Joint Reflecting and Precoding Designs for SER Minimization in Reconfigurable Intelligent Surfaces Assisted MIMO Systems

Jia Ye, Student Member, IEEE, Shuaishuai Guo, Member, IEEE, and Mohamed-Slim Alouini, Fellow, IEEE

Abstract—This paper investigates the use of a reconfigurable intelligent surface (RIS) to aid point-to-point multiple-input multiple-output (MIMO) wireless communications. We present efficient designs for both the reflecting elements at the RIS and the precoder at the transmitter, which target at minimizing the symbol error rate (SER) of the transmission. Specifically, we iteratively design the reflecting and precoding to directly minimize symbol error rate (MSER), referred as MSER-Reflecting and MSER-Precoding, which are shown to achieve favorable results but at the cost of high computational complexity. Therefore, a simplified semidefinite programming-based (SDP-) reflecting scheme and a simplified maximizing the minimum Euclidean distance (MMED-) precoding design are studied to reduce computational complexity. Furthermore, direct solutions for precoder design, e.g., maximum ratio transmission (MRT-) precoding and Eigen-Precoding are investigated in terms of SER for comparison. Simulation results show that the RIS-assisted MIMO communications combined with the proposed reflectors and precoders can offer a lower SER.

Index Terms—Large intelligent surface, phase shifts design, precoder design, symbol error rate.

I. INTRODUCTION

Better communication qualities like smaller delay, higher transmission rate, lower symbol error probability (SER), less energy consumption always attract researchers’ and users’ eyes. In the last decade, wireless networks have been greatly improved thanks to various technological advances, including massive multiple-input multiple-output (massive MIMO), millimeter wave (mmWave) communications, and ultra-dense deployments of small cells. However, some critical issues such as the hardware complexity and system update cost are still blocking their steps to the practical implementation [1]. To satisfy the growing demands with satisfying communication quality and achieve challenge goals in a green and effective way, reconfigurable intelligent surface (RIS) was proposed as a promising solution in the coming 5G or beyond era [2].

RIS is a planar array comprising of a large number of nearly passive, low-cost, reflecting elements such as positive-intrinsic-negative (PIN) diodes, which are used for altering the phase of the reflected electromagnetic wave with re-configurable parameters and smart controller. RIS can be implemented by various materials, including reflect arrays [3], liquid crystal metasurfaces [4], ferroelectric films, or even metasurfaces [5].

In the beginning, reflecting surfaces were not considered in wireless communication systems because these surfaces only had fixed phase shifters which could not adapt the phase modification in time-varying wireless propagation environments. Recently, advanced micro-electrical-mechanical systems (MEMS) and metamaterials have been investigated as a solution to this issue, which enables the real-time reconfiguration reflecting surfaces [6]. Compared to existing related technologies such as multi-antenna relay [8], backscatter communication [9] and active intelligent surface based massive MIMO [10], passive RIS does not require any dedicated energy source for either decoding, channel estimation, or transmission. It only reflects the ambient radio frequency (RF) signals in a passive way without a transmitter module. Moreover, the reflect-path signal through RIS carries the same useful information as well as the direct-path signal without any information of its own, which will not cause any additional interference.

RIS stands out among these technologies by smartly adjusting the phase shifts induced by all the elements with advantages like overcoming unfavorable propagation conditions, enriching the channel with more multi-paths, increasing the coverage area, improving the received signal power, avoiding interference, enhancing security/privacy and consuming very low energy. On the other hand, the lightweight and conformal geometry of RIS can enable the installment onto the facades of buildings in outdoor communication environments or the ceilings and walls of rooms in indoor communication environments, which provides high flexibility and superior compatibility for practical implementation [11]. Also, integrating RIS into the existing networks can be made transparent to the users without the need for any change in the hardware and software of their devices.

A. Prior Work

Due to the advantages mentioned above, RIS has attracted more and more researchers’ attention in the last few years. A lot of effort has been made to optimize the RIS parameters and systems’ structure to improve existing communication networks. Researchers compared RIS-aided communications with traditional systems in various system performance, such as outage probability, signal power, signal-to-interference-plus-noise ratio (SINR), ergodic capacity, spectral efficiency, error
probability and so on. For example, [3] enhanced the mmWave link robustness and optimized the link outage probability by deploying smart reflect-arrays when the line-of-sight (LoS) links are blocked by obstructions. The authors also investigated the optimal beam direction for randomly moving devices without any location information, incorporating the antenna sector selection at the access point (AP) and the mobile user as well as the configuration of the smart reflect-arrays. For maximizing the total received signal power, preliminary contributions appeared in [12], [13]. They considered a RIS-assisted single-user multiple-input-single-output (MISO) system and optimized the values of the phase shifts. Specifically, a centralized algorithm with the global channel state information (CSI) and a low-complexity distributed algorithm were proposed in [12]. It was shown that the RIS could provide an asymptotic power gain in the order of $O(N^2)$ in terms of the number of reflecting elements at the RIS, $N$, which is double than that of the massive MISO without the aid of RIS. With the same system setup in [12], the authors of [13] considered a more practical case when the RIS only has a finite number of discrete phase shifts in contrast to the continuous phase shifts considered in [12]. It can be seen that the asymptotic squared power gain of RIS-aided MISO shown in [12] with continuous phase shifts still holds with discrete phase shifts with a stable performance loss gap between the two conditions. The gap depends only on the number of phase-shift levels at each element, but regardless of $N$. By using the knowledge of only the channel large-scale statistics instead of the global CSI, [14] designed an optimal linear precoder and the power allocation at the base station (BS) as well as the RIS phase matrix to maximize the minimum SINR in multi-user RIS-assisted MISO communication systems. The effects of both the rank structure of the LoS channel matrix and the spatial correlation between the RIS elements were taken into account. It was shown that the RIS-assisted system can achieve power gains with a much fewer number of active antennas at the BS. Targeting for practical implementation, [15] investigated an optimal phase shift design exploiting statistical CSI to maintain an acceptable degradation of the ergodic capacity. Moreover, the authors of [15] evaluated the error performance of RIS-based communication systems by investigating the effect of the number of reflecting elements, modulation orders, and blind phase. The asymptotic results in different signal-to-noise-ratio (SNR) regimes were provided. The study of ergodic capacity of RIS under Rician fading large-scale antenna system without the LoS path appeared in [16]. The relationship between the capacity and the number of reflect-elements at RIS was revealed. [17]–[19] analyzed the approximated uplink ergodic rate of a Rician fading system and derived an optimal size of a RIS unit, where users are mapped to a limited area of the entire RIS. Besides, the asymptotic variance of the uplink rate was also derived to investigate channel hardening effects on the system performance such as system reliability, latency, and diversity. Furthermore, the performance bound of a RIS-based system was obtained by using a scaling law for the uplink SINR. It was shown that RIS can bring an improved reliability with a significantly reduced area for antenna deployment compared to massive MIMO.

Recently, the authors in [20] further investigated three conceptual RIS-based index modulation system realizations to achieve ultra-reliable transmission. A unified framework for error performance calculation was developed for RIS-space shift keying and RIS-spatial modulation schemes. The authors of [21] and [22] maximized the energy- and spectral-efficiency of a RIS-assisted multi-user MISO system. They employed zero-forcing beamforming with global CSI for designing both the transmit power allocation at the BS and the phase elements of the RIS. [23] proposed a provably convergent, low-complexity method to maximize the system sum-rate. It was shown in [23] that a nearly interference-free zone could be established in the proximity of the RIS thanks to its spatial interference nulling/cancellation capability. The authors in [24] studied the data rates for multiple wireless users in the same room accessing the same spectrum band. It demonstrated that higher spectrum-spatial efficiency and data rate can be achieved by adaptively changing the phases of the reflected wireless signals at the smart reflect-array. In summary, RIS is a promising technology to satisfy the growing demand for data rate and communication quality. The spectral- and energy-efficiency have been widely investigated. However, it should be mentioned that there is a few of literature designing the phase shifts at the RIS combining with the precoding at the transmitter based on the symbol error rate (SER) minimization criterion, which motivates us to fill this gap with the contributions listed below:

## B. Contributions

- This paper considers a RIS-enhanced point-to-point multiple-data-stream MIMO system, where a multi-antenna transmitter serves a multi-antenna receiver with the help of a RIS. Such a system can be employed to facilitate wireless information and power transfer in various 5G or beyond communication networks.
- This paper formulates a joint optimization problem for designing both phase shifts at the RIS and the precoding matrix at the transmitter to minimize the system SER. To find the optimal diagonal phase shifts matrix and the full precoding matrix, we propose a MSER-Reflecting scheme based on a given precoding matrix and develop a MSER-Precoding scheme based on given phase shifts separately. Then, we use a Two-Step iterative algorithm to get the final optimized solution.
- To reduce the computational complexity and also facilitate the implementation in a realistic system, we design simplified algorithms for both parts. We simplify the phase shifts design to be a semidefinite programming (SDP) problem and design a SDP-Reflecting scheme. Also, we transform the precoding design optimization problem to be a quadratically constrained quadratic program (QCQP) problem based on maximizing the minimum Euclidian distance (MMED) criterion and propose a MMED-Precoding scheme accordingly. Moreover, the low-complexity maximum ratio transmission (MRT-) precoding and Eigen-Precoding schemes are also investigated in terms of SER for comparison.
C. Organization

The remainder of the paper is organized as follows. Section II describes the system model. Section III introduces the problem formulation which tries to minimize the system SER and gives out the sub-optimal solution. In Section IV, we propose simplified algorithms. Numerical comparisons are presented in Section V and conclusions are drawn in Section VI.

D. Notations

In this paper, $x$ denotes a scalar; $\mathbf{x}$ represents a vector; $\mathbf{X}$ stands for a matrix. $\|\mathbf{x}\|_2$, $\|\mathbf{x}\|_p$ and $\|\mathbf{x}\|_\infty$ represents $l_2$ norm, $l_p$ norm and $l_\infty$ norm of $\mathbf{x}$ respectively. $\|\mathbf{X}\|_F$ is the Frobenius norm of $\mathbf{X}$. $\text{diag}(\mathbf{x})$ is a diagonal matrix whose diagonal entries are from vector $\mathbf{x}$. $x_{i,j}$ denotes the $i$-th entry of $\mathbf{x}$, $x_{i,j}$ is the element in $i$-th row and $j$-th column of a matrix $\mathbf{X}$. $\mathbf{x}$ denotes the $i$-th column of matrix $\mathbf{X}$. $\text{tr}(\mathbf{X})$ denotes the trace of $\mathbf{X}$. $\text{rank}(\mathbf{X})$ represents the rank of $\mathbf{X}$. $\text{vec}(\mathbf{X})$ means the vectorization of matrix $\mathbf{X}$, which is a linear transformation which converts $\mathbf{X}$ into a column vector. $\lambda_{\text{max}}(\mathbf{X})$ is the maximum eigenvalue of matrix $\mathbf{X}$. $\mathbf{X} \succeq 0$ means that matrix $\mathbf{X}$ is positive semidefinite. $\otimes$ stands for the Hadamard product, $\otimes$ denotes Kronecker product. $(\cdot)^H$ is the conjugate transpose. $\mathbb{C}$ stands for the complex domain while $\mathbb{R}$ represents the real domain. $\mathcal{CN}(\mu, \sigma^2)$ denotes the complex Gaussian distribution with mean $\mu$ and variance $\sigma^2$. $\mathcal{CN}(\mu, \Sigma)$ stands for the circularly symmetric complex Gaussian distribution with mean $\mu$ and covariance $\Sigma$. $E[\cdot]$ represents the expectation operation. $\mathbf{I}_N$ denotes an $N \times N$ identity matrix. $[x]^\top$ denotes max($x, 0$). $\text{Re}\{x\}$ and $\text{Im}\{x\}$ represent the real and imaginary part of $x$, respectively. $\nabla$ denotes the gradient of a function. $Q(\cdot)$ stands for the tail distribution function of the standard normal distribution. $\text{angle}(\cdot)$ represents the element-wise phase function.

II. System Model

In this paper, we consider a RIS-assisted MIMO system model as illustrated in Fig. 1. In the model, a transmitter equipped with $N_t$ antennas communicates with a receiver equipped with $N_r$ antennas through a RIS composed of $N$ reflecting units. The reflecting array acts as a passive relay, which is embedded in a surrounding building. There is no LoS communication between the two devices because of some blocks. The received vector $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ at the receiver can be expressed as

$$\mathbf{y} = \sqrt{\rho} \mathbf{H}_2 \Phi \mathbf{H}_1 \mathbf{s} + \mathbf{n},$$  \hspace{1cm} (1)

where $\rho$ is the SNR; $\mathbf{H}_2 \in \mathbb{C}^{N_r \times N}$ represents the channel between the RIS and the receiver; $\Phi = \text{diag}\{\phi\} \in \mathbb{C}^{N \times N}$ denotes the diagonal matrix accounting for the effective phase shifts applied by the RIS reflecting elements with $\phi = [e^{j\theta_1}, \ldots, e^{j\theta_N}]^T$; $\mathbf{H}_1 \in \mathbb{C}^{N \times N_t}$ represents the channel between the transmitter and the RIS; $\mathbf{F} \in \mathbb{C}^{N_s \times N_r}$ is the precoder to encode $N_s$ data streams; $\mathbf{s}$ is the $N_s \times 1$ transmitted data symbol vector with each entry chosen from a $M$-ary constellation $\mathcal{S}_M$ and there are totally $M^{N_s}$ legitimate symbol vectors; $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_r})$ is the additive white Gaussian noise (AWGN) vector with each entry obeying a zero-mean variance $\sigma^2$ complex Gaussian distribution. In this paper, we assume that the average power of all legitimate symbol vectors $\{\mathbf{s}\}$ is normalized. Let $\mathbf{x} = \mathbf{F} \mathbf{s}$ denote the transmitted signal vector from the multi-antenna transmitter. It is assumed that $\mathbf{x}$ satisfies the maximum average transmit power constraint:

$$E[\|\mathbf{x}\|_2^2] = E[\text{tr}(\mathbf{F} \mathbf{s}^H \mathbf{F} \mathbf{H})] = \text{tr}(\mathbf{Q}) \leq P_{\text{max}},$$  \hspace{1cm} (2)

where $\mathbf{Q} \triangleq \mathbf{F} \mathbf{F}^H$ is the signal covariance matrix and $P_{\text{max}}$ denotes the maximum average power.

It is considered that the system utilizes all $M^{N_s}$ feasible transmit vectors, the union bound on SER can thus be written as

$$P_S(\Phi, \mathbf{F}) = \frac{1}{M^{N_s}} \sum_{i=1}^{M^{N_s}} \sum_{j=1,j \neq i}^{M^{N_s}} \text{Pr}\{\mathbf{s}_i \rightarrow \mathbf{s}_j\},$$  \hspace{1cm} (3)

where $\text{Pr}\{\mathbf{s}_i \rightarrow \mathbf{s}_j\}$ denotes the pairwise SER of the vector $\mathbf{s}_i$ being erroneously detected as $\mathbf{s}_j$. By using the squared Euclidean distance $d_{ij}^2(\Phi, \mathbf{F}) = (\mathbf{H}_2 \Phi \mathbf{H}_1 \mathbf{s}_i - \mathbf{s}_j)^2$ between two vectors, $\text{Pr}\{\mathbf{s}_i \rightarrow \mathbf{s}_j\}$ can be computed as

$$\text{Pr}\{\mathbf{s}_i \rightarrow \mathbf{s}_j\} = Q\left(\sqrt{\frac{\rho d_{ij}^2(\Phi, \mathbf{F})}{2\sigma^2}}\right).$$  \hspace{1cm} (4)

III. Optimization Problem Formulation

The objective of our design is to minimize the SER. In this section, we will formulate the optimization problem with the assumption that all CSI are known. With global CSI, the optimization can be formulated as problem (P1):

$$\text{(P1)}: \text{Given: } \mathbf{H}_1, \mathbf{H}_2, \mathcal{S}_M \hspace{1cm} \text{Find: } \Phi, \mathbf{F}$$

$$\text{Minimize: } P_S(\Phi, \mathbf{F})$$

$$\text{Subject to: } \text{tr}(\mathbf{Q}) \leq P_{\text{max}}$$

$$0 \leq \theta_i \leq 2\pi, \forall i = 1, \ldots, N.$$  \hspace{1cm} (5)
Because of coupling effect between the reflecting elements in $\Phi$ and the precoder $F$, the original problem (P1) is hard to solve. To decouple them, we will first optimize $\Phi$ by fixing $F$ and then update $F$ by fixing $\Phi$ respectively. Then, we will obtain sub-optimal solutions for both $\Phi$ and $F$ by performing the process iteratively.

