Channel Shaping Using Beyond Diagonal Reconfigurable Intelligent Surface: Analysis, Optimization, and Enhanced Flexibility

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

***Abstract*—This paper investigates the capability of a passive Reconfigurable Intelligent Surface (RIS) to redistribute the singular values of point-to-point Multiple-Input Multiple-Output (MIMO) channels for achieving power and rate gains. We depart from the conventional Diagonal (D)-RIS with diagonal phase shift matrix and adopt a Beyond Diagonal (BD) architecture not restricted to diagonal scattering matrix and featuring element-wise connections for extra wave manipulation flexibility. Specifically, we first provide shaping insights by characterizing the channel singular value regions attainable by D-RIS and BD-RIS via a novel geodesic optimization. Analytical singular value bounds are then derived to explore their shaping limits in typical deployment scenarios. As a side product, we tackle BD-RIS-aided MIMO rate maximization problem by a local-optimal Alternating Optimization (AO) and then propose a shaping-inspired low-complexity approach. Results show that compared to D-RIS, BD-RIS significantly improves the dynamic range of all channel singular values, the trade-off in manipulating them, and thus the channel power gain and achievable rate. Those observations become more pronounced when MIMO and RIS dimensions increase. Of particular interest, BD-RIS is shown to activate multi-stream transmission at lower transmit power than D-RIS, hence better approaching Degree of Freedom (DoF) at low Signal-to-Noise Ratio (SNR) thanks to its higher flexibility of shaping the distribution of channel singular values.**

***Index Terms*—Reconfigurable intelligent surface, channel singular value redistribution, rate maximization, manifold optimization.**

I. Introduction

A. Background

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 can serve on demand. This is empowered by recent advances in Reconfigurable Intelligent Surface (RIS), a programmable passive planar surface that recycles and redistributes ambient electromagnetic waves for improved wireless performance. A typical RIS consists of numerous low-power sub-wavelength non-resonant scattering elements, whose response can be engineered in real-time to manipulate the amplitude, phase, frequency, and polarization of the scattered waves [1]. It enables low-noise full-duplex operation, featuring better flexibility than reflectarrays, lighter footprint than relays, and greater scalability than Multiple-Input Multiple-Output (MIMO) systems. One popular RIS research direction is *joint passive and active beamforming* design with transceivers to enhance a specific performance measure, which has attracted significant interests in wireless communication [2]–[4], sensing [5]–[7], and power transfer literature [8]–[10]. While passive beamforming at RIS suffers attenuation from double fading, it offers better asymptotic behaviors than active beamforming at transceivers (e.g., second-order array gain and fourth-order harvested power [10]). Another RIS application is *information modulation* by periodically switching its reflection pattern within the channel coherence time. This creates a free-ride message stream with dual benefits: integrating with legacy transmitter for enhanced channel capacity [11]–[13], or serving as individual information source for low-power uplink communication [14]–[16]. Different from above, *channel shaping* exploits RIS as a stand-alone device to modify the inherent properties of the wireless environment. It can compensate for the Doppler effect [17], transform frequency-selective channels into frequency-flat [18], improve the channel rank for MIMO systems [19], and create artificial time diversity for orthogonal [20] and non-orthogonal [21] multiple access schemes. Channel shaping also provides a ubiquitous solution for different wireless systems, decoupling joint beamforming design by first optimizing the channel and then the transceiver.

B. Related Works

At a specific resource block, channel shaping metrics can be classified into two categories:

- *Singular value:* The impact of RIS has been studied in terms of minimum singular value [22], effective rank [22], [23], condition number [24], [25], and degree of freedom [26]–[28]. Those are closely related to performance measures (e.g., achievable rate and harvested power [29]) but sensitive to minor perturbations of the channel matrix.
- *Power:* The impact of RIS has been studied in terms of channel power gain [2], [30]–[33] in point-to-point channels and leakage interference [34] in interference channels. Those second-order metrics are less informative in MIMO but easier to analyze and optimize.

Although above works offered inspiring glimpses into the channel shaping potential, there exist no attempts to characterize the channel singular value region manipulated by passive RIS. Besides, most relevant literature [2], [22]–[28], [34] adopted a Diagonal (D)-RIS model where each element

Yang Zhao, Hongyu Li, and Bruno Clerckx are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: {yang.zhao18, c.li21, b.clerckx}@imperial.ac.uk).

Massimo Franceschetti is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla CA 92093, USA (e-mail: massimo@ece.ucsd.edu).

is connected to a dedicated impedance and disconnected from others, such that wave impinging on one element is entirely scattered by the same element. This simple architecture is modelled by a diagonal scattering matrix with unit-magnitude diagonal entries, which only applies a phase shift to the incoming signal. The concept was soon generalized to Beyond Diagonal (BD)-RIS with group-connected architecture [30] that connects adjacent elements from the same group via passive reconfigurable components[1]. This allows wave impinging on one element to propagate within the circuit and depart partially from any element in the same group. It can thus manipulate both amplitude and phase of the scattered wave while remaining passive, generalizing the scattering matrix to block-diagonal with unitary blocks. Such a powerful model can be realized at reduced hardware cost using tree- and forest-connected architectures inspired by graph theory [32]. BD-RIS can also function in multi-sector mode [36] for full-space coverage and multi-user support. Many practical design challenges, such as channel estimation [37], mutual coupling [38], and wideband modelling [39], have also been investigated in recent publications. Its beamforming superiority over D-RIS has been studied extensively in Single-Input Single-Output (SISO) and Multiple-Input Single-Output (MISO) systems [30]–[33], [36], [40]–[42], however, the interplay between BD-RIS and MIMO is still in the infancy stage. The authors of [43] investigated the rate-optimal joint beamforming design for a fully-connected BD-RIS-aided MIMO system where the direct link is blocked. A transmitter-side BD-RIS was introduced to massive MIMO systems that exploits statistical Channel State Information (CSI) for improved spectral efficiency [44], but was again limited to blocked direct channel and fully-connected BD-RIS. Received power maximization with continuous-valued and discrete-valued BD-RIS have been tackled respectively in closed form [31] and by machine learning approach [45], but the corresponding single-stream transceiver is rate-suboptimal and the BD-RIS design problem is equivalent to SISO shaping.

C. Contributions

This paper is motivated by the critical question: *What is the singular value (and power gain) shaping capability of a passive RIS in point-to-point MIMO channels?* We aim for a comprehensive answer via theoretical analysis and numerical optimization. The contributions are summarized below.

First, we pioneer BD-RIS study in general MIMO channels and interpret its shaping potential as channel rearrangement and space alignment. Channel rearrangement refers to rearranging and recombining the entries of forward (i.e., transmitter-RIS) and backward (i.e., RIS-receiver) channel matrices by their strength. Space alignment generalizes phase matching of direct and indirect channels in SISO and MISO to the singular vector space in MIMO. The former is uniquely attributed to the off-diagonal entries of the scattering matrix of BD-RIS.

Second, we propose a novel BD-RIS design method that allows reshaping of the available channels through singular values manipulation. Our method based on Riemannian Conjugate Gradient (RCG) compares favorably with respect to existing solutions in that the updates are along the geodesics (i.e., the shortest path between two points in a Riemannian manifold) of the feasible domain to accelerate convergence. It also works for general BD-RIS optimization problems.

Third, we provide a numerical answer to the shaping question by characterizing the Pareto frontiers of channel singular values. The enclosed region generalizes most relevant metrics and provides an intuitive shaping benchmark. Results show that increasing BD-RIS group size enlarges this region, improving the dynamic range of all singular values and the trade-off in manipulating them.

Fourth, we provide an analytical answer to the shaping question in two special channel conditions. For rank-deficient forward/backward channel, we derive singular value bounds applying to D- and BD-RIS with asymptotically large number of elements. For blocked direct channel, we derive singular value bounds applying to fully-connected BD-RIS with arbitrary number of elements. Those bounds are validated by comparing with the numerical results above. Results show that for a fixed number of elements, BD-RIS can approach the asymptotic bounds better than D-RIS.

Fifth, we tackle BD-RIS-aided MIMO rate maximization problem by a local-optimal Alternating Optimization (AO) and a low-complexity shaping-inspired approach. The former iteratively updates active beamforming by eigenmode transmission and passive beamforming by geodesic RCG, until convergence. The latter exploits the BD-RIS to shape the channel for maximum power gain then performs eigenmode transmission. Interestingly, the rate deficit from the shaping-inspired approach diminishes as the BD-RIS evolves towards fully-connected. We conclude that: 1) channel shaping decouples joint beamforming for reduced design complexity; 2) the power and rate gains of BD-RIS over D-RIS increase with the number of scattering elements and MIMO dimensions; 3) at low transmit power, BD-RIS can activate more streams and better approach Degree of Freedom (DoF) than D-RIS, thanks to its higher flexibility of shaping the distribution of channel singular values.

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. $\arg(\cdot)$ gives the argument 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 (block) 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 stationary point, respectively. $(\cdot)_{[x:y]}$ is a shortcut for $(\cdot)_x,(\cdot)_{x+1},\ldots,(\cdot)_y$. $|\cdot|$, $\|\cdot\|$, and $\|\cdot\|_\mathrm{F}$ denote the absolute value, Euclidean norm, and Frobenius norm, respectively. $\odot$ represents the element-wise (Hadamard) product. $\mathcal{CN}(\mathbf{0},\mathbf{\Sigma})$ is the multivariate Circularly Symmetric Complex Gaussian

[1]Those components can be either symmetric (e.g., capacitors and inductors) or asymmetric (e.g., ring hybrids and branch-line hybrids) [35], resulting in symmetric and asymmetric scattering matrices, respectively.

(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 $N_{\rm T}$ and $N_{\rm R}$ transmit and receive antennas, respectively, and $N_{\rm S}$ scattering elements at the BD-RIS. This configuration is denoted as $N_{\rm T}\times N_{\rm S}\times N_{\rm R}$ in the following context. The BD-RIS can be modeled as an $N_{\rm S}$-port network [46] that further divides into G individual groups, each containing $L\triangleq N_{\rm S}/G$ elements interconnected by real-time reconfigurable components [30]. To simplify the analysis, we assume a lossless asymmetric network without mutual coupling between scattering elements, as have been previously considered in [36], [43], [47]. The overall scattering matrix of the BD-RIS is block-unitary[2]

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

where $\mathbf{\Theta}_g\in\mathbb{U}^{L\times L}$ is the g-th unitary block (i.e., $\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I}$) that describes the response of group $g\in\mathcal{G}\triangleq\{1,...,G\}$. Note that D-RIS can be regarded as its extreme case with group size $L=1$. Some potential physical architectures of BD-RIS are illustrated in [30, Fig. 3], [36, Fig. 5], and [32, Fig. 2], where the circuit topology has been modelled in the scattering matrix. Let $\mathbf{H}_{\rm D}\in\mathbb{C}^{N_{\rm R}\times N_{\rm T}}$, $\mathbf{H}_{\rm B}\in\mathbb{C}^{N_{\rm R}\times N_{\rm S}}$, $\mathbf{H}_{\rm F}\in\mathbb{C}^{N_{\rm S}\times N_{\rm T}}$ denote the direct (i.e., transmitter-receiver), backward (i.e., RIS-receiver), and forward (i.e., transmitter-RIS) channels, respectively. The equivalent channel is a function of the scattering matrix

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

where $\mathbf{H}_{{\rm B},g}\in\mathbb{C}^{N_{\rm R}\times L}$ and $\mathbf{H}_{{\rm F},g}\in\mathbb{C}^{L\times N_{\rm T}}$ are the backward and forward channels associated with BD-RIS group g, corresponding to the $(g-1)L+1$ to gL columns of $\mathbf{H}_{\rm B}$ and rows of $\mathbf{H}_{\rm F}$, respectively. Let $\mathbf{H}_g\triangleq\mathbf{H}_{{\rm B},g}\mathbf{\Theta}_g\mathbf{H}_{{\rm F},g}$ be the indirect channel via BD-RIS group g. Since unitary matrices constitute an algebraic group with respect to multiplication, the scattering matrix of group g can be decomposed as

$$\mathbf{\Theta}_g=\mathbf{L}_g\mathbf{R}_g^{\mathsf{H}}, \tag{3}$$

where $\mathbf{L}_g,\mathbf{R}_g\in\mathbb{U}^{L\times L}$ are two unitary factor matrices. Let $\mathbf{H}_{{\rm B},g}=\mathbf{U}_{{\rm B},g}\mathbf{\Sigma}_{{\rm B},g}\mathbf{V}_{{\rm B},g}^{\mathsf{H}}$ and $\mathbf{H}_{{\rm F},g}=\mathbf{U}_{{\rm F},g}\mathbf{\Sigma}_{{\rm F},g}\mathbf{V}_{{\rm F},g}^{\mathsf{H}}$ be the compact Singular Value Decomposition (SVD) of the backward and forward channels, respectively. The equivalent channel can thus be rewritten as

$$\mathbf{H}=\overbrace{\mathbf{H}_{\rm D}+\sum_g\mathbf{U}_{{\rm B},g}\mathbf{\Sigma}_{{\rm B},g}\underbrace{\mathbf{V}_{{\rm B},g}^{\mathsf{H}}\mathbf{L}_g\mathbf{R}_g^{\mathsf{H}}\mathbf{U}_{{\rm F},g}}_{\text{backward-forward}}\mathbf{\Sigma}_{{\rm F},g}\mathbf{V}_{{\rm F},g}^{\mathsf{H}}}^{\text{direct-indirect}}. \tag{4}$$

By analyzing (4), we conclude that the off-diagonal entries of the BD-RIS scattering matrix provide two key potentials for MIMO channel shaping:

- *Channel rearrangement:* It refers to rearranging and recombining the backward and forward channel branches (i.e., entries of $\mathbf{H}_{\rm B}$ and $\mathbf{H}_{\rm F}$) by their strength. In SISO[3], D-RIS with perfect phase matching can provide a maximum indirect channel gain $g_{\rm D}=\sum_{n=1}^{N_{\rm S}}|h_{{\rm B},n}||h_{{\rm F},n}|$, while BD-RIS can exploit in-group connections and generalize it to $g_{\rm BD}=\sum_{g=1}^{G}\sum_{l=1}^{L}|h_{{\rm B},g,\pi_{{\rm B},g}(l)}||h_{{\rm F},g,\pi_{{\rm F},g}(l)}|$, where $\pi_{{\rm B},g}$ and $\pi_{{\rm F},g}$ are arbitrary permutations of $\mathcal{L}\triangleq\{1,...,L\}$. Note the first summation is over groups and the second summation is over permuted channels. By rearrangement inequality [48, Inequality 10.2.1], for BD-RIS we have

$$g_{\rm BD}\leq\sum_{g=1}^{G}\sum_{l=1}^{L}|h_{{\rm B},g,\pi_{{\rm B},g}^{\downarrow}(l)}||h_{{\rm F},g,\pi_{{\rm F},g}^{\downarrow}(l)}|, \tag{5}$$

  where $\pi_{{\rm B},g}^{\downarrow}$ and $\pi_{{\rm F},g}^{\downarrow}$ sort $\{|h_{{\rm B},g,l}|\}_{l=1}^{L}$ and $\{|h_{{\rm F},g,l}|\}_{l=1}^{L}$ in descending orders, respectively. That is to say, when the direct channel is blocked, the maximum SISO channel gain is attained if the BD-RIS pairs the l-th strongest backward and forward branches within each group. Since the number of channels associated with each group is proportional to $N_{\rm T}N_{\rm R}$, we conclude that the advantage of BD-RIS in channel rearrangement scales with MIMO dimensions.
- *Space alignment:* It refers to aligning the singular vectors of the direct, forward, and backward channels. The BD-RIS needs to strike a balance between the alignment of backward-forward (intra-group, multiplicative) channels and direct-indirect (inter-group, additive) channels. In SISO, singular vectors become scalars and space alignment boils down to phase matching, such that the scattering matrix of group g maximizing the channel gain is

$$\mathbf{\Theta}_g^{\star}=\exp\big(\jmath\arg(h_{\rm D})\big)\mathbf{V}_{{\rm B},g}\mathbf{U}_{{\rm F},g}^{\mathsf{H}}, \tag{6}$$

  where $\mathbf{V}_{{\rm B},g}=\big[\mathbf{h}_{{\rm B},g}/\|\mathbf{h}_{{\rm B},g}\|,\ \mathbf{N}_{{\rm B},g}\big]\in\mathbb{U}^{L\times L}$, $\mathbf{U}_{{\rm F},g}=\big[\mathbf{h}_{{\rm F},g}/\|\mathbf{h}_{{\rm F},g}\|,\ \mathbf{N}_{{\rm F},g}\big]\in\mathbb{U}^{L\times L}$, and $\mathbf{N}_{{\rm B},g},\mathbf{N}_{{\rm F},g}\in\mathbb{C}^{L\times(L-1)}$ are the orthonormal bases of the null spaces of $\mathbf{h}_{{\rm B},g}$ and $\mathbf{h}_{{\rm F},g}$, respectively. It is thus apparent that D-RIS suffices for perfect phase matching in SISO. When it comes to MIMO, each D-RIS element can only apply a scalar phase shift to the associated indirect channel matrix of size $N_{\rm R}\times N_{\rm T}$ defined in (2). We thus conclude that the disadvantage of D-RIS in space alignment deteriorates with MIMO dimensions.

III. Group-Connected BD-RIS Design Frameworks

This section aims for an efficient yet universal optimization framework for group-connected BD-RIS design. Most relevant optimization problems can be formulated as

$$\max_{\mathbf{\Theta}}\quad f(\mathbf{\Theta}) \tag{7a}$$

$$\text{s.t.}\quad \mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I},\quad\forall g, \tag{7b}$$

where the objective function $f(\mathbf{\Theta})$ can be any function of the BD-RIS scattering matrix, for example, channel singular value (to be discussed in Section IV) and achievable rate

[2]Following footnote 1, we do not limit the scattering matrix to be symmetric. One can enforce symmetry over the result by $\mathbf{\Theta}\leftarrow(\mathbf{\Theta}+\mathbf{\Theta}^{\mathsf{T}})/2$ if necessary.

[3]Following previous notation, here we denote the direct channel as $h_{\rm D}$, the backward and forward channels as $\mathbf{h}_{\rm B}\in\mathbb{C}^{N_{\rm S}\times 1}$ and $\mathbf{h}_{\rm F}^{\mathsf{H}}\in\mathbb{C}^{1\times N_{\rm S}}$, whose n-th entries are $h_{{\rm B},n}$ and $h_{{\rm F},n}$, respectively. The channels associated with BD-RIS group g are also denoted as $\mathbf{h}_{{\rm B},g}\in\mathbb{C}^{L\times 1}$ and $\mathbf{h}_{{\rm F},g}^{\mathsf{H}}\in\mathbb{C}^{1\times L}$, whose l-th entries are $h_{{\rm B},g,l}$ and $h_{{\rm F},g,l}$, respectively.

(to be discussed in Section V). The feasible domain of each group is a L-dimensional Stiefel manifold $\mathbf{\Theta}_g \in \mathbb{U}^{L\times L}$, which is non-convex and non-Euclidean. Therefore, relevant optimization problems are usually solved by relax-then-project methods [40] or universal manifold RCG [36], [41], [47]. The former solves an unconstrained problem (7a) by quasi-Newton methods, then projects the solution back to the Stiefel manifold. It has no guarantee of optimality and usually suffers from poor performance. The latter generalizes the conjugate gradient methods to Riemannian manifolds, which iteratively refines the solution by addition and retraction. In the following context, we first provide an overview on the general RCG method that is widely adopted in D- and BD-RIS literature, discuss its drawbacks inherited from the non-geodesic nature, then propose a novel group-wise geodesic RCG method that operates directly on the Stiefel manifold for faster convergence.

A. General (Non-Geodesic) RCG

A geodesic is a curve representing the shortest path between two points in a Riemannian manifold, whose tangent vectors remain parallel when transporting along the curve. The universal RCG method proposed in [49], [50] is applicable to optimization problems over arbitrary manifolds. The idea is to perform additive updates along the conjugate direction guided by the Riemannian gradient, project the solution back onto the manifold, and repeat until convergence. For optimization problem (7), the steps for BD-RIS group g at iteration r are summarized below:

1) *Compute the Euclidean gradient:* The gradient of f with respect to $\mathbf{\Theta}_g^*$ in the Euclidean space is

$$\nabla_{\mathrm{E},g}^{(r)} = \frac{\partial f(\mathbf{\Theta}_g^{(r)})}{\partial \mathbf{\Theta}_g^*}; \tag{8}$$

2) *Translate to the Riemannian gradient [49]:* At point $\mathbf{\Theta}^{(r)}$, the Riemannian gradient lies in the tangent space of the Stiefel manifold $\mathcal{T}_{\mathbf{\Theta}_g^{(r)}}\mathbb{U}^{L\times L} \triangleq \{\mathbf{M}\in\mathbb{C}^{L\times L} \,|\, \mathbf{M}^{\mathsf{H}}\mathbf{\Theta}_g^{(r)} + \mathbf{\Theta}_g^{(r)\mathsf{H}}\mathbf{M} = \mathbf{0}\}$. It gives the steepest ascent direction of the objective on the manifold can be obtained by projecting the Euclidean gradient onto the tangent space:

$$\nabla_{\mathrm{R},g}^{(r)} = \nabla_{\mathrm{E},g}^{(r)} - \mathbf{\Theta}_g^{(r)}\nabla_{\mathrm{E},g}^{(r)\mathsf{H}}\mathbf{\Theta}_g^{(r)}; \tag{9}$$

3) *Determine the conjugate direction [51]:* The conjugate direction is obtained over the Riemannian gradient and previous direction as

$$\mathbf{D}_g^{(r)} = \nabla_{\mathrm{R},g}^{(r)} + \gamma_g^{(r)}\mathbf{D}_g^{(r-1)}, \tag{10}$$

where $\gamma_g^{(r)}$ is the parameter that deviates the conjugate direction from the tangent space for accelerated convergence. A popular choice is the Polak-Ribière formula

$$\gamma_g^{(r)} = \frac{\mathrm{tr}\big((\nabla_{\mathrm{R},g}^{(r)} - \nabla_{\mathrm{R},g}^{(r-1)})\nabla_{\mathrm{R},g}^{(r)\mathsf{H}}\big)}{\mathrm{tr}\big(\nabla_{\mathrm{R},g}^{(r-1)}\nabla_{\mathrm{R},g}^{(r-1)\mathsf{H}}\big)}; \tag{11}$$

4) *Perform additive update [50]:* The point is updated by moving along a straight path in the conjugate direction

$$\bar{\mathbf{\Theta}}_g^{(r+1)} = \mathbf{\Theta}_g^{(r)} + \mu\mathbf{D}_g^{(r)}, \tag{12}$$

where μ is the step size refinable by the Armijo rule [52];

5) *Retract for feasibility [47], [49]:* The resulting point needs to be projected to the closest point (in terms of Euclidean distance) on the Stiefel manifold by

$$\mathbf{\Theta}_g^{(r+1)} = \bar{\mathbf{\Theta}}_g^{(r+1)}\big(\bar{\mathbf{\Theta}}_g^{(r+1)\mathsf{H}}\bar{\mathbf{\Theta}}_g^{(r+1)}\big)^{-1/2}. \tag{13}$$

One can also combine the addition (12) and retraction (13) in one step

$$\mathbf{\Theta}_g^{(r+1)} = \big(\mathbf{\Theta}_g^{(r)} + \mu\mathbf{D}_g^{(r)}\big)\big(\mathbf{I} + \mu^2\mathbf{D}_g^{(r)\mathsf{H}}\mathbf{D}_g^{(r)}\big)^{-1/2}, \tag{14}$$

and determine the step size therein.

The method is called non-geodesic since the addition (12) and retraction (13) in each step constitute a zigzag path departing from and returning to the manifold, instead of carried directly on the manifold. It converges to stationary points of the original problem but usually requires a large number of iterations due to inefficient operations in the Euclidean space.

B. Geodesic RCG

Before introducing geodesic RCG, we revisit some basic concepts in differential geometry and Lie algebra. A Lie group is simultaneously a continuous group and a differentiable manifold. Lie algebra refers to the tangent space of the Lie group at the identity element. The exponential map acts as a bridge between the Lie algebra and Lie group, which allows one to recapture the local group structure using linear algebra techniques. The set of unitary matrices $\mathbb{U}^{L\times L}$ forms a Lie group $U(L)$ under multiplication, and the corresponding Lie algebra $\mathfrak{u}(L) \triangleq \mathcal{T}_{\mathbf{I}}\mathbb{U}^{L\times L} = \{\mathbf{M}\in\mathbb{C}^{L\times L} \,|\, \mathbf{M}^{\mathsf{H}} + \mathbf{M} = \mathbf{0}\}$ consists of skew-Hermitian matrices. A geodesic emanating from the identity with velocity $\mathbf{D}\in\mathfrak{u}(L)$ can be described by [53]

$$\mathbf{G}_{\mathbf{I}}(\mu) = \exp(\mu\mathbf{D}), \tag{15}$$

where $\exp(\mathbf{A}) = \sum_{k=0}^{\infty}(\mathbf{A}^k/k!)$ is the matrix exponential and μ is the step size (i.e., magnitude of the tangent vector). Note that the right translation is an isometry in $U(L)$. During the optimization of group g, the geodesic evaluated at the identity (15) should be translated to $\mathbf{\Theta}_g^{(r)}$ for successive updates [54]

$$\mathbf{G}_g^{(r)}(\mu) = \mathbf{G}_{\mathbf{I}}(\mu)\mathbf{\Theta}_g^{(r)} = \exp(\mu\mathbf{D}_g^{(r)})\mathbf{\Theta}_g^{(r)}, \tag{16}$$

while the Riemannian gradient evaluated at $\mathbf{\Theta}_g^{(r)}$ (9) should be translated back to the identity for exploiting the Lie algebra [54]

$$\tilde{\nabla}_{\mathrm{R},g}^{(r)} = \nabla_{\mathrm{R},g}^{(r)}\mathbf{\Theta}_g^{(r)\mathsf{H}} = \nabla_{\mathrm{E},g}^{(r)}\mathbf{\Theta}_g^{(r)\mathsf{H}} - \mathbf{\Theta}_g^{(r)}\nabla_{\mathrm{E},g}^{(r)\mathsf{H}}. \tag{17}$$

After gradient translation, the deviation parameter and conjugate direction can be determined similarly to (11) and (10)

$$\tilde{\gamma}_g^{(r)} = \frac{\mathrm{tr}\big((\tilde{\nabla}_{\mathrm{R},g}^{(r)} - \tilde{\nabla}_{\mathrm{R},g}^{(r-1)})\tilde{\nabla}_{\mathrm{R},g}^{(r)\mathsf{H}}\big)}{\mathrm{tr}\big(\tilde{\nabla}_{\mathrm{R},g}^{(r-1)}\tilde{\nabla}_{\mathrm{R},g}^{(r-1)\mathsf{H}}\big)}. \tag{18}$$

$$\mathbf{D}_g^{(r)} = \tilde{\nabla}_{\mathrm{R},g}^{(r)} + \tilde{\gamma}_g^{(r)}\mathbf{D}_g^{(r-1)}, \tag{19}$$

The solution can thus be updated along the geodesic in a multiplicative rotational manner

$$\mathbf{\Theta}_g^{(r+1)} = \mathbf{G}_g^{(r)}(\mu) = \exp(\mu\mathbf{D}_g^{(r)})\mathbf{\Theta}_g^{(r)}, \tag{20}$$

Algorithm 1: Group-wise geodesic RCG for BD-RIS design

Input: $f(\mathbf{\Theta})$, G
Output: $\mathbf{\Theta}^{\star}$

Initialize $r \leftarrow 0$, $\mathbf{\Theta}^{(0)}$
Repeat
  For $g \leftarrow 1$ to G
    $\nabla_{\mathrm{E},g}^{(r)} \leftarrow$ (8)
    $\tilde{\nabla}_{\mathrm{R},g}^{(r)} \leftarrow$ (17)
    $\tilde{\gamma}_g^{(r)} \leftarrow$ (18)
    $\mathbf{D}_g^{(r)} \leftarrow$ (19)
    If $\Re\{\mathrm{tr}(\mathbf{D}_g^{(r)\mathsf{H}} \tilde{\nabla}_{\mathrm{R},g}^{(r)})\} < 0$ ▷ not an ascent direction
      $\mathbf{D}_g^{(r)} \leftarrow \tilde{\nabla}_{\mathrm{R},g}^{(r)}$
    End If
    $\mu \leftarrow 1$
    $\mathbf{G}_g^{(r)}(\mu) \leftarrow$ (16)
    While $f(\mathbf{G}_g^{(r)}(2\mu)) - f(\mathbf{\Theta}_g^{(r)}) \geq \mu \cdot \mathrm{tr}(\mathbf{D}_g^{(r)} \mathbf{D}_g^{(r)\mathsf{H}})/2$
      $\mu \leftarrow 2\mu$
    End While
    While $f(\mathbf{G}_g^{(r)}(\mu)) - f(\mathbf{\Theta}_g^{(r)}) < \mu/2 \cdot \mathrm{tr}(\mathbf{D}_g^{(r)} \mathbf{D}_g^{(r)\mathsf{H}})/2$
      $\mu \leftarrow \mu/2$
    End While
    $\mathbf{\Theta}_g^{(r+1)} \leftarrow$ (20)
  End For
  $r \leftarrow r+1$
Until $|f(\mathbf{\Theta}^{(r)}) - f(\mathbf{\Theta}^{(r-1)})| / f(\mathbf{\Theta}^{(r-1)}) \leq \epsilon$

where an appropriate μ may be obtained by the Armijo rule. To double the step size, one can simply square the rotation matrix instead of recomputing the matrix exponential, since $\exp^2(\mu \mathbf{D}_g^{(r)}) = \exp(2\mu \mathbf{D}_g^{(r)})$. We would like to highlight that the proposed geodesic RCG method has been tailored to the Stiefel manifold.

Algorithm 1 summarizes the proposed BD-RIS design framework based on group-wise geodesic RCG. Compared to the general non-geodesic approach, it leverages the Lie group properties to replace the add-then-retract update (14) with a multiplicative rotational update (20) along geodesics of the Stiefel manifold. This leads to faster convergence and simplifies the step size tuning thanks to appropriate parameter space. Convergence to a local optimum is still guaranteed if not initialized at a stationary point. Note that the group-wise updates can be performed in parallel to facilitate large-scale BD-RIS optimization problems. Since the set of block-unitary matrices are algebraically closed under multiplication, one can also avoid group-wise updates by directly operating on $\mathbf{\Theta}$ and pinching (i.e., keeping the main block diagonal and nulling other entries) the Euclidean gradient (8), with potentially higher computational complexity and slower convergence.

IV. Channel Singular Values Redistribution

In this section, we first provide a toy example to illustrate the channel shaping advantage of BD-RIS architecture. Next, we numerically characterize the channel singular value region based on the proposed Algorithm 1. Finally, we derive some analytical singular value bounds in specific channel conditions.

A. Toy Example

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

$$\mathbf{\Theta}_{\mathrm{U}} = e^{\jmath\phi} \begin{bmatrix} e^{\jmath\alpha}\cos\psi & e^{\jmath\beta}\sin\psi \\ -e^{-\jmath\beta}\sin\psi & e^{-\jmath\alpha}\cos\psi \end{bmatrix}. \tag{21}$$

It is worth noting that ϕ has no impact on the singular value because $\mathrm{sv}(e^{\jmath\phi}\mathbf{A}) = \mathrm{sv}(\mathbf{A})$. To simplify the analysis, we also enforce symmetry $\mathbf{\Theta}_{\mathrm{U}} = \mathbf{\Theta}_{\mathrm{U}}^{\mathsf{T}}$ by $\beta = \pi/2$ such that both architectures have the same number of variables in the scattering matrix. Fig. 1 compares the achievable singular values of one typical channel realization, where the direct channel is blocked and the forward and backward channels are normalized Rayleigh fading

$$\mathbf{H}_{\mathrm{B}} = \begin{bmatrix} -0.2059+0.5914\jmath & -0.0909+0.5861\jmath \\ 0.4131+0.2651\jmath & -0.1960+0.4650\jmath \end{bmatrix},$$
$$\mathbf{H}_{\mathrm{F}} = \begin{bmatrix} -0.6362+0.1332\jmath & -0.1572+1.5538\jmath \\ 0.0196+0.4011\jmath & -0.3170-0.2303\jmath \end{bmatrix}.$$

The results are obtained by an exhaustive grid search over (θ_1, θ_2) for D-RIS and (α, ψ) for symmetric fully-connected BD-RIS. In this example, we observe that both singular values can be manipulated up to[4] $\pm 9\%$ by D-RIS (using 2 reconfigurable impedance) and $\pm 42\%$ by symmetric fully-connected BD-RIS (using 3 reconfigurable impedance). The extra performance gain comes purely from connecting both elements, and a larger gain is expected when an asymmetric reconfigurable component is available. This example shows that BD-RIS can provide a wider dynamic range of channel singular values and motivates further studies on channel shaping.

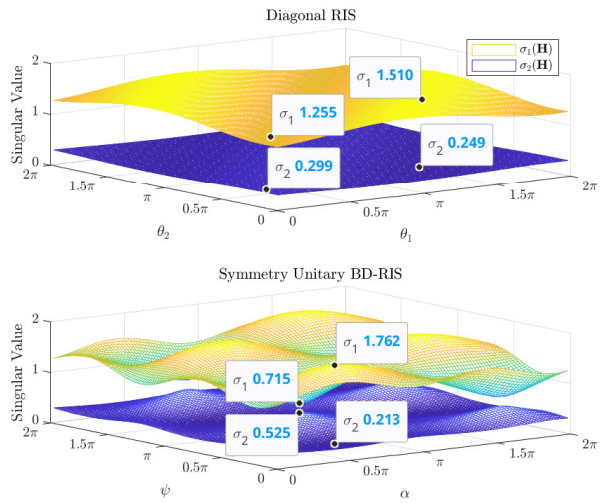


Fig. 1. $2 \times 2 \times 2$ channel singular value shaping by D-RIS and symmetric fully-connected BD-RIS when the direct channel is blocked. $\sigma_1(\mathbf{H})$ and $\sigma_2(\mathbf{H})$ refer to the most and least dominant singular values, respectively. Their maximum and minimum have been marked explicitly on the plot.

[4]The percentage for manipulating $\sigma_n(\mathbf{H})$ is calculated by $\eta_n^+ = \frac{\max\sigma_n(\mathbf{H}) - \mathrm{avg}\sigma_n(\mathbf{H})}{\mathrm{avg}\sigma_n(\mathbf{H})} \times 100\%$ and $\eta_n^- = \frac{\min\sigma_n(\mathbf{H}) - \mathrm{avg}\sigma_n(\mathbf{H})}{\mathrm{avg}\sigma_n(\mathbf{H})} \times 100\%$.

B. Pareto Frontier Characterization

We now characterize the Pareto frontier of singular values of a general $N_{\rm T}\times N_{\rm S}\times N_{\rm R}$ channel (2) by maximizing their weighted sum

$$\max_{\mathbf{\Theta}} \quad \sum_n \rho_n \sigma_n(\mathbf{H}) \tag{22a}$$

$$\text{s.t.} \quad \mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g = \mathbf{I}, \quad \forall g, \tag{22b}$$

where $n\in\mathcal{N}\triangleq\{1,...,N\}$, $N\triangleq\min(N_{\rm T},N_{\rm R})$ is the maximum channel rank, and ρ_n is the weight of the n-th singular value that can be positive, zero, or negative. Varying $\{\rho_n\}_{n\in\mathcal{N}}$ characterizes the Pareto frontier that encloses the entire singular value region. Thus, we claim problem (22) generalizes most singular value shaping problems. It can be solved optimally by Algorithm 1 with the Euclidean gradient given by Lemma 1.

Lemma 1. *The Euclidean gradient of* (22a) *with respect to BD-RIS group g is*

$$\frac{\partial\sum_n \rho_n\sigma_n(\mathbf{H})}{\partial \mathbf{\Theta}_g^*} = \mathbf{H}_{{\rm B},g}^{\mathsf{H}}\mathbf{U}\mathrm{diag}(\rho_1,...,\rho_N)\mathbf{V}^{\mathsf{H}}\mathbf{H}_{{\rm F},g}^{\mathsf{H}}, \tag{23}$$

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

Proof. Please refer to Appendix A. □

We now analyze the computational complexity of solving Pareto singular value problem (22) by Algorithm 1. To update each BD-RIS group, compact SVD of $\mathbf{H}$ requires $\mathcal{O}(NN_{\rm T}N_{\rm R})$, Euclidean gradient (23) requires $\mathcal{O}\big(LN(N_{\rm T}+N_{\rm R}+L+1)\big)$, Riemannian gradient translation (17) requires $\mathcal{O}(L^3)$, deviation parameter (18) and conjugate direction (19) together require $\mathcal{O}(L^2)$, and matrix exponential (20) requires $\mathcal{O}(L^3)$ operations [55]. The overall complexity is thus $\mathcal{O}\big(I_{\rm RCG}G\big(NN_{\rm T}N_{\rm R}+LN(N_{\rm T}+N_{\rm R}+L+1)+I_{\rm BLS}L^3\big)\big)$, where $I_{\rm RCG}$ and $I_{\rm BLS}$ are the number of iterations for geodesic RCG and backtracking line search (line 13 – 18 of Algorithm 1), respectively.

C. Some Analytical Bounds

We also provide some analytical bounds related to channel singular values, which help us to explore the channel shaping limits of passive RIS.

Definition 1 (Degree of freedom)**.** *DoF (a.k.a. multiplexing gain) refers to the maximum number of independent streams that can be transmitted in parallel over a MIMO channel, which is defined as*

$$d = \lim_{\rho\to\infty}\frac{\log\det(\mathbf{I}+\rho\mathbf{H}\mathbf{H}^{\mathsf{H}})}{\log\rho}, \tag{24}$$

where ρ is the Signal-to-Noise Ratio (SNR).

Proposition 1 (Degree of freedom)**.** *In point-to-point MIMO, BD-RIS cannot achieve a larger number of DoF than D-RIS.*

Proof. Please refer to Appendix B. □

Proposition 1 suggests that we cannot hope for a DoF gain (i.e., more parallel channels) by interconnecting the scattering elements. We thus focus on the question of reshaping the available channels through manipulation of the singular values.

Proposition 2 (Rank-deficient channel)**.** *If the minimum rank of forward and backward channels is k ($k\le N$), then for D-RIS or BD-RIS of arbitrary number of elements, the n-th singular value of the equivalent channel is bounded above and below respectively by*

$$\sigma_n(\mathbf{H}) \le \sigma_{n-k}(\mathbf{T}), \qquad \textit{if } n>k, \tag{25a}$$

$$\sigma_n(\mathbf{H}) \ge \sigma_n(\mathbf{T}), \qquad \textit{if } n<N-k+1, \tag{25b}$$

where $\mathbf{T}$ is an auxiliary matrix satisfying

$$\mathbf{T}\mathbf{T}^{\mathsf{H}} = \begin{cases} \mathbf{H}_{\rm D}(\mathbf{I}-\mathbf{V}_{\rm F}\mathbf{V}_{\rm F}^{\mathsf{H}})\mathbf{H}_{\rm D}^{\mathsf{H}}, & \textit{if } \mathrm{rank}(\mathbf{H}_{\rm F})=k, \\ \mathbf{H}_{\rm D}^{\mathsf{H}}(\mathbf{I}-\mathbf{U}_{\rm B}\mathbf{U}_{\rm B}^{\mathsf{H}})\mathbf{H}_{\rm D}, & \textit{if } \mathrm{rank}(\mathbf{H}_{\rm B})=k, \end{cases} \tag{26}$$

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

Proof. Please refer to Appendix C. □

Inequality (25a) states that if the forward and backward channels are at least rank k, then the n-th singular value of $\mathbf{H}$ can be enlarged to the $(n-k)$-th singular value of $\mathbf{T}$, or suppressed to the n-th singular value of $\mathbf{T}$. Moreover, the first k channel singular values are unbounded above[5] while the last k channel singular values can be suppressed to zero. An example is given in Corollary 2.1 for Line-of-Sight (LoS) channels (i.e., $k=1$)[6].

Corollary 2.1 (LoS channel)**.** *If at least one of forward and backward channels is LoS, then a D-RIS or BD-RIS can at most enlarge the n-th ($n\ge 2$) channel singular value to the $(n-1)$-th singular value of $\mathbf{T}$, or suppress the n-th channel singular value to the n-th singular value of $\mathbf{T}$. That is,*

$$\sigma_1(\mathbf{H})\ge\sigma_1(\mathbf{T})\ge\sigma_2(\mathbf{H})\ge...\ge\sigma_{N-1}(\mathbf{T})\ge\sigma_N(\mathbf{H})\ge\sigma_N(\mathbf{T}). \tag{27}$$

Proof. This is a direct result of (25) with $k=1$. □

We would like to highlight that Proposition 2 holds for both D- and BD-RIS with arbitrary number of elements, regardless of the presence of direct channel. The bounds are asymptotically tight when the number of scattering elements approaches infinity. On the other hand, we will show in Section VI that with a finite number of elements, BD-RIS can approach those bounds better than D-RIS. Proposition 2 evaluates the ultimate channel shaping limits of passive RIS and provides a selection guideline for scattering element number in a practical deployment scenario.

Proposition 3 (Blocked direct channel)**.** *If the direct link is blocked and the BD-RIS is fully-connected, then the channel singular values can be manipulated up to*

$$\mathrm{sv}(\mathbf{H}) = \mathrm{sv}(\mathbf{B}\mathbf{F}), \tag{28}$$

where $\mathbf{B}$ and $\mathbf{F}$ are arbitrary matrices with $\mathrm{sv}(\mathbf{B})=\mathrm{sv}(\mathbf{H}_{\rm B})$ and $\mathrm{sv}(\mathbf{F})=\mathrm{sv}(\mathbf{H}_{\rm F})$.

Proof. Please refer to Appendix D. □

[5]The energy conservation law $\sum_n\sigma_n^2(\mathbf{H})\le 1$ still has to be respected. This constraint is omitted in the following context for brevity.

[6]A similar eigenvalue result has been derived in [56] only for D-RIS.

Proposition 3 states that if the direct channel is blocked and the BD-RIS is fully-connected, the only singular value bounds on the equivalent channel are the singular value bounds on the product of unitary-transformed forward and backward channels. For example, consider a $2\times2\times2$ setup with

$$\mathbf{H}_{\mathrm{B}}=\begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix},\quad \mathbf{H}_{\mathrm{F}}=\begin{bmatrix}1 & 0\\ 0 & 4\end{bmatrix}.$$

It is apparent that any D-RIS can only achieve singular values $\mathrm{sv}(\mathbf{H})=[8,3]^{\mathsf{T}}$, but a fully connected BD-RIS with $\mathbf{\Theta}=\left[\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right]$ can achieve $\mathrm{sv}(\mathbf{H})=[12,2]^{\mathsf{T}}$. That is, fully-connected BD-RIS applies right unitary transformation on $\mathbf{H}_{\mathrm{B}}$ and left unitary transformation on $\mathbf{H}_{\mathrm{F}}$, widening the singular values to $8\leq\sigma_1(\mathbf{H})\leq12$ and $2\leq\sigma_2(\mathbf{H})\leq3$ in this case.

Proposition 3 transforms the channel shaping capability question to a mathematical question: How the singular values of matrix product are bounded by the singular values of its individual factors. Let $\bar{N}=\max(N_{\mathrm{T}},N_{\mathrm{S}},N_{\mathrm{R}})$ and $\sigma_n(\mathbf{H})=\sigma_n(\mathbf{H}_{\mathrm{F}})=\sigma_n(\mathbf{H}_{\mathrm{B}})=0$ for $N<n\leq\bar{N}$. We have the following corollaries.

Corollary 3.1 (Generic singular value bounds)**.**

$$\prod_{k\in K}\sigma_k(\mathbf{H})\leq\prod_{i\in I}\sigma_i(\mathbf{H}_{\mathrm{B}})\prod_{j\in J}\sigma_j(\mathbf{H}_{\mathrm{F}}),\tag{29}$$

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

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

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

Proof. Please refer to [57, Theorem 8]. □

Corollary (3.1) is by far the most comprehensive channel singular value bound attainable by passive RIS when the direct channel is blocked. It is also recognized as a variation of Horn's inequality [58]. We observe that the number of admissible triples (and thus bounds) grows exponentially with $\bar{N}$. For example, the number of inequalities described by (29) grows from 12 to 2062 when $\bar{N}$ increases from 3 to 7. This renders the analytic approach computationally expensive for large-scale MIMO systems. Next, we showcase some useful inequalities enclosed by (29). Readers are referred to [59, Chapter 16, 24] for more examples.

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

$$\sigma_1(\mathbf{H})\leq\sigma_1(\mathbf{H}_{\mathrm{B}})\sigma_1(\mathbf{H}_{\mathrm{F}}).\tag{30}$$

Proof. This is a direct result of (29) with $r=1$. □

Corollary 3.3 (Lower bound on the smallest singular value)**.**

$$\sigma_{\bar{N}}(\mathbf{H})\geq\sigma_{\bar{N}}(\mathbf{H}_{\mathrm{B}})\sigma_{\bar{N}}(\mathbf{H}_{\mathrm{F}}).\tag{31}$$

Proof. This can be deducted from (29) with $r_1=\bar{N}-1$ and $r_2=\bar{N}$. □

Corollary 3.4 (Upper bound on the product of first k singular values)**.**

$$\prod_{n=1}^{k}\sigma_n(\mathbf{H})\leq\prod_{n=1}^{k}\sigma_n(\mathbf{H}_{\mathrm{B}})\prod_{n=1}^{k}\sigma_n(\mathbf{H}_{\mathrm{F}}).\tag{32}$$

Proof. This is a direct result of (29) with $r=k$. □

Corollary 3.5 (Lower bound on the product of last k singular values)**.**

$$\prod_{n=\bar{N}}^{\bar{N}-k+1}\sigma_n(\mathbf{H})\geq\prod_{n=\bar{N}}^{\bar{N}-k+1}\sigma_n(\mathbf{H}_{\mathrm{B}})\prod_{n=\bar{N}}^{\bar{N}-k+1}\sigma_n(\mathbf{H}_{\mathrm{F}}).\tag{33}$$

Proof. This can be deducted from (29) with $r_1=\bar{N}-k$ and $r_2=\bar{N}$. □

Corollaries 3.3 and 3.5 are less informative when $\bar{N}\neq N$ (i.e., unequal number of transmit and receive antennas and scattering elements) as the lower bounds would coincide at zero.

Corollary 3.6 (Upper bound on the channel power gain)**.** *The channel power gain is upper bounded by the sum of sorted element-wise product of squared singular values of backward and forward channels*

$$\|\mathbf{H}\|_{\mathrm{F}}^2=\sum_{n=1}^{N}\sigma_n^2(\mathbf{H})\leq\sum_{n=1}^{N}\sigma_n^2(\mathbf{H}_{\mathrm{B}})\sigma_n^2(\mathbf{H}_{\mathrm{F}}).\tag{34}$$

Proof. Please refer to [59, Inequality 24.4.7]. □

To achieve equalities in Corollaries 3.2 – 3.6, the RIS needs to completely align the singular spaces of $\mathbf{H}_{\mathrm{B}}$ and $\mathbf{H}_{\mathrm{F}}$. The resulting scattering matrix is generally required to be unitary

$$\mathbf{\Theta}^{\star}=\mathbf{V}_{\mathrm{B}}\mathbf{U}_{\mathrm{F}}^{\mathsf{H}},\tag{35}$$

which can be concluded from (53) and (54) in Appendix D. We notice that D-RIS can attain those equalities if and only if $\mathbf{H}_{\mathrm{B}}$ and $\mathbf{H}_{\mathrm{F}}$ are both rank-1. In such case, the equivalent channel reduces to $\mathbf{H}=\sigma_{\mathrm{B}}\sigma_{\mathrm{F}}\mathbf{u}_{\mathrm{B}}\mathbf{v}_{\mathrm{B}}^{\mathsf{H}}\mathbf{\Theta}\mathbf{u}_{\mathrm{F}}\mathbf{v}_{\mathrm{F}}^{\mathsf{H}}$ and the RIS only needs to align $\mathbf{v}_{\mathrm{B}}^{\mathsf{H}}$ and $\mathbf{u}_{\mathrm{F}}$ by

$$\mathbf{\Theta}^{\star}=\mathbf{v}_{\mathrm{B}}\mathbf{u}_{\mathrm{F}}^{\mathsf{H}}\odot\mathbf{I},\tag{36}$$

which becomes a special case of (6). However, in general D-RIS is not sufficient to achieve the upper bound of channel power gain. This again highlights the benefit of BD-RIS in reshaping available channels. Finally, when $\mathbf{H}_{\mathrm{B}}$ and $\mathbf{H}_{\mathrm{F}}$ both follow Rayleigh fading, the maximum expected channel power gain $\mathbb{E}\{\|\mathbf{H}\|_{\mathrm{F}}^2\}$ can be numerically evaluated as

$$\sum_{n=1}^{N}\int_0^{\infty}f_{\lambda_n^{\min(N_{\mathrm{R}},N_{\mathrm{S}})}}(x_n)dx_n\int_0^{\infty}f_{\lambda_n^{\min(N_{\mathrm{S}},N_{\mathrm{T}})}}(x_n)dx_n,\tag{37}$$

where λ_n^K is the n-th eigenvalue of the complex $K\times K$ Wishart matrix with probability density function $f_{\lambda_n^K}(x_n)$ given by [60, Equation 51]. We notice (37) is a generalization of SISO power gain [30, Equation 58] to MIMO, which unveils how the channel power gain scales with the number of transmit and receive antennas and BD-RIS elements under double Rayleigh fading.

Proposition 3 and Corollaries 3.1 – 3.6 provide a comprehensive answer to singular value bounds and channel power gain

limits for blocked direct channel with fully-connected BD-RIS. Tight bounds are generally inapplicable since the direct-indirect channels and backward-forward channels cannot be completely aligned at the same time. In such case, we can invoke optimization approaches from a singular value perspective (Section IV-B) or a power gain perspective (Section V-A).

V. Power Gain and Achievable Rate Maximization

A. Channel Power Gain

The MIMO channel power gain maximization problem is formulated with respect to the BD-RIS scattering matrix

$$\max_{\mathbf{\Theta}} \quad \|\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}\|_{\mathrm{F}}^2 \tag{38a}$$

$$\text{s.t.} \quad \mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I}, \quad \forall g, \tag{38b}$$

which generalizes the case of SISO [30], MISO [33], [40], single-stream MIMO [31], [45], and direct link-blocked MIMO with fully-connected BD-RIS (35). The key of solving (38) is to balance the additive and multiplicative space alignments.

Remark 1. *Interestingly, in terms of maximizing the inner product* $\langle \mathbf{H}_{\mathrm{D}},\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}\rangle$, (38) *is reminiscent of the weighted orthogonal Procrustes problem [61]*

$$\min_{\mathbf{\Theta}} \quad \|\mathbf{H}_{\mathrm{D}}-\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}\|_{\mathrm{F}}^2 \tag{39a}$$

$$\text{s.t.} \quad \mathbf{\Theta}^{\mathsf{H}}\mathbf{\Theta}=\mathbf{I}, \tag{39b}$$

which relaxes the block-unitary constraint (39b) *to unitary but still has no trivial solution. One lossy transformation exploits the Moore-Penrose inverse and moves* $\mathbf{\Theta}$ *to one side of the product [62], formulating two standard orthogonal Procrustes problems*

$$\min_{\mathbf{\Theta}} \quad \|\mathbf{H}_{\mathrm{B}}^{\dagger}\mathbf{H}_{\mathrm{D}}-\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}\|_{\mathrm{F}}^2 \text{ or } \|\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\dagger}-\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\|_{\mathrm{F}}^2 \tag{40a}$$

$$\text{s.t.} \quad \mathbf{\Theta}^{\mathsf{H}}\mathbf{\Theta}=\mathbf{I}, \tag{40b}$$

which have global optimal solutions

$$\mathbf{\Theta}=\mathbf{U}\mathbf{V}^{\mathsf{H}}, \tag{41}$$

where $\mathbf{U}$ *and* $\mathbf{V}$ *are respectively the left and right compact singular matrices of* $\mathbf{H}_{\mathrm{B}}^{\dagger}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\mathsf{H}}$ *or* $\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\dagger}$ *[63].*

It is worth noting that (35) and (41) are valid fully-connected BD-RIS solutions to (38) when the direct link is blocked and present, respectively. However, the latter is neither optimal nor a generalization of the former due to the lossy transformation.

Inspired by [64], we propose an optimal solution to problem (38) with arbitrary group size. The idea is to successively approximate the quadratic objective (38a) by local Taylor expansions and solve each step in closed form by group-wise SVD.

Proposition 4. *Starting from any feasible* $\mathbf{\Theta}^{(0)}$*, the sequence*

$$\mathbf{\Theta}_g^{(r+1)}=\mathbf{U}_g^{(r)}\mathbf{V}_g^{(r)}, \quad \forall g. \tag{42}$$

converges to a stationary point of (38)*, where* $\mathbf{U}_g^{(r)}$ *and* $\mathbf{V}_g^{(r)}$ *are the left and right compact singular matrices of*

$$\mathbf{M}_g^{(r)}=\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\Big(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathrm{diag}\big(\mathbf{\Theta}_{[1:g-1]}^{(r+1)},\mathbf{\Theta}_{[g:G]}^{(r)}\big)\mathbf{H}_{\mathrm{F}}\Big)\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}} \tag{43}$$

Proof. Please refer to Appendix F. □

We now analyze the computational complexity of solving channel gain maximization problem (38) by Proposition 4. To update each BD-RIS group, matrix multiplication (43) requires $\mathcal{O}\big(N_{\mathrm{T}}N_{\mathrm{R}}+(G+1)(NL^2+N_{\mathrm{T}}N_{\mathrm{R}}L)\big)$ operations and its compact SVD requires $\mathcal{O}(L^3)$ operations. The overall complexity is thus $\mathcal{O}\big(I_{\mathrm{SAA}}G\big(N_{\mathrm{T}}N_{\mathrm{R}}+(G+1)(NL^2+N_{\mathrm{T}}N_{\mathrm{R}}L)+L^3\big)\big)$, where I_{SAA} is the number iterations for successive affine approximation.

B. Achievable Rate Maximization

We aim to maximize the achievable rate of the BD-RIS-aided MIMO system by jointly optimizing the active and passive beamforming

$$\max_{\mathbf{W},\mathbf{\Theta}} \quad R=\log\det\left(\mathbf{I}+\frac{\mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{H}\mathbf{W}}{\eta}\right) \tag{44a}$$

$$\text{s.t.} \quad \|\mathbf{W}\|_{\mathrm{F}}^2\leq P, \tag{44b}$$

$$\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I}, \quad \forall g, \tag{44c}$$

where $\mathbf{W}$ is the transmit precoder, R is the achievable rate, η is the average noise power, and P is the transmit power constraint. Problem (44) is non-convex due to the block-unitary constraint (44c) and the coupling between variables. We propose a local-optimal approach via AO and a low-complexity approach based on channel shaping.

1) Alternating Optimization: This approach updates $\mathbf{\Theta}$ and $\mathbf{W}$ iteratively until convergence. For a given $\mathbf{W}$, the passive beamforming subproblem is

$$\max_{\mathbf{\Theta}} \quad \log\det\left(\mathbf{I}+\frac{\mathbf{H}\mathbf{Q}\mathbf{H}^{\mathsf{H}}}{\eta}\right) \tag{45a}$$

$$\text{s.t.} \quad \mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I}, \quad \forall g, \tag{45b}$$

where $\mathbf{Q}\triangleq\mathbf{W}\mathbf{W}^{\mathsf{H}}$ is the transmit covariance matrix. Problem (45) can be solved optimally by Algorithm 1 with the Euclidean gradient given by Lemma 2.

Lemma 2. *The Euclidean gradient of* (45a) *with respect to BD-RIS block* g *is*

$$\frac{\partial R}{\partial \mathbf{\Theta}_g^*}=\frac{1}{\eta}\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\left(\mathbf{I}+\frac{\mathbf{H}\mathbf{Q}\mathbf{H}^{\mathsf{H}}}{\eta}\right)^{-1}\mathbf{H}\mathbf{Q}\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}. \tag{46}$$

Proof. Please refer to Appendix E. □

For a given $\mathbf{\Theta}$, the global optimal transmit precoder is given by eigenmode transmission [65]

$$\mathbf{W}^{\star}=\mathbf{V}\mathrm{diag}(\mathbf{s}^{\star})^{1/2}, \tag{47}$$

where $\mathbf{V}$ is the right singular matrix of the equivalent channel and $\mathbf{s}^{\star}$ is the optimal water-filling power allocation obtainable by the iterative method [66].

The AO algorithm is guaranteed to converge to local-optimal points of problem (44) since each subproblem is solved optimally and the objective is bounded above. Similar to the analysis in Section IV-B, the computational complexity of solving subproblem (45) by geodesic RCG is $\mathcal{O}\big(I_{\mathrm{RCG}}G(NL^2+LN_{\mathrm{T}}N_{\mathrm{R}}+N_{\mathrm{T}}^2N_{\mathrm{R}}+N_{\mathrm{T}}N_{\mathrm{R}}^2+N_{\mathrm{R}}^3+I_{\mathrm{BLS}}L^3)\big)$. On the

TABLE I
AVERAGE PERFORMANCE OF GEODESIC AND NON-GEODESIC RCG ALGORITHMS ON PROBLEM (22)

RCG path	$N_{\mathrm{S}}=16$			$N_{\mathrm{S}}=256$		
	Objective	Iterations	Time [s]	Objective	Iterations	Time [s]
Geodesic	4.359×10^{-3}	11.59	1.839×10^{-2}	1.163×10^{-2}	25.58	3.461
Non-geodesic	4.329×10^{-3}	30.92	5.743×10^{-2}	1.116×10^{-2}	61.40	13.50

other hand, the complexity of solving active beamforming subproblem by (47) is $\mathcal{O}\big(NN_{\mathrm{T}}N_{\mathrm{R}}\big)$. The overall complexity is thus $\mathcal{O}\big(I_{\mathrm{AO}}\big(I_{\mathrm{RCG}}G(NL^2+LN_{\mathrm{T}}N_{\mathrm{R}}+N_{\mathrm{T}}^2N_{\mathrm{R}}+N_{\mathrm{T}}N_{\mathrm{R}}^2+N_{\mathrm{R}}^3+I_{\mathrm{BLS}}L^3)+NN_{\mathrm{T}}N_{\mathrm{R}}\big)\big)$, where I_{AO} is the number of iterations for AO.

2) Low-Complexity Solution: We finally propose a suboptimal two-stage solution to problem (44) that decouples the joint RIS-transceiver design. The idea is to first consider channel shaping and replace the rate maximization subproblem (45) by channel power gain maximization problem (38), then proceed to conventional eigenmode transmission (47). Both steps are solved in closed form and the computational complexity is $\mathcal{O}\big(I_{\mathrm{SAA}}G\big(N_{\mathrm{T}}N_{\mathrm{R}}+(G+1)(NL^2+N_{\mathrm{T}}N_{\mathrm{R}}L)+L^3\big)+NN_{\mathrm{T}}N_{\mathrm{R}}\big)$. While suboptimal, this shaping-inspired solution avoids outer iterations and implements inner iterations more efficiently.

VI. SIMULATION RESULTS

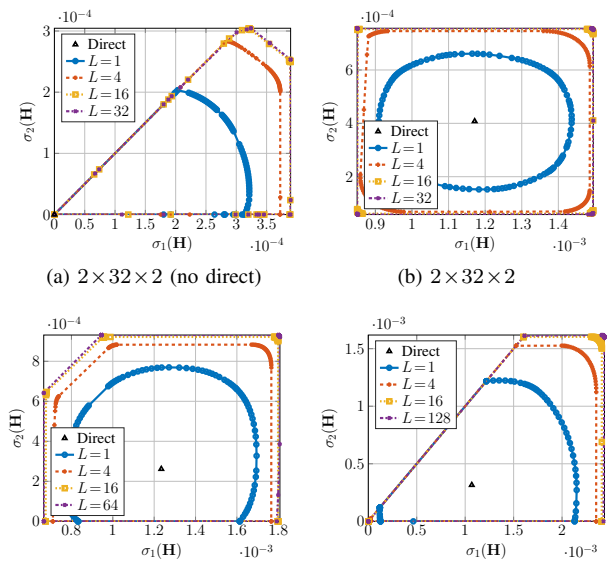
In this section, we provide numerical results to evaluate the proposed BD-RIS designs.[7] Consider a distance-dependent path loss model $\Lambda(d)=\Lambda_0 d^{-\gamma}$ where Λ_0 is the reference path loss at distance $1\,\mathrm{m}$, d is the propagation distance, and γ is the path loss exponent. 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. Rayleigh fading (i.e., $\kappa=0$) is assumed for all channels unless otherwise specified.

A. Algorithm Evaluation

We first compare in Table I the geodesic and non-geodesic RCG algorithm on problem (22) in an $N_{\mathrm{T}}=N_{\mathrm{R}}=4$ system with BD-RIS group size $L=4$. The statistics are averaged over 100 independent runs. It is observed that the geodesic RCG method achieves a slightly higher objective value with significantly (up to $3\times$) lower number of iterations and shorter (up to $4\times$) computational time than the non-geodesic method. The results demonstrate the efficiency of the proposed geodesic RCG algorithm especially in large-scale BD-RIS design problems.

B. Channel Singular Values Redistribution

[7]The simulation code is publicly available at https://github.com/snowztail/channel-shaping/.

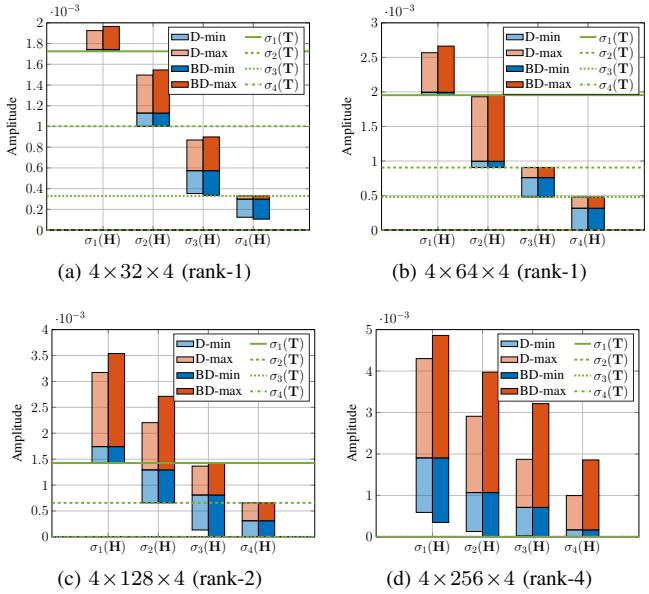


(a) $2\times32\times2$ (no direct) (b) $2\times32\times2$

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

Fig. 2. Pareto frontiers of singular values of an $N_{\mathrm{T}}=N_{\mathrm{R}}=2$ channel reshaped by a RIS.

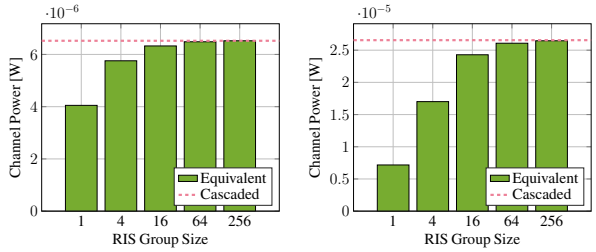
1) Pareto Frontier: Fig. 2 shows the Pareto singular values of an $N_{\mathrm{T}}=N_{\mathrm{R}}=2$ MIMO reshaped by a RIS. When the direct link is blocked, the achievable regions in Fig. 2(a) are shaped like pizza slices. This is because $\sigma_1(\mathbf{H})\geq\sigma_2(\mathbf{H})\geq0$ and there exists a trade-off between the alignment of two spaces. We observe that the smallest singular value can be enhanced up to 2×10^{-4} by D-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 improves the dynamic range of $\sigma_1(\mathbf{H})$ and $\sigma_2(\mathbf{H})$ by $22\,\%$ and $38\,\%$, but also provides a better trade-off in manipulating both singular values. It suggests the design freedom from larger group size allows better alignment of multiple spaces simultaneously. The singular value region also enlarges as the number of scattering elements N_{S} increases. In particular, Fig. 2(d) shows that the equivalent channel can be completely nulled (corresponding to the origin) by a 128-element BD-RIS thanks to its superior channel shaping capability, but not by a diagonal one. Those results demonstrate the superior channel shaping capability of BD-RIS and emphasizes the importance of reconfigurable



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

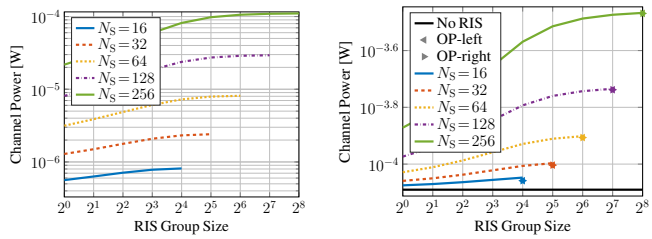
(c) $4\times128\times4$ (rank-2) (d) $4\times256\times4$ (rank-4)

Fig. 3. Achievable channel singular values: analytical bounds (green lines) and numerical optimization results (blue and red bars). The intersections of the blue and red bars denote the singular values of the direct channel. The blue (resp. red) bars are obtained by solving problem (22) with $\rho_n=-1$ (resp. $+1$) and $\rho_{n'}=0$, $\forall n'\neq n$. 'D' means D-RIS and 'BD' refers to fully-connected BD-RIS. 'rank-k' refers to the rank of the forward channel.



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

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



(a) $16\times N_S\times16$ (no direct) (b) $16\times N_S\times16$

Fig. 5. Average maximum channel power gain versus RIS configuration. 'OP-left' and 'OP-right' refer to the suboptimal solutions to problem (38) by lossy transformation (40) where $\mathbf{\Theta}$ is to the left and right of the product, respectively.

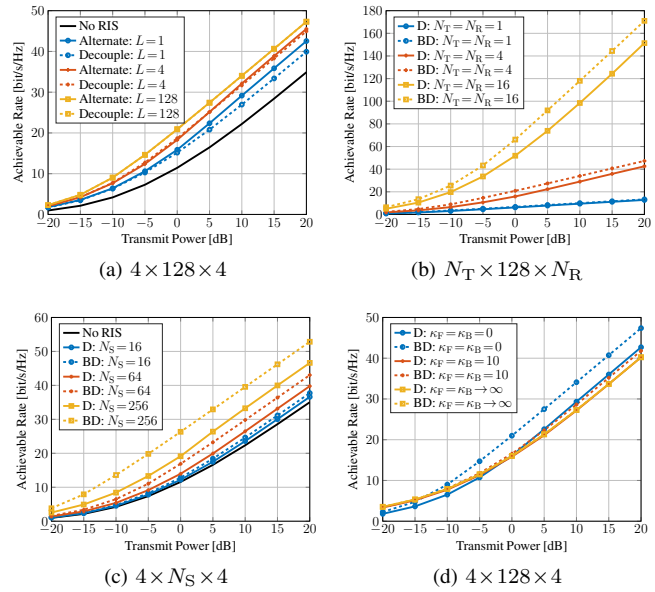
inter-connection between elements.

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 (22) with $\rho_n=\pm1$ and $\rho_{n'}=0$, $\forall n'\neq n$. Here we assume 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 2.1, 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 D-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. 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 (26) 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 D-RIS scales with forward and backward channel ranks.

Fig. 4 compares the analytical channel power bound in Corollary 3.6 and the numerical results obtained by solving problem (38) when the direct link is blocked. Here, a fully-connected BD-RIS can attain the upper bound either in closed form (35) or via optimization approach (42). For the SISO case in Fig. 4(a), the maximum channel power gain is approximately 4×10^{-6} by D-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 [30]. Interestingly, the gain surges to $270\,\%$ in $N_T=N_R=4$ MIMO as shown in Fig. 4(b). This is because space alignment boils down to phase matching in SISO such that both triangular and Cauchy-Schwarz inequalities in [30, (50)] can be simultaneously tight regardless of the group size. That is, D-RIS is sufficient for space alignment in SISO while the $62.5\,\%$ gain from BD-RIS comes purely from channel rearrangement. Now consider a D-RIS in MIMO. Each element can only apply a common phase shift to the associated rank-1 $N_R\times N_T$ indirect channel. Therefore, perfect space alignment of indirect channels through different elements is generally impossible. It means the disadvantage of D-RIS in space alignment and channel rearrangement scales with MIMO dimensions. We thus conclude that the power gain of BD-RIS scales with group size and MIMO dimensions.

C. Power Gain and Achievable Rate Maximization

We first focus on channel power gain maximization problem (38). Fig. 5 shows the maximum channel power under different RIS configurations. An interesting observation is that the relative power gain of BD-RIS over D-RIS is even larger with direct link. For example, a 64-element fully BD-RIS can almost provide the same channel power gain as a 256-element D-RIS in Fig. 5(b), but not in Fig. 5(a). This is because the RIS needs to balance the multiplicative forward-backward combining and the additive direct-indirect combining, such that the space alignment advantage of BD-RIS becomes more pronounced. We also notice that the suboptimal solutions (41) for fully-connected BD-RIS by lossy transformation (40) are very close to optimal especially for a large N_S.



(a) $4\times128\times4$ (b) $N_{\rm T}\times128\times N_{\rm R}$

(c) $4\times N_{\rm S}\times4$ (d) $4\times128\times4$

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

Fig. 6 presents the achievable rate under different MIMO and RIS configurations. At a transmit power $P=10$dB, Fig. 6(a) shows that introducing a 128-element D-RIS to $N_{\rm T}=N_{\rm R}=4$ MIMO can improve the achievable rate from 22.2 bps/Hz to 29.2 bps/Hz ($+31.5\%$). A BD-RIS of group size 4 and 128 can further elevate those to 32.1 bps/Hz ($+44.6\%$) and 34 bps/Hz ($+53.2\%$), respectively. An interesting observation is that the rate gap between the optimal AO approach (45) – (47) and the shaping-inspired solution (42), (47) narrows at larger group size and completely vanishes for a fully-connected BD-RIS. This implies that joint RIS-transceiver designs can be decoupled by first shaping the wireless channel and then optimizing the precoder and combiner, which significantly simplifies the process at marginal performance cost. Figs. 6(b) and 6(c) also show that both *absolute and relative* rate gains of BD-RIS versus D-RIS increases with the number of transmit and receive antennas and scattering elements, especially at high SNR. For $N_{\rm S}=128$ and $P=20$dB, the achievable rate ratio of BD-RIS over D-RIS is 1.04, 1.11, and 1.13 for $N_{\rm T}=N_{\rm R}=1$, 4, and 16, respectively. For $N_{\rm T}=N_{\rm R}=4$ and $P=20$ dB, this ratio amounts to 1.03, 1.08, and 1.13 for $N_{\rm S}=16$, 64, and 256, respectively. Those observations align with the power gain results in Fig. 5 and highlight the rate benefits of BD-RIS over D-RIS in large-scale MIMO systems. In the low power regime (-20 to -10 dB), we also notice that the slope of the achievable rate of BD-RIS is steeper than that of D-RIS. That is, BD-RIS can help to activate more streams and better approach the DoF at a low transmit SNR. This is particularly visible in Fig. 6(c) where the topmost curve is almost a linear function of the transmit power. This is expected from the shaping results in Fig. 2 that BD-RIS can significantly enlarge all channel singular values for higher receive SNR. Finally, Fig. 6(d) shows that the gap between D- and BD-RIS narrows as the Rician K-factor increases and becomes indistinguishable in LoS environment. The observation is expected from previous studies [30], [31], [47] and aligns with Corollary 2.1, which suggests that the BD-RIS should be deployed in rich-scattering environments to exploit its channel shaping potential.

VII. 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 D-RIS. This translates to a wider dynamic range of and better tradeoff between 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 for rank-deficient forward/backward channel and blocked direct channel. Specifically, the former is done by proposing an efficient RCG algorithm for BD-RIS optimization problems, which converges much faster 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 space alignment and channel rearrangement capability, which scales with the number of elements, group size, MIMO dimensions, and channel diversity.

It would be interesting to investigate the impact of channel shaping through RIS on other wireless applications such as sensing and far-field power transfer.

Appendix

A. Proof of Lemma 1

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

$$\sigma_n=\mathbf{u}_n^{\rm H}\mathbf{H}\mathbf{v}_n=\mathbf{u}_n^{\rm T}\mathbf{H}^*\mathbf{v}_n^*, \tag{48}$$

whose differential with respect to $\mathbf{\Theta}_g^*$ is

$$\begin{aligned}\partial\sigma_n&=\partial\mathbf{u}_n^{\rm T}\underbrace{\mathbf{H}^*\mathbf{v}_n^*}_{\sum_m\mathbf{u}_m^*\sigma_m\mathbf{v}_m^{\rm T}\mathbf{v}_n}+\mathbf{u}_n^{\rm T}\cdot\partial\mathbf{H}^*\cdot\mathbf{v}_n^*+\underbrace{\mathbf{u}_n^{\rm T}\mathbf{H}^*}_{\mathbf{u}_n^{\rm T}\sum_m\mathbf{u}_m^*\sigma_m\mathbf{v}_m^{\rm T}}\partial\mathbf{v}_n^*\\&=\underbrace{\partial\mathbf{u}_n^{\rm T}\mathbf{u}_n^*}_{\partial1=0}\cdot\sigma_n+\mathbf{u}_n^{\rm T}\cdot\partial\mathbf{H}^*\cdot\mathbf{v}_n^*+\sigma_n\cdot\underbrace{\mathbf{v}_n^{\rm T}\partial\mathbf{v}_n^*}_{\partial1=0}\\&=\mathbf{u}_n^{\rm T}\mathbf{H}_{{\rm B},g}^*\cdot\partial\mathbf{\Theta}_g^*\cdot\mathbf{H}_{{\rm F},g}^*\mathbf{v}_n^*\\&={\rm tr}(\mathbf{H}_{{\rm F},g}^*\mathbf{v}_n^*\mathbf{u}_n^{\rm T}\mathbf{H}_{{\rm B},g}^*\cdot\partial\mathbf{\Theta}_g^*).\end{aligned}$$

According to [67], the corresponding complex derivative is

$$\frac{\partial\sigma_n}{\partial\mathbf{\Theta}_g^*}=\mathbf{H}_{{\rm B},g}^{\rm H}\mathbf{u}_n\mathbf{v}_n^{\rm H}\mathbf{H}_{{\rm F},g}^{\rm H}. \tag{49}$$

A linear combination of (49) yields (23).

B. Proof of Proposition 1

The scattering matrix of BD-RIS can be decomposed as

$$\mathbf{\Theta}=\mathbf{L}\mathbf{\Theta}_{\mathrm{D}}\mathbf{R}^{\mathsf{H}}, \tag{50}$$

where $\mathbf{\Theta}_{\mathrm{D}} \in \mathbb{U}^{N_{\mathrm{S}}\times N_{\mathrm{S}}}$ corresponds to D-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 $\mathbf{\Theta}_{\mathrm{D}}$ such that

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

The same result holds if the direct link is present.

C. Proof of Proposition 2

We consider rank-k forward channel and the proof follows similarly for rank-k backward channel. Let $\mathbf{H}_{\mathrm{F}}=\mathbf{U}_{\mathrm{F}}\mathbf{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}}$ 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

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

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 $\mathbf{Z}\triangleq\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{U}_{\mathrm{F}}\mathbf{\Sigma}_{\mathrm{F}}+\mathbf{H}_{\mathrm{D}}\mathbf{V}_{\mathrm{F}}\in\mathbb{C}^{N_{\mathrm{R}}\times k}$. That is to say, $\mathbf{G}$ can be expressed as a Hermitian matrix plus k rank-1 perturbations. According to the Cauchy interlacing formula [63], the n-th eigenvalue of $\mathbf{G}$ is bounded by

$$\lambda_n(\mathbf{G})\le\lambda_{n-k}(\mathbf{Y}),\qquad \text{if } n>k, \tag{51}$$

$$\lambda_n(\mathbf{G})\ge\lambda_n(\mathbf{Y}),\qquad \text{if } n<N-k+1. \tag{52}$$

Since $\mathbf{Y}=\mathbf{T}\mathbf{T}^{\mathsf{H}}$ is positive semi-definite, taking the square roots of (51) and (52) gives (25a) and (25b).

D. Proof of Proposition 3

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

$$\mathbf{\Theta}=\mathbf{V}_{\mathrm{B}}\mathbf{X}\mathbf{U}_{\mathrm{F}}^{\mathsf{H}}, \tag{53}$$

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 $\mathbf{X}$

$$\mathbf{H}=\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}=\mathbf{U}_{\mathrm{B}}\mathbf{\Sigma}_{\mathrm{B}}\mathbf{X}\mathbf{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{\mathsf{H}}. \tag{54}$$

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

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

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 with respect to $\mathbf{\Theta}_g^*$ is [67]

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

and the corresponding complex derivative is (46).

F. Proof of Proposition 4

The differential of (38a) with respect to $\mathbf{\Theta}_g^*$ is

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

and the corresponding complex derivative is

$$\frac{\partial\|\mathbf{H}\|_{\mathrm{F}}^2}{\partial\mathbf{\Theta}_g^*}=\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}})\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}=\mathbf{M}_g. \tag{55}$$

First, we approximate the quadratic objective (38a) by its local Taylor expansion

$$\max_{\mathbf{\Theta}}\quad\sum_g 2\Re\{\mathrm{tr}(\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{M}_g)\} \tag{56a}$$

$$\text{s.t.}\quad \mathbf{\Theta}_g^{\mathsf{H}}\mathbf{\Theta}_g=\mathbf{I},\quad\forall g. \tag{56b}$$

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

$$\Re\{\mathrm{tr}(\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{M}_g)\}=\Re\{\mathrm{tr}(\mathbf{\Sigma}_g\mathbf{V}_g^{\mathsf{H}}\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{U}_g)\}\le\mathrm{tr}(\mathbf{\Sigma}_g). \tag{57}$$

The upper bound is tight when $\mathbf{V}_g^{\mathsf{H}}\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{U}_g=\mathbf{I}$, which implies the optimal solution of (56) is $\tilde{\mathbf{\Theta}}_g=\mathbf{U}_g\mathbf{V}_g^{\mathsf{H}}$, $\forall g$.

Next, we prove that solving (56) successively does not decrease (38a). Since $\tilde{\mathbf{\Theta}}$ optimal for problem (56), we have $\sum_g 2\Re\{\mathrm{tr}(\tilde{\mathbf{\Theta}}_g^{\mathsf{H}}\mathbf{M}_g)\}\ge\sum_g 2\Re\{\mathrm{tr}(\mathbf{\Theta}_g^{\mathsf{H}}\mathbf{M}_g)\}$ which is explicitly expressed by (59). On the other hand, expanding $\|\sum_g\mathbf{H}_{\mathrm{B},g}\tilde{\mathbf{\Theta}}_g\mathbf{H}_{\mathrm{F},g}-\sum_g\mathbf{H}_{\mathrm{B},g}\mathbf{\Theta}_g\mathbf{H}_{\mathrm{F},g}\|_{\mathrm{F}}^2\ge 0$ gives (60). Adding (59) and (60), we have

$$\begin{aligned}&2\Re\Big\{\mathrm{tr}(\tilde{\mathbf{\Theta}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\mathsf{H}})\Big\}+\mathrm{tr}(\mathbf{H}_{\mathrm{F}}^{\mathsf{H}}\tilde{\mathbf{\Theta}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}\tilde{\mathbf{\Theta}}\mathbf{H}_{\mathrm{F}})\\&\ge 2\Re\Big\{\mathrm{tr}(\mathbf{\Theta}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F}}^{\mathsf{H}})\Big\}+\mathrm{tr}(\mathbf{H}_{\mathrm{F}}^{\mathsf{H}}\mathbf{\Theta}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}^{\mathsf{H}}\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}\mathbf{H}_{\mathrm{F}}),\end{aligned} \tag{58}$$

which suggests that updating $\tilde{\mathbf{\Theta}}$ does not decrease (38a).

Finally, we prove that the converging point of (56), denoted by $\tilde{\mathbf{\Theta}}^{?}$, is a stationary point of (38). The Karush-Kuhn-Tucker (KKT) conditions of (38) and (56) are equivalent in terms of primal/dual feasibility and complementary slackness, while the stationary conditions are respectively, $\forall g$,

$$\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}^{\star}\mathbf{H}_{\mathrm{F}})\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}-\mathbf{\Theta}_g^{\star}\mathbf{\Lambda}_g^{\mathsf{H}}=0, \tag{61}$$

$$\mathbf{M}_g-\mathbf{\Theta}_g^{\star}\mathbf{\Lambda}_g^{\mathsf{H}}=0. \tag{62}$$

On convergence, (62) becomes $\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}(\mathbf{H}_{\mathrm{D}}+\mathbf{H}_{\mathrm{B}}\mathbf{\Theta}^{?}\mathbf{H}_{\mathrm{F}})\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}}-\mathbf{\Theta}_g^{?}\mathbf{\Lambda}_g^{\mathsf{H}}=0$ and reduces to (61). The proof is thus completed.

$$2\Re\Big\{\sum_{g}\mathrm{tr}(\tilde{\boldsymbol{\Theta}}_{g}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}})+\sum_{g_1,g_2}\mathrm{tr}(\tilde{\boldsymbol{\Theta}}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2}\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}})\Big\}\geq 2\Re\Big\{\sum_{g}\mathrm{tr}(\boldsymbol{\Theta}_{g}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g}^{\mathsf{H}}\mathbf{H}_{\mathrm{D}}\mathbf{H}_{\mathrm{F},g}^{\mathsf{H}})+\sum_{g_1,g_2}\mathrm{tr}(\boldsymbol{\Theta}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2}\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}})\Big\} \quad (59)$$

$$\sum_{g_1,g_2}\mathrm{tr}(\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}}\tilde{\boldsymbol{\Theta}}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\tilde{\boldsymbol{\Theta}}_{g_2}\mathbf{H}_{\mathrm{F},g_2})-2\Re\Big\{\sum_{g_1,g_2}\mathrm{tr}(\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}}\tilde{\boldsymbol{\Theta}}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2})\Big\}+\sum_{g_1,g_2}\mathrm{tr}(\mathbf{H}_{\mathrm{F},g_1}^{\mathsf{H}}\boldsymbol{\Theta}_{g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_1}^{\mathsf{H}}\mathbf{H}_{\mathrm{B},g_2}\boldsymbol{\Theta}_{g_2}\mathbf{H}_{\mathrm{F},g_2})\geq 0 \quad (60)$$

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.

[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, Nov 2019.

[3] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Transactions on Wireless Communications*, vol. 19, pp. 3064–3076, May 2020.

[4] Y. Liu, Y. Zhang, X. Zhao, S. Geng, P. Qin, and Z. Zhou, "Dynamic-controlled RIS assisted multi-user MISO downlink system: Joint beamforming design," *IEEE Transactions on Green Communications and Networking*, vol. 6, pp. 1069–1081, Jun 2022.

[5] Y. He, Y. Cai, H. Mao, and G. Yu, "RIS-assisted communication radar coexistence: Joint beamforming design and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 40, pp. 2131–2145, Jul 2022.

[6] H. Luo, R. Liu, M. Li, Y. Liu, and Q. Liu, "Joint beamforming design for RIS-assisted integrated sensing and communication systems," *IEEE Transactions on Vehicular Technology*, vol. 71, pp. 13 393–13 397, Dec 2022.

[7] 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, Apr 2023.

[8] Q. Wu and R. Zhang, "Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1735–1748, Aug 2020.

[9] 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.

[10] 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.

[11] R. Karasik, O. Simeone, M. D. Renzo, and S. S. Shitz, "Beyond max-SNR: Joint encoding for reconfigurable intelligent surfaces," in *2020 IEEE International Symposium on Information Theory (ISIT)*, Jun 2020, pp. 2965–2970.

[12] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, pp. 3187–3196, May 2020.

[13] J. Ye, S. Guo, S. Dang, B. Shihada, and M.-S. Alouini, "On the capacity of reconfigurable intelligent surface assisted MIMO symbiotic communications," *IEEE Transactions on Wireless Communications*, vol. 21, pp. 1943–1959, Mar 2022.

[14] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, pp. 1242–1255, Dec 2020.

[15] Y. Zhao and B. Clerckx, "RIScatter: Unifying backscatter communication and reconfigurable intelligent surface," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, Dec 2024.

[16] H. Yang, H. Ding, K. Cao, M. Elkashlan, H. Li, and K. Xin, "A RIS-segmented symbiotic ambient backscatter communication system," *IEEE Transactions on Vehicular Technology*, vol. 73, pp. 812–825, Jan 2024.

[17] E. Basar, "Reconfigurable intelligent surfaces for doppler effect and multipath fading mitigation," *Frontiers in Communications and Networks*, vol. 2, May 2021.

[18] E. Arslan, I. Yildirim, F. Kilinc, and E. Basar, "Over-the-air equalization with reconfigurable intelligent surfaces," *IET Communications*, vol. 16, pp. 1486–1497, Aug 2022.

[19] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Using intelligent reflecting surfaces for rank improvement in MIMO communications," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020, pp. 9160–9164.

[20] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Transactions on Communications*, vol. 68, pp. 4522–4535, Jul 2020.

[21] G. Chen and Q. Wu, "Fundamental limits of intelligent reflecting surface aided multiuser broadcast channel," *IEEE Transactions on Communications*, vol. 71, pp. 5904–5919, Oct 2023.

[22] 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, Feb 2021.

[23] 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," *IEEE Wireless Communications Letters*, vol. 13, pp. 437–441, Feb 2024.

[24] 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, May 2022.

[25] W. Huang, B. Lei, S. He, C. Kai, and C. Li, "Condition number improvement of IRS-aided near-field MIMO channels," in *2023 IEEE International Conference on Communications Workshops (ICC Workshops)*, May 2023, pp. 1210–1215.

[26] A. H. 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 Transactions on Communications*, vol. 70, pp. 3063–3080, May 2022.

[27] S. Zheng, B. Lv, T. Zhang, Y. Xu, G. Chen, R. Wang, and P. C. Ching, "On DoF of active RIS-assisted MIMO interference channel with arbitrary antenna configurations: When will RIS help?" *IEEE Transactions on Vehicular Technology*, Dec 2023.

[28] S. H. Chae and K. Lee, "Cooperative communication for the rank-deficient MIMO interference channel with a reconfigurable intelligent surface," *IEEE Transactions on Wireless Communications*, vol. 22, pp. 2099–2112, Mar 2023.

[29] S. Shen and B. Clerckx, "Beamforming optimization for MIMO wireless power transfer with nonlinear energy harvesting: RF combining versus DC combining," *IEEE Transactions on Wireless Communications*, vol. 20, pp. 199–213, Jan 2021.

[30] 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*, vol. 21, pp. 1229–1243, Feb 2022.

[31] M. Nerini, S. Shen, and B. Clerckx, "Closed-form global optimization of beyond diagonal reconfigurable intelligent surfaces," *IEEE Transactions on Wireless Communications*, vol. 23, pp. 1037–1051, Feb 2024.

[32] M. Nerini, S. Shen, H. Li, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces utilizing graph theory: Modeling, architecture design, and optimization," *IEEE Transactions on Wireless Communications*, pp. 1–1, May 2024.

[33] 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.

[34] ——, "Interference leakage minimization in RIS-assisted MIMO interference channels," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 39, Jun 2023, pp. 1–5.

[35] H.-R. Ahn, *Asymmetric Passive Components in Microwave Integrated Circuits*. Hoboken, NJ, USA: Wiley, 2006.

[36] 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, Aug 2023.

[37] H. Li, S. Shen, Y. Zhang, and B. Clerckx, "Channel estimation and beamforming for beyond diagonal reconfigurable intelligent surfaces," *arXiv:2403.18087*, 2024.

[38] H. Li, S. Shen, M. Nerini, M. D. Renzo, and B. Clerckx, "Beyond diag-

onal reconfigurable intelligent surfaces with mutual coupling: Modeling and optimization," *IEEE Communications Letters*, pp. 1–1, Oct 2024.

[39] H. Li, M. Nerini, S. Shen, and B. Clerckx, "Wideband modeling and beamforming for beyond diagonal reconfigurable intelligent surfaces," *arXiv:2403.12893*, 2024.

[40] T. Fang and Y. Mao, "A low-complexity beamforming design for beyond-diagonal RIS aided multi-user networks," *IEEE Communications Letters*, pp. 1–1, Jul 2023.

[41] 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.

[42] M. Soleymani, I. Santamaria, E. Jorswieck, and B. Clerckx, "Optimization of rate-splitting multiple access in beyond diagonal RIS-assisted URLLC systems," *IEEE Transactions on Wireless Communications*, pp. 1–1, Jul 2024.

[43] 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, Mar 2023.

[44] A. Mishra, Y. Mao, C. D'Andrea, S. Buzzi, and B. Clerckx, "Transmitter side beyond-diagonal reconfigurable intelligent surface for massive MIMO networks," *IEEE Wireless Communications Letters*, vol. 13, pp. 352–356, Feb 2024.

[45] M. Nerini, S. Shen, and B. Clerckx, "Discrete-value group and fully connected architectures for beyond diagonal reconfigurable intelligent surfaces," *IEEE Transactions on Vehicular Technology*, vol. 72, pp. 16 354–16 368, Dec 2023.

[46] 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, Jul 2010.

[47] 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, Apr 2023.

[48] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge, UK: Cambridge University Press, 1952.

[49] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton University Press, 2009.

[50] 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, Aug 2022.

[51] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, Sep 2006.

[52] L. Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pacific Journal of Mathematics*, vol. 16, pp. 1–3, Jan 1966.

[53] 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, Jan 1998.

[54] 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, Mar 2008.

[55] C. Moler and C. V. Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM Review*, vol. 45, pp. 3–49, Jan 2003.

[56] 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)*, Sep 2023, pp. 221–225.

[57] W. Fulton, "Eigenvalues, invariant factors, highest weights, and schubert calculus," *Bulletin of the American Mathematical Society*, vol. 37, pp. 209–249, Apr 2000.

[58] R. Bhatia, "Linear algebra to quantum cohomology: The story of alfred horn's inequalities," *The American Mathematical Monthly*, vol. 108, pp. 289–318, Apr 2001.

[59] L. Hogben, Ed., *Handbook of Linear Algebra*. Boca Raton, FL, USA: CRC press, 2013.

[60] A. Zanella, M. Chiani, and M. Win, "On the marginal distribution of the eigenvalues of wishart matrices," *IEEE Transactions on Communications*, vol. 57, pp. 1050–1060, Apr 2009.

[61] J. C. Gower and G. B. Dijksterhuis, *Procrustes Problems*. Oxford, UK: Oxford University Press, 2004.

[62] T. Bell, "Global positioning system-based attitude determination and the orthogonal procrustes problem," *Journal of Guidance, Control, and Dynamics*, vol. 26, pp. 820–822, Sep 2003.

[63] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins University Press, 2013.

[64] 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, Nov 2017.

[65] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*. Waltham, MA, USA: Academic Press, 2013.

[66] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press, May 2005.

[67] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, pp. 2740–2746, Jun 2007.