

Interference Minimization in Beyond-Diagonal RIS-Assisted MIMO Interference Channels

IGNACIO SANTAMARIA ¹ (Senior Member, IEEE), MOHAMMAD SOLEYMANI ² (Member, IEEE),
EDUARD JORSWIECK ³ (Fellow, IEEE), AND JESÚS GUTIÉRREZ ⁴ (Member, IEEE)

¹Department of Communications Engineering, Universidad de Cantabria, 39005 Santander, Spain

²Signal and System Theory Group, Universität Paderborn, 33098 Paderborn, Germany

³Institute for Communications Technology, Technische Universität Braunschweig, 38106 Braunschweig, Germany

⁴IHP - Leibniz-Institut für Innovative Mikroelektronik, 15236 Frankfurt (Oder), Germany

CORRESPONDING AUTHOR: IGNACIO SANTAMARIA (e-mail: i.santamaria@unican.es).

This work was supported by the European Commission's Horizon Europe, Smart Networks and Services Joint Undertaking, Research and Innovation Program under Grant 101139282, 6G-SENSES Project. The work of Ignacio Santamaria was supported in part by MICIU/AEI /10.13039/501100011033 under Grant PID2022-137099NB-C43 (MADDIE) and in part by FEDER, UE.

ABSTRACT This paper proposes a two-stage approach for passive and active beamforming in multiple-input multiple-output (MIMO) interference channels (ICs) assisted by a beyond-diagonal reconfigurable intelligent surface (BD-RIS). In the first stage, the passive BD-RIS is designed to minimize the aggregate interference power at all receivers, a cost function called interference leakage (IL). To this end, we propose an optimization algorithm in the manifold of unitary matrices and a suboptimal but computationally efficient solution. In the second stage, users' active precoders are designed under different criteria such as minimizing the IL (min-IL), maximizing the signal-to-interference-plus-noise ratio (max-SINR), or maximizing the sum rate (max-SR). The residual interference not cancelled by the BD-RIS is treated as noise by the precoders. Our simulation results show that the max-SR precoders provide more than 20% sum rate improvement compared to other designs, especially when the BD-RIS has a moderate number of elements and users transmit with high power, in which case the residual interference is still significant.

INDEX TERMS Beyond-diagonal reconfigurable intelligent surface, multi-antenna communications, interference leakage, interference channel, manifold optimization.

I. INTRODUCTION

A. BACKGROUND

The potential of reconfigurable intelligent surfaces (RISs) as a PHY-layer 6G enabling technology that can improve coverage, transmission rates, energy efficiency, or network latency is being recognized and assessed by academia and industry alike. A general overview of this research topic and the main applications of RISs in wireless communication systems can be found at [1]. Relevant use cases, deployment scenarios, and operational requirements for each identified use case are reported in [2]. Initial studies and implementations of RISs have been focused on passive lossless single-connected structures, leading to diagonal scattering matrices whose elements are modeled as phase shifters. The joint optimization of the precoders and the diagonal RIS in a multi-cell MIMO system to maximize the weighted sum rate is proposed in [3].

Based on a realistic model for the RIS power consumption, the problem of maximizing the energy efficiency is formulated in [4], and an algorithm for optimizing the RIS elements and the power transmitted to the users in a downlink channel is proposed. The authors in [5] consider the optimization of the active precoders and the passive reflective RIS to minimize the power transmitted by an access point or base station subject to individual quality of service (QoS) constraints for users.

More recently, the concept of beyond-diagonal RIS (BD-RIS) has been proposed as a generalization in which all RIS elements can be connected by variable reactances to each other, leading to fully-connected or group-connected (block-diagonal) scattering matrices [6], with different trade-offs between complexity and performance. Analyzing the BD-RIS as a multi-port network, in [7] it is shown that the BD-RIS

scattering matrix must be unitary and symmetric, which introduces new challenges in its optimization.

B. RELATED WORK ON BD-RIS OPTIMIZATION

Over the last few years, different algorithms have been proposed to optimize the BD-RIS elements to maximize, among other cost functions, the received signal-to-noise ratio (SNR), the rate, or the energy efficiency. Closed-form fully-connected BD-RIS solutions that maximize the equivalent channel gain exist for single-input single-output (SISO) and MISO/SIMO channels [8], [9], [10]. In [11], [12], the authors consider a multi-user multiple-input single-output (MU-MISO) downlink channel assisted by a multi-sector BD-RIS in which the RIS matrices of each cell are single-connected and therefore diagonal. The optimization of multi-sector BD-RIS coefficients is performed on the complex sphere manifold. Li et al proposed in [6] a manifold optimization algorithm to maximize the sum rate in BD-RIS-assisted MU-MISO systems when the direct channels are blocked. A similar algorithm is described in [13]. The successive convex approximation technique has been used in [14] for optimizing the spectral efficiency and energy efficiency of a multiple access system assisted by a BD-RIS with a group-connected architecture of group size two. An optimization framework for BD-RIS based on the penalty dual decomposition (PDD) method [15] for the optimization of different cost functions is proposed in [16]. Although this algorithm is rather general, it only considers single antenna receivers (MU-MISO) and has a high computational cost. Moreover, the symmetry constraint of the BD-RIS matrix is not explicitly considered in [16]. Geodesic manifold optimization algorithms have recently been proposed in [17], [18] to maximize the rate in BD-RIS-aided point-to-point MIMO channels. While [17] only considers the unitary constraint for the BD-RIS matrix, [18] also considers the symmetry constraint. In [19] a closed-form BD-RIS solution is found by projecting the Maximum Ratio Transmission (MRT) precoder (to maximize the signal power) or the zero-forcing precoder (to cancel interference) into the set of symmetric and unitary matrices. This relax-then-project (RtP) approach, also used in [10], may cause a significant performance degradation compared to solving the original constrained problem. The method presented in this paper avoids the projection on the set of symmetric matrices, thus improving the results.

C. MOTIVATION

Most of the aforementioned BD-RIS-assisted scenarios mainly consider point-to-point or MU-MISO systems. However, works on the K -user MIMO interference channel (IC) are more scarce and have been mostly limited to diagonal RIS. To achieve the maximum degrees of freedom (DoF) of the K -user IC, the interfering signals at each receiver must fall into a reduced-dimensional subspace, a technique called *interference alignment* (IA) [20]. The design of IA multi-antenna precoders uses as optimization criterion the sum of

the power of the interference leaked in the receive subspaces of all users, a cost function that is known in the literature as *interference leakage* (IL) [21], [22]. Therefore, most studies on RIS-assisted ICs aim at canceling or aligning the interference, thus maximizing the channel DoFs.

The first work in this line is [23], where the ability of an RIS to completely null the interference simultaneously on all the receivers in an IC is studied for the first time. However, the scenario discussed in [23] is rather limited, considering only a SISO-IC with blocked direct channels. In [24], the authors optimize an RIS to maximize the DoF of the K -user MIMO-IC. In [25], the K -user MIMO-IC sum rate, including mutual coupling effects, is maximized using a weighted minimum mean squared error (wMMSE) approach. In [26] the authors study the DoF of the K -user RIS-assisted SISO IC with symbol extensions (i.e., for time-varying channels). In [27] an algorithm for interference minimization in the K -user MIMO-IC assisted by a passive lossless RIS was proposed. An active (unconstrained) RIS to perform IA in the K -user MIMO-IC has been considered in [28]. The active diagonal RIS reduces the size of the interference subspace at each receiver while zero-forcing precoders are used to null the multiuser interference. In this work, we consider more realistic passive architectures.

Certain important reasons justify minimizing the IL in a BD-RIS-assisted K -user MIMO-IC. First, for the K -user MIMO-IC the min-IL precoders attain (if the IA problem is feasible) the maximum DoF of the channel. Second, if a BD-RIS can sufficiently suppress interference, the users could design their precoders independently treating the residual interference as noise (TIN) [29]. Finally, the IL yields quadratic functions that are usually easier to optimize than other cost functions commonly used, such as the sum rate or the mean-squared error (MSE). Motivated by these considerations, in this paper we consider the IL as a suitable metric for the BD-RIS design in MIMO-ICs. However, the literature has not studied algorithms that optimize the IL using BD-RIS architectures subject to different power constraints and using various types of precoders. This is the problem that motivates this work.

D. CONTRIBUTIONS

This paper proposes a two-stage approach for interference management in the K -user MIMO-IC assisted by a BD-RIS. The IL is minimized in the first stage by the BD-RIS, while the precoder design takes care in the second stage of the residual interference that the BD-RIS could not remove. The contributions of this work can be summarized as follows:

- 1) Unlike most of the works on BD-RIS that considers MU-MISO systems [6], [13], [14], [16], this paper focuses on the K -user MIMO-IC, which requires specific precoding and interference management strategies such as IA to achieve the maximum channel's DoF.
- 2) For this scenario, we consider for the first time a two-stage approach for designing the BD-RIS and the

precoders. In the first stage, the BD-RIS minimizes the IL to reduce the interference level below the noise level thus facilitating the design of the precoders in the second stage. The precoders are designed in the second stage to optimize different criteria such as maximizing the sum rate, the SINR, or minimizing the IL.

- 3) For a fully-connected BD-RIS architecture we propose a new optimization algorithm in the manifold of unitary matrices. Unlike existing alternatives in the literature, the proposed manifold optimization (MO) algorithm incorporates the symmetry constraint in the solution by taking advantage of the reparametrization of any symmetric and unitary matrix through Takagi's factorization.
- 4) We propose a relax-then-project (RtP) algorithm that provides a suboptimal solution with lower computational complexity than the MO algorithm. The RtP method first solves a relaxed version of the problem by replacing the unitary constraint on the BD-RIS matrix with a constraint on its trace. The relaxed problem incorporates the symmetry constraint of the BD-RIS into the problem through the commutation matrix. We believe this contribution might be useful in other problems involving symmetric BD-RISs. After solving the resulting convex problem, the relaxed solution is projected on the set of unitary matrices using the method proposed in [10], [30].
- 5) To reduce the computational and circuit complexity associated with the fully connected BD-RIS architecture, we adapt the IL minimization algorithm to a group-connected architecture [6], [7]. The group-connected algorithm sequentially optimizes each group using either the MO algorithm or the RtP method.
- 6) A theoretical analysis for unconstrained BD-RIS (probably using active components), allows us to estimate the minimum number of elements needed to cancel the interference simultaneously at all receivers, thus generalizing previous results for diagonal RIS.
- 7) Simulation results compare the performance of the proposed two-stage approach in terms of interference-to-noise ratio (INR) levels and achievable rates for different BD-RIS locations and precoder designs. A fully-connected BD-RIS with a moderately high number of elements and optimally positioned significantly reduces interference levels. The group-connected BD-RIS optimized with the MO algorithm and with group size $M_g = 8$ provides a reasonable tradeoff between performance and complexity. The precoders take care of the residual interference through proper design. The max-SR precoders provide more than 20% sum rate improvement compared to other designs, especially when the BD-RIS has a moderate number of elements and users transmit with high power, in which case the residual interference is still significant.

E. PAPER OUTLINE AND NOTATIONS

The rest of the paper is organized as follows. Section II presents the IL minimization problem in the K -user MIMO-IC and briefly introduces the two-stage optimization framework. Section III describes Stage I, where the BD-RIS is optimized to minimize the IL. We present a manifold optimization algorithm based on Takagi decomposition [31] for the fully-connected architecture, and a relax-then-project suboptimal solution. Optimization algorithms for the group-connected BD-RIS and the conventional diagonal RIS are also described in Section III. Section IV describes Stage II of the method, where the users' precoders are optimized according to different criteria, such as minimizing the IL, maximizing the SINR, or maximizing the sum rate. An alternating optimization (AO) algorithm to jointly design precoders and BD-RIS minimizing the IL is also described in Section IV. Section V compares the performance of the different BD-RIS and RIS designs in terms of the INR and the sum rate achieved in several scenarios. This work's main conclusions and future directions are described in Section VI.

Notation: In this paper, matrices are denoted by bold-faced upper case letters, column vectors are denoted by bold-faced lower case letters, and scalars are denoted by light-faced lower case letters. The superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote transpose, conjugate, and Hermitian conjugate, respectively. The trace and determinant of a matrix \mathbf{A} will be denoted, respectively, as $\text{tr}(\mathbf{A})$ and $\det(\mathbf{A})$. \mathbf{I}_n denotes the identity matrix of size n , $\mathcal{CN}(0, 1)$ denotes a complex proper Gaussian distribution with zero mean and unit variance, and $\mathbf{x} \sim \mathcal{CN}_n(\mathbf{0}, \mathbf{R})$ denotes a complex Gaussian vector in \mathbb{C}^n with zero mean and covariance matrix \mathbf{R} . $\mathbb{E}[\cdot]$ denotes mathematical expectation and $\text{Re}(a)$ denotes the real part of a . In addition, \odot denotes Hadamard product and \otimes denotes Kronecker product. The diag operator of a matrix, $\text{diag}(\mathbf{A})$, extracts the elements of the main diagonal of the square matrix \mathbf{A} and stores them in a column vector, and $\text{vec}(\mathbf{A})$ denotes the vectorization operator which converts the $M \times M$ matrix \mathbf{A} into an $M^2 \times 1$ column vector by stacking the columns of \mathbf{A} on top of one another, $\text{unvec}(\mathbf{a})$ undoes this operation, returning an $M \times M$ matrix from an $M^2 \times 1$ vector \mathbf{a} . $\nabla_{\mathbf{Q}} J$ denotes the gradient of the matrix-valued function J evaluated at \mathbf{Q} . Finally, $\mathcal{O}(\cdot)$ is the big-O notation for representing the computational complexity of algorithms.

II. BD-RIS-ASSISTED MIMO-IC

A. SIGNAL MODEL

As illustrated schematically in Fig. 1, we consider a K -user MIMO-IC assisted by a BD-RIS designed to suppress the aggregate interference power at all receivers. The k th link has N_{T_k} transmit antennas, N_{R_k} receive antennas, and transmits d_k streams. The BD-RIS has M elements and all Txs and Rxs are in the reflection space of the BD-RIS. According to the commonly used notation, we denote the MIMO-IC in abbreviated form as $(N_{T_k} \times N_{R_k}, d_k)^K$ [21], [32]. The equivalent MIMO

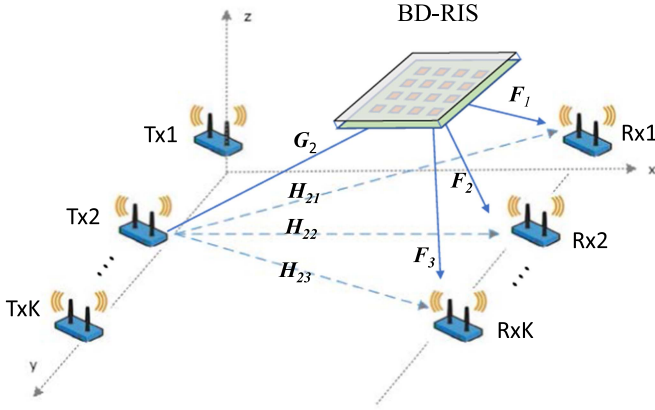


FIGURE 1. BD-RIS-assisted K -user MIMO IC. Dashed lines represent the direct Tx-Rx links; solid lines represent the channels through the BD-RIS.

channel from the l th transmitter to the k th receiver is

$$\tilde{\mathbf{H}}_{lk} = \mathbf{H}_{lk} + \mathbf{F}_k \mathbf{\Theta} \mathbf{G}_l^H, \quad (1)$$

where $\mathbf{H}_{lk} \in \mathbb{C}^{N_{R_k} \times N_{T_l}}$ is the (l, k) MIMO-IC, $\mathbf{G}_l \in \mathbb{C}^{N_{T_l} \times M}$ is the channel from the l th transmitter to the BD-RIS, $\mathbf{F}_k \in \mathbb{C}^{N_{R_k} \times M}$ is the channel from the BD-RIS to the k th receiver, and $\mathbf{\Theta}$ is the $M \times M$ BD-RIS scattering matrix. The k th user transmits a precoded signal $\mathbf{x}_k = \mathbf{V}_k \mathbf{s}_k$, where $\mathbf{V}_k \in \mathbb{C}^{N_{T_k} \times d_k}$ is a precoder scaled to satisfy the total power constraint and \mathbf{s}_k contains the information symbols, $\mathbf{s}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{d_k})$. The precoders determine the beamforming directions or beams and the allocated powers to the d_k streams. Therefore, the total power transmitted by user k is $P_k = \text{tr}(\mathbf{E}[\mathbf{x}_k \mathbf{x}_k^H]) = \text{tr}(\mathbf{V}_k \mathbf{V}_k^H)$. The signal received by the k th user is

$$\mathbf{y}_k = \tilde{\mathbf{H}}_{kk} \mathbf{x}_k + \sum_{l=1, l \neq k}^K \tilde{\mathbf{H}}_{lk} \mathbf{x}_l + \mathbf{n}_k, \quad (2)$$

where the second term is the interference caused by other users and $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_{R_k}})$ is the additive white Gaussian noise (AWGN). The k th user applies a unitary decoder $\mathbf{U}_k \in \mathbb{C}^{N_{R_k} \times d_k}$ to obtain

$$\mathbf{z}_k = \underbrace{\mathbf{U}_k^H \tilde{\mathbf{H}}_{kk} \mathbf{V}_k \mathbf{s}_k}_{\text{Signal}} + \underbrace{\mathbf{U}_k^H \left(\sum_{l=1, l \neq k}^K \tilde{\mathbf{H}}_{lk} \mathbf{V}_l \mathbf{s}_l \right)}_{\text{Interference}} + \underbrace{\mathbf{U}_k^H \mathbf{n}_k}_{\text{Noise}}.$$

As the noise is Gaussian and isotropically distributed, $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_{R_k}})$, and the decoder is unitary, $\mathbf{U}_k^H \mathbf{U}_k = \mathbf{I}_{d_k}$, then $\mathbf{U}_k^H \mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{d_k})$ [33, App. D]. In certain cases, we will find it useful to absorb the precoders and decoders into a new equivalent channel from the l th transmitter to the k th receiver as

$$\begin{aligned} \mathbf{U}_k^H \tilde{\mathbf{H}}_{lk} \mathbf{V}_l &= \mathbf{U}_k^H \mathbf{H}_{lk} \mathbf{V}_l + \mathbf{U}_k^H \mathbf{F}_k \mathbf{\Theta} \mathbf{G}_l^H \mathbf{V}_l \\ &= \bar{\mathbf{H}}_{lk} + \bar{\mathbf{F}}_k \mathbf{\Theta} \bar{\mathbf{G}}_l^H, \end{aligned} \quad (3)$$

TABLE 1. Symbols Used Frequently and Their Definitions

\mathbf{H}_{kl}	(l, k) MIMO-IC direct link
$\tilde{\mathbf{H}}_{kl}$	(l, k) Equivalent BD-RIS-assisted MIMO-IC
$\bar{\mathbf{H}}_{kl}$	(l, k) MIMO-IC after precoding+decoding
\mathbf{F}_k	Channel from BD-RIS to k th user Rx
$\bar{\mathbf{F}}_k$	Channel from BD-RIS to k th user Rx after decoding
\mathbf{G}_l	Channel from l th user Tx to BD-RIS
$\bar{\mathbf{G}}_l$	Channel from l th user Tx to BD-RIS after precoding
$\mathbf{V}_k, \mathbf{U}_k$	k th user precoder and decoder

where $\bar{\mathbf{F}}_k = \mathbf{U}_k^H \mathbf{F}_k$ is the $d_k \times M$ equivalent channel from the BD-RIS to the receiver after decoding, $\bar{\mathbf{G}}_l = \mathbf{V}_l^H \mathbf{G}_l$ is the $d_l \times M$ equivalent channel from the transmitter to the BD-RIS after precoding, and $\bar{\mathbf{H}}_{lk} = \mathbf{U}_k^H \mathbf{H}_{lk} \mathbf{V}_l$ is the $d_k \times d_l$ equivalent channel matrix after precoding-decoding. Table 1 shows the most frequently used symbols and their definitions.

B. PROPOSED TWO-STAGE OPTIMIZATION FRAMEWORK

Any cost function optimized on the K -user MIMO-IC requires the joint optimization of the precoders/decoders and the BD-RIS scattering matrix, a procedure usually carried out through AO algorithms with a high computational cost. We adopt a different approach that decouples the problem in two stages. The passive BD-RIS is designed in the first stage to reduce the total aggregate interference power at the receivers (a cost function in which the direct channels do not intervene). For the second stage, the (active) precoders and decoders are designed in the second stage, optimizing another cost function that improves the quality of the equivalent direct channels for each user. Two-stage optimization procedures have been recently proposed in BD-RIS assisted MU-MISO communication networks in [10], [19]. However, this work is the first to propose a two-stage approach for the K -user MIMO-IC.

The proposed two-stage framework entails a reduction in complexity, owing to the decoupling of passive and active beamforming problems. Another motivation is the architectural evolution from 5G to 6G systems, whereby the BD-RIS infrastructure provider might be different from the radio access network (RAN) operator in charge of the users' resource allocation strategies.

III. STAGE I: BD-RIS DESIGN TO MINIMIZE INTERFERENCE

A. PROBLEM STATEMENT

The BD-RIS is a passive multiport network, so the scattering matrix $\mathbf{\Theta}$ is modeled as a symmetric and unitary $M \times M$ matrix, $\mathbf{\Theta}^T = \mathbf{\Theta}$ and $\mathbf{\Theta}^H \mathbf{\Theta} = \mathbf{I}_M$. This architecture, in which tunable impedances connect all pairs of ports, is known as fully-connected BD-RIS [7]. Despite having a high implementation complexity, the fully-connected BD-RIS architecture provides the greatest flexibility in optimizing the various cost functions.

In the first stage, the BD-RIS minimizes the IL [21], defined as the sum of the squared Frobenius norms of all ICs. The BD-RIS optimization problem is¹

$$(\mathcal{P}_1) : \min_{\Theta} \sum_{k=1}^K \sum_{l=1, l \neq k}^K \|\mathbf{H}_{lk} + \mathbf{F}_k \Theta \mathbf{G}_l^H\|_F^2 \quad (4a)$$

$$\text{s.t. } \Theta^T = \Theta, \Theta^H \Theta = \mathbf{I}_M. \quad (4b)$$

The IL cost function (4a), minimized in Stage I by the BD-RIS, is independent of the precoders and decoders to be designed in Stage II and can be interpreted as a worst-case for the interference power observed with a given set of precoders and decoders.

B. UNCONSTRAINED BD-RIS

Although this paper focuses on passive BD-RIS, it may be illustrative to analyze the problem with an unconstrained, probably active, BD-RIS. When the number of elements is sufficiently large, an unconstrained BD-RIS can make the IL cost function in (4a) zero, thus perfectly neutralizing the interference at all receivers. The following lemma presents this result.

Lemma 1: Consider a generic² K -user MIMO-IC ($N_{T_k} \times N_{R_k}, d_k$) ^{K} assisted by an unconstrained BD-RIS, Θ , optimized to minimize the IL cost function in (4a). The only assumption is that $M \geq \max(N_{T_k}, N_{R_k}) \forall k$. An unconstrained BD-RIS can make the IL zero if the number of elements satisfies $M^2 \geq \sum_{l \neq k} N_{R_k} N_{T_l}$.

Proof: It is a simple algebra exercise to show that the IL cost function can be rewritten as the following quadratic function

$$IL(\mathbf{r}) = \text{tr}(\mathbf{T}) + \mathbf{r}^H \Sigma \mathbf{r} + 2\text{Re}(\mathbf{r}^H \mathbf{s}), \quad (5)$$

where $\mathbf{r} = \text{vec}(\Theta)$, $\mathbf{s} = \text{vec}(\sum_{l \neq k} \mathbf{F}_k^H \mathbf{H}_{lk} \mathbf{G}_l)$, $\Sigma = \sum_{l \neq k} \mathbf{G}_l^T \mathbf{G}_l^* \otimes \mathbf{F}_k \mathbf{F}_k^H$, and $\mathbf{T} = \sum_{l \neq k} \mathbf{H}_{lk}^H \mathbf{H}_{lk}$. Notice first that $\mathbf{s} \in \text{colspan}(\Sigma)$ and Σ is an $M^2 \times M^2$ positive semidefinite (psd) matrix. The matrices $\mathbf{F}_k^H \mathbf{F}_k$ and $\mathbf{G}_l^T \mathbf{G}_l^*$ are, respectively, rank- N_{R_k} and rank- N_{T_l} psd matrices. Then, Σ has rank $\min(M^2, \sum_{k \neq l} N_{R_k} N_{T_l})$ [34]. This means that without a power constraint on the BD-RIS elements, the IL can always be made zero regardless of the precoders or decoders with a BD-RIS with $M^2 > \sum_{k \neq l} N_{R_k} N_{T_l}$ active elements. If the BD-RIS is symmetric but otherwise unconstrained, we only have to replace M^2 by $M(M+1)/2$ in the inequality, as $M(M+1)/2$ is the dimension of the space of $M \times M$ symmetric matrices. ■

Remark 1: When all users have the same number of antennas and the links are $N_R \times N_T$ MIMO channels, the condition in Lemma 1 reduces to $M^2 > K(K-1)N_R N_T$, thus generalizing the zero IL condition for unrestricted (active) diagonal

RIS-assisted MIMO ICs derived in [27] and [28], which is $M > K(K-1)N_R N_T$. The condition for the K user SISO-IC $M^2 > K(K-1)$, was derived in [23]. In addition, when the precoders and decoders have been designed in a previous step and the unconstrained BD-RIS is designed afterward, the number of elements needed for zero IL relaxes to $M^2 > \sum_{k \neq l} d_k d_l$. When all users transmit d streams, the condition becomes $M^2 > K(K-1)d^2$.

C. MANIFOLD OPTIMIZATION ALGORITHM

The IL minimization problem \mathcal{P}_1 is non-convex and requires an iterative algorithm to find a (probably suboptimal) solution. In this subsection, we propose an iterative algorithm on the manifold of unitary matrices similar to that in [18] to maximize the capacity in a point-to-point MIMO link.

The main idea of the algorithm proposed in [18] is to apply Tagaki's factorization [8], [31], [35], which proves that any complex symmetric matrix \mathbf{A} can be factored as $\mathbf{A} = \mathbf{Q} \Sigma \mathbf{Q}^T$, where \mathbf{Q} belongs to the manifold of complex unitary matrices denoted here as $\mathcal{U}(M)$, and Σ is a diagonal matrix containing the singular values of \mathbf{A} along its diagonal. The BD-RIS scattering matrix, Θ , is symmetric and unitary. Therefore, all its singular values are one and it can be factored as $\Theta = \mathbf{Q} \mathbf{Q}^T$. We use this fact to formulate the IL minimization problem as follows³

$$(\mathcal{P}_{1a}) : \min_{\mathbf{Q} \in \mathcal{U}(M)} \sum_{l \neq k} \|\mathbf{H}_{lk} + \mathbf{F}_k \mathbf{Q} \mathbf{Q}^T \mathbf{G}_l^H\|_F^2. \quad (6)$$

To solve (\mathcal{P}_{1a}) , we perform the optimization on the unitary group [36]. The tangent plane at a point $\mathbf{Q} \in \mathcal{U}(M)$, denoted as $\Delta \mathbf{Q}$, comprises all $M \times M$ matrices such that $\mathbf{Q}^H \Delta \mathbf{Q}$ is skew-Hermitian. The complex matrix derivative of (6) with respect to \mathbf{Q}^* is

$$\nabla_{\mathbf{Q}^*} J = \sum_{l \neq k} \mathbf{F}_k^H (\mathbf{F}_k \mathbf{Q} \mathbf{Q}^T \mathbf{G}_l^H + \mathbf{H}_{lk}) \mathbf{G}_l \mathbf{Q}^*. \quad (7)$$

The projection of the unconstrained gradient onto the tangent space at \mathbf{Q} is $\pi_T(\nabla_{\mathbf{Q}^*} J) = \mathbf{Q} \mathbf{B}_{skew}$, where

$$\mathbf{B}_{skew} = ((\nabla_{\mathbf{Q}^*} J)^H \mathbf{Q} - \mathbf{Q}^H (\nabla_{\mathbf{Q}^*} J)) / 2 \quad (8)$$

is skew-Hermitian. Finally, to update the unitary matrix we move along the geodesic starting at \mathbf{Q} (the value at the current iteration) with direction $\nabla_{\mathbf{Q}^*} J = \mathbf{Q} \mathbf{B}_{skew}$ as

$$\mathbf{Q} = \mathbf{Q} e^{\mu \mathbf{B}_{skew}}, \quad (9)$$

where $e^{\mathbf{A}}$ denotes here the matrix exponential. The learning step size, $\mu > 0$, can be conveniently chosen and adapted using a line search procedure. A summary of the min-IL MO algorithm is shown in Algorithm 1.

a) Computational complexity: The complexity per iteration of the MO algorithm for a $(N_T \times N_R, d)^K$ MIMO-IC is as follows. The computation of the $M \times M$ unconstrained gradient

¹To simplify the notation of the remainder of the paper we will denote the double summation $\sum_{k=1}^K \sum_{l=1, l \neq k}^K$ simply as $\sum_{l \neq k}$.

²Meaning that all channels are independent of each other and their entries are also independently drawn from a continuous distribution.

³The same formulation would be valid for a given set of precoders and decoders by replacing the matrices $(\mathbf{H}_{lk}, \mathbf{F}_k, \mathbf{G}_l)$ by $(\bar{\mathbf{H}}_{lk}, \bar{\mathbf{F}}_k, \bar{\mathbf{G}}_l)$.

Algorithm 1: Min-IL MO BD-RIS.

Input: Initial $\mathbf{Q} \in \mathcal{U}(\mathcal{M})$; $\mathbf{H}_{lk}, \mathbf{F}_k, \mathbf{G}_l, \forall k \neq l$, learning rate μ
Output: Final BD-RIS $\Theta = \mathbf{Q}\mathbf{Q}^T$

```

1 while MO convergence criterion not true do
2   Calculate  $\nabla_{\mathbf{Q}} J$  as (7)
3   Find  $\mathbf{B}_{skew}$  as (8)
4   Update  $\mathbf{Q}$  as (9)

```

in (7) has complexity $\mathcal{O}(K(K-1)N_R^2N_T^2M^3)$. The computation of the skew-Hermitian matrix in (8) and the matrix exponentiation in (9) both have $\mathcal{O}(M^3)$. Hence, the overall complexity is $\mathcal{O}((K(K-1)N_R^2N_T^2 + 2)M^3)$ per iteration.

D. A RELAX-THEN-PROJECT SUBOPTIMAL SOLUTION

The computational complexity of the proposed MO algorithm, which is cubic on the number of BD-RIS elements per iteration, may be excessive for large M . Therefore, it is of interest to explore other suboptimal solutions that may be computationally more efficient. In this subsection, we solve a relaxed version of the problem (4a) that replaces the unitary constraint $\Theta^H\Theta = \mathbf{I}_M$ with a convex constraint on the trace $\text{tr}(\Theta^H\Theta) \leq M$. The solution of the relaxed problem on the set of complex and symmetric matrices admits a closed solution as a regularized least squares problem. The relaxed solution is then projected onto the set of unitary matrices to obtain the final suboptimal solution using the approach proposed in [10], [30]. We denote this algorithm as relax-then-project (RtP). The relaxed convex optimization problem is

$$\min_{\Theta} \sum_{l \neq k} \|\mathbf{H}_{lk} + \mathbf{F}_k \Theta \mathbf{G}_l^H\|_F^2 \quad (10a)$$

$$\text{s.t. } \Theta^T = \Theta, \text{tr}(\Theta^H\Theta) \leq M. \quad (10b)$$

Now we describe how to incorporate the symmetry constraint into the problem and transform it into a standard Quadratically Constrained Quadratic Program (QCQP) problem. As described in the proof of Lemma 1, the MinIL cost function can be expressed as a quadratic function of the vectorized BD-RIS coefficients $\mathbf{r} = \text{vec}(\Theta)$

$$\sum_{l \neq k} \|\mathbf{H}_{lk} + \mathbf{F}_k \Theta \mathbf{G}_l^H\|_F^2 = \text{tr}(\mathbf{T}) + \mathbf{r}^H \Sigma \mathbf{r} + 2\text{Re}(\mathbf{r}^H \mathbf{s}),$$

where \mathbf{T} , \mathbf{s} , and Σ are defined as in (5).

The linear symmetry constraint $\Theta^T = \Theta$ in (10b) can be written in terms of \mathbf{r} through the $M^2 \times M^2$ commutation matrix⁴ \mathbf{P} , which transforms $\text{vec}(\Theta) = \mathbf{r}$ into $\text{vec}(\Theta^T) = \mathbf{r}$,

$$\mathbf{r} = \mathbf{P}\mathbf{r}. \quad (11)$$

This expression implies that the vectorization of any symmetric BD-RIS matrix belongs to the nullity of the matrix $\mathbf{I}_{M^2} - \mathbf{P}$. The rank of $\mathbf{I}_{M^2} - \mathbf{P}$ is $M(M-1)/2$ and the rank of its null space is $M^2 - M(M-1)/2 = M(M+1)/2$, which

⁴A Matlab snippet to generate \mathbf{P} is $\mathbf{A} = \text{reshape}(1 : M * M, M, M)$; $\mathbf{v} = \text{reshape}(\mathbf{A}', 1, [])$; $\mathbf{P} = \text{eye}(M^2)$; $\mathbf{P} = \mathbf{P}(\mathbf{v}, :)$.

Algorithm 2: RtP MinIL BD-RIS.

1 **Input:** Initial $\mathbf{H}_{lk}, \mathbf{F}_k, \mathbf{G}_l, \forall k \neq l$
Output: Final BD-RIS Θ

```

2 Calculate  $\mathbf{s} = \text{vec}(\sum_{l \neq k} \mathbf{F}_k^H \mathbf{H}_{lk} \mathbf{G}_l)$ , and
    $\Sigma = \sum_{l \neq k} \mathbf{G}_l^T \mathbf{G}_l^* \otimes \mathbf{F}_k^H \mathbf{F}_k$ ;
3 Calculate the  $M^2 \times M^2$  commutation matrix  $\mathbf{P}$ ;
4 Calculate a basis  $\mathbf{N}$  for the nullity of  $\mathbf{I}_{M^2} - \mathbf{P}$ ;
5 Solve problem  $\mathcal{P}_2$  to obtain  $\Theta = \text{unvec}(\mathbf{N}\mathbf{x})$ ;
6 Compute  $\text{uni}(\Theta)$  projecting  $\Theta$  into the set of unitary matrices

```

is the dimension of the space of symmetric $M \times M$ matrices. Let $\mathbf{N} \in \mathbb{C}^{M^2 \times (M(M+1)/2)}$ be a basis for the nullity of $\mathbf{I}_{M^2} - \mathbf{P}$. Then, any symmetric solution of the min-IL problem is of the form

$$\mathbf{r} = \mathbf{N}\mathbf{x}, \quad (12)$$

where $\mathbf{x} \in \mathbb{C}^{(M(M+1)/2) \times 1}$. Therefore, the relaxed min-IL problem (10a) can be rewritten as the following QCQP problem

$$(\mathcal{P}_2) : \min_{\mathbf{x}} \mathbf{x}^H \mathbf{N}^H \Sigma \mathbf{N} \mathbf{x} + 2\text{Re}(\mathbf{x}^H \mathbf{N}^H \mathbf{s}) \quad (13)$$

$$\text{s.t. } \mathbf{x}^H \mathbf{x} \leq M, \quad (14)$$

whose solution is

$$\mathbf{x} = -(\mathbf{N}^H \Sigma \mathbf{N} + \lambda \mathbf{I}_{d_m})^{-1} \mathbf{N}^H \mathbf{s}, \quad (15)$$

where $d_m = M(M+1)/2$ and $\lambda \geq 0$ is a parameter selected to meet the power constraint. This regularization parameter can be found by bisection. Alternatively, we could use CVX to solve \mathcal{P}_2 . Once the solution for \mathbf{x} has been found, the optimal value of \mathbf{r} is $\mathbf{r} = \mathbf{N}\mathbf{x}$. The BD-RIS obtained as $\Theta = \text{unvec}(\mathbf{r})$ is symmetric but not unitary. The final solution is obtained by projecting Θ into the set of unitary matrices as described in [10], [30] and repeated here for completeness. Compute the SVD of the symmetric solution as $\Theta = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$, and partition the \mathbf{U} and \mathbf{V} matrices as $\mathbf{U} = [\mathbf{U}_d, \mathbf{U}_{M-d}]$ and $\mathbf{V} = [\mathbf{V}_d, \mathbf{V}_{M-d}]$, where d denotes the rank of Θ . The unitary and symmetric solution is $\text{uni}(\Theta) = [\mathbf{U}_d, \mathbf{V}_{M-d}^*] \mathbf{V}^H$.

Remark 2: An equivalent way to derive the optimal unitary projection is through Takagi's factorization [31]. The symmetric matrix Θ can be factorized as $\Theta = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$, where \mathbf{Q} is unitary and $\mathbf{\Lambda}$ is diagonal. Then, the optimal unitary projection is $\text{uni}(\Theta) = \mathbf{Q}\mathbf{Q}^T$.

A summary of the RtP BD-RIS algorithm is shown in Algorithm 2.

a) Computational complexity: To get the relaxed symmetric BD-RIS it is necessary to solve a QCQP problem of size $M(M+1)/2$, whose complexity is $\mathcal{O}((M(M+1)/2)^{3.5})$ [37], and a SVD of an $M \times M$ matrix for the unitary projection, with complexity $\mathcal{O}(M^3)$. Therefore, the total complexity of the RtP method is $(\mathcal{O}((M(M+1)/2)^{3.5}) + \mathcal{O}(M^3))$, which is dominated by the first term.

E. GROUP-CONNECTED BD-RIS

To reduce the computational and circuit complexity of the fully-connected BD-RIS architecture, a group-connected architecture was proposed in [6], [7]. The M reflective elements of a group-connected BD-RIS are partitioned into G groups, each of $M_g = M/G$ elements. The elements of each group are fully connected but disconnected from the other groups. For a group-connected architecture, the min-IL optimization problem is

$$(\mathcal{P}_{1b}) : \min_{\Theta} \sum_{l \neq k} \|\mathbf{H}_{lk} + \mathbf{F}_k \Theta \mathbf{G}_l^H\|_F^2 \quad (16a)$$

$$\text{s.t. } \Theta = \text{blkdiag}(\Theta_1, \dots, \Theta_G), \quad (16b)$$

$$\Theta_g^T = \Theta_g, \quad \Theta_g^H \Theta_g = \mathbf{I}_{M_g}. \quad (16c)$$

If the direct channel is blocked it is easy to check that the cost function (16a) decouples into G independent problems, each involving an $M_g \times M_g$ BD-RIS matrix. However, when the direct channel is not blocked the problem remains coupled among blocks, which makes it difficult to find a non-iterative solution.

The proposed procedure follows an iterative approach in which the blocks are optimized one at a time. For the optimization of the r th block Θ_r , the blocks $g = 1, \dots, r-1, r+1, \dots, G$ are considered to be fixed. This procedure decouples the optimization of the different blocks, which are now solved individually. The optimization problem for the r th block, Θ_r , is

$$\min_{\Theta_r} \sum_{l \neq k} \|\mathbf{C}_{kl} + \mathbf{F}_{kr} \Theta_r \mathbf{G}_{lr}^H\|_F^2 \quad (17a)$$

$$\text{s.t. } \Theta_r^T = \Theta_r, \quad \text{tr}(\Theta_r^H \Theta_r) \leq M_g, \quad (17b)$$

where $\mathbf{C}_{kl} = \mathbf{H}_{lk} + \sum_{g=1, g \neq r}^G \mathbf{F}_{kg} \Theta_g \mathbf{G}_{lg}^H$ denotes the fixed part of the equivalent channel, and $\mathbf{F}_{kr}, \mathbf{G}_{lr}$ are, respectively, $N_{R_k} \times M_g$ and $N_{T_l} \times M_g$ matrices formed by extracting the M_g columns of \mathbf{F}_k and \mathbf{G}_l acting on the g th BD-RIS block. To solve Problem (17), we can use either the MO algorithm described in Section III-C (cf. Algorithm 1), or the RtP version described in Section III-D (cf. Algorithm 2). The procedure is initialized to a feasible unitary and symmetric solution such as $\Theta = \mathbf{I}_M$. All groups are sequentially optimized at each iteration and the optimization stops when the IL difference in two successive iterations falls below a threshold. Note that the IL decreases or remains constant after optimizing each group; therefore, the convergence of the algorithm is guaranteed. A summary of the iterative algorithm for group-connected BD-RIS is shown in Algorithm 3.

a) *Computational complexity*: Taking the optimization of each group by the RtP method as an example, the computational cost per iteration to optimize a group-connected BD-RIS with G groups of M_g elements each is $G(\mathcal{O}((M_g(M_g + 1)/2)^{3.5}))$. The number of iterations to reach convergence is usually small (less than 20 iterations in all the simulations described in Section V). As $M_g \ll M$ the complexity to

Algorithm 3: Group-connected MinIL BD-RIS

1 **Input:** Initial $\mathbf{H}_{lk}, \mathbf{F}_k, \mathbf{G}_l, \forall k \neq l, \Theta_g = \mathbf{I}_{M_g}, g = 1, \dots, G$
Output: Final BD-RIS $\Theta = \text{blkdiag}(\Theta_1, \dots, \Theta_G)$
2 **while** Convergence criterion not true **do**
3 **for** $r = 1, \dots, G$ **do**
4 Update Θ_r with either the MO algorithm or the RtP algorithm

optimize a group-connected architecture can be significantly lower than that of a fully-connected BD-RIS.

F. MIN-IL PASSIVE LOSSLESS RIS

It may be of interest to compare the results obtained for a BD-RIS with those of a diagonal RIS $\Theta = \text{diag}(\mathbf{r})$ where $\mathbf{r} = (r_1, r_2, \dots, r_M)^T$ and r_m is the m th reflecting element. The IL cost function for a RIS has the same quadratic expression as for a BD-RIS

$$IL(\mathbf{r}) = \text{tr}(\mathbf{T}) + \mathbf{r}^H \Sigma \mathbf{r} + 2\text{Re}(\mathbf{r}^H \mathbf{s}),$$

where now $\mathbf{r} = \text{diag}(\Theta)$, $\Sigma = \sum_{l \neq k} \mathbf{G}_l^T \mathbf{G}_l^* \odot \mathbf{F}_k^H \mathbf{F}_k$ and $\mathbf{s} = \text{diag}(\sum_{l \neq k} \mathbf{F}_k^H \mathbf{H}_{lk} \mathbf{G}_l)$. Interestingly, the IL minimization problems for RIS and BD-RIS are structurally identical. Note however that the dimensions of the quadratic problem, and hence the computational complexity to solve it, are of the order of M for an RIS and of the order of M^2 for a fully-connected BD-RIS.

The minimization problem is

$$(\mathcal{P}_2) : \min_{\mathbf{r}} \mathbf{r}^H \Sigma \mathbf{r} + 2\text{Re}(\mathbf{r}^H \mathbf{s}) \quad (18a)$$

$$\text{s.t. } |r_m| = 1, \forall m. \quad (18b)$$

This is a unit-modulus quadratic programming problem. The function to be minimized is convex but the constraint is not. There are several iterative algorithms in the literature to solve \mathcal{P}_2 [38], [3], [39], [27]. The block coordinate descent (BCD) algorithm in [27] is used in this paper. At each step of the BCD algorithm, all values of \mathbf{r} but the m th $r_m = e^{j\theta_m}$ are fixed. Denoting $\bar{m} = \{1, \dots, m-1, m+1, \dots, M\}$, the IL as a function of r_m can be written as

$$IL(r_m) = \beta + 2\text{Re}(r_m^*(s_m + \Sigma_{\bar{m}}^H \mathbf{r}_{\bar{m}})), \quad (19)$$

where β is a constant, $\Sigma_{\bar{m}}$ denotes the m th column of Σ with the m th element removed, and

$$\mathbf{r}_{\bar{m}} = (r_1, \dots, r_{m-1}, r_{m+1}, \dots, r_M).$$

The optimization problem for the m th RIS coefficient is then

$$\min_{r_m} \text{Re}(r_m^*(s_m + \Sigma_{\bar{m}}^H \mathbf{r}_{\bar{m}})) \quad (20)$$

$$\text{s.t. } |z_m| = 1, \quad (21)$$

which has the following closed-form solution

$$\theta_m^* = \angle(s_m + \Sigma_{\bar{m}}^H \mathbf{r}_{\bar{m}}) - \pi. \quad (22)$$

IV. STAGE II: PRECODER DESIGN

Once the BD-RIS has been optimized to minimize the IL, the optimization of the precoders and decoders can be performed according to any of the criteria already considered in the classical MIMO literature. This section summarizes the selected criteria and the algorithms used. Although the optimization problems in this section have been studied for the MIMO-IC without BD-RIS, it is convenient to provide a summary of them for the completeness of the paper, as well as to understand the complexity analyses and simulation results in Section VI. To simplify the notation in this section we consider all users to have $N \times N$ symmetric MIMO links. In addition, remember that the equivalent channels of the BD-RIS-assisted MIMO IC after Stage I are $\tilde{\mathbf{H}}_{lk}$ (see (1)).

a) *Interference oblivious SVD-based precoders* [40]: Each user independently optimizes its precoder by considering only the receiver noise and disregarding the residual interference not eliminated by the BD-RIS. If the BD-RIS designed in Stage I reduces the interference significantly below the noise level, the MIMO-IC channel decouples into K parallel MIMO Gaussian channels and the SVD-based precoders are capacity-achieving [40]. This typically requires a very large number of BD-RIS elements. There will always be residual interference with a more limited number of elements, therefore the SVD-based solution may be rather suboptimal. For a fixed BD-RIS, the interference oblivious SVD-based transmit covariance matrix for user k is $\mathbf{R}_{xx,k} = \mathbf{V}_k \mathbf{P}_k \mathbf{V}_k^H$, where \mathbf{V}_k is the left eigenspace of the equivalent MIMO channel $\tilde{\mathbf{H}}_{kk}$ and $\mathbf{P} = \text{diag}(p_1, \dots, p_d)$ is a diagonal matrix where p_i denotes the optimal power allocated to the i th stream given by the water-filling strategy to satisfy the power constraint $\sum_i p_i = P_{t_k}$.

b) *Min-IL precoders* [21], [22]: The precoders and decoders are designed to minimize the residual IL that the BD-RIS may not have eliminated. This approach can be considered as an altruistic design of the precoders in which each user tries to generate the least interference to other users. In contrast, SVD-based precoders can be considered a selfish design in which each user tries to maximize the rate over its direct channel without considering the interference it causes to (or receives from) other users. The optimization problem is to find precoders, \mathbf{V}_k , and decoders, \mathbf{U}_k , to minimize

$$IL(\{\mathbf{V}_k\}, \{\mathbf{U}_k\}) = \sum_{l \neq k} \|\mathbf{U}_k^H \tilde{\mathbf{H}}_{lk} \mathbf{V}_l\|_F^2. \quad (23)$$

This problem can be solved by applying the AO method described in [21] that optimizes the decoders for a fixed set of precoders and then the precoders for a fixed set of decoders. Each step has a closed-form solution. Let us take the decoder of the k th user as an example. The $N \times N$ covariance matrix of the interference at the k th receiver is

$$\mathbf{S}_k = \sum_{l=1, l \neq k}^K \mathbf{S}_{kl} = \sum_{l=1, l \neq k}^K \tilde{\mathbf{H}}_{lk} \mathbf{V}_l \mathbf{V}_l^H \tilde{\mathbf{H}}_{lk}^H,$$

where $\tilde{\mathbf{H}}_{kl}$ is the equivalent MIMO channel in (1) and $\mathbf{S}_{kl} = \tilde{\mathbf{H}}_{lk} \mathbf{V}_l \mathbf{V}_l^H \tilde{\mathbf{H}}_{lk}^H$ is the interference covariance matrix from transmitter l to receiver k . The decoder \mathbf{U}_k is thus obtained as the eigenvectors of \mathbf{S}_k corresponding to their d_k smallest eigenvalues. To design the min-IL precoders, the channel reciprocity is exploited by exchanging the roles of transmitters and receivers, and transmitting along the directions where each transmitter generates less interference. Let us recall that [41] shows that there is still channel reciprocity for RIS-assisted wireless networks unless the RIS uses active nonreciprocal circuits or has structural asymmetries [42]. We will refer to the precoders designed in this way as min-IL precoders.

c) *Max-SINR precoders* [21], [43]: The min-IL precoders designed to minimize (23) do not consider the direct channels. An alternative approach that obtains a better sum rate than min-IL precoders designs the precoders and decoders to maximize the SINR [21], [43]. The max-SINR algorithm follows an AO procedure similar to the min-IL precoders. Details can be found in [21], [43].

d) *Max sum-rate precoders* [44], [45]: If interference is sufficiently reduced by the BD-RIS designed in Stage I, treating interference as noise (TIN) is the optimal decoding strategy to maximize the sum rate [46]. Utilizing TIN, the achievable rate of user k is

$$r_k = \log \det \left(\mathbf{I}_N + \left(\sigma^2 \mathbf{I}_N + \sum_{l=1, l \neq k}^K \mathbf{S}_{kl} \right)^{-1} \mathbf{S}_{kk} \right), \quad (24)$$

where $\mathbf{S}_{kl} = \tilde{\mathbf{H}}_{lk} \mathbf{V}_l \mathbf{V}_l^H \tilde{\mathbf{H}}_{lk}^H$. Thus, the sum rate maximization problem can be written as

$$\max_{\{\mathbf{V}_1, \dots, \mathbf{V}_K\}} \sum_k r_k \quad \text{s.t.} \quad \text{tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq P_{t_k}, \quad \forall k, \quad (25)$$

which is non-convex. To solve (25), we employ the majorization-minimization (MM) iterative approach, proposed in [44], [45]. To this end, we first obtain surrogate functions for the rates by leveraging the inequality in the following lemma.

Lemma 2 ([47]): Consider arbitrary $N \times N$ matrices \mathbf{A} and $\bar{\mathbf{A}}$, and $N \times N$ positive definite matrices \mathbf{Y} and $\tilde{\mathbf{Y}}$. Then, we have:

$$\begin{aligned} \ln \det(\mathbf{I}_N + \mathbf{Y}^{-1} \mathbf{A} \mathbf{A}^H) &\geq \ln \det(\mathbf{I}_N + \mathbf{Y}^{-1} \bar{\mathbf{A}} \bar{\mathbf{A}}^H) \\ &\quad - \text{tr}(\tilde{\mathbf{Y}}^{-1} \bar{\mathbf{A}} \bar{\mathbf{A}}^H) + 2 \text{Re} \{ \text{tr}(\tilde{\mathbf{Y}}^{-1} \bar{\mathbf{A}} \mathbf{A}^H) \} \\ &\quad - \text{tr}((\tilde{\mathbf{Y}}^{-1} - (\bar{\mathbf{A}} \bar{\mathbf{A}}^H + \tilde{\mathbf{Y}})^{-1})^H (\mathbf{A} \mathbf{A}^H + \mathbf{Y})) \end{aligned} \quad (26)$$

The lower bound in Lemma 2 allows us to obtain a concave lower bound for r_k as presented in the following lemma.

Lemma 3 ([45]): A concave lower bound for r_k is given by

$$\begin{aligned} r_k \geq \bar{r}_k &= a_k + 2 \sum_l \text{Re} \{ \text{tr}(\mathbf{A}_{lk} \mathbf{V}_l^H \tilde{\mathbf{H}}_{lk}^H) \} \\ &\quad - \text{tr} \left(\mathbf{B}_k \left(\sigma^2 \mathbf{I}_N + \sum_{l=1}^K \tilde{\mathbf{H}}_{lk} \mathbf{V}_l \mathbf{V}_l^H \tilde{\mathbf{H}}_{lk}^H \right) \right), \end{aligned} \quad (27)$$

where

$$\begin{aligned}
 a_k &= \ln \det (\mathbf{I}_N + \mathbf{R}_k(\bar{\mathbf{V}}_l) \tilde{\mathbf{H}}_{lk} \bar{\mathbf{V}}_k \bar{\mathbf{V}}_k^H \tilde{\mathbf{H}}_{lk}^H) \\
 &\quad - \text{tr} (\mathbf{R}_k(\bar{\mathbf{V}}_l) \tilde{\mathbf{H}}_{lk} \bar{\mathbf{V}}_k \bar{\mathbf{V}}_k^H \tilde{\mathbf{H}}_{lk}^H), \\
 \mathbf{A}_{lk} &= \mathbf{R}_k(\bar{\mathbf{V}}_l) \tilde{\mathbf{H}}_{lk} \bar{\mathbf{V}}_k, \\
 \mathbf{B}_k &= \mathbf{R}_k(\bar{\mathbf{V}}_l) - \left(\sigma^2 \mathbf{I}_N + \sum_{l=1}^K \mathbf{S}_k(\bar{\mathbf{V}}_l) \right)^{-1}, \\
 \mathbf{R}_k(\bar{\mathbf{V}}_l) &= \left(\sigma^2 \mathbf{I}_N + \sum_{l=1, l \neq k}^K \tilde{\mathbf{H}}_{lk} \bar{\mathbf{V}}_l \bar{\mathbf{V}}_l^H \tilde{\mathbf{H}}_{lk}^H \right)^{-1}, \\
 \mathbf{S}_k(\bar{\mathbf{V}}_l) &= \tilde{\mathbf{H}}_{lk} \bar{\mathbf{V}}_l \bar{\mathbf{V}}_l^H \tilde{\mathbf{H}}_{lk}^H,
 \end{aligned}$$

and $\bar{\mathbf{V}}_l$ is the initial value of \mathbf{V}_l for all $l = 1, \dots, K$.

Substituting r_k by \bar{r}_k yields the convex problem

$$\max_{\{\mathbf{V}_1, \dots, \mathbf{V}_K\}} \sum_k \bar{r}_k \quad \text{s.t.} \quad \text{tr} (\mathbf{V}_k \mathbf{V}_k^H) \leq P_{t_k}, \forall k. \quad (28)$$

Iteratively solving (28) and updating $\{\bar{\mathbf{V}}_1, \dots, \bar{\mathbf{V}}_K\}$ converges to a stationary point of (25).

Remark 3: Note that the SVD-based and max-SR precoders optimize the number of transmitted streams. However, this number is fixed and predetermined, $\sum_{k=1}^K d_k$, for the min-IL and max-SINR precoders.

e) Computational complexity: Computing the SVD of the direct MIMO links dominates the computational complexity of designing SVD-based precoders and therefore is cubic in the number of antennas $\mathcal{O}(KN^3)$. The design of min-IL or max-SINR precoders using AO involves $\mathcal{O}(2KN^3)$ floating-point operations at each AO iteration due to the computation of SVDs for optimizing the K precoders and K decoders. Finally, each iteration of the MM procedure to design the max-SR precoders involves an outer iteration to compute the surrogate function for the rates, whose computational complexity is dominated by the matrix inversion in Lemma 3, $K\mathcal{O}(N^3)$, and an inner iteration in which a convex problem has to be solved. The number of Newton iterations to solve a convex problem is proportional to the square root of the number of constraints [48], and hence it grows with \sqrt{K} . To solve at each Newton step, the calculation of the surrogate function is $\mathcal{O}((K+1)N^2)$. In summary, the computational complexity for designing max-SR precoders per MM iteration is $\mathcal{O}(KN^3 + \sqrt{K}(K+1)N^2)$.

A. JOINT MIN-IL PRECODER AND BD-RIS DESIGN

The proposed two-stage optimization approach has clear computational advantages over the joint optimization of the BD-RIS and the precoders. However, it is relatively straightforward to implement an AO algorithm that jointly optimizes the BD-RIS and the precoders. As an example, we consider the joint optimization of precoders, decoders, and BD-RIS to

Algorithm 4: Joint min-IL design.

Input: Channels, initial $\{\mathbf{U}_k\}_{k=1}^K$ and $\{\mathbf{V}_k\}_{k=1}^K$; IL_{NoRIS} (IL without RIS), IL convergence threshold $\epsilon > 0$
Output: Final BD-RIS Θ , \mathbf{U}_k and \mathbf{V}_k for $k = 1, \dots, K$

```

1 while Relative IL reduction smaller than  $\epsilon$  do
  /* Step 1: Given  $\Theta$  and  $\{\mathbf{V}_k\}_{k=1}^K$  */
2  Calculate  $\mathbf{S}_k$  at the  $k$ th receiver and obtain  $\mathbf{U}_k$  from the
   eigenvectors corresponding to its smallest  $d_k$ 
   eigenvalues  $k = 1, \dots, K$ 
  /* Step 2: Given  $\Theta$  and  $\{\mathbf{U}_k\}_{k=1}^K$  */
3  Calculate  $\mathbf{S}_k$  at the  $k$ th transmitter and obtain  $\mathbf{V}_k$  from
   the eigenvectors corresponding to its smallest  $d_k$ 
   eigenvalues  $k = 1, \dots, K$ 
  /* Step 3: Given  $\{\mathbf{V}_k\}_{k=1}^K$  and  $\{\mathbf{U}_k\}_{k=1}^K$ 
   update  $\Theta$  with the MO algorithm
   described in Algorithm 1 */

```

minimize the IL cost function

$$IL(\{\mathbf{V}_k\}, \{\mathbf{U}_k\}, \Theta) = \sum_{l \neq k} \|\mathbf{U}_k^H \mathbf{H}_{lk} \mathbf{V}_l + \mathbf{U}_k^H \mathbf{F}_k \Theta \mathbf{G}_l^H \mathbf{V}_l\|_F^2. \quad (29)$$

The minimization of (29) may be carried out through a 3-step AO process, where in each step a set of variables (decoders, precoders, or BD-RIS) is optimized while the other variables are held fixed:

- 1) Optimize $\{\mathbf{U}_k\}_{k=1}^K$ for fixed $(\{\mathbf{V}_l\}_{l=1}^K, \Theta)$.
- 2) Optimize $\{\mathbf{V}_l\}_{l=1}^K$ for fixed $(\{\mathbf{U}_k\}_{k=1}^K, \Theta)$.
- 3) Optimize Θ for fixed $(\{\mathbf{U}_k\}_{k=1}^K, \{\mathbf{V}_l\}_{l=1}^K)$.

Steps 1 and 2 can be solved by applying the min-IL design of the precoders described in Section IV, while Step 3 can be solved with the MO algorithm of Section III. Algorithm 4 details the joint min-IL optimization procedure.

V. SIMULATION RESULTS

A. SCENARIO DESCRIPTION

We consider a $(3 \times 3, 2)^3$ MIMO-IC assisted by a BD-RIS with M elements as represented in Fig. 2. For simplicity, all the MIMO links have the same number of antennas, 3×3 , and transmit the same number of streams, $d = 2$. In this scenario, IA is not feasible by optimizing only the precoders and decoders [32]. The system is strongly limited by interference and it is necessary to assist the channel with a BD-RIS optimized to minimize the interference. For this scenario, Lemma 1 states that an unconstrained (i.e., active) BD-RIS with $M = 8$ elements could achieve $IL = 0$, since in this case Σ in (5) is a 64×64 matrix with rank $\min(M^2, \sum_{k \neq l} N_{R_k} N_{T_l}) = 54$. Of course, a passive BD-RIS design will require a much larger number of elements to achieve a significant interference reduction level.

The $K = 3$ transmitters and receivers are regularly located in a square of 50 m. The coordinates (x, y, z) in meters $[m]$ of the k th transmitter and the k th receiver are $(0, 50 \frac{(k-1)}{(K-1)}, 1.5)$ and $(50, 50 \frac{(k-1)}{(K-1)}, 1.5)$ $k = 1, \dots, K$, respectively. The BD-RIS located at $(x, y, 5)$ $[m]$, where the (x, y) coordinates can vary between 5 and 45 m. Since the BD-RIS is located higher

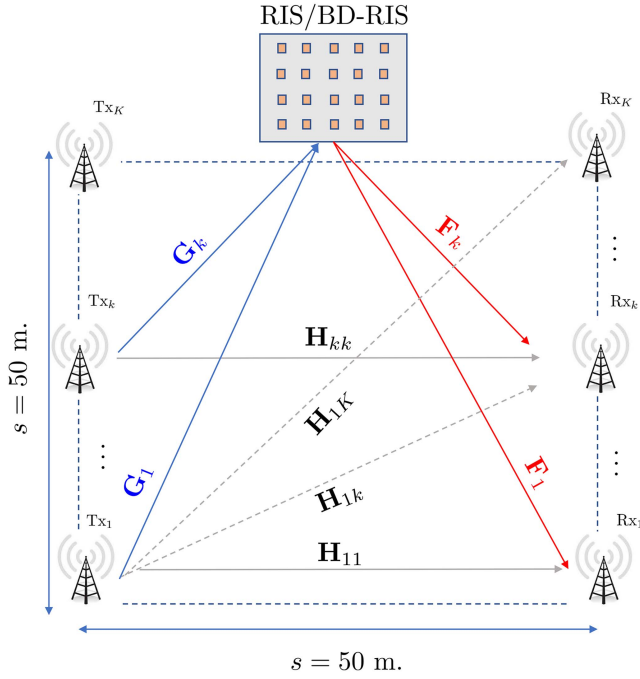


FIGURE 2. Simulation setup for the beyond diagonal RIS-assisted MIMO-IC.

than the transmitters and receivers with a direct line-of-sight (LoS) to all of them, it is assumed that all transmitters and receivers are in the reflection space of the RIS. The large-scale path loss in dB is given by

$$PL = -28 - 10\alpha \log_{10}(r),$$

where r is the link distance, and α is the path-loss exponent. The direct Tx-Rx links, \mathbf{H}_{lk} , are assumed to be non-line-of-sight (NLoS) channels, with path-loss exponent $\alpha = 3.75$ and small-scale Rayleigh fading. The Tx-RIS-Rx links, \mathbf{F}_k and \mathbf{G}_l , are assumed to be LoS channels, with path-loss exponent $\alpha = 2$ and small-scale Rice fading with Rician factor $\gamma = 3$. The noise variance is calculated as $\sigma^2 = -174(\text{dBm/Hz}) + 10 \log_{10} B(\text{Hz}) + F(\text{dB})$, where B is the bandwidth that we take in our simulations as $B = 40$ MHz and F is the noise figure of the receiver that we take as $F = 10$ dB. A new set of channels is generated in each simulation, but the positions of TxS, RxS, and BD-RIS are fixed. We average the results of 50 independent simulations. The power transmitted by all users is $P_t = 10$ dBm.

As a figure of merit, we use the difference in dBs between the IL ratios with and without BD-RIS, which is also the difference in dBs between the INR with and without BD-RIS

$$\Delta \text{INR} = 10 \log_{10} \left(\frac{\sum_{l \neq k} \|\mathbf{H}_{lk} + \mathbf{F}_k \mathbf{\Theta} \mathbf{G}_l^H\|_F^2}{\sum_{l \neq k} \|\mathbf{H}_{lk}\|_F^2} \right). \quad (30)$$

Another figure of merit used in the simulations is the sum rate

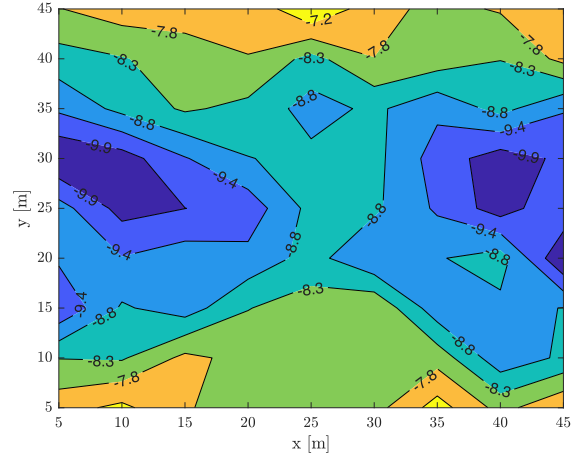


FIGURE 3. ΔINR in dB as a function of the position of a BD-RIS with $M = 40$ elements in a $(3 \times 3, 2)^3$ MIMO-IC.

of the MIMO-IC

$$R = \sum_{k=1}^K r_k, \quad (31)$$

where r_k is the rate of user k given by (24).

B. OPTIMAL BD-RIS DEPLOYMENT

The position of the reconfigurable surface or its *area of influence* is an essential aspect in optimizing system performance, as shown in [49], [50]. Therefore, we first evaluate the reduction in interference measured as in (30) achieved by varying the position in the (x, y) plane of a fully-connected BD-RIS with $M = 40$ elements. The positions in the (x, y) plane are varied on a grid of 5 [m] along each coordinate while the height remains fixed at $z = 5$ [m]. For the optimization of the BD-RIS we use the MO algorithm described in Algorithm 1. Fig. 3 shows the results obtained using a color map, where we can observe two optimal positions or hot spots for the BD-RIS at approximately (10, 25, 5) [m] and (40, 25, 5) [m], where a reduction of about 10 dB in the interference level is achieved with respect to a scenario without BD-RIS. Unless otherwise stated in the remaining experiments on Scenario 1, we will position the BD-RIS at coordinates (40, 25, 5) [m].

C. MO AND RTP CONVERGENCE

Fig. 4 illustrates the convergence of the MO algorithm used to design a min-IL fully-connected BD-RIS with $M = 40$ elements. Each curve shows the convergence for a different random initialization. The MO algorithm requires in this example 500 iterations to reach convergence. Remember that the complexity of each iteration is $\mathcal{O}(M^3)$, which may represent a difficulty when optimizing a BD-RIS with many elements. One way to reduce the computational and circuit complexities of the fully-connected BD-RIS is to employ a group-connected architecture. We study the computational complexity of Algorithm 3 when each group is optimized

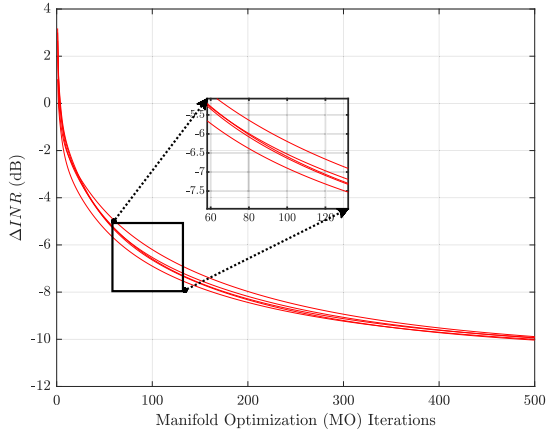


FIGURE 4. Convergence of the MO iterative algorithm for a fully-connected BD-RIS with $M = 40$ elements. Each curve shows the convergence for a different random initialization.

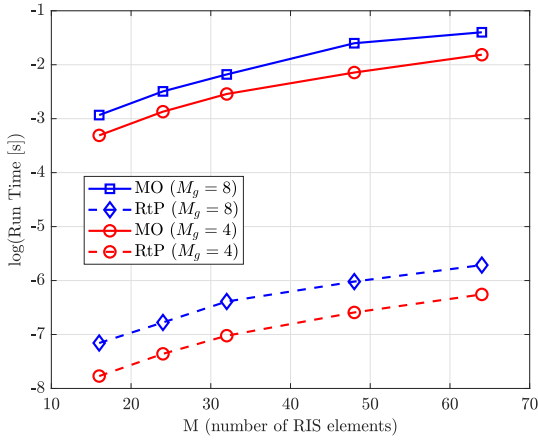


FIGURE 5. Run time of the MO and RtP algorithms for the group-connected BD-RIS architecture vs. M (total number of BD-RIS elements).

either by the MO algorithm or by the RtP method. As the number of iterations for the convergence of the two methods can be very different, we use the average run time in logarithmic scale as a figure of merit. The results for a group-connected BD-RIS with $M_g = 4$ and $M_g = 8$ for increasing values of M are shown in Fig. 5 for the MO (solid line) and the RtP method (dashed line). The computational cost of the MO algorithm is several orders of magnitude higher than that of RtP. As expected, for both the MO and RtP the computational cost grows with M_g and with the total number of elements M and therefore the number of blocks to be optimized $G = M/M_g$.

D. FULLY-CONNECTED VS. GROUP-CONNECTED BD-RIS

In this section, we compare the interference suppression capability of the fully connected and group-connected architectures for various group sizes M_g . The BD-RIS is located at (40, 25, 5) [m] and the interference suppression capability for the different architectures and optimization algorithms is measured through the ΔINR defined in (30). We employ either the MO algorithm or the RtP method

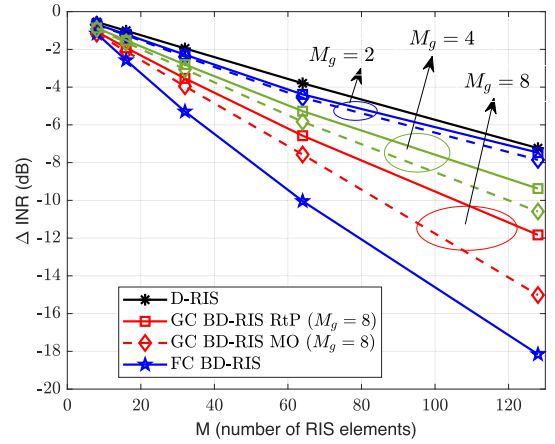


FIGURE 6. ΔINR in dB as a function of M for a fully-connected BD-RIS, a group-connected BD-RIS with group sizes $M_g = 2, 4, 8$ optimized with either MO or RtP, and a diagonal RIS.

to optimize the group-connected BD-RIS architecture. For completeness, we include the results obtained for a diagonal RIS, which can be considered a limiting case of the group-connected architecture with $M_g = 1$. The results are shown in Fig. 6. For instance, an RIS with $M = 128$ elements achieves a reduction in interference of about 7.5 dB compared to the no-RIS scenario. The interference suppression is improved by about 0.7 dB, 3 dB and 8 dB for group-connected BD-RIS of sizes $M_g = 2, 4, 8$ optimized with the MO algorithm; and more than 11 dB when a fully-connected BD-RIS is employed. We observe that to achieve a significant interference reduction at the receivers it is necessary to employ a BD-RIS with many elements. The group-connected architecture optimized with the MO algorithm with a group size $M_g = 8$ provides a reasonable tradeoff between performance and complexity. Moreover, compared to fully-connected architectures, group-connected architectures reduce not only the computational complexity of the optimization algorithms but also the number of interconnections between elements.

E. SUM RATE PERFORMANCE

In this subsection, we study the sum-rate performance of the K -user MIMO-IC assisted by a fully-connected BD-RIS. The BD-RIS is optimized in Stage I to minimize IL with the MO algorithm. In Stage II, we design precoders to optimize one of the following criteria: SVD-based, min-IL, max-SINR, or max sum-rate (labeled as max-SR in the figures). Additionally, we evaluate the joint min-IL design of the BD-RIS and the precoders as described in Algorithm 4. Fig. 7 shows the sum rate vs. the transmit power curves. The MIMO-IC is still interference-limited when the BD-RIS has only $M = 16$ elements and the transmitted power is high. Therefore, the sum-rate curves for $M = 16$ tend to saturate at high P_t . This is especially clear for the SVD-based precoders that do not take interference into account. However, a joint design of the BD-RIS and the precoders ensures that the MIMO-IC is limited mainly by noise, even at high P_t . When the number of BD-RIS

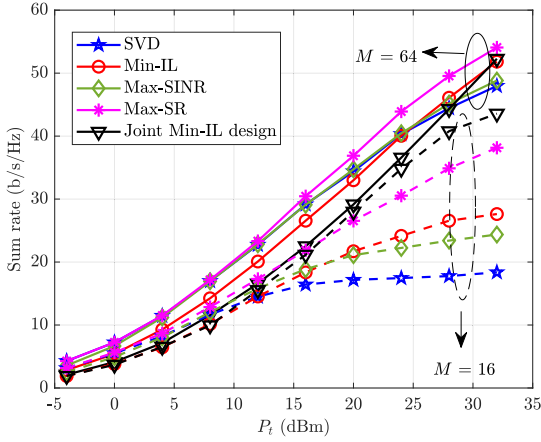


FIGURE 7. Sum rate vs. P_t for a BD-RIS assisted $(3 \times 3, 2)^3$ MIMO-IC with different precoders. The BD-RIS has $M = 16$ (dashed lines) or $M = 64$ elements (solid lines).

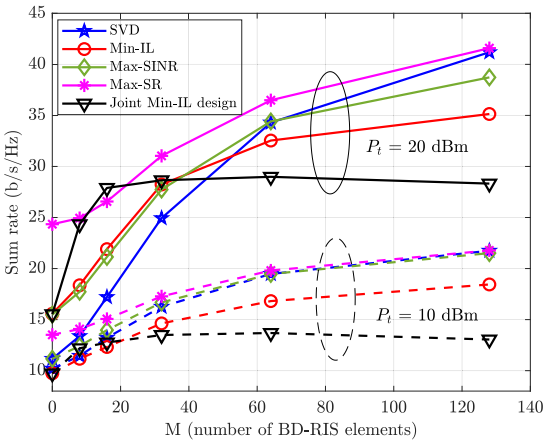


FIGURE 8. Sum rate vs. M for a BD-RIS assisted $(3 \times 3, 2)^3$ MIMO-IC with different precoders. The users transmit $P_t = 10$ dBm (dashed lines) or $P_t = 20$ dBm (solid lines).

elements increases to $M = 64$, the BD-RIS is much more effective in reducing interference, and hence the differences between the different precoders are less pronounced. Among the different precoders, the max-SR design provides the best performance, with the drawback of a higher computational cost.

Fig. 8 shows the sum rate obtained by different precoders designed in Stage II as a function of the number of BD-RIS elements for two transmission powers. When the power transmitted by the users is $P_t = 10$ dBm, the interference level is not very high. Therefore, a BD-RIS with a moderate number of elements can eliminate much of this interference thus transforming the MIMO-IC into $K = 3$ parallel MIMO Gaussian channels. This is why the differences between max-SR, max-SINR, and SVD-based precoders are negligible for $M \geq 60$. When the power transmitted by the users increases to $P_t = 20$ dBm, the residual interference not eliminated by the BD-RIS is higher and, therefore, the differences between the performance of the precoders are more remarkable. The max-SR precoders provide the best performance as expected. When

the precoders and BD-RIS are jointly designed to minimize IL, a BD-RIS with $M = 20$ elements is capable of drowning out interference below the noise level, and hence increasing M does not translate into a sum rate improvement.

VI. CONCLUSION

This work proposed a two-stage passive and active beam-forming scheme for a BD-RIS-assisted K -user MIMO-IC. The passive BD-RIS, whose scattering matrix is unitary and symmetric, is designed in the first stage to minimize the IL. A geodesic MO algorithm on the manifold of unitary matrices, which exploits Takagi's factorization, is proposed to optimize a fully connected BD-RIS. For a group-connected BD-RIS architecture, a computationally simpler approach is proposed that first obtains a symmetric BD-RIS solving a relaxed problem with a closed-form solution, and then projects the relaxed solution onto the set of unitary matrices. The (active) precoders are designed in the second stage according to different criteria that treat the residual interference as noise, such as SVD-based, min-IL, max-SINR, and max-SR precoders. Max-SR precoders are particularly effective, especially when the BD-RIS has a moderate number of elements and the residual interference is still greater than the noise power. When the BD-RIS alone can reduce the IL below the noise floor, the K -user MIMO-IC approaches K parallel Gaussian MIMO channels, and the performance of the SVD-based precoders approaches that of the max-SR precoders at a lower computational cost. The study has focused on ICs assisted by a single BD-RIS, which must be optimally located to optimize its area of influence. A future line of study considers multiple BD-RIS, which requires refined signal models accounting for the double/multiple reflections and new optimization algorithms. In addition, it would be interesting to study other criteria for designing the precoders, such as minimizing the transmit power under rate constraints.

REFERENCES

- [1] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
- [2] "Reconfigurable intelligent surfaces (RIS): Use cases, deployment scenarios and requirements," Reconfigurable Intell. Surfaces ETSI Ind. Specification Group, Tech. Rep. DGR/RIS-001, 2023.
- [3] C. Pan et al., "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.
- [4] 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, pp. 4157–4170, Aug. 2019.
- [5] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 5394–5409, Nov. 2019.
- [6] 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.
- [7] 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, pp. 1229–1243, Feb. 2022.

- [8] I. Santamaria, M. Soleymani, E. Jorswieck, and J. Gutiérrez, "SNR maximization in beyond diagonal RIS-assisted single and multiple antenna links," *IEEE Signal Process. Lett.*, vol. 30, pp. 923–926, 2023.
- [9] 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.
- [10] T. Fang and Y. Mao, "A low-complexity beamforming design for beyond-diagonal RIS aided multi-user networks," *IEEE Comm. Lett.*, vol. 28, no. 1, pp. 203–207, Jan. 2024.
- [11] H. Li, S. Shen, and B. Clerckx, "Synergizing beyond diagonal reconfigurable intelligent surface and rate-splitting multiple access," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8717–8729, Aug. 2024.
- [12] H. Li, S. Shen, and B. Clerckx, "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, Apr. 2023.
- [13] A. Mishra, Y. Mao, C. D'Andrea, S. Buzzi, and B. Clerckx, "Transmitter side beyond-diagonal reconfigurable intelligent surface for massive MIMO networks," *IEEE Comm. Lett.*, vol. 13, no. 2, pp. 352–356, Feb. 2024.
- [14] M. Soleymani, I. Santamaria, E. Jorswieck, and B. Clerckx, "Optimization of rate-splitting multiple access in beyond diagonal RIS-assisted URLLC systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 5063–5078, May 2024.
- [15] Q. Shi and M. Hong, "Penalty dual decomposition method for non-smooth nonconvex optimization- Part I: Algorithms and convergence analysis," *IEEE Trans. Signal Process.*, vol. 68, pp. 4102–4122, 2020.
- [16] 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.
- [17] Y. Zhao, H. Li, M. Franceschetti, and B. Clerckx, "Channel shaping using beyond diagonal reconfigurable intelligent surface: Analysis, optimization, and enhanced flexibility," 2024, *arXiv:2407.15196*.
- [18] I. Santamaria, M. Soleymani, E. Jorswieck, and J. Gutiérrez, "MIMO capacity maximization with beyond-diagonal RIS," in *Proc. IEEE 25th Int. Workshop Signal Process. Adv. Wireless Commun.*, Lucca, Italy, pp. 936–940, 2024.
- [19] H. Yahya, H. Li, M. Nerini, B. Clerckx, and M. Debahh, "Beyond diagonal RIS: Passive maximum ratio transmission and interference nulling enabler," *IEEE Open Jour. Comm. Soc.*, vol. 5, pp. 7613–7627, 2024.
- [20] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the k-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [21] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.
- [22] O. González, C. Lameiro, and I. Santamaria, "A quadratically convergent method for interference alignment in MIMO interference channels," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1423–1427, Nov. 2014.
- [23] T. Jiang and W. Yu, "Interference nulling using reconfigurable intelligent surface," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 5, pp. 1392–1406, May 2022.
- [24] M. Fu, Y. Zhou, and Y. Shi, "Reconfigurable intelligent surface for interference alignment in MIMO device-to-device networks," in *Proc. IEEE Int. Conf. Commun.*, 2021, pp. 1–6.
- [25] A. Abrado, D. Dardari, M. Di Renzo, and X. Qian, "MIMO interference channels assisted by reconfigurable intelligent surfaces: Mutual coupling aware sum-rate optimization based on mutual impedance channel model," *IEEE Wireless Comm. Lett.*, vol. 10, no. 12, pp. 2624–2628, Dec. 2021.
- [26] A. H. Abdollahi Bafghi, V. Jamali, M. Nasiri-Kenari, and R. Schober, "Degrees of freedom of the K-user interference channel assisted by active and passive IRSs," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3063–3080, May 2022.
- [27] I. Santamaria, M. Soleymani, E. Jorswieck, and J. Gutiérrez, "Interference leakage minimization in RIS-assisted MIMO interference channels," in *Proc. 2023 IEEE Int. Conf. Acoust., Speech Signal Process.*, Rhodes, Greece, 2023, pp. 1–5, doi: [10.1109/ICASSP49357.2023.10094656](https://doi.org/10.1109/ICASSP49357.2023.10094656).
- [28] W. Liu, A. Li, Y. Xu, Z. Zhang, Y. Zhang, and B. Jiang, "RIS-aided interference subspace alignment for k-user MIMO interference networks," *IEEE Comm. Lett.*, vol. 28, no. 8, pp. 1919–1923, Aug. 2024.
- [29] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. A. Jafar, "On the optimality of treating interference as noise," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1753–1767, Apr. 2015.
- [30] X. Zhou, T. Fang, and Y. Mao, "Joint active and passive beamforming optimization for beyond diagonal RIS-aided multi-user communications," *IEEE Comm. Lett.*, vol. 29, no. 3, pp. 517–529, Mar. 2025.
- [31] T. Takagi, "On an algebraic problem related to an analytic theorem of carathéodory and fejér and on an allied theorem of landau," *Jpn. J. Math. Trans. Abstr.*, vol. 1, pp. 83–93, 1924.
- [32] Ó. González, C. Beltrán, and I. Santamaria, "A feasibility test for linear interference alignment in MIMO channels with constant coefficients," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1840–1856, Mar. 2014.
- [33] D. Ramírez, I. Santamaria, and L. Scharf, *Coherence: In Signal Processing and Machine Learning*. Berlin, Germany: Springer, 2023.
- [34] R. A. Horn and Z. Yang, "Rank of a hadamard product," *Linear Algebra Appl.*, vol. 591, pp. 87–98, 2020.
- [35] L. Autonne, "Sur les matrices hypohermitiennes et sur les matrices unitaires," *Ann. Univ. Lyon*, vol. 38, pp. 1–77, 1915.
- [36] N. Boumal, *An Introduction to Optimization on Smooth Manifolds*. Cambridge, U.K.: Cambridge Univ. Press, 2023. [Online]. Available: <https://www.nicolasboumal.net/book>
- [37] Z.-Q. Luo, W. K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
- [38] J. Tranter, N. D. Sidiropoulos, X. Fu, and A. Swami, "Fast unit-modulus least squares with applications in beamforming," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2875–2887, Jun. 2017.
- [39] C. G. Tsinos and B. Ottersten, "An efficient algorithm for unit-modulus quadratic programs with application in beamforming for wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 25, no. 2, pp. 169–173, Feb. 2017.
- [40] R. S. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun. 2003.
- [41] W. Tang et al., "On channel reciprocity in reconfigurable intelligent surface assisted wireless networks," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 94–101, Dec. 2021.
- [42] H. Wang, Z. Han, and S. A. L., "Channel reciprocity attacks using intelligent surfaces with non-diagonal phase shifts," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1469–1485, 2024.
- [43] C. Wilson and V. Veeravalli, "A convergent version of the max SINR algorithm for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 2952–2961, Jun. 2013.
- [44] M. Soleymani, I. Santamaria, and P. J. Schreier, "Improper gaussian signaling for the K-user MIMO interference channels with hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11632–11645, Oct. 2020.
- [45] M. Soleymani, I. Santamaria, E. Jorswieck, R. Schober, and L. Hanzo, "Optimization of the downlink spectral and energy-efficiency of RIS-aided multi-user URLLC MIMO systems," *IEEE Trans. Commun.*, early access Oct. 15, 2024, doi: [10.1109/TCOMM.2024.3480988](https://doi.org/10.1109/TCOMM.2024.3480988).
- [46] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.
- [47] M. Soleymani, I. Santamaria, and P. J. Schreier, "Improper signaling for multicell MIMO RIS-assisted broadcast channels with I/Q imbalance," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 2, pp. 723–738, Jun. 2022.
- [48] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [49] G. C. Alexandropoulos et al., "RIS-enabled smart wireless environments: Deployment scenarios, network architecture, bandwidth and area of influence," *EURASIP J. Wireless Commun. Netw.*, vol. 2023, 2023, Art. no. 103.
- [50] Z. Zhang et al., "Active RIS vs. passive RIS: Which will prevail in 6G?," *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1707–1725, Mar. 2023.