A. MSER-Reflecting Scheme

By fixing $F$, we can re-express $d_{ij}^2(\Phi) = \|H_2^H (q_i - q_j)\|_2^2$ to simplify the calculation, where $q_i = H_1 F s_i$ and $q_j = H_1 F s_j$. Then, we carry out some manipulations to transform the squared Euclidean distance as

$$d_{ij}^2(\Phi) = \|H_2^H (q_i - q_j)\|_2^2 = (q_i - q_j)^H H_2^H H_2 (q_i - q_j) = \text{tr} (H_2^H R_{H2} \Delta Q_{ij}),$$

where $R_{H2} = H_2^H H_2$ and $\Delta Q_{ij} = (q_i - q_j)^H (q_i - q_j)$. Using the rule that $\text{tr} (D^H A D B^T) = x^H (A \otimes B) y$ with $D_x = \text{diag} \{ x \}$ and $D_y = \text{diag} \{ y \}$ in (23), we can re-express $d_{ij}^2(\Phi)$ as

$$d_{ij}^2(\Phi) = \text{tr} (H_2^H R_{H2} \Delta Q_{ij}) = \text{tr} (H_2^H C_{ij} \phi),$$

where $C_{ij} = R_{H2} \otimes \Delta Q_{ij}^T$.

Moreover, it is convenient to define $\phi_i = e^{i \theta_i}$. Since $|\phi_i| = 1$, accordingly we can obtain that $\text{tr} (\phi_i H^H) = N$. In order to handle the non-convex constraint of $|\phi_i| = 1$, we relax the problem (P1) into the following optimization (P2) with a convex $\ell_\infty$ constraint:

$$(\text{P2}) : \quad \text{Given} : H_1, H_2, S_M, F$$

Find : $\phi$

Minimize : $P_S(\phi)$

Subject to : $\text{tr} (\phi H^H) = N$

$$\| \phi \|_\infty \leq 1.$$  

Actually, the original feasible set is a subset of the new feasible set in (P2), i.e.,

$$\{ \phi \in \mathbb{C}^{N \times 1} : \phi_i = e^{i \theta_i}, \forall i = 1, \ldots, N \} \subset \{ \phi \in \mathbb{C}^{N \times 1} : \| \phi \|_\infty \leq 1 \},$$

which is convex due to the convexity of the $\ell_\infty$ norm.

Since the $\ell_\infty$ constraint is non-differentiable, we exploit the $\ell_p$ approximation (26) with a gradually increased large $p$, $\lim_{p \to \infty} \| \phi \|_p = \| \phi \|_\infty$, during the optimization process. To solve (P2), we utilize the barrier method to incorporate the non-negative constraint (27) with the logarithmic barrier function $I(u)$ to approximate the penalty of violating the $\ell_p$ constraint, i.e.,

$$I(u) = \begin{cases} -\frac{1}{t} \ln(u), & u > 0 \\ \infty, & u \leq 0 \end{cases}$$

where $t$ is used to scale the barrier function’s penalty. Thus, we can obtain the following optimization problem:

$$\min_{\phi \in \mathbb{C}^{N \times 1}} g(\phi, p) = P_S(\phi) + I \left( 1 - \| \phi \|_p \right),$$

(11)

To solve (11) via a gradient method, we formulate the gradient of the cost function $g(\phi, p)$ over $\phi$ as follows:

$$\nabla g(\phi, p) = \nabla P_S(\phi) + \frac{\| \phi \|_p - p}{2t} \frac{P_p}{\| \phi \|_p},$$

(12)

where $P_p \in \mathbb{C}^{N \times 1}$ is given as

$$P_p = \left[ |\phi_1|^p - 2, |\phi_2|^p - 2, \ldots, |\phi_N|^p - 2 \right]^T.$$  

Moreover, the gradient $\nabla P_S(\phi)$ can be calculated as

$$\nabla P_S(\phi) = \frac{1}{M_N} \sum_{i=1}^{M_N} \sum_{j=1, j \neq i}^{M_N} \phi_i \phi_j.$$  

According Leibniz’s integral rule, $\int_{a(x)}^{b(x)} (x, t) dt = \int_{a(x)}^{b(x)} \frac{d}{dx} f(x, t) dt = f(x, b(x)) - f(x, a(x))$. Thus, we can express $\nabla P_S(\phi)$ as

$$\nabla P_S(\phi) = \frac{1}{2\pi} \int_{a(x)}^{b(x)} (x, t) dt = f(x, b(x)) - f(x, a(x)),$$

which is convex due to the convexity of the $\ell_\infty$ norm.

By using $-\nabla g(\phi, p)$ as the search direction (28), we can search the optimized solution as listed in Algorithm 1.

Proposition 1. Algorithm 1 always guarantees $P_S(\phi_{k+1}) \leq P_S(\phi_k)$.

Proof. For any $\hat{\theta} \to 0$, based on (20), the Taylor expansion of $P_S(\phi_{k+1})$ can be derived as

$$P_S(\phi_{k+1}) = P_S(\phi_k) + g_{k+1}^H \cdot \sqrt{N} \frac{w_k}{\| w_k \|} + O(\hat{\theta}^2)$$

$$\approx P_S(\phi_k) + g_k^H \frac{\sqrt{N} w_k}{\| w_k \|}.$$  

(22)
**Algorithm 1** MSER-Reflecting Scheme

1: **Initialization:** Given a feasible initial solution $\phi_0$, $p > 0$, $\Delta p > 0$, $p_{\text{max}} > 0$, $k = 0$, halting criterion $\varepsilon_1 > 0$ and the barrier coefficient $t$.

2: **Gradient and Search direction:** Compute the gradient $g_k$ and derive the search direction as:

$$w_k = -g_k = -\nabla_{\phi} g(\phi_k, p),$$

where $\nabla_{\phi} g(\phi_k, p)$ is given in (16).

3: **Direction Projection:** Project the search direction into the tangent plane of $\text{tr}((\phi_k^H)^2) = N$

$$w_k^+ = w_k - \frac{(w_k, \phi_k)\phi_k}{\|\phi_k\|^2};$$

4: **Search for $\hat{\theta}$:** For $0 \leq \hat{\theta} \leq \pi/2$, search for it by

$$\hat{\theta} = \arg \min_{\theta} P_S(\phi_k);$$

5: **Update:** Go to step 6 if $\frac{w_k^+}{\|w_k^+\|} \leq \varepsilon_1$, else let

$$\phi_{k+1} = \cos \hat{\theta} \cdot \phi_k + \sin \hat{\theta} \cdot \sqrt{N} \frac{w_k^+}{\|w_k^+\|},$$

$k \leftarrow k + 1$ and then go to step 2.

6: **Iteration:** Go to Step 7 if $p \geq p_{\text{max}}$, else let $p \leftarrow p + \Delta p$ and then go to Step 2.

7: **Output:** The optimized reflecting elements are thus given by:

$$\phi^* = \exp \left[ \angle(\phi_k) \right].$$

We can also calculate that $g_k^H w_k^+ = (\cos^2 \alpha - 1) \cdot \|w_k\|^2 \leq 0$, where $\alpha = \arccos \frac{w_k^H w_k}{\|w_k\|^2}$ is the angle between vectors $\phi_k$ and $w_k$. Combining it with (22), we have $P_S(\phi_{k+1}) \leq P_S(\phi_k)$. The proof is completed.

**B. MSER-Precoding Scheme**

Based on the optimized $\Phi$, we can simplify (P1) as

**(P3):**

- **Given:** $H_1, H_2, S_M, \Phi$
- **Find:** $F$
- **Minimize:** $P_S(F)$
- **Subject to:** $\text{tr}(Q) \leq P_{\text{max}}.$

In this work, we consider a full rank precoding matrix since the diagonal structure might not provide the best solution. By using $H$, the received signal can be re-expressed as

$$y = HF_{s} + n$$

$$= [h_1, \cdots, h_{N_s}]$$

$$= \begin{bmatrix} f_{1,1} & \cdots & f_{1,N_s} \\ \vdots & \ddots & \vdots \\ f_{N_t,1} & \cdots & f_{N_t,N_s} \end{bmatrix} \begin{bmatrix} s_0 \\ \vdots \\ s_{N_t} \end{bmatrix} + n$$

$$= \sum_{a=1}^{N_t} \sum_{b=1}^{N_s} f_{a,b} h_a s_b + n,$$

where $H = H_2 \Phi H_1 \in \mathbb{C}^{N_r \times N_t}$.

Considering that (24) has two summations, which are not easy to handle, we rebuild the channel matrix, precoding matrix and the transmitted data stream in a new form (29). We first construct a new channel matrix as $\hat{H} = \hat{H}_1, \cdots, \hat{H}_{N_t}, \cdots, \hat{H}_{N_r} \in \mathbb{C}^{N_r \times N_t N_s}$, in which

$$\hat{H}_{N_t} = [h_{N_t}, \cdots, h_{N_t}, \cdots, h_{N_t}] = h_{N_t} \otimes 1^{1 \times N_s},$$

where each $h_{N_t}$ repeats $N_s$ times. Meanwhile, the precoding matrix entries in $F$ are collected together as

$$\hat{F} = \text{diag} \left\{ \text{vec} \left( F^T \right) \right\}$$

$$= \text{diag} \left\{ f_{1,1}, f_{1,2}, \cdots, f_{N_t,N_s} \right\}$$

$$= \text{diag} \left\{ f \right\}.$$  

Furthermore, the transmitted data stream is rearranged as

$$\hat{s} = [s, \cdots, s, \cdots, s]^T \in \mathbb{C}^{N_r N_t \times 1}.$$  

Following the procedure in (25)-(27), we can rewrite (24) to be $y = \hat{H} \hat{F} \hat{s} + n$.

Thus, we can re-express the squared Euclidean distance as $d^2_{ij} \left( \hat{F} \right) = \left\| \hat{H} \hat{F} (\hat{s}_i - \hat{s}_j) \right\|^2$. Similarly to (29), the squared Euclidean distance can be re-expressed as

$$d^2_{ij} \left( \hat{F} \right) = \left\| \hat{H} (\hat{s}_i - \hat{s}_j) \right\|^2$$

$$= (\hat{s}_i - \hat{s}_j)^H \hat{F} \hat{H}^H \hat{F} (\hat{s}_i - \hat{s}_j)$$

$$= \text{tr} \left( \hat{F}^H R_{\hat{H}} \hat{F} \Delta S_{ij} \right)$$

$$= f^H (R_{\hat{H}} \otimes \Delta S_{ij}) f$$

$$= f^H \hat{C}_{ij} f,$$

where $R_{\hat{H}} = \hat{H}^H \hat{H}$, $\Delta S_{ij} = (\hat{s}_i - \hat{s}_j) (\hat{s}_i - \hat{s}_j)^H$, $f = [f_{1,1}, f_{1,2}, \cdots, f_{N_t,N_s}]^T$ and $\hat{C}_{ij} = R_{\hat{H}} \otimes \Delta S_{ij}$. Thus, (P3) can be transformed to

**(P3-a):**

- **Given:** $\hat{H}, S_M$
- **Find:** $f$
- **Minimize:** $P_S(f)$
- **Subject to:** $\text{tr} \left( f f^H \right) \leq P_{\text{max}}.$

(29)
To solve (P3-a), we formulate the Lagrangian function as
\[
L(f, \mu) = P_S(f) + \mu \left( \text{tr}(f^H) - P_{\text{max}} \right),
\] (30)
where \( \mu \) is the Lagrangian multiplier. The optimal solution must satisfy the Karush-Kuhn-Tucker (KKT) conditions as
\[
\begin{cases}
\nabla_{f} L(f, \mu) = 0 \\
\mu \left( \text{tr}(f^H) - P_{\text{max}} \right) = 0 \\
\mu \geq 0
\end{cases}
\] (31).

Because of its monotonicity with power, \( L(f, \mu) \) is minimized when the power constraint is met with strict equality. Hence, \( \mu \left( \text{tr}(f^H) - P_{\text{max}} \right) = 0 \). The first equation in (31) can be represented as
\[
\nabla_{f} L(f, \mu) = \left[ -\frac{1}{MN_s} \Omega(f) + 2\mu I \right] f = 0,
\] (32)
where \( \Omega(f) \) is given by
\[
\Omega(f) = \sum_{i=1}^{MN_s} \sum_{j=1,j \neq i}^{MN_s} \sqrt{\frac{\rho}{4\pi \sigma^2 d_{ij}^2(f)}} \exp \left( -\frac{\rho d_{ij}^2(f)}{4\sigma^2} \right) \hat{c}_{ij}.
\] (33)

However, the closed-form solution of (32) is difficult to derive. Instead, we develop an iterative algorithm to search for the solution of (32), which can be seen in Algorithm 2.

**Proposition 2.** Algorithm 2 always guarantees \( P_S(f_{k+1}) \leq P_S(f_k) \).

**Proof.** The proof is same as the Algorithm 1. \( \square \)

Since (P3-a) is a non-convex problem, (32) is a necessary condition for global optimum and the generated vector \( f \) in Algorithm 2 is thus a critical point.

**C. Two-Step Algorithm**

By combining Algorithms 1 and 2, we therefore develop a Two-Step algorithm to minimize SER in Algorithm 3, where the reflecting elements and the precoding matrix are optimized iteratively.

**Proposition 3.** Algorithm 3 always guarantees \( P_S(\phi_{k+1}, f_{k+1}) \leq P_S(\phi_k, f_k) \).

**Proof.** From **Proposition 1**, we can obtain that \( P_S(\phi_{k+1}, f_k) \leq P_S(\phi_k, f_k) \) and from **Proposition 2**, we can further obtain that \( P_S(\phi_{k+1}, f_{k+1}) \leq P_S(\phi_{k+1}, f_k) \).

Finally, Algorithm 3 can guarantees that
\[
P_S(\phi_{k+1}, f_{k+1}) \leq P_S(\phi_{k+1}, f_k) \leq P_S(\phi_k, f_k).
\] (39)

The proof is completed. \( \square \)

**Remark 1.** In this section, we split the optimization problem (P1) into (P2) and (P3-a). It should be noted that the original problem (P1) is hard to solve due to coupling effect between the reflecting elements and the precoder. Our approach of iteratively solving (P2) and (P3-a) can provide an efficient way to reduce the SER gradually. Due to the non-convexity of the problem, our approach in Algorithm 3 only ensures a sub-optimal solution.

---

**Algorithm 2 MSER-Precoding Scheme**

1: **Initialization:** Given a feasible initial solution \( f_0, k = 0 \) and halting criterion \( \varepsilon_2 > 0 \).

2: **Gradient and Search direction:** Compute the gradient \( g_k \) and derive the search direction as:
\[
r_k = -\hat{g}_k = \Omega(f_k) f_k,
\] (34)
where \( \Omega(f_k) \) given in (33).

3: **Direction Projection:** Project the search direction into the tangent plane of \( \text{tr}(f^H) = P_{\text{max}} \)
\[
r_k^\perp = r_k - \langle r_k, f_k \rangle f_k / \| f_k \|^2;
\] (35)

4: **Search for \( \hat{\beta} \):** For \( 0 \leq \hat{\beta} \leq \pi/2 \), searching for it by
\[
\hat{\beta} = \arg \min_{\beta} P_S(f_k); \quad \beta
\] (36)

5: **Update:** Go to step 6 if \( \| r_k^\perp \| \leq \varepsilon_2 \), else let
\[
f_{k+1} = \cos \hat{\beta} \cdot f_k + \sin \hat{\beta} \cdot \sqrt{P_{\text{max}}} r_k^\perp / \| r_k^\perp \|;
\] (37)

6: **Output:** The optimized precoding matrix are thus given by:
\[
F^* = \begin{bmatrix}
f_{k1} & \cdots & f_{kN_s} \\
\vdots & \ddots & \vdots \\
f_{k(N_s-1)N_s} & \cdots & f_{kNN_s}
\end{bmatrix},
\] (38)
where \( f_{ki} \) is the \( i \)-th elements of \( f_k \).

---

**Algorithm 3 Two-Step MSER Algorithm**

1: **Initialization:** Given a feasible initial solution \( \phi_0, f_0, k = 0 \).

2: **Optimize the reflecting elements:** Based on \( f_k \), optimize the reflecting elements via Algorithm 1, which yields \( \phi_{k+1} \).

3: **Optimize the precoder:** Based on \( \phi_{k+1} \), optimize the precoder via Algorithm 2, which yields \( f_{k+1} \).

4: **Iteration:** Let \( k \leftarrow k + 1 \). Go to step 2 until convergence.

---

**D. Computational Complexity**

In this subsection, we analyze the computational complexity of the proposed Algorithms. Specifically, we analyze the complexity order of the Algorithm 1 and Algorithm 2 in the iterative procedure as follows.

1) **Complexity of Algorithm 1:** As can be seen from Algorithm 1, the computational complexity is primarily dominated by the gradient calculation, which involves i) calculating the \( M^{N_s} (M^{N_s} - 1) \) matrix multiplications \( \phi^H C_{ij} \phi \), and ii) calculating the \( \ell_p \) norm. Therefore, the complexity order of Algorithm 1 for each iteration is:
\[
C_1 = O \left[ M^{2N_s} N_s^2 + pN_s \right].
\] (40)
2) Complexity of Algorithm 2: Similar to Algorithm 1, the complexity of Algorithm 2 is also mainly consumed by the gradient calculation involving matrices $\mathbf{f}^H \mathbf{C}_1 \mathbf{f}$. Therefore, the complexity order of Algorithm 2 for each iteration is:

$$C_2 = O \left[ M^{2N_s} N_s^2 N_s^2 \right].$$

(41)

3) Complexity of Algorithm 3: Finally, by preserving the dominant terms, the overall complexity order of Algorithm 3 can be expressed as follows:

$$C_3 = O \left[ N_p M^{2N_s} N_s^2 + N \sum_{n=1}^{N_p} p^{(n)} + M^{2N_s} N_s^2 N_s^2 \right],$$

(42)

where $N_p = (p_{\text{max}} - p) / \Delta p$, $p^{(n)} = p + (n - 1) \Delta p$, and $p$, $\Delta p$, $p_{\text{max}}$ are specified in step 1 in Algorithm 1.

As will be shown in the simulations in Section V, the proposed Two-Step MSER Algorithm is able to achieve favorable error performance. However, the complexity imposes a great burden in the real-time shifter design at the RIS and precoder design at the transmitter, especially when the number of reflecting elements $N$ is large. To address this issue, we develop low-complexity solutions to solve (P1) problem in the following section.

IV. SIMPLIFIED OPTIMIZATION ALGORITHMS

To facilitate practical implementation, we investigate low-complexity algorithms in this section based on the Two-Step alternating optimization in Algorithm 3. Specifically, we will simplify the reflecting design and the precoding design separately. Then, by adopting the simplified algorithms into Algorithm 3, we can get the low-complexity solutions.

A. SDP-Reflecting Scheme

By using the exponential upper bound of $Q$-function, another SER upper bound can be given as

$$P_S(\mathbf{F}, \mathbf{F}) \leq \frac{1}{M^{N_s}} \sum_{i=1}^{M^{N_s}} \sum_{j=1, j \neq i}^{M^{N_s}} \exp \left( - \frac{\rho d_{ij}^2 (\mathbf{F}, \mathbf{F})}{4 \sigma^2} \right).$$

(43)

Since the RIS adds new supplementary links to maintain the communication link, the overall system performance could be increased by the in-direct multiple-path without additional interference leading to low required SNR. For this communication situation, \cite{25}, can be expanded by Taylor expansion and could be expressed as

$$P_S(\mathbf{F}, \mathbf{F}) \leq - \frac{\rho}{4M^{N_s} \sigma^2} \sum_{i=1}^{M^{N_s}} \sum_{j=1, j \neq i}^{M^{N_s}} d_{ij}^2 (\mathbf{F}, \mathbf{F})$$

$$+ M^{N_s} - 1.$$  

(44)

Thus, the optimization problem of (P1) is asymptotically equivalent to

(P4) : Given : $\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}_M$

Find : $\mathbf{F}, \mathbf{F}$

Maximize : $\sum_{i=1}^{M^{N_s}} \sum_{j=1, j \neq i}^{M^{N_s}} d_{ij}^2 (\mathbf{F}, \mathbf{F})$

Subject to : $\text{tr}(\mathbf{Q}) \leq F_{\text{max}}$

$0 \leq \theta_i \leq 2\pi, \forall i = 1, ..., N.$

(45)

Based on the given $\mathbf{F}$, the optimization problem can be written as

(P4-a) : Given : $\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}_M, \mathbf{F}$

Find : $\mathbf{\phi}$

Maximize : $\mathbf{\phi}^H \mathbf{\Gamma} \mathbf{\phi}$

Subject to : $|\phi_i| = 1, \forall i = 1, ..., N,$

where $\mathbf{\Gamma} = \mathbf{R}_{\mathbf{H}_2} \otimes \left( \sum_{i=1}^{M^{N_s}} \sum_{j=1, j \neq i}^{M^{N_s}} \Delta \mathbf{Q}_{ij} \right).$ However, this problem is NP-hard in general. Define $\mathbf{V} = \mathbf{\phi} \mathbf{\phi}^H$ with constraints $\mathbf{V} \succeq 0$ and $\text{rank}(\mathbf{V}) = 1$. The objective function can be re-expressed as $\mathbf{\phi}^H \mathbf{\Gamma} \mathbf{\phi} = \text{tr}(\mathbf{\Gamma} \mathbf{\phi} \mathbf{\phi}^H) = \text{tr}(\mathbf{\Gamma} \mathbf{V})$. Since the rank-one constraint is non-convex, semidefinite relaxation is used to relax this constraint which is similar to \cite{12}. As a result, problem is reduced to

(P4-b) : Given : $\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}_M, \mathbf{F}$

Find : $\mathbf{\phi}$

Maximize : $\text{tr}(\mathbf{\Gamma} \mathbf{V})$

Subject to : $\mathbf{V}_{i,i} = 1, \forall i = 1, ..., N$

$\mathbf{V} \succeq 0,$

(46)

which is a standard convex SDP problem. It can be optimally solved by existing convex optimization solvers such as CVX \cite{51}. Generally, the relaxed problem may not lead to a rank-one solution, i.e., $\text{rank}(\mathbf{V}) \neq 1$, which implies that the optimal objective value of the problem is an upper bound of (P4-b). Thus, additional steps are needed to construct a rank-one solution from the optimal higher-rank solution to problem. Specifically, we first obtain the eigenvalue decomposition of $\mathbf{V}$ as $\mathbf{V} = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^H,$ where $\mathbf{U} = [e_1, \ldots, e_N] \in \mathbb{C}^{N \times N}$ and $\mathbf{\Sigma} = \text{diag}(\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^{N \times N}$ are a unitary matrix and a diagonal matrix, respectively. Then, a suboptimal solution to (P4-b) as $\mathbf{\phi}^* = \mathbf{U} \mathbf{\Sigma}^{1/2} r$ can be obtained, where $r \in \mathbb{C}^{N \times 1}$ is a random vector generated according to $r \in \mathcal{CN}(0, \mathbf{I}_N)$. With independently generated Gaussian random vector $\mathbf{r}$, the objective value of (P4-b) is approximated as the maximum one attained by the best $\mathbf{\phi}$ among $\mathbf{r}$’s. Finally, the solution $\mathbf{\phi}^*$ can be recovered by $\mathbf{\phi}^* = e^{j \text{arsin}(\mathbf{r})}.$
B. MMED-Precoding Scheme

In the high SNR regime, the SER for a given channel can be simplified as

\[ P_S(\Phi, F) \approx \frac{\lambda}{MN_s} Q \left( \frac{\rho d_{\min}^2(\Phi, F)}{2\sigma^2} \right), \tag{48} \]

where \( \lambda \) is the number of neighbor points, and \( d_{\min}^2(\Phi, F) = \min_{i,j,i \neq j} d_{ij}^2(\Phi, F) \) is the minimum squared Euclidean distance between the noise-free received signal vectors. As \( P_S(\Phi, F) \) is a monotonically decreasing function of \( d_{\min}^2(\Phi, F) \), problem (P1) can be formulated as maximizing the minimum Euclidean distance problem [22].

\[
\text{(P5)}: \quad \text{Given: } H_1, H_2, S_M \\
\text{Find: } \Phi, F \\
\text{Maximize: } d_{\min}^2(\Phi, F) \\
\text{Subject to: } \text{tr}(Q) \leq P_{\text{max}} \\
0 \leq \theta_i \leq 2\pi, \forall i = 1, ..., N. 
\]

Based on given \( \Phi \) and [22], (P5) can be rewritten as

\[
\text{(P5-a)}: \quad \text{Given: } H_1, H_2, S_M, \Phi \\
\text{Find: } f \\
\text{Maximize: } \min_i \phi_i(\hat{C}_{ij})f \\
\text{Subject to: } \text{tr}(Q) \leq P_{\text{max}}. 
\]

By introducing an auxiliary variable \( r \), we have the equivalent epigraph of (P5-a) as

\[
\text{(P5-b)}: \quad \text{Given: } H, S_M \\
\text{Find: } f \\
\text{Maximize: } \phi_i(\hat{C}_{ij})f \geq r \forall i, j, i \neq j \\
\text{Subject to: } \text{tr}(f^H) \leq P_{\text{max}}. 
\]

However, we can review (P5-b) in another equivalent way [29] as

\[
\text{(P5-c)}: \quad \text{Given: } H, S_M \\
\text{Find: } f \\
\text{Minimize: } \|f\|^2 \\
\text{Subject to: } \phi_i(\hat{C}_{ij})f \geq d_{\min} \forall i, j, i \neq j, 
\]

where \( d_{\min} \) is the desired squared minimum distance. The rationale behind (P5-c) is to guarantee the minimum squared distance, while pursuing the minimum power usage as the objective. It can be seen that (P5-c) is a large-scale non-convex QCQP problem. Similar to [29], we use a new dual ascent method in terms of augmented Lagrangian capable of robustness to the dual ascent method. Firstly, we convert (P5-c) to a real-valued form by setting \( v = [\text{Re}\{f\}^T \text{Im}\{f\}^T]^T \in \mathbb{R}^{2N_s N_s \times 1} \) and a real matrix

\[
M_{ij} = \begin{bmatrix} \text{Re}\{C_{ij}\} & -\text{Im}\{C_{ij}\} \\ \text{Im}\{C_{ij}\} & \text{Re}\{C_{ij}\} \end{bmatrix}. \tag{53} 
\]

In this case, (P5-c) can be equivalently represented as

\[
\text{(P5-d)}: \quad \text{Given: } M_{ij} \\
\text{Find: } v \\
\text{Minimize: } v^T v \\
\text{Subject to: } v^T M_{ij} v \geq d_{\min} \forall i, j, i \neq j. \tag{54} 
\]

By transforming the inequality constraints \( v^T M_{ij} v \geq d_{\min} \) to the equality constraints by introducing a slack vector \( s = \{s_{ij}\} \) and having \( v^T M_{ij} v = d_{\min} + s_{ij} \) with \( s_{ij} \geq 0 \), the augmented Lagrangian for (P5-d) can be written as

\[
L(v, s, \lambda, \kappa) = v^T v - \sum_{i=1}^{M_s} \sum_{j=1,j \neq i}^{M_s} \lambda_{ij} (v^T M_{ij} v - s_{ij} - d_{\min}) \\
+ \frac{\kappa}{2} \sum_{i=1}^{M_s} \sum_{j=1,j \neq i}^{M_s} \lambda_{ij} (v^T M_{ij} v - s_{ij} - d_{\min})^2, \tag{55} 
\]

where \( \lambda = \{\lambda_{ij}\} \) is the dual vector, \( \sum_{i=1}^{M_s} \sum_{j=1,j \neq i}^{M_s} \lambda_{ij} (v^T M_{ij} v - s_{ij} - d_{\min}) \) is the penalty parameter. Thus, the problem can be transformed as minimizing the augmented Lagrangian \( L(v, s, \lambda, \kappa) \) with respect to \( v \) and \( s \) at the \( k \)-th iteration, which is given by

\[
\text{(P5-e)}: \quad \text{Given: } M_{ij} \\
\text{Find: } v, s \\
\text{Minimize: } L(v, s, \lambda_k, \kappa_k) \\
\text{Subject to: } s \geq 0. \tag{56} 
\]

It is clear that for each iteration, \( s_{ij} \) occurs in just two terms of (55), which is in fact a convex quadratic function with respect to each of these slack variables. Therefore, the minimization process in (P5-e) can be carried out with respect to each of the \( s_{ij} \) separately. With \( \nabla_s L(v, s, \lambda, \kappa) = 0 \), we have

\[
s_{ij} = v^T M_{ij} v - d_{\min} - \frac{\lambda_{ij,k}}{\kappa_k}. \tag{57} 
\]

If this unconstrained minimizer is smaller than the lower bound of 0, then the optimal value of \( s_{ij} \) is 0 since \( s_{ij} \geq 0 \). The optimal value of \( s_{ij} \) is therefore given by

\[
s_{ij} = \left[ v^T M_{ij} v - d_{\min} - \frac{\lambda_{ij,k}}{\kappa_k} \right]^+. \tag{58} 
\]

Substituting (55) into \( L(v, s, \lambda, \kappa) \), we can obtain an equivalent form for \( L(v, \lambda_k, \kappa_k) \) without \( s \), which can be rewritten at the top of the next page, where the function \( \phi(z, a, b) \) is
\[ L(v, \lambda_k, \kappa_k) = \begin{cases} v^T v - \sum_{i=1}^{M_N} \sum_{j=1,j \neq i}^{M_N} \lambda_{ij} (v_j M_{ij} v - d_{min}) + \frac{\gamma}{2} \sum_{i=1}^{M_N} \sum_{j=1,j \neq i}^{M_N} \lambda_{ij} (v_j M_{ij} v - d_{min})^2, & s_{ij} \leq 0 \\ v^T v + \sum_{i=1}^{M_N} \sum_{j=1,j \neq i}^{M_N} \phi (v_j M_{ij} v - d_{min}, \lambda_{ij,k}, \kappa_k), & \text{otherwise} \end{cases} \]

(59)

\[ \phi(z, a, b) \Delta \begin{cases} -az + \frac{b}{2} z^2, z - \frac{a}{b} \leq 0 \\ -\frac{a^2}{2b}, \text{otherwise} \end{cases} \] (60)

Algorithm 4 Minimizing the SER Based on Given \( \Phi \)

1. **Initialization**: Given a feasible initial solution \( S_0, v_0, \lambda_{ij,k} > 0, \kappa_0 > 0, \beta > 0, l = 0, k = 0 \) and halting criterion \( \varepsilon_4 > 0 \).
2. **Gradient and Search direction**: Compute the gradient \( \bar{g}_l = \nabla_v L(v_l, \lambda_k, \kappa_k) \) and derive the search direction as:

\[ d_l = -S_l \bar{g}_l - S_l \nabla_v L(v_l, \lambda_k, \kappa_k); \] (66)

3. **Update and define vectors**: Compute \( v_{l+1} \) as

\[ v_{l+1} = v_l + \alpha_l d_l, \] (67)

where \( \alpha_l \) is computed by a line search procedure to satisfy the Wolfe conditions [30]. Compute \( \bar{g}_{l+1} \) and define vectors as \( \delta_l = v_{l+1} - v_l \) and \( \gamma_l = \bar{g}_{l+1} - \bar{g}_l \).
4. **Iteration**: Go to step 5 if \( \bar{g}_{l+1} = 0 \), else compute the approximate inverse Hessian matrix.

\[ S_{l+1} = S_l + \left( 1 + \frac{\gamma_l^2 \delta_l}{\gamma_l^T \delta_l} \right) \frac{\delta_l \delta_l^T - \delta_l \gamma_l \gamma_l^T \delta_l}{\gamma_l^T \delta_l}, \] (68)

\[ l \leftarrow l + 1 \) and then go to step 2.
5. **Update**: Let \( v^* = v_{l+1} \) and go to step 6 if \( \|v_{l+1} - v_0\|_2 < \varepsilon_4 \), else compute the Dual vector \( \lambda_{k+1} \) and penalty parameter respectively as:

\[ \lambda_{ij,k+1} = [\lambda_{ij,k} - \kappa_k (v_{l+1} M_{ij} v_{l+1} - d_{min})]^+ \] (69)

and

\[ \kappa_{k+1} = \beta \kappa_k, \] (70)

and let \( v_0 = v_{l+1}, l = 0, k \leftarrow k + 1 \) and go to step 2.
6. **Output**: The optimized precoding matrix are thus given by:

\[ F^* = \begin{bmatrix} v_1^* & iv_{N_x} N_{x+1} & \cdots & iv_{N_x} N_{(N_x+1)} \\ \vdots & \ddots & \cdots & \vdots \\ v_{(N_t-1)N_x}^* + iv_{(2N_t-1)N_x} & \cdots & v_{N_x N_{N_t}} + iv_{2N_x N_{N_t}} \end{bmatrix}, \] (71)

where \( v_i^* \) is the \( i \)-th elements of \( v^* \).

