Beyond Diagonal RIS Aided Power Minimization Beamforming for MIMO Systems

Junwen Yang*, Ang Li*, Xuewen Liao*, Christos Masouros†, and A. Lee Swindlehurst§
School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, China*
Department of Electronic and Electrical Engineering, University College London, London, UK†
Center for Pervasive Communications and Computing, University of California, Irvine, CA 92697, USA§
Email: jwyang@stu.xjtu.edu.cn, ang.li.2020@xjtu.edu.cn, yeplos@mail.xjtu.edu.cn, c.masouros@ucl.ac.uk, swindle@uci.edu

***Abstract*—Beyond diagonal reconfigurable intelligent surfaces (BD-RIS) with interconnected reflecting elements present an emerging technology for manipulating the propagation environment, and their new structure requires careful investigation. In this paper, we explore BD-RIS-aided power minimization beamforming, where the BD-RIS scattering matrix and transmit beamforming are jointly optimized under nonconvex constraints related to signal-to-interference-plus-noise ratio (SINR) thresholds and the structure of the scattering matrix. To efficiently solve the problem, we propose a single-loop algorithm, where we adopt a variable splitting strategy with an auxiliary variable to split the scattering matrix, and then alternately update the resulting variables. Through further derivations, we show that each nonconvex subproblem can be solved efficiently. Simulation results demonstrate the high efficiency of our proposed single-loop algorithm and the effectiveness of BD-RIS in improving performance.**

***Index Terms*—Beyond diagonal reconfigurable intelligent surfaces (BD-RIS), power minimization, beamforming, MIMO.**

I. Introduction

RECONFIGURABLE intelligent surfaces (RIS) have emerged as a pivotal technique in the evolution of future ubiquitous and intelligent wireless communications [1], [2]. Composed of a large number of passive reflecting elements made from metamaterials, RIS are known for their ability to adaptively control the reflection of impinging electromagnetic signals with extremely low power consumption and without the need for radio frequency (RF) chains.

Owing to their intelligent reflection capabilities and low-cost operation, RIS have been widely studied in wireless communications for channel shaping [3]–[5] and signal enhancement [6]–[8]. For example, [3] investigates passive beamforming for multi-user multiple-input multiple-output (MIMO) systems with one-bit analog-to-digital converters, where the cascaded channel is designed to minimize the symbol error rate. The authors of [4] study joint optimization of the transmit power and RIS reflection coefficients for orthogonal frequency division multiplexing (OFDM) systems, aiming to maximize the achievable rate. In [5], an RIS is used to increase the rank of the channel matrix, and a closed-form solution is derived to increase the multiplexing gain and channel capacity. In these channel-shaping studies, the RIS functions as a standalone device for programming the propagation environment. For signal enhancement scenarios, the RIS collaborates with active beamforming at the transmitter to more effectively optimize the received signal. In [6], the authors propose minimizing transmit power through the joint optimization of active beamforming at the transmitter and passive beamforming at the RIS, and solve the problem using semidefinite relaxation (SDR). The authors of [7] consider joint optimization of the transmit power allocation for all users and the reflection coefficients of the RIS to maximize energy efficiency. An RIS-aided weighted sum-rate maximization problem is investigated in [8] using fractional programming.

The work cited above assumes a conventional RIS consisting of isolated reflecting elements that result in a diagonal reflection matrix. Such RIS can be referred to as diagonal RIS (D-RIS). Building on this, the concept of beyond diagonal RIS (BD-RIS) with interconnected reflecting elements has recently been proposed for enhanced design [9], [10]. The elements of a BD-RIS are connected together via reconfigurable impedances that provide additional degrees of freedom (DoFs) for manipulation of the reflected signal. These DoFs can be exploited to provide significant performance advantages compared to conventional D-RIS [11]. However, the use of BD-RIS also introduces new challenges for scattering matrix optimization, necessitating new signal processing approaches. The authors in [12] address the problem of optimizing the BD-RIS passive beamforming to maximize received signal power, and they derive a corresponding closed-form solution. However, this solution cannot be extended to other BD-RIS scenarios and does not consider coordination with the active transmit beamforming. Although [13] proposes a unified approach to optimize the BD-RIS configuration based on the penalty dual decomposition (PDD) method, it includes an outer alternating optimization layer and an inner PDD layer with two nested layers, which results in high computational complexity that limits its practical implementation.

In this paper, we consider BD-RIS-aided power minimization beamforming in a MIMO system, where the BD-RIS scattering matrix and transmit beamforming are jointly optimized to minimize transmit power while adhering to individual signal-to-interference-plus-noise ratio (SINR) constraints, as well as constraints on the BD-RIS scattering matrix. In addition to the SINR performance, the BD-RIS scattering matrix must be constrained to be a symmetric and orthogonal matrix. This means that the feasible region for the solution is the intersection of three different subspaces or manifolds, which

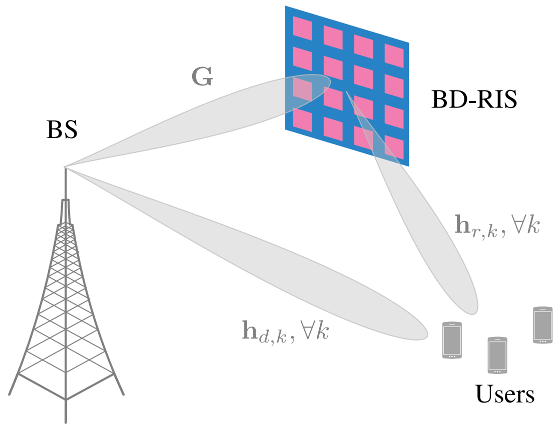

Fig. 1. The BD-RIS-aided multi-antenna beamforming system.

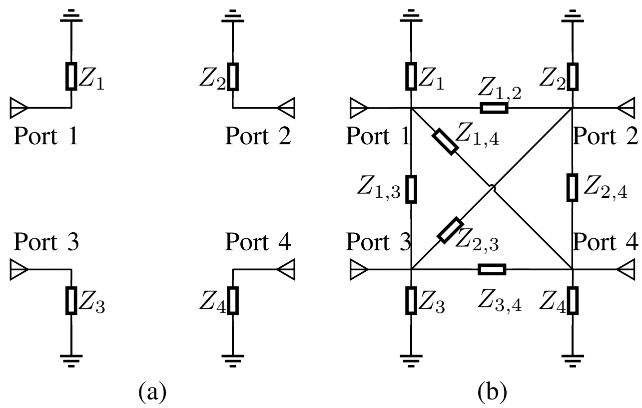

Fig. 2. (a) 4-element/port single-connected RIS. (b) 4-element/port fully-connected RIS.

accounts for the main obstacle in solving the problem. To solve this joint optimization problem with multiple variables and constraints, we split the scattering matrix into two parts by introducing an auxiliary variable that is equal to the scattering matrix itself. We allocate different constraints to the resultant variables: one for the SINR and symmetric constraints, and the other for the orthogonality constraint. We propose a novel single-loop algorithm to solve this nonconvex problem. By leveraging the alternating direction method of multipliers (ADMM), we decompose the problem into three subproblems, each of which can be solved efficiently using the phase invariance of the SINR and the singular value decomposition (SVD). We present simulation results demonstrating the high efficiency of the proposed single-loop algorithm. These results also show that the BD-RIS-aided system requires significantly lower transmit power compared to conventional D-RIS.

II. System Model and Problem Formulation

Consider a BD-RIS-aided multi-user multiple-input single-output (MU-MISO) communication system, where a base station (BS) with N_t transmit antennas serves N_r single-antenna users in the same resource block, and an N_i-element BD-RIS is deployed as part of the communication link to enhance the performance, as shown in Fig. 1. We use $\mathbf{\Phi} \in \mathbb{C}^{N_i \times N_i}$ to denote the scattering matrix of the BD-RIS. The reconfigurable impedance network of a 4-element/port single-connected and fully-connected RIS are compared in Fig. 2 [9]. While the single-connected D-RIS has only 4 reconfigurable impedances, the fully-connected BD-RIS has 9. Therefore, the fully-connected BD-RIS clearly provides increased DoFs. In this paper, we consider a lossless fully-connected BD-RIS model [9], [10], which has a symmetric and orthogonal scattering matrix, i.e.,

$$\mathbf{\Phi} = \mathbf{\Phi}^T, \mathbf{\Phi}^H \mathbf{\Phi} = \mathbf{I}. \tag{1}$$

Single- and group-connected BD-RIS can be regarded as special cases of the architecture.

The wireless channels between the BS and user k, between the BS and the BD-RIS, and between the BD-RIS and user k are denoted as $\mathbf{h}_{d,k}$, $\mathbf{G}$, and $\mathbf{h}_{r,k}$, respectively. The stacked channels are expressed as $\mathbf{H}_d \triangleq [\mathbf{h}_{d,1}, \cdots, \mathbf{h}_{d,N_r}]$ and $\mathbf{H}_r \triangleq [\mathbf{h}_{r,1}, \cdots, \mathbf{h}_{r,N_r}]$. Given the modulated data symbol vector $\mathbf{s} \triangleq [s_1, \cdots, s_{N_r}]^T$, the received signal for user k is given by:

$$y_k = \left(\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G}\right) \sum_{i=1}^{N_r} \mathbf{w}_i s_i + n_k, \tag{2}$$

where $\mathbf{w}_k$ denotes the beamforming vector for user k, and $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ denotes additive white Gaussian noise. Assuming that $|s_k| = 1, \forall k$, the SINR for user k is given by

$$\text{SINR}_k \triangleq \frac{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{\sum_{j \neq k}^{N_r} \left|\mathbf{h}_k^H \mathbf{w}_j\right|^2 + \sigma_k^2}, \tag{3}$$

where $\mathbf{h}_k^H \triangleq \mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G}$ denotes the cascaded channel.

In this paper, we focus on the problem of optimizing the active and passive beamforming to minimize the transmit power while ensuring that the received SINR of each user is no less than a given threshold. Meanwhile, the scattering matrix of the BD-RIS is subject to symmetric and orthogonality constraints. The considered optimization problem is formulated as follows:

$$\min_{\mathbf{W}, \mathbf{\Phi}} \sum_{i=1}^{N_r} \left\|\mathbf{w}_i\right\|^2 \tag{4}$$

$$\text{s.t.} \frac{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{\sum_{j \neq k}^{N_r} \left|\mathbf{h}_k^H \mathbf{w}_j\right|^2 + \sigma_k^2} \geq \gamma_k, \forall k, \tag{5}$$

$$\mathbf{\Phi} = \mathbf{\Phi}^T, \mathbf{\Phi}^H \mathbf{\Phi} = \mathbf{I}, \tag{6}$$

where γ_k denotes the prescribed SINR threshold for user k, and $\mathbf{W} \triangleq [\mathbf{w}_1, \cdots, \mathbf{w}_{N_r}]$. Note that (4) involves the joint optimization of $\mathbf{W}$ and $\mathbf{\Phi}$, which makes it difficult to achieve a reasonable trade-off between performance and complexity. Moreover, this problem is nonconvex not only because of the fractional SINR constraint on $\mathbf{W}$ and $\mathbf{\Phi}$, but also due to the nonconvex orthogonality constraint on $\mathbf{\Phi}$.

III. Single Loop Algorithm

In this section, we address the difficulties inherent in the problem structure of (4). We propose a single-loop algorithm to iteratively solve the joint optimization problem. Specifically, we begin by incorporating the SINR, symmetric, and orthogonality constraints into the objective using indicator functions, resulting in a more tractable formulation. The steps for solving the reformulated problem using the single-loop algorithm are then detailed.

Since the scattering matrix $\mathbf{\Phi}$ is subject to three different kinds of constraints, the feasible region of $\mathbf{\Phi}$ is complicated to enforce. Therefore, we use a variable splitting strategy to address this issue. We introduce a copy of $\mathbf{\Phi}$ as an auxiliary variable, denoted as $\mathbf{\Psi}$. The auxiliary variable essentially serves as a splitting operator for $\mathbf{\Phi}$. The new problem with the variable splitting strategy can then be formulated as:

$$\min_{\mathbf{W},\mathbf{\Phi},\mathbf{\Psi}} \sum_{i=1}^{N_r} \|\mathbf{w}_i\|^2 \tag{7}$$

$$\text{s.t.} \frac{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{\sum_{j\neq k}^{N_r} \left|\mathbf{h}_k^H \mathbf{w}_j\right|^2 + \sigma_k^2} \geq \gamma_k, \forall k, \tag{8}$$

$$\mathbf{\Phi} = \mathbf{\Phi}^T, \mathbf{\Psi}^H \mathbf{\Psi} = \mathbf{I}, \tag{9}$$

$$\mathbf{\Phi} = \mathbf{\Psi}. \tag{10}$$

In the next step, we retain the newly introduced equality constraint $\mathbf{\Phi} = \mathbf{\Psi}$ and incorporate the other constraints into the objective function. Let the feasible region of $\mathbf{W}$ and $\mathbf{\Phi}$, defined by the SINR and symmetric constraints, be denoted as $\mathcal{C}$, and let the feasible region of $\mathbf{\Psi}$, defined by the orthogonality constraint, be denoted as $\mathcal{D}$. The indicator function associated with the two feasible regions can then be expressed as follows:

$$\mathbb{I}_{\mathcal{C}}(\mathbf{W},\mathbf{\Phi}) = \begin{cases} 0, \text{if } \mathbf{W},\mathbf{\Phi} \in \mathcal{C}, \\ +\infty, \text{otherwise}, \end{cases} \tag{11}$$

$$\mathbb{I}_{\mathcal{D}}(\mathbf{\Psi}) = \begin{cases} 0, \text{if } \mathbf{\Psi} \in \mathcal{D}, \\ +\infty, \text{otherwise}. \end{cases} \tag{12}$$

Using (11) and (12), the problem in (7) can be equivalently written as:

$$\min_{\mathbf{W},\mathbf{\Phi},\mathbf{\Psi}} \sum_{i=1}^{N_r} \|\mathbf{w}_i\|^2 + \mathbb{I}_{\mathcal{C}}(\mathbf{W},\mathbf{\Phi}) + \mathbb{I}_{\mathcal{D}}(\mathbf{\Psi}) \tag{13}$$

$$\text{s.t. } \mathbf{\Phi} = \mathbf{\Psi}. \tag{14}$$

The problem in (13) has only one equality constraint, and thus has a simpler structure compared to its original form in (7). In the following, we propose a single-loop algorithm based on the ADMM framework to address the reformulated problem. By adopting ADMM, we introduce a Lagrange multiplier and a penalty parameter to decompose the reformulated problem into three subproblems, followed by a closed-form update of the Lagrange multiplier.

According to Lagrangian duality [14], the augmented Lagrangian for (13) is defined as

$$\mathcal{L}(\mathbf{W},\mathbf{\Phi},\mathbf{\Psi},\mathbf{\Lambda}) = \sum_{i=1}^{N_r} \|\mathbf{w}_i\|^2 + \mathbb{I}_{\mathcal{C}}(\mathbf{W},\mathbf{\Phi}) + \mathbb{I}_{\mathcal{D}}(\mathbf{\Psi}) + \Re\left\{\text{tr}\left\{\mathbf{\Lambda}^H (\mathbf{\Phi}-\mathbf{\Psi})\right\}\right\} + \frac{\rho}{2}\|\mathbf{\Phi}-\mathbf{\Psi}\|_F^2, \tag{15}$$

where $\mathbf{\Lambda}$ and ρ denote the Lagrange multiplier and the penalty parameter associated with the equality constraint, respectively. It is observed that if the reformulated problem in (7) is solved, then plugging the optimal variables into the augmented Lagrangian will produce an optimal value equal to the objective function in (7). Conversely, by minimizing the augmented Lagrangian with respect to $\mathbf{W}$, $\mathbf{\Phi}$, $\mathbf{\Psi}$, and $\mathbf{\Lambda}$, we can obtain an approximate solution to (7). The accuracy of this solution is determined by the penalty parameter ρ and the number of iterations. Next, we outline the decomposition procedure and elaborate on the solution to the resulting subproblems. The subproblem for the scattering matrix is transformed into an equivalent problem using the phase invariance inherent in the SINR constraint, resulting in a convex formulation. The update of the auxiliary variable is formulated as a quadratic programming problem with an orthogonality constraint and solved using the SVD.

The four optimization variables in the augmented Lagrangian are updated in an alternating manner, as follows:

$$\mathbf{W}^{t+1} = \arg\min_{\mathbf{W}} \mathcal{L}(\mathbf{W},\mathbf{\Phi}^t,\mathbf{\Psi}^t,\mathbf{\Lambda}^t), \tag{16}$$

$$\mathbf{\Phi}^{t+1} = \arg\min_{\mathbf{\Phi}} \mathcal{L}(\mathbf{W}^{t+1},\mathbf{\Phi},\mathbf{\Psi}^t,\mathbf{\Lambda}^t), \tag{17}$$

$$\mathbf{\Psi}^{t+1} = \arg\min_{\mathbf{\Psi}} \mathcal{L}(\mathbf{W}^{t+1},\mathbf{\Phi}^{t+1},\mathbf{\Psi},\mathbf{\Lambda}^t), \tag{18}$$

$$\mathbf{\Lambda}^{t+1} = \mathbf{\Lambda}^t + \rho\left(\mathbf{\Phi}^{t+1} - \mathbf{\Psi}^{t+1}\right), \tag{19}$$

where the superscripts t and $t+1$ denote iteration indices. It is evident that the updates for $\mathbf{W}$, $\mathbf{\Phi}$, and $\mathbf{\Psi}$ require further investigation because the indicator functions used in the augmented Lagrangian cannot be evaluated directly. We address the three subproblems in the remainder of this section, deriving a convex reformulation of (16) and (17), as well as a closed-form solution to problem (18).

A. Update of the Beamforming Matrix

To update the beamforming matrix $\mathbf{W}$, we can fix the other variables and substitute the augmented Lagrangian into (16). By rewriting the indicator function for $\mathbf{W}$ in terms of the original SINR constraint, we obtain the following subproblem for updating $\mathbf{W}$:

$$\min_{\mathbf{W}} \sum_{i=1}^{N_r} \|\mathbf{w}_i\|^2 \tag{20}$$

$$\text{s.t.} \frac{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{\sum_{j\neq k}^{N_r} \left|\mathbf{h}_k^H \mathbf{w}_j\right|^2 + \sigma_k^2} \geq \gamma_k, \forall k. \tag{21}$$

Note that the fractional SINR constraint is nonconvex. Although this problem cannot be solved directly due to its nonconvex nature, a fixed-point algorithm is proposed in [15] using Lagrangian duality. Additionally, thanks to the phase invariance in the SINR formulation, it can be reformulated as a convex second-order cone constraint [16]:

$$\min_{\mathbf{W}} \sum_{i=1}^{N_r} \|\mathbf{w}_i\|^2 \tag{22}$$

$$\text{s.t.} \left\| \begin{bmatrix} \mathbf{W}^H \mathbf{h}_k \\ \sigma_k \end{bmatrix} \right\| - \sqrt{1+\frac{1}{\gamma_k}} \Re\left\{\mathbf{h}_k^H \mathbf{w}_k\right\} \leq 0, \forall k, \tag{23}$$

$$\Im\left\{\mathbf{h}_k^H \mathbf{w}_k\right\} = 0, \forall k. \tag{24}$$

Therefore, this problem can be readily solved using standard convex optimization tools.

B. Update of the Scattering Matrix

Following the procedure for updating $\mathbf{W}$, the scattering matrix $\mathbf{\Phi}$ can be updated as follows:

$$\min_{\mathbf{\Phi}} \Re\left\{\text{tr}\left\{\mathbf{\Lambda}^H\left(\mathbf{\Phi}-\mathbf{\Psi}\right)\right\}\right\} + \frac{\rho}{2}\|\mathbf{\Phi}-\mathbf{\Psi}\|_F^2 \tag{25}$$

$$\text{s.t.} \frac{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{\sum_{j\neq k}^{N_r}\left|\mathbf{h}_k^H \mathbf{w}_j\right|^2 + \sigma_k^2} \geq \gamma_k, \forall k, \tag{26}$$

$$\mathbf{\Phi} = \mathbf{\Phi}^T. \tag{27}$$

The SINR constraint in (26) is nonconvex with respect to $\mathbf{\Phi}$. Below, we leverage the phase invariance inherent in the SINR formulation to rewrite the fractional SINR constraint as a quadratic constraint. Since a phase rotation of the precoding matrix is irrelevant to the SINR, we can choose the phase such that the desired signal, $\mathbf{h}_k^H \mathbf{w}_k$, is real-valued and positive. To this end, the SINR constraint with respect to $\mathbf{\Phi}$ can be equivalently written as:

$$\left\|\left[\mathbf{h}_{d,k}^H \mathbf{W} + \mathbf{h}_{r,k}^H \mathbf{\Phi}\mathbf{G}\mathbf{W} \quad \sigma_k\right]\right\| - \sqrt{1+\frac{1}{\gamma_k}} \Re\left\{\mathbf{h}_{d,k}^H \mathbf{w}_k + \mathbf{h}_{r,k}^H \mathbf{\Phi}\mathbf{G}\mathbf{w}_k\right\} \leq 0, \forall k, \tag{28}$$

$$\Im\left\{\mathbf{h}_{d,k}^H \mathbf{w}_k + \mathbf{h}_{r,k}^H \mathbf{\Phi}\mathbf{G}\mathbf{w}_k\right\} = 0, \forall k. \tag{29}$$

While this leads to a convex problem for updating the matrix $\mathbf{\Phi}$, to solve the problem using standard optimization tools, we derive an equivalent formulation using a vector-valued optimization variable for problem (25). The symmetry constraint on $\mathbf{\Phi}$ indicates that there are $N_i(N_i+1)/2$ independent entries in $\mathbf{\Phi}$. Based on this, $\mathbf{\Phi}$ can be vectorized as follows:

$$\boldsymbol{\phi} = \text{vec}\left\{\mathbf{\Phi}\right\} = \mathbf{K}\boldsymbol{\varphi}, \tag{30}$$

where $\boldsymbol{\varphi}$ is obtained by vectorizing the upper triangular part of $\mathbf{\Phi}$, and $\mathbf{K} \in \{0,1\}^{N_i^2 \times N_i(N_i+1)/2}$ serves as a reshaping matrix that links $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$. Based on this relationship and the Kronecker product, we rewrite the update of $\mathbf{\Phi}$ in a new form with the vector optimization variable:

$$\min_{\boldsymbol{\varphi}} \Re\left\{\boldsymbol{\lambda}^H\left(\boldsymbol{\varphi}-\boldsymbol{\psi}\right)\right\} + \frac{\rho}{2}\|\boldsymbol{\varphi}-\boldsymbol{\psi}\|^2 \tag{31}$$

$$\text{s.t.} \left\| \begin{bmatrix} \text{vec}\left\{\mathbf{h}_{d,k}^H \mathbf{W}\right\} + \mathbf{F}_{k,1}\mathbf{K}\boldsymbol{\varphi} \\ \sigma_k \end{bmatrix} \right\| - \sqrt{1+\frac{1}{\gamma_k}} \Re\left\{\mathbf{h}_{d,k}^H \mathbf{w}_k + \mathbf{f}_{k,2}\mathbf{K}\boldsymbol{\varphi}\right\} \leq 0, \forall k, \tag{32}$$

$$\Im\left\{\mathbf{h}_{d,k}^H \mathbf{w}_k + \mathbf{f}_{k,2}\mathbf{K}\boldsymbol{\varphi}\right\} = 0, \forall k, \tag{33}$$

where $\boldsymbol{\psi} = \text{vec}\left\{\mathbf{\Psi}\right\}$, $\boldsymbol{\lambda} = \text{vec}\left\{\mathbf{\Lambda}\right\}$, $\mathbf{F}_{k,1} = \left(\mathbf{G}\mathbf{W}\right)^T \otimes \mathbf{h}_{r,k}^H$, and $\mathbf{f}_{k,2} = \left(\mathbf{G}\mathbf{w}_k\right)^T \otimes \mathbf{h}_{r,k}^H$. Thus we obtain a standard quadratic programming problem with quadratic constraints, which can be solved using conventional algorithms.

C. Update of the Auxiliary Matrix

Substituting the augmented Lagrangian into (18) and fixing all other variables except $\mathbf{\Psi}$, we obtain the following subproblem for updating the auxiliary matrix:

$$\min_{\mathbf{\Psi}} \left\|\mathbf{\Phi} - \mathbf{\Psi} + \frac{1}{\rho}\mathbf{\Lambda}\right\|_F^2$$

$$\text{s.t. } \mathbf{\Psi}^H\mathbf{\Psi} = \mathbf{I}. \tag{34}$$

Note that the orthogonality constraint in (34) is nonconvex. Insights can be gained by examining (34) from the perspective of projection. Essentially, the problem in (34) is equivalent to projecting the matrix $\mathbf{\Phi} + \frac{1}{\rho}\mathbf{\Lambda}$ onto the feasible region defined by the orthogonality constraint. Assume the SVD of $\mathbf{\Phi} + \frac{1}{\rho}\mathbf{\Lambda}$ is expressed as:

$$\mathbf{\Phi} + \frac{1}{\rho}\mathbf{\Lambda} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{35}$$

where $\mathbf{U}$ and $\mathbf{V}$ denote unitary matrices composed of the left and right singular vectors, respectively, and $\mathbf{\Sigma}$ is a diagonal matrix containing the singular values of $\mathbf{\Phi} + \frac{1}{\rho}\mathbf{\Lambda}$ on the diagonal. To satisfy the orthogonality constraint by means of a projection, we set the singular values of the candidate solution to 1. Therefore, the optimal closed-form solution to (34) is given by [17]

$$\mathbf{\Psi} = \mathbf{U}\mathbf{V}^H. \tag{36}$$

After obtaining the three aforementioned variables, we can update the Lagrange multiplier according to (19) to complete one iteration. It is evident that the proposed algorithm features a single-loop structure, making it capable of achieving a better trade-off between complexity and performance. The proposed single-loop algorithm based on ADMM is summarized in Algorithm 1.

D. Algorithm Initialization

Since the joint optimization problem in (4) is nonconvex, a good initialization for the proposed single-loop algorithm is crucial to avoid convergence to a bad local minimum. To address this issue, we adopt an approximate solution to the effective channel gain maximization problem [18]. Specifically, the initial value of $\mathbf{\Phi}$ is given by:

$$\mathbf{\Phi} = \dot{\mathbf{U}}\dot{\mathbf{V}}^H, \tag{37}$$

where $\dot{\mathbf{U}}$ and $\dot{\mathbf{V}}$ are comprised of the left and right singular vectors of $\mathbf{H}_r\mathbf{H}_d^H\mathbf{G}^H + \left(\mathbf{H}_r\mathbf{H}_d^H\mathbf{G}^H\right)^T$, respectively. This initialization is a feasible solution to the joint optimization problem in (4), with each impedance configured to maximize the effective channel gain, which is intuitively beneficial for transmit power minimization beamforming. In addition, the complexity of this method is low [18].

Algorithm 1 Single-Loop Algorithm for BD-RIS-Aided Power Minimization Beamforming (4)

Input: $\mathbf{h}_{d,k}, \mathbf{h}_{r,k}, \mathbf{G}, \sigma_k, \gamma_k, \forall k$
Output: $\mathbf{W}, \mathbf{\Phi}$
  Initialize $\mathbf{\Phi}^0$, $\mathbf{\Psi}^0$, and $\mathbf{\Lambda}^0$
  Set ρ
  Set $t \leftarrow 0$
  repeat
    Update the beamforming matrix $\mathbf{W}^{t+1}$ by solving (22)
    Update the scattering matrix $\mathbf{\Phi}^{t+1}$ by solving (31)
    Update the auxiliary matrix $\mathbf{\Psi}^{t+1}$ by (36)
    Update the Lagrange multiplier matrix $\mathbf{\Lambda}^{t+1}$ by (19)
    Set $t \leftarrow t+1$
  until the convergence criterion is met

IV. Simulation Results

In this section, we present simulation results to demonstrate the performance of the proposed single-loop algorithm. The noise variance and SINR threshold for each user are assumed to be the same, i.e., $\sigma_k^2 = -70$ dBm, $\gamma_k = \gamma, \forall k$. The simulation settings are as follows: a BS with $N_t = 4$ transmit antennas is located at $(0, 0, 4)$ m, and a BD-RIS with N_i reflecting elements is positioned at $(0, 52, 2)$. Meanwhile, $N_r = 4$ single-antenna users are randomly distributed within a horizontal circular area with a radius of 5 m, centered at $(0, 55, 1)$. The distance-dependent path loss is modeled as

$$L_{ij}\left(d_{ij}\right) = L_0 \left(\frac{d_{ij}}{d_0}\right)^{-\alpha_{ij}}, \tag{38}$$

where $L_0 = -30$ dB is the path loss at the reference distance $d_0 = 1$ m, d_{ij} denotes the distance, and α_{ij} denotes the path loss exponent for a direct signal link from i to j, where i and j can represent the BS, a user, or the BD-RIS. We set the path loss exponents for the links between the BS and users, BS and RIS, and RIS and users as $\alpha_{RT} = 3.5$, $\alpha_{IT} = 2$, and $\alpha_{RI} = 2.8$, respectively. The channel vector between i and j is given by:

$$\mathbf{h}_{ij} = \sqrt{L_{ij}}\left(\sqrt{\frac{K_F}{1+K_F}}\mathbf{h}_{ij}^{LoS} + \sqrt{\frac{1}{1+K_F}}\mathbf{h}_{ij}^{NLoS}\right), \tag{39}$$

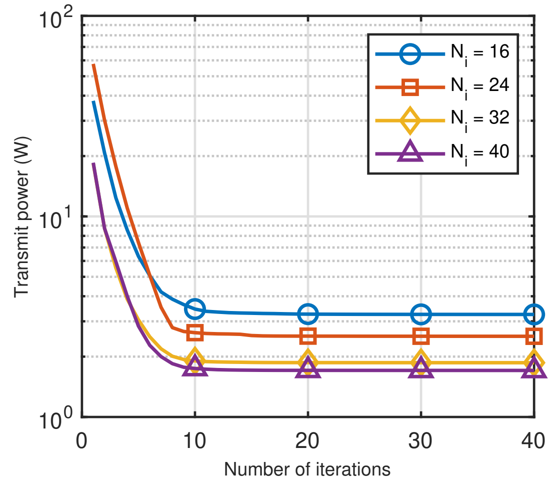

Fig. 3. Transmit power versus the number of iterations.

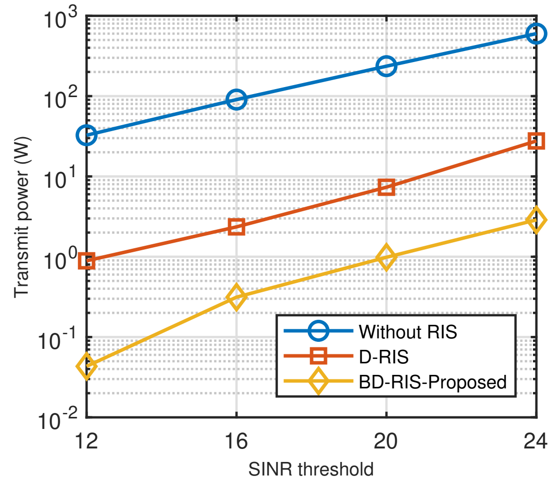

Fig. 4. Transmit power versus SINR threshold.

where $K_F = 3$ dB denotes the Rician factor, $\tilde{\mathbf{h}}_{ij}^{LoS}$ represents the line-of-sight (LoS) component, and $\tilde{\mathbf{h}}_{ij}^{NLoS} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{I}\right)$ denotes the Rayleigh fading non-line-of-sight (NLoS) component. The penalty parameter is set to $\rho = 0.5$.

In Fig. 3, we illustrate the convergence of the optimized transmit power for different scenarios with various BD-RIS sizes, for a case with $\gamma = 12$ dB. We observe that in all scenarios, the transmit power converges within about 10 iterations. Fig. 3 demonstrates the efficiency of the proposed single-loop algorithm for various BD-RIS sizes, and illustrates that the required transmit power decreases as the BD-RIS size increases. Clearly, a larger BD-RIS provides more DoFs for designing the reflected signal, which enables less transmit power to be consumed.

In Fig. 4, we evaluate the proposed single-loop algorithm under various SINR thresholds, with the number of BD-RIS elements set to $N_i = 32$. The results are compared to the case without RIS, and the case with conventional D-RIS as proposed in [6]. Fig. 4 illustrates that transmit power increases with the SINR threshold and that the proposed single-loop

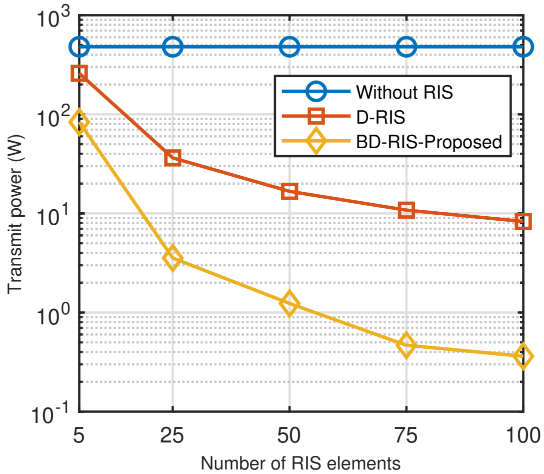

Fig. 5. Transmit power versus the number of RIS elements.

algorithm is effective for different SINR thresholds. The results also highlight the significant advantage of the BD-RIS architecture compared with conventional D-RIS.

In Fig. 5, we plot the transmit power as a function of the number of RIS elements for an SINR threshold of $\gamma = 24$ dB. The results validate that BD-RIS-aided beamforming is the most efficient approach for transmit power reduction. Similar to Fig. 3, we see that the transmit power decreases as the number of RIS elements increases. The figure also validates the effectiveness of the proposed single-loop algorithm for various BD-RIS sizes.

V. Conclusion

In this paper, we investigated BD-RIS-aided power minimization beamforming, and formulated the problem as a transmit power minimization task with SINR constraints, as well as symmetric and orthogonality constraints on the BD-RIS scattering matrix. To solve this joint optimization problem, we proposed a single-loop algorithm based on a splitting strategy that introduces an auxiliary variable to split the scattering matrix. We adopted the ADMM framework to alternately update the relevant variables. The nonconvex subproblem for the scattering matrix is reformulated as a convex problem by leveraging the invariance of the SINR to a common phase rotation of the transmit signal. Additionally, we presented a closed-form solution based on the SVD for updating the auxiliary matrix. Simulation results demonstrate the superiority of the BD-RIS compared with conventional single-connected D-RIS.

References

[1] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[2] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 14–20, Jun. 2021.

[3] Y. Zheng, T. Lin, and Y. Zhu, "Passive beamforming for IRS-assisted MU-MIMO systems with one-bit ADCs: An SER minimization design approach," *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1101–1105, May 2022.

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

[5] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Using intelligent reflecting surfaces for rank improvement in MIMO communications," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2020, pp. 9160–9164.

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

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

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

[9] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1229–1243, Feb. 2022.

[10] H. Li, S. Shen, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces: From transmitting and reflecting modes to single-, group-, and fully-connected architectures," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2311–2324, Apr. 2023.

[11] ——, "Beyond diagonal reconfigurable intelligent surfaces: A multi-sector mode enabling highly directional full-space wireless coverage," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2446–2460, Aug. 2023.

[12] M. Nerini, S. Shen, and B. Clerckx, "Closed-form global optimization of beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1037–1051, Feb. 2024.

[13] Y. Zhou, Y. Liu, H. Li, Q. Wu, S. Shen, and B. Clerckx, "Optimizing power consumption, energy efficiency, and sum-rate using beyond diagonal RIS — A unified approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 7423–7438, Jul. 2024.

[14] 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, 2011.

[15] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[16] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.

[17] J. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.

[18] T. Fang and Y. Mao, "A low-complexity beamforming design for beyond-diagonal RIS aided multi-user networks," *IEEE Commun. Lett.*, vol. 28, no. 1, pp. 203–207, Jan. 2024.