IEEE Open Journal of the Communications Society

Received 7 January 2026; revised 3 February 2026; accepted 18 February 2026. Date of publication 23 February 2026; date of current version 27 February 2026.

Digital Object Identifier 10.1109/OJCOMS.2026.3666997

Optimizing the Achievable Sum-Rate in OFDM-Based Multi-User MIMO Systems Assisted by Multiple Beyond-Diagonal RISs

DIOGO MENDES [1,2] (Graduate Student Member, IEEE), NUNO SOUTO [1,2] (Senior Member, IEEE), JOÃO PEDRO PAVIA [3,4] (Member, IEEE), AND JOÃO SILVA [1,4]

[1]Department of Information Science and Technology, ISCTE—University Institute of Lisbon, 1649-026 Lisbon, Portugal
[2]Instituto de Telecomunicações (IT), 1049-001 Lisbon, Portugal
[3]Department of Applied Digital Technologies, ISCTE—University Institute of Lisbon (Sintra), 2710-569 Sintra, Portugal
[4]Centro de Investigação em Ciências da Informação, Tecnologias e Arquitetura (ISTAR), 1649-026 Lisbon, Portugal

CORRESPONDING AUTHOR: D. MENDES (Diogo_Roque@iscte-iul.pt)

This work was supported by the Fundação para a Ciência e a Tecnologia (FCT)/Ministério da Ciência, Tecnologia e Ensino Superior (MCTES) through National Funds co-funded European Union (EU) Funds through the Instituto de Telecomunicações and Centro de Investigação em Ciências da Informação, Tecnologias e Arquitetura (ISTAR), under Grant UID/50008/2020, Grant UIDB/04466/2025, and Grant UIDP/04466/2025.

ABSTRACT Massive multiple-input, multiple-output (MIMO) systems operating in the millimeter wave (mmWave) and terahertz (THz) frequency bands offer high data rates and spatial multiplexing, yet they face significant propagation challenges. Reconfigurable intelligent surfaces (RISs) and their different architectures have emerged as a promising solution to these challenges, having the potential to enhance system performance. This paper addresses the joint sum-rate maximization problem in multi-user, multi-stream, multi-carrier MIMO systems aided by multiple parallel RIS panels. To minimize the inter-user interference, we adopt a problem formulation that adds a regularization term. To solve the resulting problem, we then propose a regularized cyclic block proximal gradient (MU-RCBPG) algorithm, which can jointly optimize precoders and RIS phase shifts without increasing the complexity compared to traditional single-valued decomposition (SVD)-based methods. The resulting algorithm has a flexible design that allows it to support configurations with beyond-diagonal RIS (BD-RIS), conventional diagonal RIS (D-RIS), and active D-RIS. Numerical results demonstrate that the MU-RCBPG algorithm outperforms existing RIS-aided schemes in various scenarios.

INDEX TERMS BD-RIS, beyond-5G/6G, mmWave–THz links, multi-carrier communications, MU-MIMO.

I. INTRODUCTION

MASSIVE multiple-input, multiple-output (MIMO) systems, enabled by the emergence of millimeter-wave (mmWave) and terahertz (THz) frequency bands, represent a pivotal advancement in the evolution of multiuser wireless networks [1].

The impetus for operating in the mmWave/THz band derives from the critical requirements of 5G and 6G networks for ultrahigh data rates, ultra-low latency, and ubiquitous connectivity [2], [3], [4]. These requirements propel the exploration and adoption of higher frequencies, where broader bandwidths become available. The mmWave/THz frequency bands offer vast, underutilized spectral resources [5]. However, they also introduce new propagation challenges due to the severe path loss and molecular absorption, limited diffraction, and increased sensitivity to blockage. In order to capitalize on the full potential of these bands, the implementation of large-scale antenna arrays is imperative for directional beamforming and spatial multiplexing. Additionally, the dense packing of antennas enabled by short mmWave/THz wavelengths supports spatial multiplexing at an unprecedented scale, thereby enhancing the potential of massive and ultra-massive MIMO (UM-MIMO) systems [1], [5], [6]. In comparison to lower frequency bands, mmWave/THz communications offer substantial advantages, including ultra-wide bandwidth, enhanced spatial resolution, and superior directionality. These characteristics are essential for high-capacity, dense urban environments. The use of ultra-narrow beams, enabled by mmWave/THz bands, has been demonstrated to enhance security, mitigate interference

among users, and combat path loss by concentrating the signal on a specific user direction. However, these advantages are accompanied by several disadvantages in terms of propagation that limit their practical deployment. Due to their short wavelengths, these bands are highly susceptible to scattering and dispersion, particularly when interacting with rough surfaces. Furthermore, mmWave and THz signals experience significant attenuation due to atmospheric absorption, resulting in substantial power loss over distance. Although ultra-narrow beams help to mitigate the effects of higher attenuation, signal blockage by obstacles such as buildings, foliage, or human bodies can severely impair connectivity. To ensure robust and high-performance communications, these challenges necessitate the implementation of complementary technologies to maintain reliable links [7].

Reconfigurable intelligent surfaces (RISs) have emerged as a key enabler for enhancing signal propagation, overcoming mmWave/THz propagation impairments. In environments where there is no direct line-of-sight (LoS) between users and BSs, RISs can effectively reconfigure the wireless channel by creating virtual LoS paths. This enhances signal coverage and reliability. RISs can steer, focus, or scatter incident waves in the desired direction, compensating for the absence of natural multipath propagation in high-frequency bands [8]. They can be broadly categorized into diagonal RIS (D-RIS) and beyond-diagonal RIS (BD-RIS), with the latter encompassing group-, and fully-connected topologies. D-RISs represent the most explored configuration, where each reflecting element operates independently, forming a diagonal phase shift matrix. Its simplicity comes with limited control over wavefronts. BD-RISs have been shown to offer greater flexibility and superior performance, allowing inter-element connectivity, forming symmetric and unitary scattering matrices. Fully-connected BD-RISs provide the highest degrees of freedom (DoF), at the cost of increased hardware and computational burden. The RISs can be classified into reflective architectures, including passive, active, single-, group-, and fully-connected. While passive RIS architectures have a lower cost of implementation, making them more energy-efficient, active RISs offer signal amplification, improving channel capacity. However, this is accompanied by increased complexity and power consumption [9].

A substantial body of literature has emerged on the subject of systems comprising D-RISs, such as [10], [11], [12], [13], [14], [15], and [16]. These works have emphasized the value of joint optimization in RIS-aided MIMO systems and demonstrated the increase in performance and system capabilities that results from the addition of RISs to MIMO systems. The authors in [14] proposed a multiuser, multistream algorithm for maximizing the sum-rate using the Gaussian multiple-access and broadcast (MAC-BC) duality, where they have shown the impact of adding multiple parallel RISs to enhance MU-MIMO system capabilities. This work assumes a prior implementation of dirty paper coding (DPC), which is a significantly complex operation. The authors also compared the results with those of a linear precoding algorithm first presented in [16], which does not assume any prior coding and represents a less complex, more cost-effective solution. However, both [14] and [16] assume the transmission of a single subcarrier, which is not consistent with the 5G standard of orthogonal frequency-division multiplexing (OFDM) waveform. Consequently, there is a need for more adaptive, flexible, and versatile algorithms that optimize multistream RIS-aided MU-MIMO communication over multiple subcarriers multiplexed in frequency.

More recently, various optimization strategies have been developed to handle the increased dimensionality and constraints of BD-RIS, such as [9], [17], [18], [19], [20], [21], [22], [23], and [24], which underline the relevance and the need for unified algorithms that scale reliably across BD-RIS-aided MIMO networks, and significantly enhance spectral efficiency and channel capacity. This paradigm aims to improve achievable rates while also reducing the complexity of joint optimization frameworks under realistic system constraints. The authors in [17] and [20] present different algorithms and variations, which aim to maximize the achievable sum-rate of a multiuser MIMO system. The authors in [17] consider a BD-RIS-assisted multi-user multiple-input single-output (MU-MISO) system and propose a two-stage, low-complexity beamforming design framework. This framework separates the optimization of passive (at the BD-RIS) and active (at the BS) beamforming using a closed-form method. They present a closed-form approximation to the optimal passive beamforming matrix that maximizes the sum of effective channel gains. This approach achieves significant computational gains over classical alternating optimization methods. The authors in [20] address the weighted sum-rate (WSR) maximization problem in a similar MU-MISO downlink system employing a fully-connected BD-RIS. They introduce an efficient alternating optimization algorithm, which they named FP-PSLA, which iteratively optimizes the active (BS) and passive (BD-RIS) beamforming matrices using closed-form solutions within a fractional programming framework (FP). Their results demonstrate that the proposed method substantially reduces computational complexity while achieving nearly optimal WSR performance, particularly in large-scale networks.

The existing body of literature on the subject indicates that RISs, particularly BD-RISs, present a promising avenue for scalable and flexible system designs. This consensus lends further credence to the growing interest and effectiveness of RISs in complex scenarios. However, the majority of system models are constrained to simplified topologies, as summarized in Table 1.

This paper contributes to the literature by addressing a broader setting and by proposing a unified algorithm for joint optimization of precoding and RIS matrices. The objective of the algorithm is to maximize the sum-rate performance of the multi-parallel RIS-aided MU-MIMO network, comprising multi-stream and multi-carrier communications at mmWave/THz frequency band. Since the optimization

TABLE 1. Comparison of supported configurations of the proposed algorithm with those of the relevant literature.

Reference	Optimization objective	Multi-user	BD-RIS	D-RIS	Multi-RIS	Multi-rx antennas	Multi-stream	Multi-carrier
[10]	Rate	-	-	✓	✓	✓	✓	-
[12]	Rate	-	-	✓	-	✓	✓	✓
[13]	Weighted Sum-rate	✓	-	✓	✓	✓	-	✓
[14]	Sum-rate	✓	-	✓	✓	✓	✓	-
[16]	Weighted Sum-rate	✓	-	✓	-	✓	✓	-
[17]	Sum-rate	✓	✓	✓	-	-	-	-
[20]	Weighted Sum-rate	✓	✓	✓	-	-	-	-
[21]	Rx Signal Power	✓	✓	✓	✓	-	-	-
[22]	Sum-rate	✓	✓	✓	-	-	-	-
Proposed	Sum-rate	✓	✓	✓	✓	✓	✓	✓

problem is non-convex and involves large-scale systems, the block proximal gradient (BPG) method is adopted. Although this approach does not guarantee finding global optimal solutions for non-convex problems, it enables the development of algorithms that strike a good balance between complexity and performance [25], [26].

The main contributions of the paper can be summarized as follows:

- We formulate the joint achievable sum-rate maximization problem, defined over the precoders and the phase shifts of all RIS elements, considering a multi-stream MU-MIMO communication system aided by multiple parallel RIS panels. Since eliminating the inter-user interference directly using block-diagonalization (BD) precoding would require a singular value decomposition (SVD) per user and subcarrier, an operation with considerable computational cost, we proposed an alternative problem formulation that adds a regularization term to the objective function, which drives it towards a solution with reduced inter-user interference.
- In order to determine a solution to the adopted problem formulation, an iterative algorithm (MU-RCBPG) based on regularized cyclic block proximal gradient (CBPG) is derived. Using the CBPG method allows us to simplify the joint optimization process by alternately optimizing the precoders and the RIS matrices. The exact expressions for the required gradients and proximal operators are presented. As CBPG is a first-order optimization method, it is particularly appealing to large-scale problems involving a large number of transmit/receive antennas, RIS elements, users, and subcarriers, such as the one we are addressing here. This makes our approach more scalable. While initially designed for BD-RISs, the algorithm was also extended to less complex scenarios where the RIS elements are independent from each other, assuming a D-RIS setup. Furthermore, we present an extension of the proposed algorithm to active (amplified) D-RISs.
- We present numerical results that show the capabilities of the MU-RCBPG in scenarios comprising multiple users, multiple streams, multiple subcarriers, and multiple parallel D-, and BD-RIS panels, and we compare them with a standard BD precoding algorithm with passive and active blind RISs, and algorithms from the literature.

This paper is organized as follows: section II introduces the receive signal model taken into account in the multi-stream, multicarrier RIS-assisted mmWave/THz MU-MIMO system, section III presents the problem formulation and the proposed joint optimization algorithm of precoders and RISs matrices, section IV presents the numerical results of the link level simulations, and section V outlines the conclusions of our study.

Notation: Matrices and vectors are represented in boldface letters in upper and lower case letters, respectively. $\mathbb{C}^{m\times n}$ denotes the space of $m\times n$ complex matrices. $\mathbf{I}_n$ is the identity matrix of dimension $n\times n$. The superscript $(\cdot)^{-1}$, $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote the inverse, transpose, complex conjugate, and Hermitian transpose, respectively. $\ln(\cdot)$ represents the natural logarithm. $\mathbb{E}\{\cdot\}$ denotes the expectation operator. $|x|$ represents the modulus of the complex number x and $|\mathbf{X}|$ represents the determinant of the matrix $\mathbf{X}$. $diag(\cdot)$ denotes the diagonal matrix with the elements of a vector on the diagonal. $\mathrm{Tr}\{\cdot\}$ denotes the trace, $\|\cdot\|_F$ the Frobenius norm, and $\mathcal{N}(\cdot)$ the null space. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts. Symbol $\oslash$ denotes the Hadamard division.

II. SYSTEM MODEL

This section aims to describe the system model of a multicarrier, multiuser MIMO (MU-MIMO) system that operates in the mmWave/THz band and is capable of serving N_u users simultaneously, each with N_{rx} receive antennas. The multiple links are established with the aid of N_{pan} RIS panels, each comprising N_{ris} elements, as illustrated in figure 1.

The channel is modeled on the assumption that the system uses OFDM to cope with frequency-selective fading. At the receivers, the resulting vector for the k^{th} subcarrier of the u^{th} user, $\mathbf{r}_{k,u}$, can be written using

$$\mathbf{r}_{k,u} = \sqrt{\rho_k}\mathbf{H}_{k,u}\mathbf{F}_k\mathbf{s}_k + \mathbf{n}_{k,u} \tag{1}$$

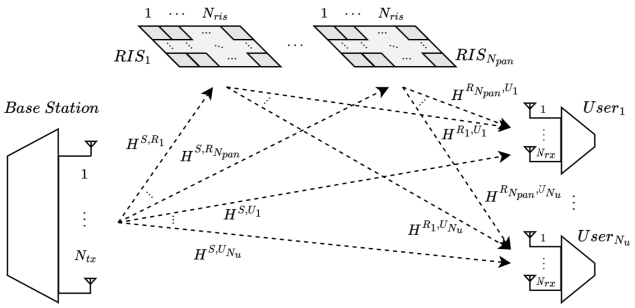

FIGURE 1. Layout considered for the RIS-assisted MU-MIMO communication system.

where $u = 1, \ldots, N_u$ and $k = 1, \ldots, N_c$, N_c being the total number of subcarriers allocated to each user. $\mathbf{s}_k \in \mathbb{C}^{N_u N_s \times 1}$ represents the symbol vector containing the amplitude and phase modulated symbols with $\mathbb{E}\{\|\mathbf{s}_k\|_2^2\} = N_u N_s$, where N_s is the number of streams transmitted simultaneously. $\mathbf{F}_k \in \mathbb{C}^{N_{tx} \times N_u N_s}$ represents the precoder and ρ_k the power per stream and per subcarrier. Although we are assuming the BS precoder, $\mathbf{F}$, is fully digital, it is straightforward to transition to hybrid structures by finding a product of digital-analog precoder matrices that minimizes the Frobenius distance to matrix $\mathbf{F}$, as outlined in [27], [28], [29]. Vector $\mathbf{n}_{k,u} \in \mathbb{C}^{N_{rx} \times 1}$ represents the N_{rx} complex noise components at the receiver, which are assumed to be independent zero-mean circularly symmetric Gaussian samples with covariance $\sigma_n^2 \mathbf{I}_{N_{rx}}$ [30], [31]. The matrix $\mathbf{H}_{k,u} \in \mathbb{C}^{N_{rx} \times N_{tx}}$ is the k^{th} subcarrier frequency domain channel between the BS and the u^{th} user, and can be defined for each subcarrier k and user u as

$$\mathbf{H}_{k,u} = \mathbf{H}_k^{S,U_u} + \sum_{i=0}^{N_{pan}} \mathbf{H}_k^{R_i,U_u} \mathbf{\Phi}_i \mathbf{H}_k^{S,R_i} \tag{2}$$

where $\mathbf{H}_k^{S,U_u} \in \mathbb{C}^{N_{rx} \times N_{tx}}$ represents the direct channel between the BS and the user u, $\mathbf{H}_k^{R_i,U_u} \in \mathbb{C}^{N_{rx} \times N_{ris}}$, the channel between the i^{th} RIS and the user u, $\mathbf{H}_k^{S,R_i} \in \mathbb{C}^{N_{ris} \times N_{tx}}$ the channel between the BS and the i^{th} RIS and $\mathbf{\Phi}_i \in \mathbb{C}^{N_{ris} \times N_{ris}}$ is the i^{th} RIS matrix. The definition in (2) is based on the assumption that the RISs are used in scenarios involving significant blockage losses and are positioned far apart, with LOS to the BS, but not to each other. Furthermore, under the assumption of effective RIS beamforming, signal reflections will not be directed towards other RISs. In scenarios where inter-RIS reflections could occur, optimization and performance estimation would become even more complex. Therefore, as in [10], [13], [14], and [21], we ignore the effects of coupling between different panels and neglect multiple reflections when defining the total channel matrix. The considered mmWave/THz model includes channels with one LoS and N_{ray} non-LoS (NLoS) components, where the rays arrive in clusters [32]. If the distance between the BS and the u^{th} user, d_{S,U_u}, is greater than the Fraunhofer distance, defined as $D_F \triangleq 2L_{array}^2/\lambda$ [33], L_{array} being the largest dimension of the array, we can assume a far-field propagation model with planar wavefronts and write the channel frequency response between the BS and the u^{th} user, at subcarrier k, as

$$\begin{aligned}\mathbf{H}_k^{S,U_u} &= \sqrt{\beta_{LOS}^{S,U_u}} \mathbf{a}_{U_u}(\phi_0^{U_u \leftarrow S}, \theta_0^{U_u \leftarrow S}) \\ &\quad \times \mathbf{a}_S^H(\phi_0^{S \rightarrow U_u}, \theta_0^{S \rightarrow U_u}) e^{-j2\pi \frac{d_{S,U_u}}{\lambda_k}} \\ &\quad + \frac{1}{\sqrt{K_{Rice}}} \sum_{l=1}^{N_{ray}} \sqrt{\beta_{NLOS}^{S,U_u}(l)} \alpha_l^{S,U_u} \mathbf{a}_{U_u}(\phi_l^{U_u \leftarrow S}, \theta_l^{U_u \leftarrow S}) \\ &\quad \times \mathbf{a}_S^H(\phi_l^{S \rightarrow U_u}, \theta_l^{S \rightarrow U_u}) e^{-j2\pi \tau_l f_k}\end{aligned} \tag{3}$$

which describes $\mathbf{H}_k^{S,U_u}$ as the sum of one LoS and N_{ray} NLoS components. The subcarrier frequency, f_k, can be calculated as $f_k = f_c + \frac{B}{N_c}(k - 1 - \frac{N_c - 1}{2})$, where B is the bandwidth and f_c the carrier frequency. λ_k is the subcarrier wavelength according to $\lambda_k = c/f_k$, c being the speed of light in vacuum. τ_l is the delay of the l^{th} path, and K_{Rice} is the ratio between the LoS and NLoS components, which indicates the quality of the channel and the relative fading the path is subjected to [34]. $\alpha_l^{S,U_u} \in \mathbb{C}^{N_{ray} \times 1}$, is the l^{th} NLoS complex gain between the BS and the user u, while β_{LOS}^{S,U_u} and β_{NLOS}^{S,U_u} represent the path loss of the LoS and NLoS channels, respectively. An approximation of the path loss between the BS and the u^{th} user can be computed as follows [35]

$$\beta^{S,U_u} = \frac{G_{tx} G_{rx}}{(\frac{4\pi}{\lambda})^2 (d_{S,U_u})^{\eta}} e^{k_{abs}(f) d_{S,U_u}} \tag{4}$$

where G_{tx} and G_{rx} are the gains of the transmit and receive antennas, respectively, η is the path loss exponent, and $k_{abs}(f)$ is the coefficient of molecular absorption at the frequency f [35]. The clusters and the rays within each cluster are modeled to have specific angles of arrival (AoA) and departure (AoD). $\mathbf{a}_S(\phi_l^{S \rightarrow U_u}, \theta_l^{S \rightarrow U_u})$ and $\mathbf{a}_{U_u}(\phi_l^{U_u \leftarrow S}, \theta_l^{U_u \leftarrow S})$ represent the array response vectors at the AoD and AoA from the large transmitter antenna array at the BS to the u^{th} receiver antenna array, respectively, for the considered azimuth ϕ_l and elevation θ_l. Under the assumption of a uniform planar array (UPA) implementation, the array response vectors at the BS and user u can be calculated, respectively, as [36], [37]

$$\begin{aligned}&\mathbf{a}_S(\phi_l^{S \rightarrow U_u}, \theta_l^{S \rightarrow U_u}) \\ &= \left[1, \ldots, e^{j\frac{2\pi}{\lambda} d_{tx}(p \sin\phi_l^{S \rightarrow U_u} \sin\theta_l^{S \rightarrow U_u} + q \cos\theta_l^{S \rightarrow U_u})}, \ldots\right]^T,\end{aligned} \tag{5}$$

and

$$\begin{aligned}&\mathbf{a}_{U_u}(\phi_l^{U_u \leftarrow S}, \theta_l^{U_u \leftarrow S}) \\ &= \left[1, \ldots, e^{j\frac{2\pi}{\lambda} d_{rx}(p' \sin\phi_l^{U_u \leftarrow S} \sin\theta_l^{U_u \leftarrow S} + q' \cos\theta_l^{U_u \leftarrow S})}, \ldots\right]^T,\end{aligned} \tag{6}$$

where $p, q = 0, \ldots, \sqrt{N_{tx}} - 1$ and $p', q' = 0, \ldots, \sqrt{N_{rx}} - 1$ are the relative positions of the $(p, q)^{\text{th}}$ transmit antenna and $(p', q')^{\text{th}}$ receive antenna. d_{tx} and d_{rx} are the spacing between antennas that make up the rectangular arrays at the transmitter and receiver, respectively. The channel frequency

response at subcarrier k, between the BS and the i^{th} RIS, $\mathbf{H}_k^{S,R_i}$, and between the i^{th} RIS and the u^{th} user, $\mathbf{H}_k^{R_i,U_u}$, can both be written in a similar way to (3), by replacing the user angles of arrival with the RIS angles of arrival. The pathloss coefficients should also consider the distances between the considered pair and the gain of the RIS. The array response vectors at the AoD and AoA from the large transmitter antenna array at the BS to the surface of the i^{th} RIS, and from the surface of the i^{th} RIS to the u^{th} user, can also be calculated, similar to (5) and (6).

In specific cases where the distances d_{S,U_u} (BS-User u), d_{S,R_i} (BS-RIS i), or d_{R_i,U_u} (RIS i-User u) are less than the Fraunhofer distance, a spherical wavefront near-field (NF) propagation model should be assumed and (3) must be adapted, to include the effect of all the individual path distances between each transmitting-receiving antenna pair [38]. Therefore, assuming an NF propagation model and unit-normalized power radiation patterns along the directions of interest for both the transmitting and receiving antennas, we can rewrite (3), for subcarrier k, as [10]

$$\mathbf{H}_k^{S,U_u}(n,m) = \sqrt{\beta_{LOS}^{m,n}} e^{-j2\pi \frac{d_{m,n}}{\lambda_k}} + \frac{1}{\sqrt{K_{Rice}}} \sum_{l=1}^{N_{ray}} \sqrt{\beta_{NLOS}^{m,n}(l)} |\alpha_l^{m,n}| e^{-j2\pi \tau_l^{m,n} f_k}, \tag{7}$$

thus, representing the channel frequency response between the m^{th} transmit antenna at the BS to the n^{th} receive antenna at user u. $\tau_l^{m,n}$ is the delay of the l^{th} path between elements m and n, and $\alpha_l^{m,n}$ is the complex gain of the l^{th} path between elements m and n normalized as $\sum_{l=1}^{N_{ray}} E[|\alpha_l^{m,n}|^2] = 1$. In order to compute the path loss of the LOS and NLOS channels, $\beta_{LOS}^{m,n}$ and $\beta_{NLOS}^{m,n}$, respectively, we consider the gains of the m^{th} transmit antenna at the BS and the n^{th} receive antenna at user u, and the distance of the individual (m,n) pair, $d_{m,n}$. For the NF model, the path loss coefficient between the m^{th} transmit antenna and the n^{th} receive antenna can be computed as [10], [35]

$$\beta^{m,n} = \frac{G_m G_n}{(\frac{4\pi}{\lambda})^2 (d_{m,n})^{\eta}} e^{k_{abs}(f) d_{m,n}}. \tag{8}$$

III. PROBLEM FORMULATION AND PROPOSED ALGORITHM

In this section, we present the problem formulation and introduce the proposed algorithm for joint optimization of the precoders and RIS elements. The main objective is to maximize the achievable sum-rate of the multi-stream, multi-carrier, multi-user and multi-parallel RIS MIMO link, previously described in section II.

A. PROBLEM FORMULATION

Assuming perfect channel knowledge, the achievable sum-rate, in bits per second per Hertz (bits/s/Hz), can be given by [39]

$$R = \sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \ln \left| \mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma_n^2} \mathbf{F}_{k,u}^H \mathbf{H}_{k,u}^H \mathbf{\Psi}_{k,u}^{-1} \mathbf{H}_{k,u} \mathbf{F}_{k,u} \right|, \tag{9}$$

where $\rho_{k,u}/\sigma_n^2$ is the signal-to-noise ratio (SNR) of the k^{th} subcarrier from the u^{th} user, $\mathbf{F}_{k,u} \in \mathbb{C}^{N_{tx} \times N_s}$ the precoder of the u^{th} user in subcarrier k with $\mathbf{F}_k = [\mathbf{F}_{k,1}, \ldots, \mathbf{F}_{k,N_u}]$, and $\Psi_{k,u}$ represents the interference, i.e., the amount of power that user u receives in subcarrier k that was intended for the other users. The total interference experienced by user u in subcarrier k is given by

$$\Psi_{k,u} = \mathbf{I}_{N_{rx}} + \sum_{\substack{u'=1 \\ u' \neq u}}^{N_u} \frac{\rho_{k,u'}}{\sigma_n^2} \mathbf{H}_{k,u} \mathbf{F}_{k,u'} \mathbf{F}_{k,u'}^H \mathbf{H}_{k,u}^H. \tag{10}$$

In order to maximize (9), we formulate the optimization problem as

$$\min_{\mathbf{F},\mathbf{\Phi}} \quad f(\mathbf{F},\mathbf{\Phi}) = -\sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \ln \left| \mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma_n^2} \mathbf{F}_{k,u}^H \mathbf{H}_{k,u}^H \mathbf{H}_{k,u} \mathbf{F}_{k,u} \right| \tag{11a}$$

$$\text{s.t.} \quad \sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \|\mathbf{F}_{k,u}\|_F^2 \leq N_c N_u N_s \tag{11b}$$

$$\sum_{k=1}^{N_c} \sum_{\substack{u'=1 \\ u' \neq u}}^{N_u} \mathbf{H}_{k,u'} \mathbf{F}_{k,u} = \mathbf{0} \qquad u = 1, \ldots, N_u \tag{11c}$$

$$\mathbf{\Phi}_i \in \mathcal{R} \qquad i = 1, \ldots, N_{pan} \tag{11d}$$

where $\mathcal{R}$ represents the set of feasible RIS matrices, whose specific definition depends on the type of RIS. The restriction (11c) is equivalent to enforcing other users' channels, $\mathbf{H}_{k,u'}$, to lie in the null space of the u^{th} user's precoder, $\mathbf{F}_{k,u}$, thus eliminating inter-user interference.

In order to ease the complexity of imposing the constraint (11c), we redefine the optimization problem using

$$\min_{\mathbf{F},\mathbf{\Phi}} \quad f(\mathbf{F},\mathbf{\Phi}) = -\sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \ln \left| \mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma_n^2} \mathbf{F}_{k,u}^H \mathbf{H}_{k,u}^H \mathbf{H}_{k,u} \mathbf{F}_{k,u} \right| + \gamma \left\| \frac{\sqrt{\rho_{k,u}}}{\sigma_n} \overline{\mathbf{H}}_{k,u} \mathbf{F}_{k,u} \right\|_F^2 \tag{12a}$$

$$\text{s.t.} \quad \sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \|\mathbf{F}_{k,u}\|_F^2 \leq N_c N_u N_s \tag{12b}$$

$$\mathbf{\Phi}_i \in \mathcal{R} \qquad i = 1, \ldots, N_{pan} \tag{12c}$$

where we remove constraint (11c) and add the regularization component, $\sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \gamma \left\| \frac{\sqrt{\rho_{k,u}}}{\sigma_n} \overline{\mathbf{H}}_{k,u} \mathbf{F}_{k,u} \right\|_F^2$. The constant γ represents the regularization parameter and matrix $\overline{\mathbf{H}}_{k,u} \in \mathbb{C}^{(N_u-1)N_{rx} \times N_{tx}}$ represents all user channels in subcarrier k, $[\mathbf{H}_{k,1} \ldots \mathbf{H}_{k,N_u}]^T$, with the u^{th} user's N_{rx} lines removed, thus

representing the channel associated with the interference generated by the transmission of the signal directed at user u on all other users.

The proposed objective function employs a squared Frobenius norm penalty, which serves as a differentiable surrogate for the aggregate multi-user interference leakage. Theoretically, this formulation transforms the traditionally non-convex and tightly coupled SINR-based sum-rate maximization into a Lagrangian-style unconstrained optimization suitable for CBPG methods. In multi-user MIMO systems, traditional interference-nulling techniques, such as block diagonalization, act as a "hard" constraint and force the precoders into the null space of the matrix $\overline{\mathbf{H}}_{k,u}$, which can be computationally expensive. This often comes at a significant cost to the desired signal power, especially in rank-deficient or highly correlated channels. Our regularization term shifts this paradigm to a soft-interference management framework. This transforms the sum-rate maximization problem from a constrained interference-nulling formulation into an unconstrained trade-off between signal enhancement and interference suppression. By minimizing this term, the algorithm solving this problem inherently seeks a balance: it favors orthogonality when interference is high but allows for controlled leakage if the resulting gain in the user's own link budget disproportionately increases the total sum-rate. By allowing controlled interference leakage, the system can exploit stronger signal subspaces when they provide a net sum-rate gain. Since the regularization term is embedded directly in the objective function, it tends to guide both precoder and RIS updates toward solutions with reduced interference leakage across all optimization blocks. Later in section IV, we demonstrate that a moderate regularization weight significantly improves the achievable sum-rate, while excessively large values overly prioritize orthogonality and degrade performance. The regularization factor γ therefore serves as a tunable parameter that governs the trade-off between inter-user interference suppression and sum-rate maximization.

B. ALGORITHM DERIVATION

In order to find a proximal solution to (12a), we use the cyclic block proximal gradient method (CBPG) [26]. As a first-order optimization method, the CBPG is well-suited to dealing with large-scale problems (i.e. large number of transmit/receive antennas, RIS elements, users, and subcarriers), such as the one we are addressing here. To implement the CBPG, we can use the indicator function, defined for a generic set $\mathcal{A}$ as $\mathcal{I}_{\mathcal{A}}(x)$, and which returns 0 if $x \in \mathcal{A}$ and $+\infty$ otherwise. This allows us to rewrite the problem as

$$\min_{\mathbf{F},\mathbf{\Phi}} \quad f(\mathbf{F},\mathbf{\Phi}) = \sum_{k=1}^{N_c}\sum_{u=1}^{N_u}\left[-\ln\left|\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma_n^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right|\right. \left.+\gamma\,\|\frac{\sqrt{\rho_{k,u}}}{\sigma_n}\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\|_F^2+\mathcal{I}_{\mathcal{P}}(\mathbf{F}_{k,u})\right]+\sum_{i=1}^{N_{pan}}\mathcal{I}_{\mathcal{R}}(\mathbf{\Phi}_i) \tag{13}$$

where $\mathcal{I}_{\mathcal{P}}$ and $\mathcal{I}_{\mathcal{R}}$ are the indicator functions for sets $\mathcal{P}$ and $\mathcal{R}$, respectively, which are defined as

$$\mathcal{P}=\left\{\mathbf{P}\in\mathbb{C}^{N_{tx}\times N_s}:\|\mathbf{P}\|_F^2\le N_s\right\}$$

and

$$\mathcal{R}=\left\{\mathbf{R}\in\mathbb{C}^{N_{ris}\times N_{ris}}:\mathbf{R}^H\mathbf{R}=\mathbf{I}_{N_{ris}}\right\},$$

for the case of BD-RIS.

The CBPG consists of cyclically picking a block and performing a proximal gradient step with respect to the chosen block [26]. For a given iteration, q, and gradient step, t, the precoders and RIS matrices can be updated using

$$\mathbf{F}_{k,u}^{(q+1)}=\mathbf{F}_{k,u}^{(q)}-t\nabla_{\mathbf{F}_{k,u}^*}f(\mathbf{F}_{k,u}^{(q)},\mathbf{\Phi}^{(q+1)}) \tag{14}$$

and

$$\mathbf{\Phi}_i^{(q+1)}=\mathbf{\Phi}_i^{(q)}-t\nabla_{\mathbf{\Phi}_i^*}f(\mathbf{F}^{(q)},\mathbf{\Phi}_i^{(q)}), \tag{15}$$

respectively, where $\nabla_{\mathbf{F}_{k,u}^*}f$ and $\nabla_{\mathbf{\Phi}_i^*}f$ are the gradients of f with respect to the conjugate of the precoders, and RISs, respectively.

Lemma 1: Considering (12a), the gradients of f w.r.t. the conjugate of the precoders, and BD-RIS matrices, $\nabla_{\mathbf{F}_{k,u}^*}f$ and $\nabla_{\mathbf{\Phi}_i^*}f$, can be defined, respectively, as

$$\nabla_{\mathbf{F}_{k,u}^*}f=-\frac{\rho_{k,u}}{\sigma^2}\Big[\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}+\gamma\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\Big]. \tag{16}$$

and

$$\nabla_{\mathbf{\Phi}_i^*}f=-\sum_{k=1}^{N_c}\sum_{u=1}^{N_u}\frac{\rho_{k,u}}{\sigma^2}\Big[\left(\mathbf{H}_k^{R_i,U_u}\right)^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H+\gamma\left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H\Big]. \tag{17}$$

with $\overline{\mathbf{H}}_k^{R_i,U_u}\in\mathbb{C}^{(N_u-1)N_{rx}\times N_{ris}}$ representing all channels in subcarrier k between the i^{th} RIS panel and the users, $[\mathbf{H}_k^{R_i,U_1},\ldots,\mathbf{H}_k^{R_i,U_{N_u}}]^T$, with the u^{th} user's N_{rx} lines removed, representing the interference generated by the transmission of the signal directed at user u to all other users.

Proof: see Appendix A.

After the proximal gradient step, the algorithm will also update the estimate optimization variables by applying a proximal operator, $prox_{\propto g}$, to impose constraints (12b) and (12c). The proximal operator for an arbitrary function g can be defined as $prox_{\propto g}(\mathbf{z})=\underset{\hat{\mathbf{x}}}{\text{argmin}}\, g(\hat{\mathbf{x}})+\frac{1}{2}\|\hat{\mathbf{x}}-\mathbf{z}\|^2$ which, for (12b), can be calculated as an Euclidean projection, where g corresponds to the indicator function, $\mathcal{I}$, of set $\mathcal{P}$, and can be computed as

$$prox_{\propto\mathcal{I}_{\mathcal{P}}}(\mathbf{Z})=\begin{cases}\mathbf{Z} & ,\ \|\mathbf{Z}\|_F^2\le N_s\\ \frac{\sqrt{N_s}}{\|\mathbf{Z}\|_F}\mathbf{Z} & ,\ \text{otherwise}\end{cases}. \tag{18}$$

Regarding constraint (12c), the proximal operator for set $\mathcal{R}$, $prox_{\propto \mathcal{I}_{\mathcal{R}}}$, can be described as

$$prox_{\propto \mathcal{I}_{\mathcal{R}}}(\mathbf{Z}) = \mathbf{U}\mathbf{V}^H, \tag{19}$$

where $\mathbf{U}$ and $\mathbf{V}$ are obtained from the singular value decomposition (SVD) of $\mathbf{Z}$, with $\mathbf{Z} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$. It is important to note that applying proximal mapping to indicator functions such as (18) and (19) essentially results in a block projected gradient algorithm, as clarified in [26], [40], [41], and [42].

To find the most suitable gradient step, we use the backtracking line-search method [40], [41]. For a given gradient step, t, decreasing by a factor, β, and a linear extrapolation slope, α, the stop condition can be generically defined, for complex optimization variables, as $f(x + t\Delta x) \leq f(x) + 2\alpha t\, \Re\left\{\nabla f(x)^H \Delta x\right\}$ [40]. Aiming to minimize the objective by jointly optimizing the precoder, $\mathbf{F}$, and RIS matrices, $\mathbf{\Phi}$, and considering a descent direction for the gradients, the backtracking line search should run until

$$\begin{aligned} f(\mathbf{F}^{candidate}, \mathbf{\Phi}^{candidate}) \leq f(\mathbf{F}, \mathbf{\Phi}) \\ + 2\alpha t\, \Re\left\{\Delta F^H(-\Delta F)\right\} \\ + 2\alpha t\, \Re\left\{\Delta \Phi^H(-\Delta \Phi)\right\} \end{aligned} \tag{20}$$

holds, where $\Delta F = \mathbf{F}^{candidate} - \mathbf{F}$ and $\Delta\Phi = \mathbf{\Phi}^{candidate} - \mathbf{\Phi}$, to take into account projections after the gradient descent.

The resulting algorithm consists of a total of Q CBPG iterations, where, for each main iteration, q, the precoder matrices of all N_c subcarriers from all N_u users, and the elements of all N_{pan} RIS panels are updated in two alternating cycles, with the backtracking line search used to find an adequate gradient step according to (20). This step has initial values of t_Φ and t_F for the RIS panels and precoders, respectively. The algorithm is designed to stop either when the number of iterations reaches its limit, Q, or when the relative difference between consecutive objective function values reaches a threshold ε (in the simulations we adopted $\varepsilon = 10^{-5}$). The resulting multiuser regularized CBPG (MU-RCBPG) algorithm is summarized in algorithm 1.

Due to the reciprocal nature of certain passive BD-RIS networks, some algorithm implementations consider projections of symmetric, unitary RIS matrices. This symmetry constraint can easily be incorporated in our proposed algorithm using the projection defined in [17]. In this scenario, the set R must be redefined as

$$\mathcal{R} = \left\{\mathbf{R} \in \mathbb{C}^{N_{ris} \times N_{ris}} : \mathbf{R}^H\mathbf{R} = \mathbf{I}_{N_{ris}}, \mathbf{R} = \mathbf{R}^T\right\},$$

effectively restricting the solution space to the set of complex symmetric unitary matrices. This approach ensures that the optimized matrix satisfies the losslessness and reciprocity conditions for passive BD-RIS architectures.

The proposed CBPG-based algorithm with backtracking line search is known to converge under standard conditions, namely that the objective function is bounded, admits block-wise Lipschitz continuous gradients, each block update corresponds to a well-defined proximal subproblem, the feasible sets associated with the individual blocks are closed and convex, and a sufficient descent condition is enforced [26]. The objective function, defined in (12a), is smooth with respect to each optimization block when the remaining variables are fixed. For fixed BD-RIS coefficients, for example, the objective is smooth in the precoders because its gradient only involves matrix multiplications and inverses of positive definite covariance matrices. There are no singularities due to strictly positive noise variance. Under finite transmit power constraints, the feasible precoder set is compact, and the corresponding Hessian is bounded. This implies the gradient is Lipschitz continuous on the feasible set. Similarly, when the precoders are fixed, the objective function is smooth with respect to the BD-RIS coefficients. However, although it is closed and bounded, the set of feasible BD-RIS matrices is not convex. Therefore, we cannot guarantee pointwise convergence. Furthermore, a backtracking line search is used to adaptively enforce the required descent condition without needing to know the closed-form Lipschitz constants explicitly. These constants are analytically intractable to derive because they rely on nested matrix products, log-determinant terms, and covariance matrix inverses that depend on the precoders and BD-RIS coefficients. Since the feasible set of BD-RIS matrices is nonconvex, the only conclusion that can be drawn is that the sequence of objective function values, $f(\mathbf{F}^{(q)}, \mathbf{\Phi}^{(q)})$, $q \geq 0$, is nonincreasing and bounded below, and therefore converges. Proving that the associated sequence of points $f(\mathbf{F}^{(q)}, \mathbf{\Phi}^{(q)})$ converges to a stationary point for our formulation is highly nontrivial, but it may be something to explore further as future work. However, the functional convergence of the objective function is sufficient for our purposes since we can ensure that the objective function is decreasing. Under these assumptions, the proposed algorithm with backtracking line search generates a sequence of feasible iterates with monotonically decreasing objective values.

Algorithm 1 MU-RCBPG - Multiuser Regularized CBPG for Joint Optimization of Precoders and RIS Panels

Input: $R^{(0)}$, $\mathbf{H}^{S,D}$, $\mathbf{H}^{S,i}$, $\mathbf{H}^{i,D}$, $\mathbf{F}^{(0)}$, $\mathbf{\Phi}^{(0)}$, $\mathbf{Q}$, t_F, t_Φ, α, β
Output: $\mathbf{F}^{(Q)}$, $\mathbf{\Phi}^{(Q)}$

repeat
  Compute $\nabla_{\mathbf{\Phi}_i^*} f$ using (17), for all i
  repeat
    Update $\mathbf{\Phi}_i^{(q+1)}$ using (15), for all i
    Project $\mathbf{\Phi}_i^{(q+1)}$ using (19), for all i
  until condition (20) holds
  Compute $\nabla_{\mathbf{F}_{k,u}^*} f$ using (16), for all k, u
  repeat
    Update $\mathbf{F}_{k,u}^{(q+1)}$ using (14), for all k, u
    Project $\mathbf{F}_{k,u}^{(q+1)}$ using (18), for all k, u
  until condition (20) holds
until $q > Q$ or $\left|\frac{f(\mathbf{F}^{(q+1)}, \mathbf{\Phi}^{(q+1)}) - f(\mathbf{F}^{(q)}, \mathbf{\Phi}^{(q)})}{f(\mathbf{F}^{(q+1)}, \mathbf{\Phi}^{(q+1)})}\right| < \varepsilon$

Therefore, any point obtained at the end of the algorithm, $f(\mathbf{F}^{(q)}, \boldsymbol{\Phi}^{(q)})$, will potentially be a good solution.

The complexity of the BD-RIS matrix design is primarily influenced by the SVD required for projection (19) (line 5 of algorithm 1), which has a complexity order of $\mathcal{O}(N_{ris}^3)$. Regarding the complexity of the precoder matrix design, it is primarily influenced by the computation of its gradient (line 7 of algorithm 1), which has a complexity of $\mathcal{O}(N_{tx}^2 N_{rx} N_u)$. By considering I_Φ as the number of line search loops for the RISs, algorithm 1 has a complexity order of $\mathcal{O}(Q(I_\Phi N_{pan} N_{ris}^3 + N_c N_u^2 N_{tx}^2 N_{rx}))$, which means that the complexity of the MU-RCBPG grows linearly with the number of RIS panels, subcarriers and receive antennas per user. An increase in complexity stems from the quadratic dependency on the number of users and transmit antennas at the BS. Due to the required projection of the BD-RIS matrices, the MU-RCBPG exhibits a cubic dependency on the number of RIS elements. This demonstrates that our approach does not increase the computational complexity compared to traditional SVD-based methods, yet it achieves improved performance, as will be seen later in section IV.

C. SPECIAL CASE OF D-RIS

D-RISs are a specific configuration of BD-RISs, also known as single-connected BD-RIS [7], where each element operates independently. Therefore, the i^{th} RIS matrix can be represented as $\boldsymbol{\Phi}_i = diag(\boldsymbol{\varphi}_i)$, with $\boldsymbol{\varphi}_i = [\varphi_{i,1}, \ldots, \varphi_{i,N_{ris}}]^T$ as the vector of phase shift values of each RIS element. In order to use the MU-RCBPG for a scenario with a D-RIS, we first need to redefine the set $\mathcal{R}$ as

$$\mathcal{R} = \left\{ \mathbf{r} \in \mathbb{C}^{N_{ris} \times 1} : |r_j| = a \quad , \quad j = 1, \ldots, N_{ris} \right\},$$

which no longer forces the RIS matrix to be a unitary matrix, but instead, given that it is a diagonal matrix, forces its diagonal to consist only of complex values with amplitude a.

For a given iteration, q, and gradient step, t, the RIS array response vector, $\boldsymbol{\varphi}_i$, can be updated using

$$\boldsymbol{\varphi}_i^{(q+1)} = \boldsymbol{\varphi}_i^{(q)} - t \nabla_{\boldsymbol{\varphi}_i^*} f(\mathbf{F}^{(q)}, \boldsymbol{\varphi}_i^{(q)}), \tag{21}$$

where $\nabla_{\boldsymbol{\varphi}_i^*} f$ is the gradient of f with respect to the conjugate of the RIS diagonal vector.

Lemma 2: Considering (12a), the gradient of f w.r.t. the conjugate of the i^{th} D-RIS vector, $\nabla_{\boldsymbol{\varphi}_i^*} f$, can be defined as

$$\begin{aligned} \nabla_{\boldsymbol{\varphi}_i^*} f = -\sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \frac{\rho_{k,u}}{\sigma^2} \Bigg(diag\Bigg[\left(\mathbf{H}_k^{R_i,U_u}\right)^H \mathbf{H}_{k,u} \mathbf{F}_{k,u} \\ \times \left(\mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma^2} \mathbf{F}_{k,u}^H \mathbf{H}_{k,u}^H \mathbf{H}_{k,u} \mathbf{F}_{k,u} \right)^{-1} \mathbf{F}_{k,u}^H \left(\mathbf{H}_k^{S,R_i}\right)^H \Bigg] \\ + \gamma \; diag \left[\left(\overline{\mathbf{H}}_k^{(R_i,U_u)}\right)^H \overline{\mathbf{H}}_{k,u} \mathbf{F}_{k,u} \mathbf{F}_{k,u}^H \left(\mathbf{H}_k^{(S,R_i)}\right)^H \right] \Bigg). \end{aligned} \tag{22}$$

Proof: see Appendix B.

The proximal operator for the redefined $\mathcal{R}$ can be calculated as an Euclidean projection described as

$$prox_{\propto \mathcal{I}_\mathcal{R}}(\mathbf{z}) = a(\mathbf{z} \oslash |\mathbf{z}|). \tag{23}$$

In summary, lines 2, 4, and 5 of algorithm 1 must be modified to reflect the new gradient and projection definitions using (22), (21), and (23), respectively.

The complexity of the D-RIS matrix design is primarily influenced by the computation of its gradient, as in (22), which has a complexity order of $\mathcal{O}(N_c N_u N_{ris} N_{tx} (N_u - 1) N_{rx})$. Therefore, the D-RIS version of algorithm 1 has complexity order of $\mathcal{O}(Q(N_{pan} N_c N_u N_{ris} N_{tx} (N_u - 1) N_{rx} + N_c N_u^2 N_{tx}^2 N_{rx}))$, which shows a linear dependency on the number of RIS panels, RIS elements, subcarriers and receive antennas, but a quadratic dependency on the number of users and transmit antennas at the BS.

D. SPECIAL CASE OF ACTIVE D-RIS

Conventional passive D-RISs are inherently limited by the absence of signal amplification, which presents challenges when it comes to enhancing link quality, particularly over long distances. Active D-RISs, which integrate active circuit components such as negative resistance elements or amplifiers within each reflecting unit, have been introduced as an advanced paradigm to overcome this limitation. These components enable the phase and amplitude of incident signals to be dynamically controlled. This active design not only enables adaptive beamforming and channel reconfiguration, but also compensates for path loss attenuation, significantly improving the signal-to-noise ratio (SNR) and spectral efficiency [43].

In the presence of active D-RISs, we must consider the noise amplification caused by the active elements themselves and redefine the model of the received signal in (1) as

$$\mathbf{r}_{k,u} = \sqrt{\rho_k} \mathbf{H}_{k,u} \mathbf{F}_k \mathbf{s}_k + \sum_{i=0}^{N_{pan}} \mathbf{H}_k^{R_i,U_u} \boldsymbol{\Phi}_i \mathbf{n}_{k,u}^{R_i} + \mathbf{n}_{k,u} \tag{24}$$

where vector $\mathbf{n}_{k,u}^{R_i} \in \mathbb{C}^{N_{ris} \times 1}$ represents the N_{ris} complex noise components at the i^{th} RIS. $\mathbf{n}_{k,u}^{R_i}$ is composed of two components: one related to the thermal noise captured by the RIS panel, which is assumed to have a power identical to the one at the receiver (σ_n^2); the other represents the noise generated internally by the RIS element components (e.g. amplifiers). The noise at the RIS is modeled by the following distribution

$$\mathbf{n}_{k,u}^{R_i} \sim \mathcal{CN}(0, F\sigma_n^2 \mathbf{I}_{N_{ris}}), \tag{25}$$

where F is the noise factor associated with the RIS. Therefore, the matrix given in (10), which contains the total noise plus interference (normalized by the noise σ_n^2), should be redefined as

$$\begin{aligned} \Psi_{k,u} = \mathbf{I}_{N_{rx}} + \sum_{\substack{u'=1 \\ u' \neq u}}^{N_u} \frac{\rho_{k,u'}}{\sigma_n^2} \mathbf{H}_{k,u} \mathbf{F}_{k,u'} \mathbf{F}_{k,u'}^H \mathbf{H}_{k,u}^H \\ + F \sum_{i=0}^{N_{pan}} \mathbf{H}_k^{R_i,U_u} \boldsymbol{\Phi}_i \boldsymbol{\Phi}_i^H \left(\mathbf{H}_k^{R_i,U_u}\right)^H. \end{aligned} \tag{26}$$

Considering active RISs with a fixed amplitude gain we have $\mathbf{\Phi}_i = diag(\boldsymbol{\varphi}_i)$, with $|\varphi_{i,1}| = \ldots = |\varphi_{i,N_{ris}}| = a$, where $a > 1$, therefore resulting in

$$\Psi_{k,u} = \mathbf{I}_{N_{rx}} + \sum_{\substack{u'=1 \\ u' \neq u}}^{N_u} \frac{\rho_{k,u'}}{\sigma_n^2} \mathbf{H}_{k,u} \mathbf{F}_{k,u'} \mathbf{F}_{k,u'}^H \mathbf{H}_{k,u}^H + a^2 F \sum_{i=1}^{N_{pan}} \mathbf{H}_k^{R_i,U_u} \left(\mathbf{H}_k^{R_i,U_u} \right)^H . \quad (27)$$

This adopted definition of the total interference plus noise matrix means that the system's sum-rate, as defined in (9), remains unchanged.

Disregarding inter-user interference, we can rewrite (27) as

$$\Psi_{k,u} = \mathbf{I}_{N_{rx}} + a^2 F \sum_{i=1}^{N_{pan}} \mathbf{H}_k^{R_i,U_u} \left(\mathbf{H}_k^{R_i,U_u} \right)^H , \quad (28)$$

and redefine the problem formulation in (11a) as

$$\min_{\mathbf{F},\mathbf{\Phi}} \quad - \sum_{k=1}^{N_c} \sum_{u=1}^{N_u} \ln \left| \mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma_n^2} \mathbf{F}_{k,u}^H \widetilde{\mathbf{H}}_{k,u}^H \widetilde{\mathbf{H}}_{k,u} \mathbf{F}_{k,u} \right| , \quad (29)$$

with $\widetilde{\mathbf{H}}_{k,u} = \mathbf{K}_{k,u}^H \mathbf{H}_{k,u}$, where $\mathbf{K}_{k,u}$ is a lower triangular matrix obtained from the Cholesky decomposition of $\mathbf{\Psi}^{-1}$, which is given as $\mathbf{\Psi}_{k,u}^{-1} = \mathbf{K}_{k,u} \mathbf{K}_{k,u}^H$. Given the problem representation in (29), the proposed algorithm can be applied directly by replacing the total channel matrix of each user in each subcarrier, $\mathbf{H}_{k,u}$, with the modified matrix, $\widetilde{\mathbf{H}}_{k,u}$.

This formulation adopted for active RIS corresponds to highly relevant practical scenarios, particularly when it is assumed that the power supplies to the BS and RISs are independent. However, it is important to note that other, more complex formulations may be adopted, particularly those in which it is assumed that the power supplied to the BS and RIS is shared. In this case, incorporating power budget and hardware consumption constraints [44] results in a different problem formulation, meaning that the proposed algorithm cannot be extended directly. Nevertheless, it is possible to apply an approach similar to that followed in this paper, resulting in a different, more complex algorithm that accounts for the additional constraints and variables.

IV. SIMULATION RESULTS

This section presents the simulation results and performance of the proposed algorithm and compares them with the performance of several algorithms from the literature in terms of achievable sum-rate. Additionally, we emphasize the potential performance gains of adding non-reciprocal fully connected BD-RIS and active/passive D-RIS panels to the UM-MIMO system by also benchmarking the results against two simpler transmission schemes: BD-precoding with no RIS panel and BD-precoding with one or multiple RIS panels without reconfiguration capabilities (''blind RIS''). Although channel estimation (CE) is a highly relevant topic in RIS-aided communications, several studies have already addressed CE schemes for conventional RIS-aided environments, such as [45], [46], [47], [48], and [49]. However, for BD-RIS-aided systems, the task of estimating the full channel matrices is considerably more complex and challenging, as well as being computationally expensive. Several works on BD-RIS-aided CE have recently been developed, including [50], [51], [52], and [53]. Therefore, we assume prior full channel state information (CSI) acquisition, and our main focus is on designing BD-RIS-assisted beamforming schemes that maximize the system's sum-rate.

The simulated scenario consists of one BS transmitting to a small area, such as a room or office, where one or more RISs are deployed. The RISs are more advantageous in scenarios involving LoS blockage and the strict presence of NLoS components. Therefore, it is assumed that the RISs are positioned so as to guarantee LoS links to users deployed in a region with obstructed direct links to the BS. It was also assumed that all RISs served all users within the zone without LoS to the BS. The operating frequency is set to 100GHz, with an occupied bandwidth of 480kHz per subcarrier. We apply path loss exponents of 2.05 and 4.6 for LOS and NLOS propagation, respectively, based on [54]. The number of propagation paths is set to $N_{ray} = 3$, and the azimuth and elevation AoD/AoA are uniformly random distributed in $(-\pi, \pi)$ and $(-\frac{\pi}{2}, \frac{\pi}{2})$, respectively. For the indirect link through each RIS, we apply $K_{rice} = 10$, while the direct link has no LOS component ($K_{rice} = 0$). The averaging of all simulated results is done over 100 random channel realizations. Simplification of the setup is achieved through simulation in a two-dimensional coordinate system, where the third coordinate is assumed to be equal and constant for the BS, RIS panels, and users. Considering $d_{S,U}$ as the distance between the BS and the center of the 3×3m room, the BS is assumed to be at $(0, 0)$, the 4 RIS panels at $(d_{S,U} - 0.5, 1.5)$, $(d_{S,U} - 0.5, -1.5)$, $(d_{S,U} + 0.5, 1.5)$ and $(d_{S,U} + 0.5, -1.5)$, and the users are randomly distributed throughout the room, either closer to or farther from the RISs. However, they always maintain the same distance from the BS. For instance, for a user at position (x, y), $d_{S,U} = \sqrt{x^2 + y^2}$ with $-1.5 < y < 1.5$ must hold.

Since the performance of our algorithm relies on the value chosen for the regularization term, γ, Fig. 2 aims to provide an overview of how the achievable sum rate varies in multi-carrier, multi-user, and multi-RIS scenarios with different γ values. Fig. 2a shows the evolution of the achievable sum-rate as the regularization factor varies from 0 to 10. As can be seen, using a regularization factor of 0 (i.e., not applying the regularization) results in a low achievable sum-rate, which increases substantially with the introduction of the regularization term; however, performance decreases considerably for values of γ above 0.5. This is an expected effect, as it implies an increase in the weight of the regularization term applied to each algorithm iteration and, consequently, placing greater importance on finding values that keep the precoders as orthogonal as possible, even if sacrificing some of the achievable sum-rate. Using a smaller sample interval when $0 \leq \gamma \leq 1$ enables us to conclude that, for this scenario,

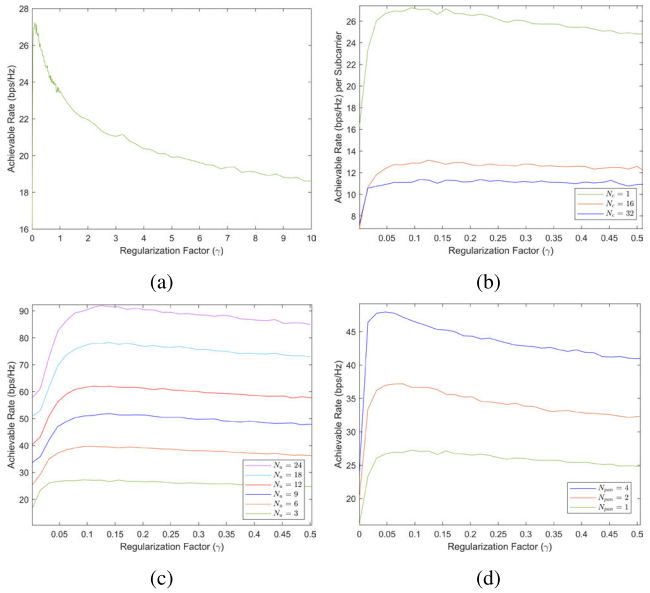

FIGURE 2. Performance of the proposed algorithm with different γ values. Scenario: $N_S = 2$, $N_{tx} = 256$, $N_{rx} = 4$, $N_{ris} = 128$, $P_{tx} = 30$ dBm, $d_{S,U} = 100$ m. (a) Benchmark with $N_U = 3$, $N_{pan} = 1$, $N_C = 1$, (b) Multi-carrier, (c) Multi-user, (d) Multi-RIS.

there is a ceiling of around 27 bps/Hz that is obtained with γ values between 0.05 and 0.17. Therefore, it is reasonable to assume that using a regularization factor value within this range will optimize the performance of this algorithm in this scenario. Fig. 2b, 2c, and 2d show the performance dependency on γ in multi-carrier, multi-user and multi-RIS scenarios, respectively. Fig. 2b shows the difference in performance with 1, 16, and 32 subcarriers. Note that the bandwidth of the transmitted signal varies with the number of subcarriers (480 kHz/subcarrier), and so does the power allocated for each subcarrier. Therefore, the direct comparison of the sum-rates between the three scenarios is unfair. However, as the number of subcarriers increases, it shows that the increase in the value of γ seems to have less impact. Assuming $0.05 \leq \gamma \leq 0.17$ (as previously estimated), it is relatively safe to conclude that the number of subcarriers is not a significant factor in selecting a suitable value for γ. Fig. 2c shows the difference in performance with various numbers of users, ranging from 3 to 24 users. As expected, the greater the number of users, the greater the impact of the value of γ on the achievable sum-rate. Fig. 2d shows that the range of gamma selections that produced the best results in this multi-user scenario tends to fall between 0.1 and 0.17. Fig. 2b shows the difference in performance with 1, 2, or 4 RIS panels. As expected, the achievable sum-rate differs, but the behavior of the three curves is generally the same. We mentioned that the best observed values of γ for the multi-carrier and multi-user scenarios with a single RIS panel tended to be between 0.1 and 0.17. However, this is not the case when we increase the number of panels. For scenarios with 2 and 4 RIS panels, performance is higher when selecting γ from the intervals [0.03, 0.09] and [0.03, 0.07], respectively. Since the single-RIS window, [0.1, 0.17], has no interception with the two multi-RIS windows, we can generalize γ by assigning a value that varies linearly with the number of RIS panels (N_{pan}). In this case, we propose the adoption of the value of $\gamma = 0.16/N_{pan}$ which yields 0.16, 0.08, and 0.04 for 1-, 2-, and 4-RIS scenarios, respectively, which fall within the best ranges [0.1, 0.17], [0.03, 0.09], and [0.03, 0.07], respectively. This behavior was observed in several other scenarios, so we defined the values of γ as $\gamma = 0.16/N_{pan}$. The same approach can be used to define γ in settings where these may not be appropriate.

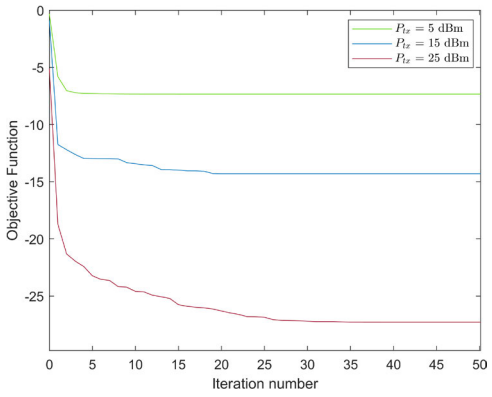

FIGURE 3. Convergence performance of the proposed algorithm. Scenario: $N_U = 3$, $N_S = 2$, $N_{tx} = 256$, $N_{rx} = 4$, $N_{ris} = 128$, $N_{pan} = 2$, $N_C = 1$, $d_{S,U} = 120$ m.

Fig. 3 illustrates the convergence performance of the proposed algorithm in the first 50 iterations. This scenario includes 3 users, 256 transmitting antennas at the BS, 2 streams, 4 receive antennas, 1 subcarrier per user, 2 BD-RIS panels with 128 reflecting elements, and a distance of 120 meters between the BS and the users. Note that the objective function value at iteration 0 is the result of computing the objective function with the initial values $\mathbf{F}^{(0)}$ and $\mathbf{\Phi}^{(0)}$. Fig. 3 confirms that the proposed algorithm converges in a relatively small number of iterations. This is significant given that the complexity of the MU-RCBPG grows linearly with the number of panels but cubically with the number of RIS elements. Furthermore, as seen in Fig. 3, the number of iterations required for convergence increases with the SNR. This behavior is expected in interference-limited multi-user systems [55], [56]. This is a consequence of the system transitioning from a noise-limited to an interference-limited regime. At a noise-limited regime (low SNR), the objective function is dominated by the desired signal power, leading to a relatively well-conditioned and 'smooth' optimization landscape. However, as the transmit power (and thus the SNR) increases, the noise floor ceases to act as a natural regularizer for the objective function's Hessian, and we consider the system to be in an interference-limited regime [55], [56]. The algorithm must then perform much higher-precision suppression of multi-user interference to reach a stationary point, as even marginal leakage becomes significant relative to the low noise power. This increased coupling between the user precoders and the resulting sensitivity of the sum-rate

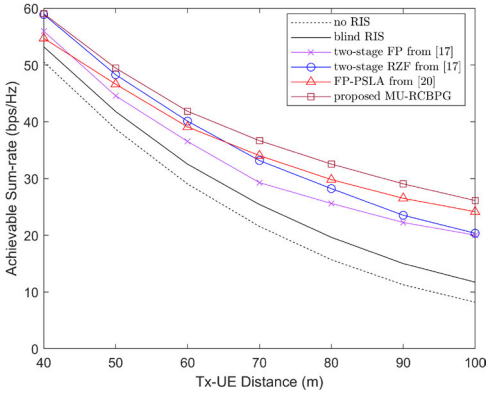

FIGURE 4. Comparison of BD-RIS algorithms' performances. Scenario: $N_u = 8$, $N_s = 1$, $N_{tx} = 256$, $N_{rx} = 1$, $N_{ris} = 128$, $N_{pan} = 1$, $N_c = 1$, $P_{tx} = 25$ dBm.

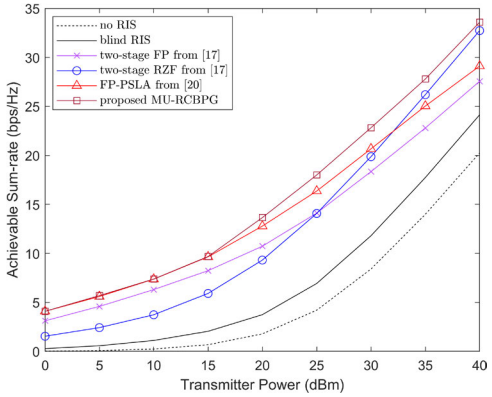

FIGURE 5. Comparison of BD-RIS algorithms' performances. Scenario: $N_u = 4$, $N_s = 1$, $N_{tx} = 256$, $N_{rx} = 1$, $N_{ris} = 128$, $N_{pan} = 1$, $N_c = 1$, $d_{S,U} = 100$ m.

to small phase shifts at the multiple RIS panels necessitates a higher number of iterative steps to achieve convergence. It is important to note that the convergence behavior of the proposed algorithm can depend on several system parameters, such as the number of users, transmit and receive antennas, RIS elements and panels, and subcarriers. Increasing these dimensions generally enlarges the optimization space and strengthens the coupling among optimization variables. This results in a more complex, non-convex landscape and potentially slower convergence. While larger antenna arrays and RIS deployments potentially increase the available degrees of freedom and achievable sum-rate, they also increase the dimensionality of the block-coordinate updates. Additionally, multi-carrier transmission introduces frequency coupling in the RIS optimization, which may impact convergence speed despite the fact that the per-iteration computational complexity scales linearly with the number of subcarriers.

Fig. 4 shows how the performance of the proposed approach varies with the distance between the BS and the users, compared to BD-precoding, the heuristic two-stage solutions to the WSR maximization problem, proposed in [17], which use fractional programming (FP) and regularized zero forcing (RZF), and the FP-based projected successive linear approximation (FP-PSLA) algorithm, proposed in [20]. These algorithms only consider single-RIS, single-antenna, and single-carrier users. Therefore, in order to compare the proposed algorithm against those benchmark algorithms, we use a simpler scenario which includes 8 users, 256 transmitting antennas at the BS, one receive antenna and one subcarrier, a single BD-RIS panel at $(d_{S,U} - 0.5, 1.5)$ with 128 reflecting elements, and a total transmitting power of 25 dBm. Even though this is a scenario for which the benchmark algorithms were designed, our approach is able to achieve higher sum-rates for all simulated distances. As the distance between the BS and the users increases, the gain resulting from our approach also increases compared to the BD-precoding and the two algorithms from [17], thus proving its effectiveness for medium and long transmission distances. The gain resulting from using our algorithm appears to be consistent with that of the FP-PSLA from [20], indicating similar behavior despite our solution's superior performance. Adding a single BD-RIS panel represents an 8% and 43% gain, respectively, for distances of 50m and 100m when considering a link with a blind RIS. For the same distances, our approach can reach values representing gains of 28% and 218%, respectively, when compared to a scenario without RIS panels, which is a considerable performance increase.

Fig. 5 illustrates how the performance of the proposed approach varies with the total transmitted power, compared to the same benchmark cases of the previous figure. This scenario includes 4 users, 256 transmitting antennas at the BS, one receive antenna, one subcarrier, a single BD-RIS panel with 128 reflecting elements, and a distance of 100 meters between the BS and the users. Adding a blind RIS to the system enables the sum-rate to reach values representing a 40% gain, for 30 dBm of transmitted power, compared to the scenario without RIS panels. Considering the performance of the BD-RIS algorithms with 25 dBm of total transmitted power, the algorithms from [17], [20], and the proposed MU-RCBPG achieve gains of 103%, 136% and 160%, respectively, compared to the blind RIS. In low-power scenarios, the FP-PSLA and the proposed approach demonstrate equivalent performance. However, our approach outperforms the FP-PSLA when power increases, possibly due to MU-RCBPG's ability to mitigate inter-user interference, which becomes more challenging in scenarios with increased signal power. Overall, the proposed approach demonstrates superior performance in the simulated scenario.

Fig. 6 shows how the multi-carrier performance of the proposed approach with a single BD-RIS and a single D-RIS varies with the total transmitted power per stream and per subcarrier compared to BD-precoding. This scenario includes 4 users, 256 transmitting antennas at the BS, 4 streams, 8 receive antennas, and 16 subcarriers per user, a single RIS panel with 128 reflecting elements, and a distance of 120 meters between the BS and the users. For scenarios in which multiple subcarriers are transmitted simultaneously, our approach demonstrates significant improvements compared to the blind RIS case. With 0 dBm of total transmitted power per stream and per subcarrier, the proposed

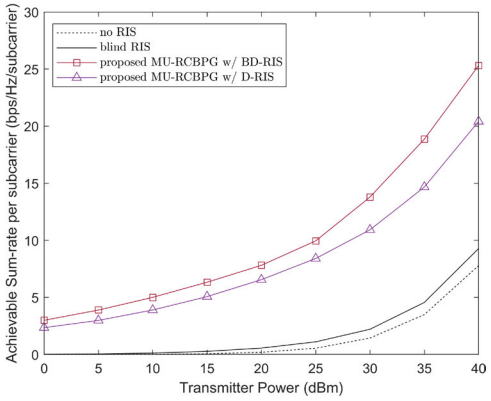

FIGURE 6. Comparison of proposed algorithm performance with multiple subcarriers using BD-RIS and D-RIS. Scenario: $N_U = 4$, $N_S = 4$, $N_{tx} = 256$, $N_{rx} = 8$, $N_{ris} = 128$, $N_{pan} = 1$, $N_C = 16$, $d_{S,U} = 120$ m.

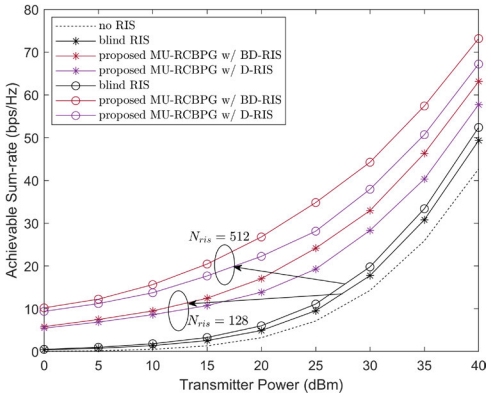

FIGURE 7. Comparison of proposed algorithm performance with BD-RIS and D-RIS. Scenario: $N_U = 4$, $N_S = 4$, $N_{tx} = 256$, $N_{rx} = 4$, $N_{ris} = [128 - 512]$, $N_{pan} = 1$, $N_C = 1$, $d_{S,U} = 100$ m.

MU-RCBPG can achieve 2.35 and 2.99 bps/Hz/subcarrier, for scenarios assisted by a single D-RIS and BD-RIS panel, respectively. With a more realistic power constraint of 25 dBm, the MU-RCBPG achieves rates that are 6.6x and 9x higher using a D-RIS and a BD-RIS, respectively, when compared to the blind RIS. As expected, MU-RCBPG with BD-RIS achieves higher sum-rates than with D-RIS, demonstrating the benefits of using fully connected element structures, like BD-RISs, over independent reflecting elements, such as D-RISs, in multi-carrier scenarios. However, it must be emphasized that this performance improvement comes at the cost of an increase in implementation complexity and computing costs. This trade-off between performance gains and complexity when using BD-RIS over D-RIS has also been addressed in many recent works as [57], [58], and [59].

Fig. 7 shows how the performance of the proposed approach with a single BD-RIS and a single D-RIS varies with the total transmitted power per stream. This scenario includes 4 users, 256 transmitting antennas at the BS, 4 streams, 4 receive antennas, and one subcarrier per user, and a distance of 100 meters between the BS and the users.

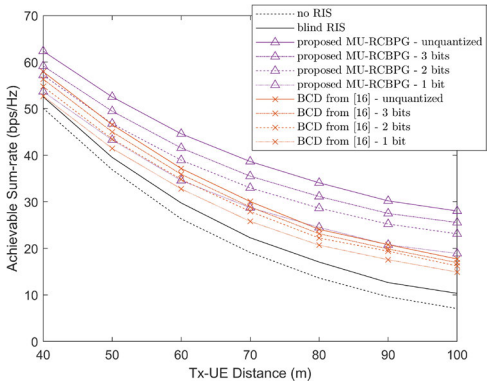

FIGURE 8. Comparison of D-RIS algorithms' performances. Scenario: $N_U = 3$, $N_S = 3$, $N_{tx} = 256$, $N_{rx} = 8$, $N_{ris} = 512$, $N_{pan} = 1$, $N_C = 1$, $P_{tx} = 25$ dBm.

To analyze the impact of having a larger number of RIS elements in each RIS panel, we added curves with 128 and 512 elements. As expected, the more elements that comprise the RISs, the better the system's performance in terms of achievable sum-rate. Although it achieves better results than the scenario without RIS panels, the blind RIS yields low gains compared to those attained by the MU-RCBPG. For settings with higher transmitted power per stream, the MU-RCBPG with a 512-element D-RIS achieves a sum-rate 5% higher than that of the 128-element BD-RIS. This demonstrates that a BD-RIS with fewer elements can perform similarly to a larger D-RIS. It also underscores the importance of considering the trade-off between performance and complexity when selecting the type and size of RISs.

Fig. 8 demonstrates how the performance of the proposed approach with a single D-RIS varies with the distance between the BS and the users, compared to the no RIS case, the blind RIS case, and the classic block coordinate descent (BCD) algorithm proposed in [16]. This scenario includes 3 users, 256 transmitting antennas at the BS, 3 streams, 8 receive antennas, and one subcarrier per user, a single D-RIS panel with 512 reflecting elements, and a total transmitting power of 25 dBm. To further analyze the performance of both algorithms, curves were added for unquantized D-RIS and 1-, 2-, and 3-bit quantized D-RIS. For N_b bits, we assume that each element has 2^{N_b} possible values to represent the optimized phase shift, which is calculated by the algorithms. As expected, the greater the distance between the BS and the users, the greater the gain and advantage of adding RISs to the system. However, the systems' performance decreases as the distance between the BS and the users increases, but the gains associated with the use of joint optimization algorithms become more pronounced. As can be seen in Fig. 8, quantizing the phases of the RIS elements can have some impact on the system's performance both for MU-RCBPG and BCD. For example, 3 quantization bits can result in losses above 10%, even for smaller distances. As expected, the greater the number of bits available for quantization, the greater the achievable sum-rate. A good trade-off between performance and complexity should also be considered when designing

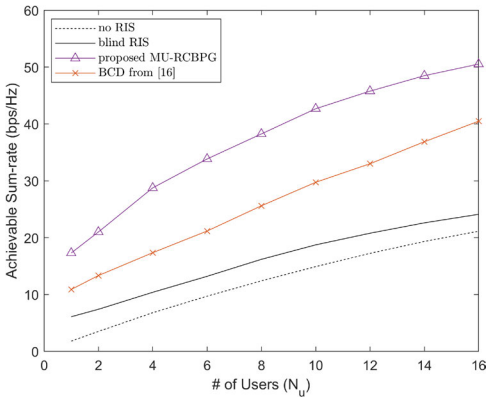

FIGURE 9. Comparison of D-RIS algorithms' performances. Scenario: $N_S = 2$, $N_{tx} = 256$, $N_{rx} = 4$, $N_{ris} = 512$, $N_{pan} = 1$, $N_C = 1$, $P_{tx} = 25$ dBm, $d_{S,U} = 100$ m.

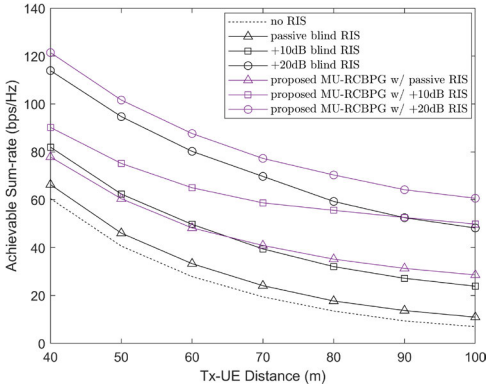

FIGURE 10. Comparison of proposed algorithm performance with passive and active D-RISs. Scenario: $N_U = 4$, $N_S = 4$, $N_{tx} = 256$, $N_{rx} = 4$, $N_{ris} = 512$, $N_{pan} = 1$, $N_C = 1$, $P_{tx} = 25$ dBm.

quantized RIS-aided multiuser MIMO systems. Nevertheless, figure 8 shows that the proposed MU-RCBPG outperforms the unquantized benchmark cases, even with 2- or 3-bit quantization.

Fig. 9 demonstrates how the performance of the proposed approach with a single D-RIS varies with the number of users. This scenario includes 256 transmitting antennas at the BS, 2 streams, 4 receive antennas, and one subcarrier per user, a single D-RIS panel with 512 reflecting elements, a total transmitting power of 25 dBm, and a distance of 100 meters between the BS and the users. Fig. 9 not only demonstrates how the system achieves higher sum-rates when using MU-RCBPG than BCD in multi-user scenarios, but it also shows that it outperforms the benchmark cases in a single-user scenario. With a blind RIS and without RIS, the system achieves rates of 6.09 and 1.80 bps/Hz, respectively. The BCD algorithm and the proposed MU-RCBPG achieve 10.89 and 17.30 bps/Hz, representing gains of 79% and 184%, respectively, compared to the blind RIS. The greater the number of users, the greater the gain from using the joint optimization algorithms compared to no RIS and blind RIS implementations.

Fig. 10 demonstrates how the performance of the proposed approach varies with the distance between the BS and the

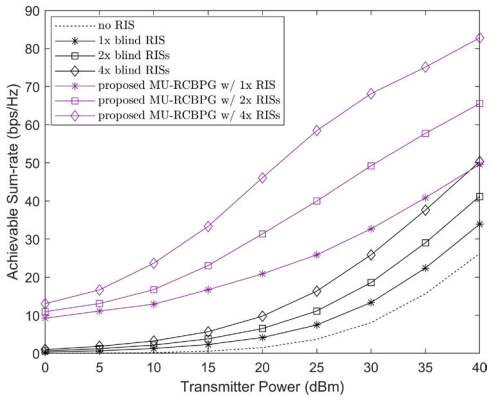

FIGURE 11. Comparison of proposed algorithm performance with multiple D-RISs. Scenario: $N_U = 4$, $N_S = 2$, $N_{tx} = 256$, $N_{rx} = 4$, $N_{ris} = 512$, $N_{pan} = 1, 2, 4$, $N_C = 1$, $d_{S,U} = 120$ m.

users when assisted by a single active D-RIS with gains of +10 and +20 dB. The amplification of the active RIS is assumed to be element-wise, meaning every RIS element has a gain of 10 or 20 dB between the incident (incoming) and the reflected (outgoing) signals. The noise model used is the one described in III-III-D with a noise factor of 5 dB. To evaluate the performance increase, we include the no RIS, blind RIS, and passive RIS cases. This scenario includes 4 users, 256 transmitting antennas at the BS, 4 streams, 4 receive antennas, and one subcarrier per user, a single passive/active D-RIS panel with 512 reflecting elements, and a total transmitting power of 25 dBm. As expected, the greater the gain in each RIS element, the greater the system's performance. For a distance of 100 meters between the BS and the users, the proposed MU-RCBPG achieves sum-rates that are 2.61x (passive), 2.09x (10 dB active), and 1.26x (20 dB active) higher than those of blind RIS cases. Although greater amplification results in greater noise introduced by the RIS, the proposed approach outperforms blind RIS cases in every considered scenario. Additionally, the performance of MU-RCBPG with a passive RIS is similar to that of a 10 dB amplified blind RIS, which further emphasizes the benefits of using MU-RCBPG to simultaneously optimize both the precoder at the BS and the phase shift matrix of the RISs.

Fig. 11 shows how the performance of the proposed approach with multiple D-RISs varies with the total transmitted power per stream. To analyze the impact of having multiple RIS panels, we added curves with one, two, and four RISs. In the single RIS case, the RIS panel is assumed to be located at $(d_{S,U} - 0.5, 1.5)$. Then, we added a second RIS at $(d_{S,U} - 0.5, -1.5)$, followed by two more panels at $(d_{S,U} + 0.5, 1.5)$ and $(d_{S,U} + 0.5, -1.5)$. This scenario includes 4 users, 256 transmitting antennas at the BS, 2 streams, 4 receive antennas, and one subcarrier, and a distance of 120 meters between the BS and the users. It is clear that adding more RIS panels increases the system's performance. With 20 dBm of transmitted power per stream, the transmission without a RIS achieves a sum-rate of 1.51 bps/Hz. The case with one, two, and four blind RIS panels achieves, 4.15,

6.50, and 9.76 bps/Hz, respectively. For the same scenario, the MU-RCBPG achieves 20.86, 31.35, and 46.07 bps/Hz, representing gains of 403%, 382%, and 372%, respectively. Furthermore, the MU-RCBPG with two and four RIS panels achieves a sum-rate that is 1.50x and 2.21x higher than the single D-RIS scenario, respectively. It is important to note that the MU-RCBPG with a single D-RIS outperforms the approach with four blind RISs. This again shows the massive impact of using the MU-RCBPG algorithm, instead of mirror RISs with no reconfiguration capabilities.

V. CONCLUSION

In this paper, we addressed the joint design of the transmit precoder and RIS phases in the context of multi-stream, multi-carrier, multi-user MIMO communications aided by multiple parallel BD-RIS panels. Aiming at the maximization of the sum-rate while minimizing inter-user interference, we proposed a problem formulation that includes a regularization term and derived a joint optimization algorithm based on the CBPG method. The algorithm can be tailored to multiple BD-RISs, D-RISs, and passive and active D-RISs. The regularization term was shown to drive the algorithm toward a suboptimal solution with minimal inter-user interference. We presented a detailed performance analysis that showed that the proposed MU-RCBPG algorithm often achieves higher sum-rates than other benchmarked approaches. Additionally, it was demonstrated that the fully interconnected elements structure (BD-RIS) considerably outperforms the independent elements structure (D-RIS). Furthermore, the use of multiple parallel RIS panels and adding amplification to RIS elements was shown to substantially impact the system's performance. Finally, the proposed framework offers a scalable and efficient solution for next-generation wireless networks operating in mmWave/THz frequency bands.

APPENDIX
PROOF OF LEMMA 1 AND 2
A. PROOF OF LEMMA 1

As stated before, (12a) can be used to approximately maximize the total achievable sum-rate comprising N_u users transmitting over N_c subcarriers. Its calculation depends on N_cN_u contributions, totally independent of each other, which means that for each pair of subcarriers and users, (k, u), there is a distinct gradient of the function f with respect to the precoder matrix, $\mathbf{F}_{k,u}$. To simplify the derivation, we can represent the gradient of f as the sum of the gradients of the two terms in (12a), $\nabla_{\mathbf{F}_{k,u},\mathbf{F}_{k,u}^*} f = \nabla_{\mathbf{F}_{k,u},\mathbf{F}_{k,u}^*} f_1 + \nabla_{\mathbf{F}_{k,u},\mathbf{F}_{k,u}^*} f_2$. By starting with the first term, f_1, and considering the equality $d\left[\ln\det(X)\right] = \mathrm{Tr}\left\{X^{-1}dX\right\}$ [60], the differential $df_{1\mathbf{F}_{k,u},\mathbf{F}_{k,u}^*}$ can be described as

$$df_{1\mathbf{F}_{k,u},\mathbf{F}_{k,u}^*} = -\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times \frac{\rho_{k,u}}{\sigma^2}\left(\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}d\mathbf{F}_{k,u} + d\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)\right\}.$$

By separating both terms according to $\mathrm{Tr}\{A+B\} = \mathrm{Tr}\{A\} + \mathrm{Tr}\{B\}$, the two differentials can be calculated separately as

$$df_{1\mathbf{F}_{k,u}} = -\frac{\rho_{k,u}}{\sigma^2}\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times \mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}d\mathbf{F}_{k,u}\right\},$$

and

$$\begin{aligned} df_{1\mathbf{F}_{k,u}^*} &= -\frac{\rho_{k,u}}{\sigma^2}\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times d\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= -\frac{\rho_{k,u}}{\sigma^2}\mathrm{Tr}\left\{\left[\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \times \left(\mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}\right]^T d\mathbf{F}_{k,u}^*\right\}. \end{aligned}$$

Using the definition $df_{Z,Z^*} = \mathrm{Tr}\left\{\mathbf{A}_0^T d\mathbf{Z} + \mathbf{A}_1^T d\mathbf{Z}^*\right\}$ [60], the gradient of f_1 with respect to $\mathbf{F}_{k,u}^*$ can be written as

$$\nabla_{\mathbf{F}_{k,u}^*} f_1 = -\frac{\rho_{k,u}}{\sigma^2}\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \times \left(\mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}. \tag{30}$$

Regarding the second term of (12a), referred to as f_2, we can separate both differentials with respect to $\mathbf{F}_{k,u}$ and $\mathbf{F}_{k,u}^*$, and obtain

$$df_{2\mathbf{F}_{k,u}} = \gamma\frac{\rho_{k,u}}{\sigma^2}\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}d\mathbf{F}_{k,u}\right\},$$

and

$$\begin{aligned} df_{2\mathbf{F}_{k,u}^*} &= \gamma\frac{\rho_{k,u}}{\sigma^2}\mathrm{Tr}\left\{d\mathbf{F}_{k,u}^H\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}\mathrm{Tr}\left\{\left[\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right]^T d\mathbf{F}_{k,u}^*\right\}. \end{aligned}$$

Using the definition $df_{Z,Z^*} = \mathrm{Tr}\left\{\mathbf{A}_0^T d\mathbf{Z} + \mathbf{A}_1^T d\mathbf{Z}^*\right\}$ [60], the gradient of f_2 with respect to $\mathbf{F}_{k,u}^*$ can be written as

$$\nabla_{\mathbf{F}_{k,u}^*} f_2 = \gamma\frac{\rho_{k,u}}{\sigma^2}\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}. \tag{31}$$

By summing (30) and (31) we obtain the following equality:

$$\nabla_{\mathbf{F}_{k,u}^*} f = -\frac{\rho_{k,u}}{\sigma^2}\left[\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \times \left(\mathbf{I}_{N_s} + \frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} + \gamma\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right]. \tag{32}$$

The idea of subcarrier- and user-independent derivatives does not hold for the optimization of RIS matrices, because it relies on the contribution of all precoders from all the subcarriers of all users. Therefore, the gradient of f with respect to each RIS matrix $\mathbf{\Phi}_i$ can be represented as $\nabla_{\mathbf{\Phi}_i,\mathbf{\Phi}_i^*} f = \sum_{k=1}^{N_c}\sum_{u=1}^{N_u}\left[\nabla_{\mathbf{\Phi}_i,\mathbf{\Phi}_i^*} f_1 + \nabla_{\mathbf{\Phi}_i,\mathbf{\Phi}_i^*} f_2\right]$. By starting with the

first term, f_1, and considering the equality $d\left[\ln\det\left(X\right)\right] = \mathrm{Tr}\left\{X^{-1}dX\right\}$ [60], the differential $df_{1\,\boldsymbol{\Phi},\boldsymbol{\Phi}_i^*}$ can be written as

$$df_{1\,\boldsymbol{\Phi}_i,\boldsymbol{\Phi}_i^*} = -\,\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\frac{\rho_{k,u}}{\sigma^2}\left(\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H d\mathbf{H}_{k,u}\mathbf{F}_{k,u}\ +\mathbf{F}_{k,u}^H d\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)\right\}.$$

By separating both terms according to $\mathrm{Tr}\{A+B\} = \mathrm{Tr}\{A\} + \mathrm{Tr}\{B\}$, both differentials can be calculated separately as

$$\begin{aligned} df_{1\,\boldsymbol{\Phi}_i} &= -\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\left(\mathbf{H}_k^{R_i,U_u} d\boldsymbol{\Phi}_i\mathbf{H}_k^{S,R_i}\right)\mathbf{F}_{k,u}\right\} \\ &= -\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\Bigg\{\mathbf{H}_k^{S,R_i}\mathbf{F}_{k,u} \times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_k^{R_i,U_u} d\boldsymbol{\Phi}_i\Bigg\}, \end{aligned}$$

and

$$\begin{aligned} df_{1\,\boldsymbol{\Phi}_i^*} &= -\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\mathbf{F}_{k,u}^H\left(\left(\mathbf{H}_k^{S,R_i}\right)^H d\boldsymbol{\Phi}_i^H\left(\mathbf{H}_k^{R_i,U_u}\right)^H\right)\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= -\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\Bigg\{\left[\left(\mathbf{H}_k^{R_i,U_u}\right)^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H\right]^T d\boldsymbol{\Phi}_i^*\Bigg\}, \end{aligned}$$

thus resulting in

$$\nabla_{\boldsymbol{\Phi}_i^*} f_1 = -\frac{\rho_{k,u}}{\sigma^2}\left(\mathbf{H}_k^{R_i,U_u}\right)^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H. \tag{33}$$

The gradient of the second term, f_2, with respect to $\boldsymbol{\Phi}_i$ and $\boldsymbol{\Phi}_i^*$, $\nabla_{\boldsymbol{\Phi}_i,\boldsymbol{\Phi}_i^*} f_2$, can be represented as

$$\nabla_{\boldsymbol{\Phi}_i,\boldsymbol{\Phi}_i^*} f_2 = \frac{df_2}{d_{\boldsymbol{\Phi}_i}}\left(\gamma\|\frac{\sqrt{\rho_{k,u}}}{\sigma}\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\|_F^2\right) + \frac{df_2}{d_{\boldsymbol{\Phi}_i^*}}\left(\gamma\|\frac{\sqrt{\rho_{k,u}}}{\sigma}\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\|_F^2\right),$$

and separating both differentials, we have

$$\begin{aligned} df_{2\,\boldsymbol{\Phi}_i} &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H\overline{\mathbf{H}}_{k,u}^H d\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H\overline{\mathbf{H}}_{k,u}^H\left(\overline{\mathbf{H}}_k^{R_i,U_u} d\boldsymbol{\Phi}_i\overline{\mathbf{H}}_k^{S,R_i}\right)\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\overline{\mathbf{H}}_k^{S,R_i}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_k^{R_i,U_u} d\boldsymbol{\Phi}_i\right\}, \end{aligned}$$

and

$$\begin{aligned} df_{2\,\boldsymbol{\Phi}_i^*} &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H d\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H \times\left(\left(\overline{\mathbf{H}}_k^{S,R_i}\right)^H d\boldsymbol{\Phi}_i{}^H\left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\right)\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\Bigg\{\left[\left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\overline{\mathbf{H}}_{k,u} \times\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\left(\overline{\mathbf{H}}_k^{S,R_i}\right)^H\right]^T d\boldsymbol{\Phi}_i{}^*\Bigg\}, \end{aligned}$$

which shows that

$$\nabla_{\boldsymbol{\Phi}_i{}^*} f_2 = \gamma\frac{\rho_{k,u}}{\sigma^2}\left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\left(\overline{\mathbf{H}}_k^{S,R_i}\right)^H. \tag{34}$$

By considering every contribution of every user and subcarrier, and by joining (33) and (34), we obtain the following equality:

$$\nabla_{\boldsymbol{\Phi}_i^*} f = -\sum_{k=1}^{N_c}\sum_{u=1}^{N_u}\frac{\rho_{k,u}}{\sigma^2}\Bigg[\left(\mathbf{H}_k^{R_i,U_u}\right)^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H +\gamma\ \left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H\Bigg]. \tag{35}$$

B. PROOF OF LEMMA 2

In the case of D-RISs, the RIS elements are not connected and, therefore, the gradient of f can be defined with respect to each vector $\boldsymbol{\varphi}_i$. Separating both terms of (12a), the gradient of f with respect to each vector $\boldsymbol{\varphi}_i$ can be represented as the sum of the gradients of both terms, $\nabla_{\boldsymbol{\varphi}_i,\boldsymbol{\varphi}_i^*} f = \sum_{k=1}^{N_c}\sum_{u=1}^{N_u}\left[\nabla_{\boldsymbol{\varphi}_i,\boldsymbol{\varphi}_i^*} f_1 + \nabla_{\boldsymbol{\varphi}_i,\boldsymbol{\varphi}_i^*} f_2\right]$. By starting with the first term, f_1, and considering the equality $d\left[\ln\det\left(X\right)\right] = \mathrm{Tr}\left\{X^{-1}dX\right\}$ [60], $df_{1\,\boldsymbol{\varphi},\boldsymbol{\varphi}_i^*}$ can be written as

$$df_{1\,\boldsymbol{\varphi}_i,\boldsymbol{\varphi}_i^*} = -\,\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\left(\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H d\mathbf{H}_{k,u}\mathbf{F}_{k,u}\ +\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H d\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)\right\}.$$

By separating both terms according to $\mathrm{Tr}\{A+B\} = \mathrm{Tr}\{A\} + \mathrm{Tr}\{B\}$, both differentials can be calculated separately as

$$\begin{aligned} df_{1\,\boldsymbol{\varphi}_i} &= -\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\left(\mathbf{H}_k^{R_i,U_u} diag(d\boldsymbol{\varphi}_i)\mathbf{H}_k^{S,R_i}\right)\mathbf{F}_{k,u}\right\} \\ &= -\frac{\rho_{k,u}}{\sigma^2} diag\Bigg[\mathbf{H}_k^{S,R_i}\mathbf{F}_{k,u} \times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \times\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_k^{R_i,U_u}\Bigg]^T d\boldsymbol{\varphi}_i, \end{aligned}$$

and

$$\begin{aligned} df_{1\boldsymbol{\varphi}_i^*} &= -\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}\right. \\ &\quad\times \mathbf{F}_{k,u}^H\left(\left(\mathbf{H}_k^{S,R_i}\right)^H diag(d\boldsymbol{\varphi}_i^*)\left(\mathbf{H}_k^{R_i,U_u}\right)^H\right) \\ &\quad\left.\times \mathbf{H}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= -\frac{\rho_{k,u}}{\sigma^2}diag\Bigg[\left(\mathbf{H}_k^{R_i,U_u}\right)^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \\ &\quad\times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1} \\ &\quad\times\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H\Bigg]^T d\boldsymbol{\varphi}_i^*, \end{aligned}$$

thus resulting in

$$\begin{aligned} \nabla_{\boldsymbol{\varphi}_i^*}f_1 = -\frac{\rho_{k,u}}{\sigma^2}diag\Bigg[&\left(\mathbf{H}_k^{R_i,U_u}\right)^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \\ &\times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H\Bigg]. \end{aligned} \tag{36}$$

The gradient of the second term, f_2, with respect to $\boldsymbol{\varphi}_i$ and $\boldsymbol{\varphi}_i^*$, $\nabla_{\boldsymbol{\varphi}_i,\boldsymbol{\varphi}_i^*}f_2$, can be represented as

$$\begin{aligned} \nabla_{\boldsymbol{\varphi}_i,\boldsymbol{\varphi}_i^*}f_2 = &\frac{df_2}{d_{\boldsymbol{\varphi}_i}}\left(\gamma\|\frac{\sqrt{\rho_{k,u}}}{\sigma}\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\|_F^2\right) \\ &+\frac{df_2}{d_{\boldsymbol{\varphi}_i^*}}\left(\gamma\|\frac{\sqrt{\rho_{k,u}}}{\sigma}\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\|_F^2\right), \end{aligned}$$

and separating both differentials, we have

$$\begin{aligned} df_{2\boldsymbol{\varphi}_i} &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H\overline{\mathbf{H}}_{k,u}^H d\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H\overline{\mathbf{H}}_{k,u}^H\right. \\ &\quad\left.\times\left(\overline{\mathbf{H}}_k^{R_i,U_u}diag(d\boldsymbol{\varphi}_i)\overline{\mathbf{H}}_k^{S,R_i}\right)\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}diag\left[\overline{\mathbf{H}}_k^{S,R_i}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\right. \\ &\quad\left.\times\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_k^{R_i,U_u}\right]^T d\boldsymbol{\varphi}_i, \end{aligned}$$

and

$$\begin{aligned} df_{2\boldsymbol{\varphi}_i^*} &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\left\{\mathbf{F}_{k,u}^H d\overline{\mathbf{H}}_{k,u}^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\right\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}\,\mathrm{Tr}\Bigg\{\mathbf{F}_{k,u}^H \\ &\quad\times\left(\left(\overline{\mathbf{H}}_k^{S,R_i}\right)^H diag(d\boldsymbol{\varphi}_i{}^*)\left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\right) \\ &\quad\times\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\Bigg\} \\ &= \gamma\frac{\rho_{k,u}}{\sigma^2}diag\Bigg[\left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\overline{\mathbf{H}}_{k,u} \\ &\quad\times\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\left(\overline{\mathbf{H}}_k^{S,R_i}\right)^H\Bigg]^T d\boldsymbol{\varphi}_i{}^*, \end{aligned}$$

which shows that

$$\begin{aligned} \nabla_{\boldsymbol{\varphi}_i{}^*}f_2 = \gamma\frac{\rho_{k,u}}{\sigma^2}diag\Bigg[&\left(\overline{\mathbf{H}}_k^{R_i,U_u}\right)^H\overline{\mathbf{H}}_{k,u} \\ &\times\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\left(\overline{\mathbf{H}}_k^{S,R_i}\right)^H\Bigg]. \end{aligned} \tag{37}$$

By considering every contribution of every user and subcarrier, and by joining (36) and (37), we obtain the following equality:

$$\begin{aligned} \nabla_{\boldsymbol{\varphi}_i^*}f = -\sum_{k=1}^{N_c}\sum_{u=1}^{N_u}\frac{\rho_{k,u}}{\sigma^2}\Bigg(&diag\Bigg[\left(\mathbf{H}_k^{R_i,U_u}\right)^H\mathbf{H}_{k,u}\mathbf{F}_{k,u} \\ &\times\left(\mathbf{I}_{N_s}+\frac{\rho_{k,u}}{\sigma^2}\mathbf{F}_{k,u}^H\mathbf{H}_{k,u}^H\mathbf{H}_{k,u}\mathbf{F}_{k,u}\right)^{-1}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{S,R_i}\right)^H\Bigg] \\ &+\gamma\ diag\left[\left(\overline{\mathbf{H}}_k^{(R_i,U_u)}\right)^H\overline{\mathbf{H}}_{k,u}\mathbf{F}_{k,u}\mathbf{F}_{k,u}^H\left(\mathbf{H}_k^{(S,R_i)}\right)^H\right]\Bigg). \end{aligned} \tag{38}$$

REFERENCES

[1] I. F. Akyildiz, C. Han, and S. Nie, "Combating the distance problem in the millimeter wave and terahertz frequency bands," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 102–108, Jun. 2018.

[2] (2015). *Recommendation ITU-R M.2083-0 IMT Vision-Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond M Series Mobile, Radiodetermination, Amateur and Related Satellite Services*. [Online]. Available: http://www.itu.int/ITU-R/go/patents/en

[3] S. R. Sabuj et al., "Machine-type communications in NOMA-based terahertz wireless networks," *Int. J. Intell. Netw.*, vol. 3, pp. 31–47, May 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666603022000045

[4] (2020). *The Next Hyper Connected Experience for All*. [Online]. Available: https://research.samsung.com/next-generation-communications#6gPop

[5] S. Dash, C. Psomas, I. Krikidis, I. F. Akyildiz, and A. Pitsillides, "Active control of THz waves in wireless environments using graphene-based RIS," *IEEE Trans. Antennas Propag.*, vol. 70, no. 10, pp. 8785–8797, Oct. 2022.

[6] C. Han, Y. Wu, Z. Chen, and X. Wang, "Terahertz communications (teracom): Challenges and impact on 6G wireless systems," 2019, *arXiv:1912.06040*.

[7] O. Maraqa, M. H. Khoshafa, O. O. Oyerinde, and T. M. N. Ngatched, "Beyond diagonal RIS-aided wireless communications systems: State-of-the-art and future research directions," 2025, *arXiv:2503.08826*.

[8] Y. Liu et al., "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd Quart., 2021.

[9] Z. Wu and B. Clerckx, "Optimization of beyond diagonal RIS: A universal framework applicable to arbitrary architectures," 2024, *arXiv:2412.15965*.

[10] N. Souto and J. C. Silva, "Joint beamforming algorithm for multi-stream MIMO systems assisted by multiple reconfigurable intelligent surfaces," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1317–1333, 2023.

[11] N. Souto, "Joint active and passive beamforming for RIS-aided MIMO communications with low-resolution phase shifts," *IEEE Commun. Lett.*, vol. 27, no. 6, pp. 1604–1608, Jun. 2023.

[12] V. R. J. Velez, J. P. C. B. B. Pavia, N. M. B. Souto, P. J. A. Sebastião, and A. M. C. Correia, "Performance assessment of a RIS-empowered post-5G/6G network operating at the mmWave/THz bands," *IEEE Access*, vol. 11, pp. 49625–49638, 2023.

[13] Z. Zhang and L. Dai, "A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network," *IEEE Trans. Signal Process.*, vol. 69, pp. 4085–4101, 2021.

[14] N. S. Perovic, L.-N. Tran, M. D. Renzo, and M. F. Flanagan, "On the maximum achievable sum-rate of the RIS-aided MIMO broadcast channel," *IEEE Trans. Signal Process.*, vol. 70, pp. 6316–6331, 2022.

[15] H. D. