Channel Shaping Using Reconfigurable Intelligent Surfaces: From Diagonal to Beyond

Yang Zhao, *Member, IEEE,* Hongyu Li, *Graduate Student Member, IEEE,*
Massimo Franceschetti, *Fellow, IEEE,* and Bruno Clerckx, *Fellow, IEEE*

***Abstract*—This paper investigates how a passive Reconfigurable Intelligent Surface (RIS) can reshape the Multiple-Input Multiple-Output (MIMO) point-to-point channel in terms of singular values. We depart from the widely-adapted diagonal phase shift model to a general Beyond-Diagonal (BD) architecture, which provides superior shaping capability thanks to in-group connections between elements. An efficient Riemannian Conjugate Gradient (RCG) algorithm is tailored for smooth optimization problems of asymmetric BD-RIS with arbitrary group size, then invoked for the Pareto frontier of channel singular values. To understand the gain from off-diagonal entries, we also derive analytical singular value bounds in rank-deficient and fully-connected scenarios. As a side product, we tackle MIMO rate maximization problem by alternating between active beamformer (eigenmode transmission) and passive beamformer (RCG algorithm) until convergence. A low-complexity suboptimal solution based on channel shaping is also proposed, where the decoupled problem is formulated as channel power maximization and solved in closed form iteratively. Theoretical analysis and numerical evaluation reveal that the shaping advantage of BD-RIS increases with group size and MIMO dimensions, stemming from stronger subchannel rearrangement and subspace alignment capabilities.**

***Index Terms*—Reconfigurable intelligent surface, multi-input multi-output, manifold optimization, singular value control, rate maximization.**

I. Introduction

Today we are witnessing a paradigm shift from connectivity to intelligence, where the wireless environment is no longer a chaotic medium but a conscious agent that serves on demand. This is empowered by the recent advances in Reconfigurable Intelligent Surface (RIS), a real-time programmable metasurface of numerous non-resonant sub-wavelength scattering elements. It can manipulate the amplitude, phase, frequency, and polarization of the scattered waves [1] with a higher energy efficiency, lower cost, lighter footprint, and greater scalability than relays. Using RIS for passive beamforming has attracted significant interest in wireless communication [2]–[5], backscatter [6], [7], sensing [8], [9], and power transfer literature [10]–[12], reporting a second-order array gain and fourth-order power scaling law (with proper waveform). On the other hand, RIS also enables backscatter modulation by dynamically switching between different patterns, as already investigated [13]–[15] and prototyped [16], [17]. Despite fruitful outcomes, one critical unanswered question is the channel shaping capability: *To what extent can a passive RIS reshape the wireless channel?*

The answer indeed depends on the hardware architecture and scattering model. In conventional (a.k.a. diagonal) RIS, each scattering element is tuned by a dedicated impedance and acts as an *individual* phase shifter [18]. The concept is generalized to Beyond-Diagonal (BD)-RIS [19], [20] which groups adjacent elements using passive components. This allows *cooperative* scattering — wave impinging on one element can propagate within the circuit and depart partially from any element in the same group. BD-RIS can thus control both amplitude and phase of the reflected wave, generalizing the scattering matrix from diagonal with unit-magnitude entries to block diagonal with unitary blocks. Its benefit has been recently shown in receive power maximization [21]–[24], transmit power minimization [25], and rate maximization [24]–[28]. Practical issues such as channel estimation [29] and mutual coupling [30] have also been investigated. Therefore, BD-RIS is envisioned as the next generation channel shaper with stronger signal processing flexibility [31].

Channel shaping is different from passive beamforming as it seeks to modify the inherent properties of the channel itself. This allows one to decouple the RIS-transceiver design and explore the fundamental limits of channel manipulation. For example, diagonal RIS has been proved useful for improving channel power [32], degree of freedom [33], [34], condition number [35], [36], and effective rank [37], [38] in Multiple-Input Multiple-Output (MIMO). In contrast, BD-RIS can provide a higher channel power but existing results are limited to Single-Input Single-Output (SISO)[1]. [21] and Multiple-Input Single-Output (MISO) [22]. While these studies offer promising glimpses into the channel shaping potential, a comprehensive understanding of the capabilities and limitations is desired, and a universal design framework is missing. This paper aims to answer the channel shaping question through theoretical analysis and numerical optimization. The contributions are summarized below.

First, we quantify the capability of a BD-RIS to reshape the MIMO point-to-point channel in terms of singular values. The *Pareto frontiers* are characterized by optimizing the weighted sum of singular values, where the weights can be positive, zero, or negative. The resulting singular value region generalizes most relevant metrics and provides an intuitive channel shaping benchmark. We then discuss some analytical singular value bounds in rank-deficient and fully-connected scenarios, which help to demystify the gain from off-diagonal entries. This is the first paper to answer the channel shaping question and highlight the BD-RIS gain from a Pareto perspective.

Second, we propose a Riemannian Conjugate Gradient (RCG) algorithm adapted from [39], [40] for smooth optimization problems of asymmetric BD-RIS with arbitrary

[1]In terms of channel shaping, single-stream MIMO with given precoder and combiner [21] is equivalent to SISO.

group size. Specifically, block-wise update is performed along the geodesics[2] of the Stiefel manifold, which are expressed compactly by the exponential map [41]. It features lower complexity and faster convergence than general manifold optimization [42], [43], and solves the Pareto singular value problem. This is the first paper to tailor an efficient optimization framework for asymmetric BD-RIS.

Third, we tackle BD-RIS MIMO rate maximization with two solutions: a local-optimal approach through Alternating Optimization (AO) and a low-complexity approach over channel shaping. The former updates active and passive beamformers by eigenmode transmission and RCG algorithm, respectively. The latter suboptimally decouples both blocks, recasts the shaping problem as channel power maximization, and solves it iteratively in closed form. Interestingly, the gap in between vanishes as BD-RIS evolves from diagonal (single-connected) to unitary (fully-connected). It suggests channel shaping offers a promising low-complexity solution for joint RIS-transceiver designs.

Fourth, extensive simulations reveal that the performance gain from BD-RIS increases with group size and MIMO dimensions. In terms of channel power, fully-connected BD-RIS boosts up to 62% and 270% over diagonal RIS in 1×1 and 4×4 MIMO under independent Rayleigh fading, respectively. The superiority stems from stronger *subchannel rearrangement* and *subspace alignment* capabilities empowered by in-group cooperation. It emphasizes the importance of using BD-RIS in large-scale MIMO systems.

Notation: Italic, bold lower-case, and bold upper-case letters indicate scalars, vectors and matrices, respectively. $\jmath$ denotes the imaginary unit. $\mathbb{C}$ represents the set of complex numbers. $\mathbb{H}^{n\times n}$ and $\mathbb{U}^{n\times n}$ denotes the set of $n\times n$ Hermitian and unitary matrices, respectively. $\mathbf{0}$ and $\mathbf{I}$ are the all-zero and identity matrices with appropriate size, respectively. $\Re\{\cdot\}$ takes the real part of a complex number. $\mathrm{tr}(\cdot)$ and $\det(\cdot)$ evaluates the trace and determinant of a square matrix, respectively. $\mathrm{diag}(\cdot)$ constructs a square matrix with arguments on the main diagonal and zeros elsewhere. $\mathrm{sv}(\cdot)$ returns the singular value vector. $\sigma_n(\cdot)$ and $\lambda_n(\cdot)$ is the n-th largest singular value and eigenvalue, respectively. $(\cdot)^*$, $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, $(\cdot)^\dagger$ $(\cdot)^{(r)}$, $(\cdot)^\star$ denote the conjugate, transpose, conjugate transpose (Hermitian), Moore-Penrose inverse, r-th iterated point, and final solution, respectively. $(\cdot)_{[x:y]}$ is a shortcut for $(\cdot)_x,(\cdot)_{x+1},...,(\cdot)_y$. $|\cdot|$ denotes the absolute value. $\|\cdot\|$ means the Euclidean norm. $\|\cdot\|_{\mathrm{F}}$ represents the Frobenius norm. $\mathcal{CN}(\mathbf{0},\mathbf{\Sigma})$ is the multivariate Circularly Symmetric Complex Gaussian (CSCG) distribution with mean $\mathbf{0}$ and covariance $\mathbf{\Sigma}$. $\sim$ means "distributed as".

II. BD-RIS Model

Consider a BD-RIS aided point-to-point MIMO system with N_{T}, N_{S}, N_{R} transmit, scatter, and receive antennas, respectively. This configuration is denoted as $N_{\mathrm{T}}\times N_{\mathrm{S}}\times N_{\mathrm{R}}$ in the following context. The BD-RIS can be modeled as an N_{S}-port network [44] that divides into G individual groups, each containing $L\triangleq N_{\mathrm{S}}/G$ elements interconnected by real-time reconfigurable components [19]. To simplify the analysis, we assume a lossless asymmetric network [3] without mutual coupling between scatter elements, as previously considered in [20], [27], [28]. The overall scattering matrix of the BD-RIS is

$$\mathbf{\Theta}=\mathrm{diag}(\mathbf{\Theta}_1,...,\mathbf{\Theta}_G), \tag{1}$$

where $\mathbf{\Theta}_g\in\mathbb{U}^{L\times L}$ is a unitary matrix describing the scatter response of group g. Note that diagonal (single-connected) and unitary (fully-connected) RIS can be regarded as its extreme cases with group size 1 and N_{S}, respectively. The potential physical architectures of BD-RIS are shown in [19, Fig. 3], [28, Fig. 5], and [24, Fig. 2], where the antenna radiation patterns and circuit network structures are taken into account.

Let $\mathbf{H}_{\mathrm{D}}\in\mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}}$, $\mathbf{H}_{\mathrm{F}}\in\mathbb{C}^{N_{\mathrm{S}}\times N_{\mathrm{T}}}$, $\mathbf{H}_{\mathrm{B}}\in\mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{S}}}$ denote the direct (transmitter-receiver), forward (transmitter-RIS), and backward (RIS-receiver) channels, respectively. The equivalent channel is

$$\mathbf{H}=\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}=\mathbf{H}_{\mathrm{D}}+\sum_g\mathbf{H}_{\mathrm{B},g}\mathbf{\Theta}_g\mathbf{H}_{\mathrm{F},g}, \tag{2}$$

where $\mathbf{H}_{\mathrm{B},g}\in\mathbb{C}^{N_{\mathrm{R}}\times L}$ and $\mathbf{H}_{\mathrm{F},g}\in\mathbb{C}^{L\times N_{\mathrm{T}}}$ are the backward and forward channels of RIS group g, respectively.

Remark 1. *BD-RIS reduces to diagonal RIS and unitary RIS with group size L 1 and N_{S}, respectively.*

Remark 2. *Individual forward and backward Channel State Information (CSI) are required for BD-RIS designs. This is different from diagonal RIS where estimating their product is usually sufficient.*

III. Channel Singular Values Redistribution

A. A Toy Example

We first illustrate the channel shaping capabilities of different RIS by a toy example. Consider a $2\times 2\times 2$ setup where the direct link is blocked. The diagonal RIS is modeled by $\mathbf{\Theta}_{\mathrm{D}}=\mathrm{diag}(e^{\jmath\theta_1},e^{\jmath\theta_2})$ while the unitary BD-RIS has 4 independent angular parameters

$$\mathbf{\Theta}_{\mathrm{U}}=e^{\jmath\phi}\begin{bmatrix}e^{\jmath\alpha}\cos\psi & e^{\jmath\beta}\sin\psi\\ -e^{-\jmath\beta}\sin\psi & e^{-\jmath\alpha}\cos\psi\end{bmatrix}. \tag{3}$$

In particular, ϕ has no impact on the singular value because $\mathrm{sv}(e^{\jmath\phi}\mathbf{A})=\mathrm{sv}(\mathbf{A})$. We also enforce symmetry by $\beta=\pi/2$ such that both architectures have the same number of angular parameters. Fig. 1 shows the channel singular values achieved by an exhaustive grid search over (θ_1,θ_2) for diagonal RIS and (α,ψ) for symmetric unitary RIS. It is observed that both singular values can be manipulated up to 9% using diagonal RIS and 42% using symmetric BD-RIS, despite both architectures have the same number of scattering elements and design parameters. A larger performance gap is expected when asymmetric BD-RIS is available. This example shows BD-RIS can provide a wider dynamic range of channel singular values and motivates further studies on channel shaping.

[2]A geodesic refers to the shortest path between two points in a Riemannian manifold.

[3]While symmetric BD-RIS (involving capacitors and inductors) is often considered in the literature [19], [21]–[27], asymmetric BD-RIS can be built over reconfigurable asymmetric passive components (e.g., ring hybrids and branch-line hybrids) [45].

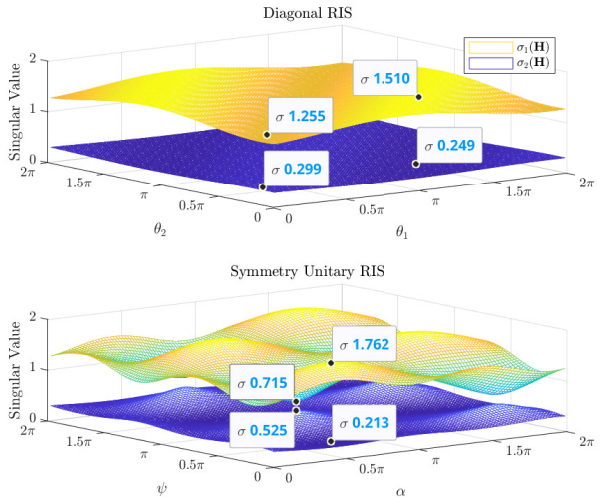


Fig. 1. $2\times2\times2$ (no direct) channel singular value shaping by diagonal and symmetry unitary RIS.

B. Pareto Frontier Characterization

We then characterize the Pareto frontier of channel singular values by maximizing their weighted sum

$$\max_{\boldsymbol{\Theta}} \quad \sum_n \rho_n \sigma_n(\mathbf{H}) \tag{4a}$$

$$\text{s.t.} \quad \boldsymbol{\Theta}_g^{\mathsf{H}}\boldsymbol{\Theta}_g = \mathbf{I}, \quad \forall g, \tag{4b}$$

where $n\in\{1,\ldots,N\triangleq\min(N_{\mathrm{T}},N_{\mathrm{R}})\}$ and ρ_n is the weight of the n-th singular value that can be positive, zero, or negative. Varying $\{\rho_n\}$ unveils the entire achievable singular value region. Thus, the Pareto frontier problem (4) generalizes most relevant metrics and provides a powerful shaping framework. The objective (4a) is smooth in $\boldsymbol{\Theta}$ and the feasible domain (4b) for group g corresponds to the Stiefel manifold. Next, we zoom out to general smooth maximization problems of asymmetric BD-RIS.

Inspired by [39], [40], we propose a block-wise RCG algorithm along the geodesics on the Lie group of unitary matrices $\mathbb{U}^{L\times L}$. It leverages the fact that unitary matrices are closed under multiplication. At iteration r, the gradient is computed in the Euclidean space and translated to the Riemannian manifold [42]

$$\nabla_{\mathrm{E},g}^{(r)} = \frac{\partial f(\boldsymbol{\Theta}_g^{(r)})}{\partial \boldsymbol{\Theta}_g^{*}}, \tag{5}$$

$$\nabla_{\mathrm{R},g}^{(r)} = \nabla_{\mathrm{E},g}^{(r)}\boldsymbol{\Theta}_g^{(r)\mathsf{H}} - \boldsymbol{\Theta}_g^{(r)}\nabla_{\mathrm{E},g}^{(r)\mathsf{H}}. \tag{6}$$

The Polak-Ribierre parameter [46] is approximated as [40]

$$\gamma_g^{(r)} = \frac{\mathrm{tr}\big((\nabla_{\mathrm{R},g}^{(r)} - \nabla_{\mathrm{R},g}^{(r-1)})\nabla_{\mathrm{R},g}^{(r)\mathsf{H}}\big)}{\mathrm{tr}\big(\nabla_{\mathrm{R},g}^{(r-1)}\nabla_{\mathrm{R},g}^{(r-1)\mathsf{H}}\big)}, \tag{7}$$

and the conjugate direction is

$$\mathbf{D}_g^{(r)} = \nabla_{\mathrm{R},g}^{(r)} + \gamma_g^{(r)}\mathbf{D}_g^{(r-1)}. \tag{8}$$

Algorithm 1: Block-wise geodesic RCG for asymmetric BD-RIS

Input: $f(\boldsymbol{\Theta})$, G
Output: $\boldsymbol{\Theta}^{\star}$
1: **Initialize** $r\leftarrow0$, $\boldsymbol{\Theta}^{(0)}$
2: **Repeat**
3: **For** $g\leftarrow1$ to G
4: $\nabla_{\mathrm{E},g}^{(r)}\leftarrow$ (5)
5: $\nabla_{\mathrm{R},g}^{(r)}\leftarrow$ (6)
6: $\gamma_g^{(r)}\leftarrow$ (7)
7: $\mathbf{D}_g^{(r)}\leftarrow$ (8)
8: **If** $\Re\{\mathrm{tr}(\mathbf{D}_g^{(r)\mathsf{H}}\nabla_{\mathrm{R},g}^{(r)})\}<0$ ▷ not an ascent direction
9: $\mathbf{D}_g^{(r)}\leftarrow\nabla_{\mathrm{R},g}^{(r)}$
10: **End If**
11: $\mu\leftarrow1$
12: $\mathbf{G}_g^{(r)}(\mu)\leftarrow$ (9)
13: **While** $f\big(\mathbf{G}_g^{(r)}(2\mu)\big)-f(\boldsymbol{\Theta}_g^{(r)})\geq\mu\cdot\mathrm{tr}(\mathbf{D}_g^{(r)}\mathbf{D}_g^{(r)\mathsf{H}})/2$
14: $\mu\leftarrow2\mu$
15: **End While**
16: **While** $f\big(\mathbf{G}_g^{(r)}(\mu)\big)-f(\boldsymbol{\Theta}_g^{(r)})<\mu/2\cdot\mathrm{tr}(\mathbf{D}_g^{(r)}\mathbf{D}_g^{(r)\mathsf{H}})/2$
17: $\mu\leftarrow\mu/2$
18: **End While**
19: $\boldsymbol{\Theta}_g^{(r+1)}\leftarrow$ (10)
20: **End For**
21: $r\leftarrow r+1$
22: **Until** $|f(\boldsymbol{\Theta}^{(r)})-f(\boldsymbol{\Theta}^{(r-1)})|/f(\boldsymbol{\Theta}^{(r-1)})\leq\epsilon$

In the Stiefel manifold, the geodesic emanating from $\boldsymbol{\Theta}_g^{(r)}$ with velocity $\mathbf{D}_g^{(r)}$ and step size μ is described compactly by the exponential map [41]

$$\mathbf{G}_g^{(r)}(\mu) = \exp(\mu\mathbf{D}_g^{(r)})\boldsymbol{\Theta}_g^{(r)}. \tag{9}$$

An appropriate $\mu^{\star}$ can be obtained by the Armijo rule [47].[4] Finally, the scattering matrix is updated along the geodesic as

$$\boldsymbol{\Theta}_g^{(r+1)} = \mathbf{G}_g^{(r)}(\mu^{\star}). \tag{10}$$

Algorithm 1 summarizes the proposed block-wise geodesic RCG method for smooth maximization problems of asymmetric BD-RIS. Convergence to stationary points is guaranteed.

Remark 3. *Compared with universal manifold optimization [42], [43], Algorithm 1 inherits a trifold benefit from [39], [40]:*

1) *No retraction thanks to rotational update* (9), (10)*;*
2) *Lower computational complexity per iteration;*
3) *Faster convergence thanks to proper parameter space.*

Lemma 1. *The Euclidean gradient of* (4a) *w.r.t. BD-RIS group g is*

$$\frac{\partial\sum_n\rho_n\sigma_n(\mathbf{H})}{\partial\boldsymbol{\Theta}_g^{*}} = \mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\mathbf{U}\mathrm{diag}(\rho_1,\ldots,\rho_N)\mathbf{V}^{\mathsf{H}}\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}, \tag{11}$$

where $\mathbf{U}$ *and* $\mathbf{V}$ *are the left and right singular matrices of* $\mathbf{H}$*, respectively.*

Proof. Please refer to Appendix A. □

Algorithm 1 can thus be invoked for the Pareto singular value problem (4) where line 4 uses (11) explicitly.

[4]To double the step size, one only need to square the rotation matrix instead of recomputing the matrix exponential, i.e., $\exp(2\mu\mathbf{D}_g^{(r)}) = \exp^2(\mu\mathbf{D}_g^{(r)})$.

C. Some Analytical Bounds

We then discuss some analytical bounds related to channel singular values.

Proposition 1 (degree of freedom)**.** *In point-to-point MIMO, BD-RIS cannot achieve a higher Degree of Freedom (DoF) than diagonal RIS.*

Proof. Please refer to Appendix B. □

Proposition 2 (rank-deficient channel)**.** *If the forward or backward channel is rank-k, then regardless of the RIS size and architecture, the n-th singular value of the equivalent channel is bounded by*

$$\sigma_n(\mathbf{H}) \le \sigma_{n-k}(\mathbf{T}), \qquad \text{if } n>k, \tag{12}$$

$$\sigma_n(\mathbf{H}) \ge \sigma_n(\mathbf{T}), \qquad \text{if } n<N-k+1, \tag{13}$$

where

$$\mathbf{T}\mathbf{T}^{\mathsf{H}} = \begin{cases} \mathbf{H}_{\mathrm{D}}(\mathbf{I}-\mathbf{V}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}})\mathbf{H}_{\mathrm{D}}^{\mathsf{H}}, & \text{if } \mathrm{rank}(\mathbf{H}_{\mathrm{F}})=k, \\ \mathbf{H}_{\mathrm{D}}^{\mathsf{H}}(\mathbf{I}-\mathbf{U}_{\mathrm{B}}\mathbf{U}_{\mathrm{B}}^{\mathsf{H}})\mathbf{H}_{\mathrm{D}}, & \text{if } \mathrm{rank}(\mathbf{H}_{\mathrm{B}})=k, \end{cases} \tag{14}$$

and $\mathbf{V}_{\mathrm{F}}$ and $\mathbf{U}_{\mathrm{B}}$ are the right and left compact singular matrices of $\mathbf{H}_{\mathrm{F}}$ and $\mathbf{H}_{\mathrm{B}}$, respectively.

Proof. Please refer to Appendix C. □

Corollary 2.1 (extreme singular values)**.** *With a sufficiently large RIS, the first k channel singular values are unbounded above while the last k channel singular values can be suppressed to zero.*

Corollary 2.2 (Line-of-Sight (LoS) channel[5])**.** *If the forward or backward channel is LoS, then a RIS can at most enlarge (resp. suppress) the n-th ($n\ge 2$) channel singular value to the $(n-1)$-th (resp. n-th) singular value of $\mathbf{T}$, that is,*

$$\sigma_1(\mathbf{H}) \ge \sigma_1(\mathbf{T}) \ge \sigma_2(\mathbf{H}) \ge \ldots \ge \sigma_{N-1}(\mathbf{T}) \ge \sigma_N(\mathbf{H}) \ge \sigma_N(\mathbf{T}). \tag{15}$$

In Section V, we will show by simulation that a finite-size BD-RIS can approach those bounds better than diagonal RIS.

Proposition 3 (fully-connected RIS without direct link)**.** *If the BD-RIS is fully-connected and the direct link is absent, then the channel singular values can be manipulated up to*

$$\mathrm{sv}(\mathbf{H}) = \mathrm{sv}(\mathbf{B}\mathbf{F}), \tag{16}$$

where $\mathbf{B}$ and $\mathbf{F}$ are arbitrary matrices with the same singular values as $\mathbf{H}_{\mathrm{B}}$ and $\mathbf{H}_{\mathrm{F}}$, respectively,

Proof. Please refer to Appendix D. □

The problem now becomes how the singular values of matrix product are bounded by the singular values of its individual factors. Let $N'=\max(N_{\mathrm{T}},N_{\mathrm{S}},N_{\mathrm{R}})$ and $\sigma_n(\mathbf{H})=\sigma_n(\mathbf{H}_{\mathrm{F}})=\sigma_n(\mathbf{H}_{\mathrm{B}})=0$ for $N<n\le N'$. We have the following corollaries.

Corollary 3.1 (generic singular value bounds [49])**.**

$$\prod_{k\in K}\sigma_k(\mathbf{H}) \le \prod_{i\in I}\sigma_i(\mathbf{H}_{\mathrm{B}})\prod_{j\in J}\sigma_j(\mathbf{H}_{\mathrm{F}}), \tag{17}$$

[5]A similar result has been derived for diagonal RIS in [48].

for all admissible triples $(I,J,K)\in T_r^{N'}$ with $r<N'$, where

$$T_r^{N'} \triangleq \Big\{(I,J,K)\in U_r^{N'} \,|\, \forall p<r, (F,G,H)\in T_p^r, \sum_{f\in F} i_f + \sum_{g\in G} j_g \le \sum_{h\in H} k_h + p(p+1)/2\Big\},$$

$$U_r^{N'} \triangleq \Big\{(I,J,K)\,|\sum_{i\in I} i + \sum_{j\in J} j = \sum_{k\in K} k + r(r+1)/2\Big\}.$$

Corollary 3.2 (upper bound on the largest singular value)**.**

$$\sigma_1(\mathbf{H}) \le \sigma_1(\mathbf{H}_{\mathrm{B}})\sigma_1(\mathbf{H}_{\mathrm{F}}). \tag{18}$$

Corollary 3.3 (upper bound on the product of first k singular values)**.**

$$\prod_{n=1}^{k}\sigma_n(\mathbf{H}) \le \prod_{n=1}^{k}\sigma_n(\mathbf{H}_{\mathrm{B}})\prod_{n=1}^{k}\sigma_n(\mathbf{H}_{\mathrm{F}}). \tag{19}$$

Corollary 3.4 (upper bound on the sum of first k singular values to the power of p)**.**

$$\sum_{n=1}^{k}\sigma_n^p(\mathbf{H}) \le \sum_{n=1}^{k}\sigma_n^p(\mathbf{H}_{\mathrm{B}})\sigma_n^p(\mathbf{H}_{\mathrm{F}}), \quad p>0. \tag{20}$$

When $k=N'$ and $p=2$, (20) suggests the channel power is upper bounded by the sum of (sorted) element-wise power product of backward and forward subchannels.

Remark 4. *From (37) and (38) in the proof of Proposition 3, we notice that (17)–(20) are simultaneously tight when*

$$\mathbf{\Theta} = \mathbf{V}_{\mathrm{B}}\mathbf{U}_{\mathrm{F}}^{\mathsf{H}}. \tag{21}$$

An interpretation is that the off-diagonal entries can enhance the capabilities of

- *subspace alignment: $\mathbf{V}_{\mathrm{B}}$ and $\mathbf{U}_{\mathrm{F}}^{\mathsf{H}}$ in (37) fully align the subspaces of $\mathbf{H}_{\mathrm{B}}$ and $\mathbf{H}_{\mathrm{F}}$ by rotation;*
- *subchannel rearrangement: $\mathbf{X}=\mathbf{I}$ in (38) pairs the subchannels of $\mathbf{H}_{\mathrm{B}}$ and $\mathbf{H}_{\mathrm{F}}$ from strongest to weakest, which attains the maximal in rearrangement inequality.*

Tight bounds are inapplicable when a MIMO direct link is present, as the RIS needs to balance the direct-indirect (additive) and backward-forward (multiplicative) subspace alignments. Such a balance often involves optimization approaches and another example will be discussed in Section IV-B.

IV. Achievable Rate Maximization

The MIMO achievable rate maximization problem is formulated w.r.t. joint active and passive beamforming

$$\max_{\mathbf{W},\mathbf{\Theta}} \quad R = \log\det\left(\mathbf{I} + \frac{\mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{H}\mathbf{W}}{\eta}\right) \tag{22a}$$

$$\text{s.t.} \quad \|\mathbf{W}\|_{\mathrm{F}}^2 \le P, \tag{22b}$$

$$\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g = \mathbf{I}, \quad \forall g, \tag{22c}$$

where $\mathbf{W}$ is the transmit precoder, R is the achievable rate, η is the noise power, and P is the transmit power budget. Two methods are proposed below to solve problem (22).

A. Alternating Optimization

Consider an AO approach that updates $\mathbf{\Theta}$ and $\mathbf{W}$ iteratively. For a given $\mathbf{W}$, the passive beamforming subproblem is

$$\max_{\mathbf{\Theta}} \quad \mathrm{logdet}\left(\mathbf{I}+\frac{\mathbf{HQH}^{\mathsf{H}}}{\eta}\right) \tag{23a}$$

$$\text{s.t.} \quad \mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I}, \quad \forall g, \tag{23b}$$

where $\mathbf{Q}\triangleq\mathbf{WW}^{\mathsf{H}}$ is the transmit covariance matrix.

Lemma 2. *The Euclidean gradient of* (23a) *w.r.t. BD-RIS block g is*

$$\frac{\partial R}{\partial \mathbf{\Theta}_g^*}=\frac{1}{\eta}\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\left(\mathbf{I}+\frac{\mathbf{HQH}^{\mathsf{H}}}{\eta}\right)^{-1}\mathbf{HQH}_{\mathrm{F},g}^{\mathsf{H}}. \tag{24}$$

Proof. Please refer to Appendix E. □

Algorithm 1 is then invoked to solve problem (22) where line 4 uses (24) explicitly. Since (23a) is a concave function of $\mathbf{\Theta}$, convergence to local-optimal points is guaranteed. On the other hand, the global optimal transmit precoder for a fixed $\mathbf{\Theta}$ is given by the eigenmode transmission [50]

$$\mathbf{W}^{\star}=\mathbf{VS}^{\star 1/2}, \tag{25}$$

where $\mathbf{V}$ is the right channel singular matrix and $\mathbf{S}^{\star}$ is the optimal water-filling power allocation matrix. The overall AO algorithm converges to local-optimal points of problem (22) since each subproblem is solved optimally and the objective is bounded above.

B. Low-Complexity Solution

We then propose a low-complexity solution to problem (22) based on channel shaping. The passive beamforming subproblem (23) involves transmit covariance matrix $\mathbf{Q}$ and thus requires iterative RCG update. Instead, we decouple the joint RIS-transceiver design by recasting (23) as channel power maximization

$$\max_{\mathbf{\Theta}} \quad \|\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}\|_{\mathrm{F}}^2 \tag{26a}$$

$$\text{s.t.} \quad \mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I}, \quad \forall g. \tag{26b}$$

Remark 5. *As mentioned in Section III-C, the key of solving* (26) *is to balance the additive and multiplicative subspace alignments. Problem* (26) *is very similar (in terms of maximizing the inner product of $\mathbf{H}_{\mathrm{D}}$ and $\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}$) to the weighted orthogonal Procrustes problem [51]*

$$\min_{\mathbf{\Theta}} \quad \|\mathbf{H}_{\mathrm{D}}-\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}\|_{\mathrm{F}}^2 \tag{27a}$$

$$\text{s.t.} \quad \mathbf{\Theta}^{\mathsf{H}}\mathbf{\Theta}=\mathbf{I}, \tag{27b}$$

which has no trivial solution. One lossy transformation, by moving $\mathbf{\Theta}$ to one side [52], formulates standard orthogonal Procrustes problems

$$\min_{\mathbf{\Theta}} \quad \|\mathbf{H}_{\mathrm{B}}^{\dagger}\mathbf{H}_{\mathrm{D}}-\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}\|_{\mathrm{F}}^2 \text{ or } \|\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\dagger}-\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\|_{\mathrm{F}}^2 \tag{28a}$$

$$\text{s.t.} \quad \mathbf{\Theta}^{\mathsf{H}}\mathbf{\Theta}=\mathbf{I}, \tag{28b}$$

which has global optimal solutions

$$\mathbf{\Theta}^{\star}=\mathbf{UV}^{\mathsf{H}} \tag{29}$$

where $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular matrices of $\mathbf{H}_{\mathrm{B}}^{\dagger}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\mathsf{H}}$ or $\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\dagger}$ [53]. This suboptimal solution only applies to fully-connected BD-RIS.

Inspired by [54], we propose a general solution to problem (26) with arbitrary group size. The idea is to successively approximate the quadratic objective (26a) by local Taylor expansions and solve each step in closed form.

Proposition 4. *Starting from any $\mathbf{\Theta}^{(0)}\in\mathbb{U}^{N_{\mathrm{S}}\times N_{\mathrm{S}}}$, the sequence*

$$\mathbf{\Theta}_g^{(r+1)}=\mathbf{U}_g^{(r)}\mathbf{V}_g^{(r)}, \quad \forall g. \tag{30}$$

converges to a stationary point of (26)*, where $\mathbf{U}_g^{(r)}$ and $\mathbf{V}_g^{(r)}$ are the left and right compact singular matrix of*

$$\mathbf{M}_g^{(r)}=\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\left(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathrm{diag}\big(\mathbf{\Theta}_{[1:g-1]}^{(r+1)},\mathbf{\Theta}_{[g:G]}^{(r)}\big)\mathbf{H}_{\mathrm{F}}\right)\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}} \tag{31}$$

Proof. Please refer to Appendix F. □

Once the channel shaping problem (26) is solved, the transmit precoder can be obtained by (25). This two-stage approach decouples both blocks and is computationally efficient.

V. Simulation Results

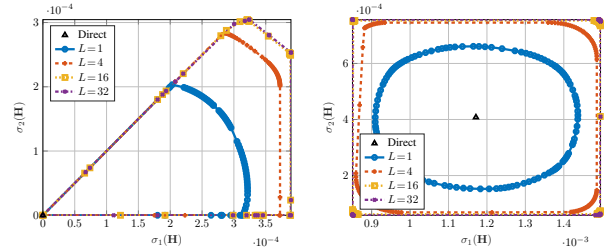
In this section, we provide numerical results to evaluate the proposed BD-RIS designs. Consider a distance-dependent path loss model $\Lambda(d)=\Lambda_0 d^{-\gamma}$ where Λ_0 is the reference path loss at distance $1\,\mathrm{m}$, d is the propagation distance, and γ is the path loss exponent. The small-scale fading model is $\mathbf{H}=\sqrt{\kappa/(1+\kappa)}\mathbf{H}_{\mathrm{LoS}}+\sqrt{1/(1+\kappa)}\mathbf{H}_{\mathrm{NLoS}}$, where κ is the Rician K-factor, $\mathbf{H}_{\mathrm{LoS}}$ is the deterministic LoS component, and $\mathbf{H}_{\mathrm{NLoS}}\sim\mathcal{CN}(\mathbf{0},\mathbf{I})$ is the Rayleigh component. We set $\Lambda_0=-30\,\mathrm{dB}$, $d_{\mathrm{D}}=14.7\,\mathrm{m}$, $d_{\mathrm{F}}=10\,\mathrm{m}$, $d_{\mathrm{B}}=6.3\,\mathrm{m}$, $\gamma_{\mathrm{D}}=3$, $\gamma_{\mathrm{F}}=2.4$ and $\gamma_{\mathrm{B}}=2$ for reference, which corresponds to a typical indoor environment with $\Lambda_{\mathrm{D}}=-65\mathrm{dB}$, $\Lambda_{\mathrm{F}}=-54\mathrm{dB}$, $\Lambda_{\mathrm{B}}=-46\mathrm{dB}$. The indirect path via RIS is thus $35\,\mathrm{dB}$ weaker than the direct path. $\kappa\to\infty$ is assumed for all channels unless otherwise specified.

A. Channel Singular Values Redistribution

1) Pareto Frontier: Fig. 2 shows the Pareto singular values of a 2T2R MIMO reshaped by a RIS. When the direct link is absent, the achievable regions in Fig. 2(a) are shaped like pizza slices. This is because $\sigma_1(\mathbf{H})\geq\sigma_2(\mathbf{H})\geq 0$ and there exists a tradeoff between aligning the two subspaces. We observe that the smallest singular value is enhanced up to 2×10^{-4} by diagonal RIS and 3×10^{-4} by fully-connected BD-RIS, corresponding to a $50\,\%$ gain. When the direct link is present, the shape of the singular value region depends heavily on the relative strength of the indirect link. In Fig. 2(b), a 32-element RIS is insufficient to compensate the $35\,\mathrm{dB}$ path loss imbalance and results in a limited singular value region that is symmetric around the direct point. As the group size L increases, the shape of the region evolves from elliptical to square. This transformation not only provides a better tradeoff in subchannel manipulation but also improves the dynamic range of $\sigma_1(\mathbf{H})$ and $\sigma_2(\mathbf{H})$ by $22\,\%$ and $38\,\%$, respectively. The achievable singular value region also enlarges as the number of scattering

TABLE I
AVERAGE PERFORMANCE OF BD-RIS DESIGNS

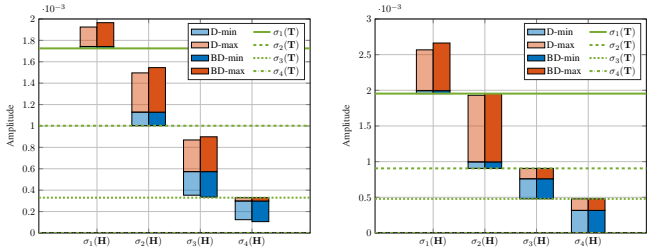
RCG path	$N_S = 16$			$N_S = 256$		
	Objective	Iterations	Time [s]	Objective	Iterations	Time [s]
Geodesic	4.36×10^{-3}	11.61	2.04×10^{-2}	1.16×10^{-2}	25.78	3.21
Non-geodesic	4.17×10^{-3}	169.50	1.42×10^{-1}	8.87×10^{-3}	278.08	27.81



(a) $2\times32\times2$ (no direct) (b) $2\times32\times2$

(c) $2\times64\times2$ (d) $2\times128\times2$

Fig. 2. Pareto frontiers of singular values of a 2T2R channel reshaped by a RIS.



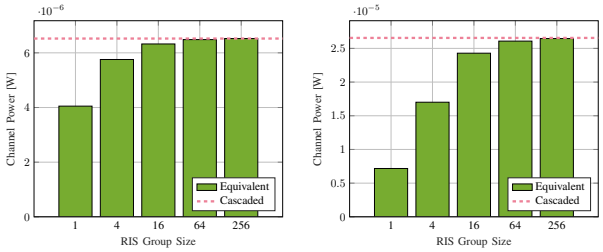
(a) $4\times32\times4$ (rank-1) (b) $4\times64\times4$ (rank-1)

(c) $4\times128\times4$ (rank-2) (d) $4\times256\times4$ (rank-4)

Fig. 3. Achievable channel singular values: analytical bounds (green lines) and numerical optimization results (blue and red bars). 'D' means diagonal RIS and 'BD' means fully-connected BD-RIS. 'rank-k' refers to the forward channel.

elements N_S increases. In particular, Fig. 2(d) shows that the equivalent channel can be completely nulled by a 128-element BD-RIS but not by a diagonal one. Those results demonstrate the superior channel shaping capability of BD-RIS for better signal enhancement and interference suppression.

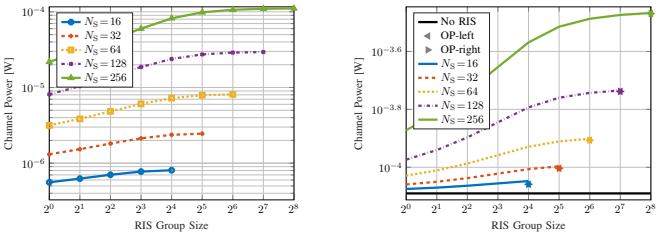
2) Analytical Bounds and Numerical Results: Fig. 3 illustrates the analytical singular value bounds in Proposition 2 and the numerical results obtained by solving problem (4) with $\rho_n = \pm1$ and $\rho_{n'} = 0, \forall n' \neq n$. Here we assme a rank-$k$ forward channel without loss of generality. When the RIS is in the vicinity of the transmitter, Figs. 3(a) and 3(b) show that the achievable channel singular values indeed satisfy Corollary 2.2, namely $\sigma_1(\mathbf{H}) \geq \sigma_1(\mathbf{T})$, $\sigma_2(\mathbf{T}) \leq \sigma_2(\mathbf{H}) \leq \sigma_1(\mathbf{T})$, etc. It is obvious that BD-RIS can approach those bounds better than diagonal RIS especially for a small N_S. Another example is given in Fig. 3(c) with rank-2 forward channel. The first two channel singular values are unbounded above and bounded below by the first two singular values of $\mathbf{T}$, while the last two singular values can be suppressed to zero and bounded above by the first two singular values of $\mathbf{T}$. Those observations align with Proposition 2 and Corollary 2.1. Finally, Fig. 3(d) confirms there are no extra singular value bounds when both forward and backward channels are full-rank. This can be predicted from (14) where the compact singular matrix $\mathbf{V}_F$ becomes unitary and $\mathbf{T} = \mathbf{0}$. The numerical results are consistent with the analytical bounds, and we conclude that the channel shaping advantage of BD-RIS over diagonal RIS scales with forward and backward channel ranks.



(a) $1\times256\times1$ (no direct) (b) $4\times256\times4$ (no direct)

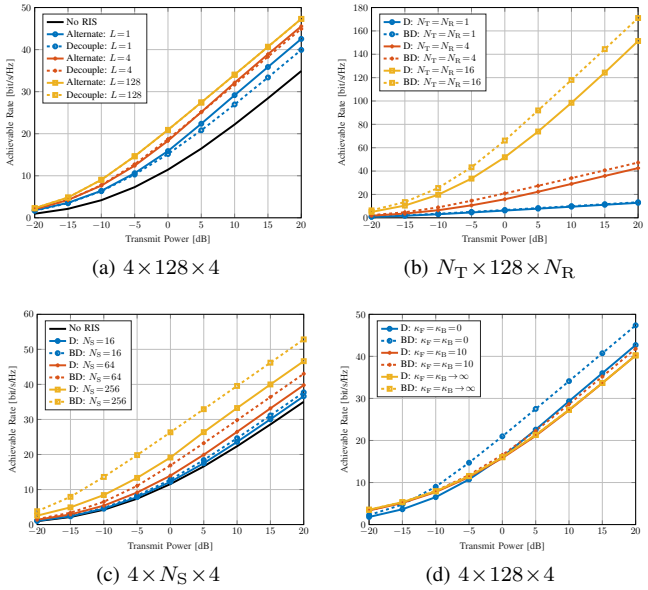
Fig. 4. Average maximum channel power versus BD-RIS group size and MIMO dimensions. 'Cascaded' refers to the available power of the cascaded channel, i.e., the sum of (sorted) element-wise power product of backward and forward subchannels.

Fig. 4 compares the analytical channel power bound in Corollary 3.4 with $k = N'$, $p = 2$ and the numerical results obtained by solving problem (26) when the direct link is absent. Here, a fully-connected BD-RIS can attain the upper bound either in closed form (21) or via optimization approach (30). For the SISO case in Fig. 4(a), the maximum channel power is approximately 4×10^{-6} by diagonal RIS and 6.5×10^{-6} by fully-connected BD-RIS, corresponding to a $62.5\,\%$ gain. This aligns with the asymptotic BD-RIS scaling law derived for SISO in [19]. Interestingly, the gain surges to $270\,\%$ in 4T4R MIMO as shown in Fig. 4(b). This is because subspace



(a) $16\times N_{\mathrm{S}}\times 16$ (no direct) (b) $16\times N_{\mathrm{S}}\times 16$

Fig. 5. Average maximum channel power versus RIS configuration. 'OP-left' and 'OP-right' refer to the suboptimal solutions to problem (26) by lossy transformation (28) where $\boldsymbol{\Theta}$ is to the left and right of the product, respectively.



(a) $4\times 128\times 4$ (b) $N_{\mathrm{T}}\times 128\times N_{\mathrm{R}}$

(c) $4\times N_{\mathrm{S}}\times 4$ (d) $4\times 128\times 4$

Fig. 6. Average achievable rate versus MIMO and RIS configurations. The noise power is $\eta=-75\,\mathrm{dB}$, corresponding to a direct SNR of -10 to $30\,\mathrm{dB}$. 'Alternate' refers to the alternating optimization and 'Decouple' refers to the low-complexity design. 'D' means diagonal RIS and 'BD' means fully-connected BD-RIS.

alignment boils down to phase matching in SISO such that both triangular and Cauchy-Schwarz inequalities in [19, (50)] can be simultaneously tight regardless of the group size. That is, diagonal RIS is sufficient for subspace alignment in SISO while the $62.5\,\%$ gain from BD-RIS comes purely from subchannel rearrangement (i.e., pairing the forward and backward channels from strongest to weakest). Now consider a diagonal RIS in MIMO. Each element can only apply a common phase shift to the associated rank-1 $N_{\mathrm{R}}\times N_{\mathrm{T}}$ indirect channel. Therefore, perfect subspace alignment of indirect channels through different elements is generally impossible. It means the disadvantage of diagonal RIS in subspace alignment and subchannel rearrangement scales with MIMO dimensions. We thus conclude that the power gain of BD-RIS scales with group size and MIMO dimensions.

B. Achievable Rate Maximization

We first focus on channel shaping subproblem (40). Fig. 5 shows the achievable channel power under different RIS configurations. An interesting observation is that the relative power gain of BD-RIS over diagonal RIS is even larger with direct link. For example, a 64-element fully BD-RIS can almost provide the same channel power as a 256-element diagonal RIS in Fig. 5b, but not in Fig. 5a. This is because the RIS needs to balance the multiplicative forward-backward combining and the additive direct-indirect combining, such that the subspace alignment advantage of BD-RIS becomes more pronounced. We also notice that the suboptimal solutions (29) for fully-connected BD-RIS by lossy transformation (28) are very close to optimal especially for a large N_{S}.

Fig. 6 presents the achievable rate under different MIMO and RIS configurations. At a transmit power of $10\,\mathrm{dB}$, Fig. 6(a) shows that introducing a 128-element diagonal RIS to 4T4R MIMO can improve the achievable rate from $22.2\,\mathrm{bps/Hz}$ to $29.2\,\mathrm{bps/Hz}$ $(+31.5\,\%)$. In contrast, a BD-RIS of group size 4 and 128 can further improve the rate to $32.1\,\mathrm{bps/Hz}$ $(+44.6\%)$ and $34\,\mathrm{bps/Hz}$ $(+53.2\%)$, respectively. Interestingly, the gap between the optimal AO approach (23)–(25) and the low-complexity solution (30) and (25) narrows as the group size increases, and completely vanishes for a fully-connected BD-RIS. This implies that the RIS-transceiver design can be completely decoupled via channel shaping with marginal performance loss. Figs. 6(b) and 6(c) also confirm the advantage of BD-RIS grows with the number of transmit, scatter, and receive antennas. In the low power regime (-20 to $-10\,\mathrm{dB}$), the slope of the achievable rate is significantly larger with BD-RIS, suggesting that multiple streams can be activated at a much lower SNR. This is because BD-RIS not only spreads the channel singular values to a wider range, but also provides a better tradeoff between subchannels (c.f. Fig. 2). Finally, Fig. 6(d) shows that the gap between diagonal and BD-RIS narrows as the Rician K-factor increases and becomes indistinguishable in LoS environment. The observation is expected from previous studies [19]–[21] and aligns with Corollary 2.2, which suggests that the BD-RIS should be deployed in rich-scattering environments to exploit its channel shaping potential.

VI. Conclusion

This paper analyzes the channel shaping capability of RIS in terms of singular values redistribution. We consider a general BD architecture that allows elements within the same group to interact, enabling more sophisticated manipulation than diagonal RIS. This translates to a wider dynamic range (with better tradeoff) of singular values and significant power and rate gains, especially in large-scale MIMO systems. We characterize the Pareto frontiers of channel singular values via optimization approach and provide analytical bounds in rank-deficient and fully-connected scenarios. An efficient RCG algorithm is proposed for smooth BD-RIS optimization problems, which offers lower computation complexity and faster convergence than existing methods. We also present two beamforming designs for rate maximization problem, one based on alternating optimization for optimal performance and the other decouples the RIS-transceiver design for lower complexity.

Extensive simulations show that the advantage of BD-RIS stems from its superior subspace alignment and subchannel rearrangement capability, which scales with the number of elements, group size, MIMO dimensions, and channel diversity.