C. MRT-Precoding Scheme

Actually, the system SER is related to the received SNR. Specifically, SER decreases with the increase of the received SNR. Thus, this problem can be transformed into maximizing

\[ L(v, \lambda_k, \kappa_k) \]
the received signal power subject to the maximum transmit power constraint at the transmitter, which can be formulated as

\[(P6) \quad \text{Given : } H_1, H_2, S_M, \Phi \]
\[\text{Find : } F \]
\[\text{Maximize : } \|HF\|^2_F \]
\[\text{Subject to : } \text{tr}(Q) \leq P_{\text{max}}. \]

(72)

For any given phase shifts $H$, it can be verified that MRT is one of the widely-adopted transmit beamforming solutions to problem (P5) \[^{[35]}\], i.e. $F = \sqrt{P_{\text{max}}}H_{[1:N_s]}$, where $H(:,[1 : N_s])$ denotes the matrix that contains the first $N_s$ columns in $H$. Note that for this case, $N_s$ should satisfy that $N_s \leq N_r$.

\[D. \text{Eigen-Precoding Scheme} \]

To obtain another solution to (P6), we first obtain the eigenvalue decomposition of $H^H H$ as $H^H H = U \Lambda U^H$, where $U = [e_1, \ldots, e_{N_r}]$ and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{N_r})$ are a unitary matrix and a diagonal matrix conditioned with $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_{N_r}$, respectively. Assuming $N_r \geq N_t \geq N_s$, $\|HF\|^2_F$ can be re-expressed as

$$\|HF\|^2_F = \text{tr} \left\{ H^H H \right\} = \text{tr} \left\{ H \Lambda H^H \right\},$$

(73)

We notice that (P6) is maximized when the power constraint is met with strict equality. Hence $\text{tr}(Q) = \text{tr} (FF^H) = \|F\|^2_F = P_{\text{max}}$. Without considering power allocation, we assume that $F = \sqrt{P_{\text{max}}}W$, where $F$ has orthonormal columns. Therefore, (73) can be re-expressed as

$$\|HF\|^2_F = \frac{P_{\text{max}}}{N_s} \text{tr} \left\{ U \Lambda \Lambda U^H \right\}$$

$$= \frac{P_{\text{max}}}{N_s} \sum_{k=1}^{N_r} \lambda_k \left( U^H \Lambda U \right)_{kk}$$

$$= \frac{P_{\text{max}}}{N_s} \sum_{k=1}^{N_r} \lambda_k e_k^H \Lambda e_k.$$ 

(74)

$$\text{tr} \left\{ U^H \Lambda U \right\} = \text{tr} (W^H W) = \|W\|^2_F = N_s, \quad (P6)$$

can be transformed as

\[(P6-a) \quad \text{Given : } H_1, H_2, S_M, \Phi \]
\[\text{Find : } W \]
\[\text{Maximize : } \sum_{k=1}^{N_r} \lambda_k e_k^H \Lambda e_k \]
\[\text{Subject to : } \sum_{k=1}^{N_r} e_k^H \Lambda e_k = N_s. \]

(75)

**Proposition 4.** The $k$th column of the solution to (P6-a), $W$ is the $k$th eigenvector of $H^H H$ corresponding to $k$-th eigenvalue with $\lambda_k \geq \lambda_{k+1}$.

**Proof.** Since $W$ has orthonormal columns. We can find that $e_k^H W W^H e_k = e_k^H W (W^H W)^{-1} W^H e_k$, where $W (W^H W)^{-1} W^H$ is orthogonal projection matrix with eigenvalues 0 or 1. Thus, $e_k^H W W^H e_k \leq \|e_k\|^2 \lambda_{\text{max}} \left( W (W^H W)^{-1} W^H \right) = 1$. It is obvious that the solution to (P6-a) is $e_k^H W W^H e_k = 1$, $\forall k = 1, ..., N_s$ and (P6-a) is $e_k^H W W^H e_k = 0$, $\forall k = N_s, ..., N_r$. Accordingly, the $k$th columns of $W$ is the $k$th eigenvector of $H^H H$.

The proof is completed. □

\[E. \text{Computational Complexity Analysis} \]

In this subsection, we analysis the computational complexity of the simplified algorithms.

1) **Complexity of SDP-Reflecting Scheme:** For the simplified Algorithm for $\Phi$, the computational complexity mainly incurs the complexity of computing the target function $\Gamma \text{V}$ and find every diagonal elements by standard Lagrangian algorithm, which is $\mathcal{O}[N^4]$.

2) **Complexity of MMED-Precoding Scheme:** As can be seen from Algorithm 4, the computational complexity is also mainly consumed by the gradient calculation involving matrices $F^H C_i f$, which is same as Algorithm 2 for each iteration. However, it should be mentioned that Algorithm 4 does not need the SNR information. In other words, MMED-Precoding are applicable for any SNR values, while MSER-Precoding needs to be re-designed as SNR varies. Besides, the number of iterations required for Algorithm 4 is much less owing to the use of BFGS method \[^{[29]}\]. Thus, in general, the complexity of MMED-Precoding scheme is much less than the MMSR-Precoding scheme.

3) **Complexity of MRT-Precoding Scheme:** For the MRT-Precoding scheme in Section IV-C, the computational complexity can be easily analyzed to be $\mathcal{O}[N_r N_s]$.

4) **Complexity of Eigen-Precoding Scheme:** The computational complexity of the Eigen-Precoding scheme in Section IV-D is consumed by computing the eigenvectors, which is $\mathcal{O}[N_r^3]$.

\[V. \text{Numerical Results} \]

In this section, we first investigate the performance of the proposed phase shifts design schemes, precoder design schemes and Two-Step algorithm with different reflecting scheme and precoding scheme combinations in variously configured $(N_r, N, N_t, N_s, M, K)$ systems, where $K$ is the Rician fading parameters. Then, we demonstrate the SER performance with various large $N$ reflecting elements to investigate the impact of the size of the surface on the performance. Finally, to show the robustness of the proposed schemes in the presence of CSI estimation errors, the system performance using perfect CSI and imperfect CSI is compared.
A. Superiority of the Proposed Schemes

Firstly, we compare our proposed phase shifts design with random phase shifts and fixed phase shifts in Fig. 2 over 1000 Rician fading channel realizations in a \((2, 3, 2, 2, 2, 2)\) RIS-assisted MIMO system. We can see that the performance under random phase shifts is close to that with fixed phase shifts while the other two optimized phase shifts provide apparent favorable SER performance. This implies that the phase shifts controller could make a significant influence on the system performance. Moreover, the proposed high-complexity MMSE-Reflecting scheme is 2-4 dB better than the low-complexity SDP-Reflecting scheme. In the depicted SNR regime, the performance gap increases with the increasing SNR.

Secondly, we compared precoding design schemes over 1000 Rician fading channel realizations in a \((3, 5, 3, 2, 4, 3)\) system as illustrated in Fig. 3. It is obvious that all the proposed schemes could provide lower SER than random precoding. We can see that the proposed MSER-Precoding scheme has a slightly better performance than MMED-Precoding scheme. As analyzed in Section V, the MMED-Precoding is of lower complexity than MSER-Precoding. Thus, in realistic systems, MMED-Precoding is a more appealing scheme compared to MSER-Precoding. Moreover, it is demonstrated that the MRT-Precoding is about 1 dB better than the Eigen-Precoding, while both are about 2-3 dB worse than MSER-Precoding and MMER-Precoding with significant lower computation complexity.

Thirdly, we substitute all the proposed phase shift design schemes and precoding schemes into the Two-Step algorithm (i.e., Algorithm 3) leading to 8 different Reflecting-Precoding combinations over 1000 Rician fading channel realizations in a \((3, 10, 3, 2, 2, 5)\) system. It is observed that MSER-MSER is almost the same as MSER-MMED. This is because there is only a very small gap between the performance of MSER-Precoding scheme and MMED-Precoding scheme as observed in Fig. 3. Similarly, we can see that SDP-MMSE almost overlaps with SDP-MMED. MSER-MRT is the third-optimal among the eight combinations since MSER-Reflecting is much better than other reflecting schemes. Moreover, it is found that the performance gap between MSER-Reflecting and SDP-Reflecting is smaller than the performance gap between MMED-Precoding and MRT-Precoding.

From the simulation results in Figs. 2-4, we can conclude that proposed shift design schemes, precoding schemes, and various Two-Step combination schemes can notably improve the SER in RIS-assisted MIMO systems. There is an obvious trade-off between complexity and performance. MSER schemes could achieve the best system performance at the cost...
of high computation complexity. MMED-Precoding could gain similar system performance with lower computation complexity compared to MSER-Precoding. MRT-Precoding and Eigen-Precoding have a worse performance but with significant lower computation complexity. In real system design, the trade-off can be made based on the users’ quality of the service (QoS) requirements.

### B. Impact of the Number of Reflecting Elements

In this subsection, we compare the performance in RIS-assisted MIMO systems with various large $N$s. In Fig. 5, we show the actual SER and SER bound for MSER-Reflecting and SDP-Reflecting over 1000 Rician fading channel realizations in $(3, 20, 3, 2, 2, 2)$ and $(3, 60, 3, 2, 2, 2)$ systems. It is clear that the union bound on SER in (3) can perfectly match the simulation ones in low SER region. We can see that the SER is greatly improved about 7-10 dB with more 40 RIS reflecting elements. The SER of SDP-Reflecting with $N = 60$ is also much better than the SER of MSER-Reflecting with $N = 20$.

Similarly, in $(5, 25, 3, 2, 4, 3)$ and $(5, 100, 3, 2, 4, 3)$ systems as illustrated in Fig. 6, we observe that the system SER can be greatly decreased by increasing the reflecting elements at RIS. The rational behind is that more reflecting elements lead to more channel gains. Moreover, we can see that Eigen-Precoding has better SER gain than MRT-Precoding with large $N$, which is opposite to small $N$ condition in Fig. 3.

In Fig. 7, we show the influence of $N$ on the phase shift design, precoding design and joint design. From Fig. 7, we can see that phase shift design will make a more important role than precoding design on affecting SER when $N$ is large. There is no obvious SER gain with larger $N$ if the reflecting is not carefully designed. Obviously in Fig. 7, the joint design could gain the optimal performance at the cost of more iteration to gain better solution. There is 1 dB SER gain with 10 more reflecting elements if both the reflecting and precoding are optimized.

In conclusion, the system SER can be greatly decreased by equipping more reflecting elements at the RIS and using the proposed reflecting design. $N$ will not provide any SER gain without phase shift design.

### C. Robustness in Presence of CSI Estimation Errors

To show the robustness of the proposed schemes in the presences of CSI estimation errors, we further evaluate the error performance of our schemes with imperfect CSI. SDP-Eigen scheme and MSER-MSER scheme are chosen as examples. Such imperfection originates from channel estimation...
VI. CONCLUSION

Joint designs for phase shifts at RIS and full rank precoder at the transmitter was proposed to minimize the SER in point-to-point RIS-assisted MIMO systems. The proposed MSER-Reflecting and MSER-Precoding algorithm can achieve a favorable system performance while simplified solutions, i.e., MSER-Reflecting, MMED-Precoding, MRT-Precoding and Eigen-Precoding could also improve system performance with lower computation complexity. Simulation results demonstrate that our proposed algorithm can significantly improve SER performance with different scheme combinations.

REFERENCES

[1] S. Zhang, Q. Wu, S. Xu, and G. Y. Li, “Fundamental green tradeoffs: Progresses, challenges, and impacts on 5G networks,” IEEE Commun. Surveys Tuts., vol. 19, no. 1, pp. 33-56, First Quarter 2017.
[2] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, and H. Gacanin, “Smart radio environments empowered by AI reconfigurable meta-surfaces: An idea whose time has come,” EURASIP J. Wireless Commun. Netw., vol. 2019, p. 129, May 2019.
[3] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, “Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays,” in Proc. IEEE INFOCOM, Honolulu, USA, Apr. 2018, pp. 270-278.
[4] S. V. Hum and J. Perrussseau-Carrier, “Reconfigurable reflectarrays and array lenses for dynamic antenna beam control: A review,” IEEE Trans. Antennas Propagat., vol. 62, no. 1, pp. 183-198, Jan. 2014.
[5] S. Foo, “Liquid-crystal reconfigurable metasurface reflectors,” in Proc. IEEE ISAP, California, USA, July 2017, pp. 2069-2070.
[6] C. Liaskos, A. Tsiosis, A. Pitsillides, S. Ioannidis, and I. F. Akyildiz, “Using any surface to realize a new paradigm for wireless communications,” CoRR, vol. abs/1806.04585, 2018. [Online]. Available: http://arxiv.org/abs/1806.04585
[7] J. Jun Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, “Coding metamaterials, digital metamaterials and programmable metamaterials,” Light, Sci. Appl., vol. 3, no. 10, p. e218, Oct. 2014.
[8] B. Sainath and N. B. Mehta, “Generalizing the amplify-and-forward relay gain model: An optimal SEP perspective,” IEEE Trans. Wireless Commun., vol. 11, no. 11, pp. 4118-4127, Nov. 2012.
[9] G. Yang, C. K. Ho and Y. L. Guan, “Multi-antenna wireless energy transfer for backscatter communication systems,” IEEE J. Sel. Areas Commun., vol. 33, no. 12, pp. 2974-2987, Dec. 2015.
[10] S. Hu, F. Rusek, and O. Edfors, “Beyond massive MIMO: The potential of data transmission with large intelligent surfaces,” IEEE Trans. Signal Process., vol. 66, no. 10, pp. 2746-2758, May 2018.
[11] L. Subrt and P. Pechac, “Intelligent walls as autonomous parts of smart indoor environments,” IET Commun., vol. 6, no. 8, pp. 1004-1010, May 2012.
[12] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design,” in Proc. IEEE GLOBECOM, Abu Dhabi, UAE, Dec. 2018, pp. 1-6.
[13] Q. Wu and R. Zhang, “Beamforming optimization for intelligent reflecting surface with discrete phase shifts,” in Proc. IEEE ICASSP, Brighton, UK, May 2019, pp. 7830-7833.
[14] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini “Asymptotic analysis of large intelligent surface assisted MIMO communication,” submitted to IEEE Trans. Wireless Commun., [Online]. Available: [http://arxiv.org/abs/1803.08127]
[15] E. Basar, “Transmission through large intelligent surfaces: A new frontier in wireless communications,” CoRR, vol. abs/1902.08463, 2019. [Online]. Available: [http://arxiv.org/abs/1902.08463]
[16] Y. Han, W. Tang, S. Jin and X. Ma, “Large intelligent surface-assisted wireless communication exploiting statistical CSI,” CoRR, vol. abs/1812.05429, 2018. [Online]. Available: [http://arxiv.org/abs/1812.05429]
[17] M. Jung, S. Wald, R. J. Young, K. Gyueyole and C. Sooyong, “Performance analysis of large intelligent surfaces (LISs): Asymptotic data rate and channel hardening effects,” CoRR, vol. abs/1810.05667, 2018. [Online]. Available: [http://arxiv.org/abs/1810.05667]
[18] M. Jung, S. Wald and K. Gyueyole, “Performance analysis of large intelligent surfaces (LISs): Uplink spectral efficiency and pilot training,” CoRR, vol. abs/1904.00453, 2019. [Online]. Available: [http://arxiv.org/abs/1904.00453]
[19] M. Jung, W. Saad, Y. Jiang, G. Kong, and S. Choi, “Uplink data rate in large intelligent surfaces: Asymptotic analysis under channel estimation errors,” in Proc. IEEE SPAWC, Cannes, France, July 2019, pp. 1-6.
[20] E. Basar, “Large intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G,” CoRR, vol. abs/1904.06704, 2019. [Online]. Available: [http://arxiv.org/abs/1904.06704]
[21] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” CoRR, vol. abs/1810.06934, 2018. [Online]. Available: [http://arxiv.org/abs/1810.06934]
[23] C. Huang, A. Zappone, M. Debbah, and C. Yuen, “Achievable rate maximization by passive intelligent mirrors,” in Proc. IEEE ICASSP, Calgary, Canada, Apr. 2018, pp. 1-6.

[24] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, “Increasing indoor spectrum sharing capacity using smart reflect-array,” in Proc. IEEE ICC, Kuala Lumpur, Malaysia, May 2016, pp. 1-6.

[25] X. Zhang, Matrix Analysis and Applications. Tsinghua, China: Tsinghua Univ. Press, 2004.

[26] L. He, J. Wang and J. Song, “Spatial modulation for more spatial multiplexing: RF-chain-limited generalized spatial modulation aided MM-Wave MIMO With hybrid precoding,” IEEE Trans. Commun., vol. 66, no. 3, pp. 986-998, Mar. 2018.

[27] S. P. Boyd and L. Vandenberghe, Convex Optimization. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[28] W. Wang and W. Zhang, “Diagonal precoder designs for spatial modulation,” in Proc. IEEE ICC, London, UK, June 2015, pp. 2411-2415.

[29] P. Cheng, Z. Chen, J. A. Zhang, Y. Li and B. Vucetic, “A unified precoding scheme for generalized spatial modulation,” IEEE Trans. Commun., vol. 66, no. 6, pp. 2502-2514, June 2018.

[30] A. Antoniou and W.-S. Lu, Practical Optimization: Algorithms and Engineering Applications. New York, NY, USA: Springer, 2007.

[31] M. Grant and S. Boyd, “CVX: MATLAB software for disciplined convex programming.” 2016. [Online] Available: http://cvxr.com/cvx

[32] S. Guo, H. Zhang, P. Zhang, S. Dang, C. Liang and M. S. Alouini, “Signal shaping for generalized spatial modulation and generalized quadrature spatial modulation,” to appear in IEEE Trans. Wireless Commun., [Online]. Available: https://ieeexplore.ieee.org/document/8734877.

[33] A. Skajaa, “Limited Memory BFGS for Non-smooth Optimization,” M.S. Thesis, Dept. Comput. Sci. Math., Courant Inst. Math. Sci., New York, NY, USA, 2010.

[34] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Found. Trends Mach. Learn., vol. 3, no. 1, pp. 1-122, Jan. 2011.

[35] D. Tse and P. Viswanath, Fundamentals of Wireless Communication. Cambridge, U.K.: Cambridge Univ. Press, 2005.