{S}_{\mathrm{b}}}\left\{\mathrm{Tr}\big(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\big)\right\} &= \mathbb{E}_{\mathbf{W},\mathbf{S}_{\mathrm{b}}}\Big\{\mathrm{Tr}\big(\mathbf{Y}_{\mathrm{b}}^{\mathsf{H}}\mathbf{Y}_{\mathrm{b}} + \mathbf{Y}_{\mathrm{b}}^{\mathsf{H}}\mathbf{Y}_{\mathrm{s}} + \mathbf{Y}_{\mathrm{b}}^{\mathsf{H}}\mathbf{W} + \mathbf{Y}_{\mathrm{s}}^{\mathsf{H}}\mathbf{Y}_{\mathrm{b}} \\ &\quad + \mathbf{Y}_{\mathrm{s}}^{\mathsf{H}}\mathbf{Y}_{\mathrm{s}} + \mathbf{Y}_{\mathrm{s}}^{\mathsf{H}}\mathbf{W} + \mathbf{W}^{\mathsf{H}}\mathbf{Y}_{\mathrm{b}} + \mathbf{W}^{\mathsf{H}}\mathbf{Y}_{\mathrm{s}} + \mathbf{W}^{\mathsf{H}}\mathbf{W}\big)\Big\} \\ &= \mathrm{Tr}\big(\mathbf{Y}_{\mathrm{e}}^{\mathsf{H}}\mathbf{Y}_{\mathrm{e}} + \mathbf{Y}_{\mathrm{s}}^{\mathsf{H}}\mathbf{Y}_{\mathrm{s}}\big) + LB\sigma_{\mathrm{w}}^2\alpha_{\mathrm{b}}^2 + LR\sigma_{\mathrm{w}}^2\alpha_{\mathrm{s}}^2,\end{aligned} \tag{B.10}$$

$$\begin{aligned}\mathbb{E}_{\mathbf{Y},\mathbf{S}_{\mathrm{b}}}\Big\{\mathrm{Tr}\Big(\big[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R}\big]\mathbf{Y}\Big)\Big\} &= \mathbb{E}_{\mathbf{W},\mathbf{S}_{\mathrm{b}}}\Big\{\mathrm{Tr}\Big(\big[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R}\big]\mathbf{Y}_{\mathrm{b}} + \big[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R}\big]\mathbf{Y}_{\mathrm{s}} \\ &\quad + \big[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R}\big]\mathbf{W}\Big)\Big\} \\ &= \mathrm{Tr}\Big(\Big[\big[\mathbf{I}_{B,1},\ \mathbf{0}_{B\times R,1}\big]^{\mathsf{T}}, \cdots, \big[\mathbf{I}_{B,L},\ \mathbf{0}_{B\times R,L}\big]^{\mathsf{T}}\Big]^{\mathsf{H}}\mathbf{Y}_{\mathrm{e}}\Big),\end{aligned} \tag{B.11}$$

$$\begin{aligned}\mathbb{E}_{\mathbf{Y},\mathbf{S}_{\mathrm{b}}}\Big\{\mathrm{Tr}\Big(\big[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}}\big]\mathbf{Y}\Big)\Big\} &= \mathbb{E}_{\mathbf{W},\mathbf{S}_{\mathrm{b}}}\Big\{\mathrm{Tr}\Big(\big[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}}\big]\mathbf{Y}_{\mathrm{b}} + \big[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}}\big]\mathbf{Y}_{\mathrm{s}} \\ &\quad + \big[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}}\big]\mathbf{W}\Big)\Big\} \\ &= \mathrm{Tr}\Big(\big[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{H}}\big]\mathbf{Y}_{\mathrm{s}}\Big),\end{aligned} \tag{B.12}$$

and

$$\begin{aligned}\mathbb{E}_{\mathbf{S}_{\mathrm{b}}}\Big\{\mathrm{Tr}\Big(\big[\mathbf{S}_{\mathrm{b}}^{\mathsf{H}},\ \mathbf{0}_{L\times R}\big]\big[\mathbf{S}_{\mathrm{b}}^{\mathsf{T}},\ \mathbf{0}_{L\times R}\big]^{\mathsf{T}}\Big)\Big\} &= \mathrm{Tr}\Big(\Big[\big[\mathbf{I}_{B,1},\ \mathbf{0}_{B\times R,1}\big]^{\mathsf{T}}, \cdots, \big[\mathbf{I}_{B,L},\ \mathbf{0}_{B\times R,L}\big]^{\mathsf{T}}\Big]^{\mathsf{H}} \\ &\quad \times \Big[\big[\mathbf{I}_{B,1},\ \mathbf{0}_{B\times R,1}\big]^{\mathsf{T}}, \cdots, \big[\mathbf{I}_{B,L},\ \mathbf{0}_{B\times R,L}\big]^{\mathsf{T}}\Big]\Big).\end{aligned} \tag{B.13}$$

The remaining two terms in (B.5) can be computed by following (B.11) and (B.12). Finally, by substituting (B.10)−(B.13) into (B.5) and expressing the function in the form of norms we obtain:

$$
\begin{aligned}
f(\alpha_{\mathrm{s}}, \alpha_{\mathrm{b}}, \mathbf{F}, \boldsymbol{\Upsilon}) = & \Big\| \alpha_{\mathrm{b}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \Big] \\
& - \Big[[\mathbf{I}_{B,1},\ \mathbf{0}_{B \times R,1}]^{\mathsf{T}}, \cdots, [\mathbf{I}_{B,L},\ \mathbf{0}_{B \times R,L}]^{\mathsf{T}} \Big] \Big\|_{\mathrm{F}}^2 \\
& + \Big\| \alpha_{\mathrm{s}} \sqrt{P_{\mathrm{s}}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_1) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{\upsilon}_L) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}} \Big] - \big[\mathbf{0}_{L \times B}\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}} \big]^{\mathsf{T}} \Big\|_{\mathrm{F}}^2 \\
& + LB\sigma_{\mathrm{w}}^2 \alpha_{\mathrm{b}}^2 + LR\sigma_{\mathrm{w}}^2 \alpha_{\mathrm{s}}^2.
\end{aligned}
\tag{B.14}
$$

Appendix C

We have the following two norms:

$$\begin{aligned} f_1 &= \Big\| \alpha_{\mathrm{b}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_1) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_L) \mathbf{H}_{\text{b-s}} \mathbf{F} + \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \Big] \\ &- \Big[[\mathbf{I}_B,\ \mathbf{0}_{B\times R,1}]^{\mathsf{T}}, \cdots, [\mathbf{I}_B,\ \mathbf{0}_{B\times R,L}]^{\mathsf{T}} \Big] \Big\|_{\mathrm{F}}^2, \end{aligned} \tag{C.1}$$

$$f_2 = \Big\| \alpha_{\mathrm{s}} \sqrt{P_{\mathrm{s}}} \Big[\mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_1) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}}, \cdots, \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \mathrm{Diag}(\boldsymbol{v}_L) \mathbf{H}_{\text{b-s}} \mathbf{v}_{\mathrm{b}} \Big] - \big[\mathbf{0}_{L\times B}\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}} \big]^{\mathsf{T}} \Big\|_{\mathrm{F}}^2. \tag{C.2}$$

By defining the matrices, $\mathbf{A} = \alpha_{\mathrm{b}} \mathbf{H}_{\text{s-u}}^{\mathsf{H}} \in \mathbb{C}^{M\times K}$ and $\mathbf{B} = (\mathbf{H}_{\text{b-s}} \mathbf{F})^{\mathsf{T}} \in \mathbb{C}^{B\times K}$, we rewrite (C.1) as follows:

$$\begin{aligned} f_1 &= \Big\| \Big[\mathbf{A} \mathrm{Diag}(\boldsymbol{v}_1) \mathbf{B}^{\mathsf{T}}, \cdots, \mathbf{A} \mathrm{Diag}(\boldsymbol{v}_L) \mathbf{B}^{\mathsf{T}} \Big] \\ &- \Big[[\mathbf{I}_B,\ \mathbf{0}_{B\times R,1}]^{\mathsf{T}} - \alpha_{\mathrm{b}} \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F}, \cdots, [\mathbf{I}_B,\ \mathbf{0}_{B\times R,L}]^{\mathsf{T}} - \alpha_{\mathrm{b}} \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \Big] \Big\|_{\mathrm{F}}^2. \end{aligned} \tag{C.3}$$

We then define a column-wise Khatri-Rao matrix, $\mathbf{D} \in \mathbb{C}^{MB\times K}$ and another matrix, $\mathbf{X} \in \mathbb{C}^{MB\times L}$, as follows:

$$\mathbf{D} = [\mathbf{b}_1 \otimes \mathbf{a}_1, \cdots, \mathbf{b}_K \otimes \mathbf{a}_K], \tag{C.4}$$

$$\mathbf{X} = \big[\mathrm{vec}\left([\mathbf{I}_B,\ \mathbf{0}_{B\times R,1}]^{\mathsf{T}} - \alpha_{\mathrm{b}} \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \right), \cdots, \mathrm{vec}\left([\mathbf{I}_B,\ \mathbf{0}_{B\times R,L}]^{\mathsf{T}} - \alpha_{\mathrm{b}} \mathbf{H}_{\text{b-u}}^{\mathsf{H}} \mathbf{F} \right) \big]. \tag{C.5}$$

Through vectorization, we have the following relation for the norm of a matrix:

$$\|\mathbf{D}\boldsymbol{v} - \mathrm{vec}(\mathbf{C})\|_2^2 = \left\| \mathbf{A} \mathrm{Diag}(\boldsymbol{v}) \mathbf{B}^{\mathsf{T}} - \mathbf{C} \right\|_{\mathrm{F}}^2. \tag{C.6}$$

By using the relation in (C.6) and substituting the matrices $\mathbf{D}$ and $\mathbf{X}$ into (C.3), we get:

$$f_1 = \|\mathbf{D}[\boldsymbol{\upsilon}_1,\ \boldsymbol{\upsilon}_2, \cdots, \boldsymbol{\upsilon}_L] - \mathbf{X}\|_{\mathrm{F}}^2\,, \tag{C.7}$$

or equivalently:

$$f_1 = \|\mathbf{D}\boldsymbol{\Upsilon} - \mathbf{X}\|_{\mathrm{F}}^2\,. \tag{C.8}$$

Similarly, by defining the matrices:

$$\mathbf{M} = \sqrt{P_{\mathrm{s}}}\alpha_{\mathrm{s}}\mathbf{H}_{\text{s-u}}^{\mathsf{H}}\mathrm{Diag}(\mathbf{H}_{\text{b-s}}\mathbf{v}_{\mathrm{b}}), \tag{C.9}$$

$$\mathbf{Z} = \left[\mathbf{0}_{L\times B},\ \mathbf{S}_{\mathrm{s}}^{\mathsf{T}}\right]^{\mathsf{T}}, \tag{C.10}$$

and then substituting them into (C.2) one obtains:

$$f_2 = \|\mathbf{M}\boldsymbol{\Upsilon} - \mathbf{Z}\|_{\mathrm{F}}^2\,. \tag{C.11}$$

Bibliography

[1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[4] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A Survey of Energy-Efficient Techniques for 5G Networks and Challenges Ahead," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 697–709, 2016.

[5] S. Zhang, Q. Wu, S. Xu, and G. Y. Li, "Fundamental Green Tradeoffs: Progresses, Challenges, and Impacts on 5G Networks," *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 33–56, 2017.

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

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

[8] S. Hu, F. Rusek, and O. Edfors, "Beyond Massive MIMO: The Potential of Data Transmission With Large Intelligent Surfaces," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2746–2758, 2018.

[9] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO Communications Relying on Intelligent Reflecting Surfaces," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5218–5233, 2020.

[10] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable Intelligent Surfaces: Principles and Opportunities," *IEEE Communications Surveys Tutorials*, pp. 1–1, 2021.

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

[12] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable Intelligent Surfaces for Energy Efficiency in Wireless Communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

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

[14] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A Framework of Robust Transmission Design for IRS-Aided MISO Communications With Imperfect Cascaded Channels," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5092–5106, 2020.

[15] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep Reinforcement Learning-Based Intelligent Reflecting Surface for Secure Wireless Communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 375–388, 2021.

[16] X. Yu, D. Xu, and R. Schober, "Optimal Beamforming for MISO Communications via Intelligent Reflecting Surfaces," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020, pp. 1–5.

[17] P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li, "Intelligent Reflecting Surface-Assisted Millimeter Wave Communications: Joint Active and Passive Precoding Design," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14 960–14 973, 2020.

[18] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent Reflecting Surface: Practical Phase Shift Model and Beamforming Optimization," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5849–5863, 2020.

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

[20] W. Yan, X. Yuan, Z.-Q. He, and X. Kuai, "Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser mimo systems," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1793–1808, 2020.

[21] H. Zhao, Y. Shuang, M. Wei, T. J. Cui, P. D. Hougne, and L. Li, "Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals," *Nature Communications*, vol. 11, no. 3926, pp. 1–10, 2020.

[22] S. Li, B. Duo, X. Yuan, Y.-C. Liang, and M. Di Renzo, "Reconfigurable Intelligent Surface Assisted UAV Communication: Joint Trajectory Design and Passive Beamforming," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 716–720, 2020.

[23] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Intelligent Reflecting Surface Enhanced Indoor Robot Path Planning: A Radio Map-Based Approach," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4732–4747, 2021.

[24] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications," 2021.

[25] Y. Li, M. Jiang, Q. Zhang, and J. Qin, "Joint beamforming design in multi-cluster MISO NOMA reconfigurable intelligent surface-aided downlink communication networks," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 664–674, 2021.

[26] C. Xu, L. Yang, and P. Zhang, "Practical Backscatter Communication Systems for Battery-Free Internet of Things: A Tutorial and Survey of Recent Research," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 16–27, 2018.

[27] D. P. Bertsekas, "Nonlinear programming," *Journal of the Operational Research Society*, vol. 48, no. 3, pp. 334–334, 1997.

[28] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6664–6684, 2019.

[29] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network: Joint Active and Passive Beamforming Design," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[30] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, 2005.

[31] H. Jedda, A. Mezghani, A. L. Swindlehurst, and J. A. Nossek, "Precoding under instantaneous per-antenna peak power constraint," in *2017 25th European Signal Processing Conference (EUSIPCO)*, 2017, pp. 863–867.

[32] R. W. Heath Jr. and A. Lozano, *Foundations of MIMO Communication*. Cambridge University Press, 2018.

[33] "approximate message-passing for convex optimization with non-separable penalties."

[34] J. Barbier, F. Krzakala, N. Macris, L. Miolane, and L. Zdeborová, "Optimal errors and phase transitions in high-dimensional generalized linear models," *Proceedings of the National Academy of Sciences*, vol. 116, no. 12, pp. 5451–5460, 2019.

[35] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *IEEE Information Theory Workshop on Information Theory (ITW 2010, Cairo)*, 2010, pp. 1–5.

[36] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer-Verlag New York, 2006.

[37] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 764–785, 2011.

[38] Q. Li, Z. Zhu, and G. Tang, "Alternating Minimizations Converge to Second-Order Optimal Solutions," in *Proceedings of the 36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, 09–15 Jun 2019, pp. 3935–3943. [Online]. Available: http://proceedings.mlr.press/v97/li19n.html

[39] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-Rank Matrix Completion Using Alternating Minimization," in *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*. New York, NY, USA: Association for Computing Machinery, 2013, p. 665–674. [Online]. Available: https://doi.org/10.1145/2488608.2488693

[40] X. Yi, C. Caramanis, and S. Sanghavi, "Alternating Minimization for Mixed Linear Regression," in *Proceedings of the 31st International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, E. P. Xing and T. Jebara, Eds., vol. 32, no. 2. Bejing, China: PMLR, 22–24 Jun 2014, pp. 613–621. [Online]. Available: http://proceedings.mlr.press/v32/yia14.html

[41] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase Retrieval Using Alternating Minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 18, pp. 4814–4826, 2015.

[42] A. Sinha, P. Malo, and K. Deb, "A review on bilevel optimization: From classical to evolutionary approaches and applications," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 276–295, 2018.

[43] P. Hannan, "The element-gain paradox for a phased-array antenna," *IEEE Transactions on Antennas and Propagation*, vol. 12, no. 4, pp. 423–433, 1964.

[44] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent Reflecting Surface-Assisted Multi-User MISO Communication: Channel Estimation and Beamforming Design," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 661–680, 2020.

[45] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface assisted wireless communication: Modeling and channel estimation," *arXiv preprint arXiv:1906.02360*, 2019.

[46] W. Gifford, M. Win, and M. Chiani, "Diversity with practical channel estimation," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1935–1947, 2005.

[47] R. Annavajjala, P. C. Cosman, and L. B. Milstein, "Performance Analysis of Linear Modulation Schemes With Generalized Diversity Combining on Rayleigh Fading Channels With Noisy Channel Estimates," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4701–4727, 2007.

[48] J. Zhang, M. Kountouris, J. G. Andrews, and R. W. Heath, "Multi-Mode Transmission for the MIMO Broadcast Channel with Imperfect Channel State Information," *IEEE Transactions on Communications*, vol. 59, no. 3, pp. 803–814, 2011.

[49] J. Yeh, *Real Analysis*, 3rd ed. WORLD SCIENTIFIC, 2014. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/9037

[50] 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*, pp. 1–1, 2021.