One future direction is introducing BD-RIS to MIMO interference channel for interference alignment or cancellation. Another open issue is to exploit different groups of BD-RIS to enhance the channel response (and possibly ride extra information) at different frequencies. Incorporating a RIS at both transmitter and receiver sides provides even stronger manipulation that potentially align both direct-indirect and forward-backward subspaces simultaneously.

APPENDIX

A. Proof of Lemma 1

Let $\mathbf{H}=\sum_n \mathbf{u}_n\sigma_n\mathbf{v}_n^{\mathsf{H}}$ be the compact Singular Value Decomposition (SVD) of the equivalent channel. Since the singular vectors are orthonormal, the n-th singular value can be expressed as

$$\sigma_n=\mathbf{u}_n^{\mathsf{H}}\mathbf{H}\mathbf{v}_n=\mathbf{u}_n^{\mathsf{T}}\mathbf{H}^*\mathbf{v}_n^*, \tag{32}$$

whose differential w.r.t. $\boldsymbol{\Theta}_g^*$ is

$$\begin{aligned}\partial\sigma_n=&\partial\mathbf{u}_n^{\mathsf{T}}\underbrace{\mathbf{H}^*\mathbf{v}_n^*}_{\sum_m\mathbf{u}_m^*\sigma_m\mathbf{v}_m^{\mathsf{T}}\mathbf{v}_n}+\mathbf{u}_n^{\mathsf{T}}\cdot\partial\mathbf{H}^*\cdot\mathbf{v}_n^*+\underbrace{\mathbf{u}_n^{\mathsf{T}}\mathbf{H}^*}_{\mathbf{u}_n^{\mathsf{T}}\sum_m\mathbf{u}_m^*\sigma_m\mathbf{v}_m^{\mathsf{T}}}\partial\mathbf{v}_n^*\\ =&\underbrace{\partial\mathbf{u}_n^{\mathsf{T}}\mathbf{u}_n^*}_{\partial 1=0}\cdot\sigma_n+\mathbf{u}_n^{\mathsf{T}}\cdot\partial\mathbf{H}^*\cdot\mathbf{v}_n^*+\sigma_n\cdot\underbrace{\mathbf{v}_n^{\mathsf{T}}\partial\mathbf{v}_n^*}_{\partial 1=0}\\ =&\mathbf{u}_n^{\mathsf{T}}\mathbf{H}_{\mathrm{B},g}^*\cdot\partial\boldsymbol{\Theta}_g^*\cdot\mathbf{H}_{\mathrm{F},g}^*\mathbf{v}_n^*\\ =&\mathrm{tr}(\mathbf{H}_{\mathrm{F},g}^*\mathbf{v}_n^*\mathbf{u}_n^{\mathsf{T}}\mathbf{H}_{\mathrm{B},g}^*\cdot\partial\boldsymbol{\Theta}_g^*).\end{aligned}$$

According to [55], the corresponding complex derivative is

$$\frac{\partial\sigma_n}{\partial\boldsymbol{\Theta}_g^*}=\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\mathbf{u}_n\mathbf{v}_n^{\mathsf{H}}\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}. \tag{33}$$

A linear combination of (33) yields (11).

B. Proof of Proposition 1

The scattering matrix of BD-RIS can be decomposed as[6]

$$\boldsymbol{\Theta}=\mathbf{L}\boldsymbol{\Theta}_{\mathrm{D}}\mathbf{R}^{\mathsf{H}}, \tag{34}$$

where $\boldsymbol{\Theta}_{\mathrm{D}}\in\mathbb{U}^{N_{\mathrm{S}}\times N_{\mathrm{S}}}$ corresponds to diagonal RIS and $\mathbf{L},\mathbf{R}\in\mathbb{U}^{N_{\mathrm{S}}\times N_{\mathrm{S}}}$ are block diagonal matrices of $L\times L$ unitary blocks. Manipulating $\mathbf{L}$ and $\mathbf{R}$ rotates the linear spans of $\bar{\mathbf{H}}_{\mathrm{B}}\triangleq\mathbf{H}_{\mathrm{B}}\mathbf{L}$ and $\bar{\mathbf{H}}_{\mathrm{F}}\triangleq\mathbf{R}^{\mathsf{H}}\mathbf{H}_{\mathrm{F}}$ and maintains their rank. On the other hand, there exists a $\boldsymbol{\Theta}_{\mathrm{D}}$ such that

$$\begin{aligned}\mathrm{rank}(\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}})=&\min\big(\mathrm{rank}(\mathbf{H}_{\mathrm{B}}),\mathrm{rank}(\boldsymbol{\Theta}_{\mathrm{D}}),\mathrm{rank}(\mathbf{H}_{\mathrm{F}})\big)\\ =&\min\big(\mathrm{rank}(\bar{\mathbf{H}}_{\mathrm{B}}),N_{\mathrm{S}},\mathrm{rank}(\bar{\mathbf{H}}_{\mathrm{F}})\big)\\ =&\max_{\boldsymbol{\Theta}}\ \mathrm{rank}(\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{F}})\end{aligned}$$

The same result holds if the direct link is present.

[6]This is because (block) unitary matrices are closed under multiplication.

C. Proof of Proposition 2

We consider rank-k forward channel and the proof follows similarly for rank-k backward channel. Let $\mathbf{H}_{\mathrm{F}}=\mathbf{U}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}}$ be the compact SVD of the forward channel. The channel Gram matrix $\mathbf{G}\triangleq\mathbf{H}\mathbf{H}^{\mathsf{H}}$ can be written as

$$\begin{aligned}\mathbf{G}=&\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{D}}^{\mathsf{H}}+\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{U}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}^{\mathsf{H}}\mathbf{U}_{\mathrm{F}}^{\mathsf{H}}\boldsymbol{\Theta}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\\ &+\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{U}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}^{\mathsf{H}}+\mathbf{H}_{\mathrm{D}}\mathbf{V}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}\mathbf{U}_{\mathrm{F}}^{\mathsf{H}}\boldsymbol{\Theta}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\\ =&\mathbf{H}_{\mathrm{D}}(\mathbf{I}-\mathbf{V}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}})\mathbf{H}_{\mathrm{D}}^{\mathsf{H}}\\ &+(\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{U}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}+\mathbf{H}_{\mathrm{D}}\mathbf{V}_{\mathrm{F}})(\boldsymbol{\Sigma}_{\mathrm{F}}\mathbf{U}_{\mathrm{F}}^{\mathsf{H}}\boldsymbol{\Theta}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}+\mathbf{V}_{\mathrm{F}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}^{\mathsf{H}})\\ =&\mathbf{Y}+\mathbf{Z}\mathbf{Z}^{\mathsf{H}},\end{aligned}$$

where we define $\mathbf{Y}\triangleq\mathbf{H}_{\mathrm{D}}(\mathbf{I}-\mathbf{V}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}})\mathbf{H}_{\mathrm{D}}^{\mathsf{H}}\in\mathbb{H}^{N_{\mathrm{R}}\times N_{\mathrm{R}}}$ and $\mathbf{Z}\triangleq\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{U}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}+\mathbf{H}_{\mathrm{D}}\mathbf{V}_{\mathrm{F}}\in\mathbb{C}^{N_{\mathrm{R}}\times k}$. That is to say, $\mathbf{G}$ can be expressed as a Hermitian matrix plus k rank-1 perturbations. According to the Cauchy interlacing formula [53], the n-th eigenvalue of $\mathbf{G}$ is bounded by

$$\lambda_n(\mathbf{G})\le\lambda_{n-k}(\mathbf{Y}),\qquad\text{if } n>k, \tag{35}$$

$$\lambda_n(\mathbf{G})\ge\lambda_n(\mathbf{Y}),\qquad\text{if } n<N-k+1. \tag{36}$$

Since $\mathbf{Y}=\mathbf{T}\mathbf{T}^{\mathsf{H}}$ is positive semi-definite, taking the square roots of (35) and (36) gives (12) and (13).

D. Proof of Proposition 3

Let $\mathbf{H}_{\mathrm{B}}=\mathbf{U}_{\mathrm{B}}\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{V}_{\mathrm{B}}^{\mathsf{H}}$ and $\mathbf{H}_{\mathrm{F}}=\mathbf{U}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}}$ be the SVD of the backward and forward channels, respectively. The scattering matrix of fully-connected RIS can be decomposed as

$$\boldsymbol{\Theta}=\mathbf{V}_{\mathrm{B}}\mathbf{X}\mathbf{U}_{\mathrm{F}}^{\mathsf{H}}, \tag{37}$$

where $\mathbf{X}\in\mathbb{U}^{N_{\mathrm{S}}\times N_{\mathrm{S}}}$ is a unitary matrix to be designed. The equivalent channel is thus a function of $\mathbf{X}$

$$\mathbf{H}=\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{F}}=\mathbf{U}_{\mathrm{B}}\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{X}\boldsymbol{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}}. \tag{38}$$

Since $\mathrm{sv}(\mathbf{U}\mathbf{A}\mathbf{V}^{\mathsf{H}})=\mathrm{sv}(\mathbf{A})$ for unitary $\mathbf{U}$ and $\mathbf{V}$, we have

$$\begin{aligned}\mathrm{sv}(\mathbf{H})=&\mathrm{sv}(\mathbf{U}_{\mathrm{B}}\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{X}\boldsymbol{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}})\\ =&\mathrm{sv}(\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{X}\boldsymbol{\Sigma}_{\mathrm{F}})\\ =&\mathrm{sv}(\bar{\mathbf{U}}_{\mathrm{B}}\boldsymbol{\Sigma}_{\mathrm{B}}\bar{\mathbf{V}}_{\mathrm{B}}^{\mathsf{H}}\bar{\mathbf{U}}_{\mathrm{F}}\boldsymbol{\Sigma}_{\mathrm{F}}\bar{\mathbf{V}}_{\mathrm{F}}^{\mathsf{H}})\\ =&\mathrm{sv}(\mathbf{B}\mathbf{F}),\end{aligned}$$

where $\bar{\mathbf{U}}_{\mathrm{B/F}}$ and $\bar{\mathbf{V}}_{\mathrm{B/F}}$ are arbitrary unitary matrices.

E. Proof of Lemma 2

The differential of R w.r.t. $\boldsymbol{\Theta}_g^*$ is [55]

$$\begin{aligned}\partial R=&\frac{1}{\eta}\mathrm{tr}\left\{\partial\mathbf{H}^*\cdot\mathbf{Q}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\Big(\mathbf{I}+\frac{\mathbf{H}^*\mathbf{Q}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}}{\eta}\Big)^{-1}\right\}\\ =&\frac{1}{\eta}\mathrm{tr}\left\{\mathbf{H}_{\mathrm{B},g}^*\cdot\partial\boldsymbol{\Theta}_g^*\cdot\mathbf{H}_{\mathrm{F},g}^*\mathbf{Q}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\Big(\mathbf{I}+\frac{\mathbf{H}^*\mathbf{Q}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}}{\eta}\Big)^{-1}\right\}\\ =&\frac{1}{\eta}\mathrm{tr}\left\{\mathbf{H}_{\mathrm{F},g}^*\mathbf{Q}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}\Big(\mathbf{I}+\frac{\mathbf{H}^*\mathbf{Q}^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}}{\eta}\Big)^{-1}\mathbf{H}_{\mathrm{B},g}^*\cdot\partial\boldsymbol{\Theta}_g^*\right\},\end{aligned}$$

and the corresponding complex derivative is (24).

$$2\Re\Big\{\sum_g \mathrm{tr}(\tilde{\boldsymbol{\Theta}}_g^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}) + \sum_{g_1,g_2}\mathrm{tr}(\tilde{\boldsymbol{\Theta}}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2}\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}})\Big\} \geq 2\Re\Big\{\sum_g \mathrm{tr}(\boldsymbol{\Theta}_g^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}) + \sum_{g_1,g_2}\mathrm{tr}(\boldsymbol{\Theta}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2}\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}})\Big\} \quad (43)$$

$$\sum_{g_1,g_2}\mathrm{tr}(\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}}\tilde{\boldsymbol{\Theta}}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\tilde{\boldsymbol{\Theta}}_{g_2}\mathbf{H}_{\mathrm{F},g_2}) - 2\Re\Big\{\sum_{g_1,g_2}\mathrm{tr}(\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}}\tilde{\boldsymbol{\Theta}}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2})\Big\} + \sum_{g_1,g_2}\mathrm{tr}(\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}}\boldsymbol{\Theta}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2}) \geq 0 \quad (44)$$

F. Proof of Proposition 4

The differential of (26a) w.r.t. $\boldsymbol{\Theta}_g^*$ is

$$\begin{aligned}\partial\|\mathbf{H}\|_{\mathrm{F}}^2 &= \mathrm{tr}\big(\mathbf{H}_{\mathrm{B},g}^*\cdot\partial\boldsymbol{\Theta}_g^*\cdot\mathbf{H}_{\mathrm{F},g}^*(\mathbf{H}_{\mathrm{D}}^{\mathsf{T}}+\mathbf{H}_{\mathrm{F}}^{\mathsf{T}}\boldsymbol{\Theta}^{\mathsf{T}}\mathbf{H}_{\mathrm{B}}^{\mathsf{T}})\big)\\ &= \mathrm{tr}\big(\mathbf{H}_{\mathrm{F},g}^*(\mathbf{H}_{\mathrm{D}}^{\mathsf{T}}+\mathbf{H}_{\mathrm{F}}^{\mathsf{T}}\boldsymbol{\Theta}^{\mathsf{T}}\mathbf{H}_{\mathrm{B}}^{\mathsf{T}})\mathbf{H}_{\mathrm{B},g}^*\cdot\partial\boldsymbol{\Theta}_g^*\big)\end{aligned}$$

and the corresponding complex derivative is

$$\frac{\partial\|\mathbf{H}\|_{\mathrm{F}}^2}{\partial\boldsymbol{\Theta}_g^*} = \mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{F}})\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}} = \mathbf{M}_g. \quad (39)$$

First, we approximate the quadratic objective (26a) by its local Taylor expansion

$$\max_{\boldsymbol{\Theta}} \quad \sum_g 2\Re\big\{\mathrm{tr}(\boldsymbol{\Theta}_g^{\mathsf{H}}\mathbf{M}_g)\big\} \quad (40a)$$

$$\text{s.t.} \quad \boldsymbol{\Theta}_g^{\mathsf{H}}\boldsymbol{\Theta}_g = \mathbf{I}, \quad \forall g. \quad (40b)$$

Let $\mathbf{M}_g = \mathbf{U}_g\boldsymbol{\Sigma}_g\mathbf{V}_g^{\mathsf{H}}$ be the compact SVD of $\mathbf{M}_g$. We have

$$\Re\big\{\mathrm{tr}(\boldsymbol{\Theta}_g^{\mathsf{H}}\mathbf{M}_g)\big\} = \Re\big\{\mathrm{tr}(\boldsymbol{\Sigma}_g\mathbf{V}_g^{\mathsf{H}}\boldsymbol{\Theta}_g^{\mathsf{H}}\mathbf{U}_g)\big\} \leq \mathrm{tr}(\boldsymbol{\Sigma}_g). \quad (41)$$

The upper bound is tight when $\mathbf{V}_g^{\mathsf{H}}\boldsymbol{\Theta}_g^{\mathsf{H}}\mathbf{U}_g = \mathbf{I}$, which implies the optimal solution of (40) is $\tilde{\boldsymbol{\Theta}}_g = \mathbf{U}_g\mathbf{V}_g^{\mathsf{H}}$, $\forall g$.

Next, we prove that solving (40) successively does not decrease (26a). Since $\tilde{\boldsymbol{\Theta}}$ optimal for problem (40), we have $\sum_g 2\Re\big\{\mathrm{tr}(\tilde{\boldsymbol{\Theta}}_g^{\mathsf{H}}\mathbf{M}_g)\big\} \geq \sum_g 2\Re\big\{\mathrm{tr}(\boldsymbol{\Theta}_g^{\mathsf{H}}\mathbf{M}_g)\big\}$ which is explicitly expressed by (43). On the other hand, expanding $\|\sum_g \mathbf{H}_{\mathrm{B},g}\tilde{\boldsymbol{\Theta}}_g\mathbf{H}_{\mathrm{F},g} - \sum_g \mathbf{H}_{\mathrm{B},g}\boldsymbol{\Theta}_g\mathbf{H}_{\mathrm{F},g}\|_{\mathrm{F}}^2 \geq 0$ gives (44). Adding (43) and (44), we have

$$\begin{aligned}&2\Re\Big\{\mathrm{tr}(\tilde{\boldsymbol{\Theta}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\mathsf{H}})\Big\} + \mathrm{tr}(\mathbf{H}_{\mathrm{F}}^{\mathsf{H}}\tilde{\boldsymbol{\Theta}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}\tilde{\boldsymbol{\Theta}}\mathbf{H}_{\mathrm{F}})\\ &\geq 2\Re\Big\{\mathrm{tr}(\boldsymbol{\Theta}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\mathsf{H}})\Big\} + \mathrm{tr}(\mathbf{H}_{\mathrm{F}}^{\mathsf{H}}\boldsymbol{\Theta}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{F}}), \quad (42)\end{aligned}$$

which suggests that updating $\tilde{\boldsymbol{\Theta}}$ does not decrease (26a).

Finally, we prove that the converging point of (40), denoted by $\tilde{\boldsymbol{\Theta}}^?$, is a stationary point of (26). The Karush-Kuhn-Tucker (KKT) conditions of (26) and (40) are equivalent in terms of primal/dual feasibility and complementary slackness, while the stationary conditions are respectively, $\forall g$,

$$\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}^\star\mathbf{H}_{\mathrm{F}})\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}} - \boldsymbol{\Theta}_g^\star\boldsymbol{\Lambda}_g^{\mathsf{H}} = 0, \quad (45)$$

$$\mathbf{M}_g - \boldsymbol{\Theta}_g^\star\boldsymbol{\Lambda}_g^{\mathsf{H}} = 0. \quad (46)$$

On convergence, (46) becomes $\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\boldsymbol{\Theta}^?\mathbf{H}_{\mathrm{F}})\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}} - \boldsymbol{\Theta}_g^?\boldsymbol{\Lambda}_g^{\mathsf{H}} = 0$ and reduces to (45). The proof is thus completed.

REFERENCES

[1] E. Basar, M. D. Renzo, J. D. Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8796365/

[2] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 5394–5409, 11 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8811733/

[3] ——, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Transactions on Communications*, vol. 68, pp. 1838–1851, 3 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8930608/

[4] Y. Yang, S. Zhang, and R. Zhang, "Irs-enhanced ofdma: Joint resource allocation and passive beamforming optimization," *IEEE Wireless Communications Letters*, vol. 9, pp. 760–764, 6 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8964457/

[5] B. Zheng, C. You, and R. Zhang, "Double-irs assisted multi-user mimo: Cooperative passive beamforming design," *IEEE Transactions on Wireless Communications*, vol. 20, pp. 4513–4526, 7 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9362274/

[6] X. Jia, J. Zhao, X. Zhou, and D. Niyato, "Intelligent reflecting surface-aided backscatter communications," vol. 2020-Janua. IEEE, 12 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9348003/

[7] Y. C. Liang, Q. Zhang, J. Wang, R. Long, H. Zhou, and G. Yang, "Backscatter communication assisted by reconfigurable intelligent surfaces," *Proceedings of the IEEE*, 2022.

[8] R. Liu, M. Li, Y. Liu, Q. Wu, and Q. Liu, "Joint transmit waveform and passive beamforming design for ris-aided dfrc systems," *IEEE Journal of Selected Topics in Signal Processing*, pp. 1–1, 5 2022.

[9] M. Hua, Q. Wu, C. He, S. Ma, and W. Chen, "Joint active and passive beamforming design for irs-aided radar-communication," *IEEE Transactions on Wireless Communications*, vol. 22, pp. 2278–2294, 4 2023.

[10] Q. Wu, X. Zhou, W. Chen, J. Li, and X. Zhang, "Irs-aided wpcns: A new optimization framework for dynamic irs beamforming," *IEEE Transactions on Wireless Communications*, pp. 1–1, 12 2021.

[11] Z. Feng, B. Clerckx, and Y. Zhao, "Waveform and beamforming design for intelligent reflecting surface aided wireless power transfer: Single-user and multi-user solutions," *IEEE Transactions on Wireless Communications*, 2022.

[12] Y. Zhao, B. Clerckx, and Z. Feng, "Irs-aided swipt: Joint waveform, active and passive beamforming design under nonlinear harvester model," *IEEE Transactions on Communications*, vol. 70, pp. 1345–1359, 2022.

[13] R. Karasik, O. Simeone, M. D. Renzo, and S. S. Shitz, "Beyond max-snr: Joint encoding for reconfigurable intelligent surfaces," vol. 2020-June. IEEE, 6 2020, pp. 2965–2970. [Online]. Available: https://ieeexplore.ieee.org/document/9174060/

[14] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond mimo paradigm for 6g," *IEEE Transactions on Communications*, vol. 68, pp. 3187–3196, 5 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8981888/

[15] Y. Zhao and B. Clerckx, "Riscatter: Unifying backscatter communication and reconfigurable intelligent surface," 12 2022. [Online]. Available: http://arxiv.org/abs/2212.09121

[16] W. Tang, J. Y. Dai, M. Chen, X. Li, Q. Cheng, S. Jin, K. Wong, and T. J. Cui, "Programmable metasurface-based rf chain-free 8psk wireless transmitter," *Electronics Letters*, vol. 55, pp. 417–420, 4 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1049/el.2019.0400

[17] J. Y. Dai, W. Tang, L. X. Yang, X. Li, M. Z. Chen, J. C. Ke, Q. Cheng, S. Jin, and T. J. Cui, "Realization of multi-modulation schemes for wireless communication by time-domain digital coding metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 68, pp. 1618–1627, 3 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8901437/

[18] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, pp. 106–112, 1 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8910627/

[19] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis," *IEEE Transactions on Wireless Communications*, pp. 1–1, 11 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9514409/

[20] 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 Transactions on Wireless Communications*, vol. 22, pp. 2311–2324, 4 2023.

[21] M. Nerini, S. Shen, and B. Clerckx, "Closed-form global optimization of beyond diagonal reconfigurable intelligent surfaces," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10155675/

[22] I. Santamaria, M. Soleymani, E. Jorswieck, and J. Gutiérrez, "Snr maximization in beyond diagonal ris-assisted single and multiple antenna links," *IEEE Signal Processing Letters*, vol. 30, pp. 923–926, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10187688/

[23] T. Fang and Y. Mao, "A low-complexity beamforming design for beyond-diagonal ris aided multi-user networks," *IEEE Communications Letters*, pp. 1–1, 7 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10319662/

[24] M. Nerini, S. Shen, H. Li, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces utilizing graph theory: Modeling, architecture design, and optimization," 5 2023. [Online]. Available: http://arxiv.org/abs/2305.05013

[25] 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 Transactions on Wireless Communications*, pp. 1–1, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10364738/

[26] H. Li, S. Shen, and B. Clerckx, "A dynamic grouping strategy for beyond diagonal reconfigurable intelligent surfaces with hybrid transmitting and reflecting mode," *IEEE Transactions on Vehicular Technology*, 12 2023.

[27] G. Bartoli, A. Abrardo, N. Decarli, D. Dardari, and M. D. Renzo, "Spatial multiplexing in near field mimo channels with reconfigurable intelligent surfaces," *IET Signal Processing*, vol. 17, 3 2023. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/sil2.12195

[28] H. Li, S. Shen, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces: A multi-sector mode enabling highly directional full-space wireless coverage," *IEEE Journal on Selected Areas in Communications*, vol. 41, pp. 2446–2460, 8 2023.

[29] H. Li, Y. Zhang, and B. Clerckx, "Channel estimation for beyond diagonal reconfigurable intelligent surfaces with group-connected architectures," 7 2023. [Online]. Available: http://arxiv.org/abs/2307.06129

[30] H. Li, S. Shen, M. Nerini, M. D. Renzo, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces with mutual coupling: Modeling and optimization," 10 2023. [Online]. Available: http://arxiv.org/abs/2310.02708

[31] H. Li, S. Shen, M. Nerini, and B. Clerckx, "Reconfigurable intelligent surfaces 2.0: Beyond diagonal phase shift matrices," 1 2023. [Online]. Available: http://arxiv.org/abs/2301.03288

[32] B. Ning, Z. Chen, W. Chen, and J. Fang, "Beamforming optimization for intelligent reflecting surface assisted mimo: A sum-path-gain maximization approach," *IEEE Wireless Communications Letters*, vol. 9, pp. 1105–1109, 7 2020.

[33] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Using intelligent reflecting surfaces for rank improvement in mimo communications." IEEE, 5 2020, pp. 9160–9164. [Online]. Available: https://ieeexplore.ieee.org/document/9052904/

[34] G.-H. Li, D.-W. Yue, and S.-N. Jin, "Spatially correlated rayleigh fading characteristics of ris-aided mmwave mimo communications," *IEEE Communications Letters*, vol. 27, pp. 2222–2226, 8 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10164200/

[35] 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 Communications Letters*, vol. 26, pp. 1101–1105, 5 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9706177/

[36] W. Huang, B. Lei, S. He, C. Kai, and C. Li, "Condition number improvement of irs-aided near-field mimo channels." IEEE, 5 2023, pp. 1210–1215. [Online]. Available: https://ieeexplore.ieee.org/document/10283534/

[37] M. A. ElMossallamy, H. Zhang, R. Sultan, K. G. Seddik, L. Song, G. Y. Li, and Z. Han, "On spatial multiplexing using reconfigurable intelligent surfaces," *IEEE Wireless Communications Letters*, vol. 10, pp. 226–230, 2 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9200661/

[38] S. Meng, W. Tang, W. Chen, J. Lan, Q. Y. Zhou, Y. Han, X. Li, and S. Jin, "Rank optimization for mimo channel with ris: Simulation and measurement," 7 2023. [Online]. Available: http://arxiv.org/abs/2307.13237

[39] T. E. Abrudan, J. Eriksson, and V. Koivunen, "Steepest descent algorithms for optimization under unitary matrix constraint," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1134–1147, 3 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4436033/

[40] T. Abrudan, J. Eriksson, and V. Koivunen, "Conjugate gradient algorithm for optimization under unitary matrix constraint," *Signal Processing*, vol. 89, pp. 1704–1714, 9 2009. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0165168409000814

[41] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, pp. 303–353, 1 1998. [Online]. Available: http://epubs.siam.org/doi/10.1137/S0895479895290954

[42] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009. [Online]. Available: https://books.google.co.uk/books?id=NSQGQeLN3NcC

[43] C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu, Y. Pan, H. Ren, M. D. Renzo, A. L. Swindlehurst, R. Zhang, and A. Y. Zhang, "An overview of signal processing techniques for ris/irs-aided wireless systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, pp. 883–917, 8 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9847080/

[44] M. T. Ivrlac and J. A. Nossek, "Toward a circuit theory of communication," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, pp. 1663–1683, 7 2010. [Online]. Available: https://ieeexplore.ieee.org/document/5446312/

[45] H.-R. Ahn, *Asymmetric Passive Components in Microwave Integrated Circuits*. Wiley, 2006. [Online]. Available: https://books.google.co.uk/books?id=X6WdLbOuSNQC

[46] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées," *Revue française d'informatique et de recherche opérationnelle. Série rouge*, vol. 3, pp. 35–43, 1969.

[47] L. Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pacific Journal of Mathematics*, vol. 16, pp. 1–3, 1 1966. [Online]. Available: http://msp.org/pjm/1966/16-1/p01.xhtml

[48] D. Semmler, M. Joham, and W. Utschick, "High snr analysis of ris-aided mimo broadcast channels," in *2023 IEEE 24th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 9 2023, pp. 221–225. [Online]. Available: https://ieeexplore.ieee.org/document/10304487/

[49] W. Fulton, "Eigenvalues, invariant factors, highest weights, and schubert calculus," *Bulletin of the American Mathematical Society*, vol. 37, pp. 209–249, 4 2000. [Online]. Available: https://www.ams.org/bull/2000-37-03/S0273-0979-00-00865-X/

[50] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*. Elsevier Science, 2013. [Online]. Available: https://books.google.co.uk/books?id=drEX1J7jHUIC

[51] J. C. Gower and G. B. Dijksterhuis, *Procrustes Problems*. OUP Oxford, 2004. [Online]. Available: https://books.google.co.uk/books?id=kRRREAAAQBAJ

[52] T. Bell, "Global positioning system-based attitude determination and the orthogonal procrustes problem," *Journal of Guidance, Control, and Dynamics*, vol. 26, pp. 820–822, 9 2003. [Online]. Available: https://arc.aiaa.org/doi/10.2514/2.5117

[53] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Johns Hopkins University Press, 2013. [Online]. Available: https://jhupbooks.press.jhu.edu/title/matrix-computations

[54] F. Nie, R. Zhang, and X. Li, "A generalized power iteration method for solving quadratic problem on the stiefel manifold," *Science China Information Sciences*, vol. 60, p. 112101, 11 2017. [Online]. Available: http://link.springer.com/10.1007/s11432-016-9021-9

[55] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, pp. 2740–2746, 6 2007. [Online]. Available: http://ieeexplore.ieee.org/document/4203075/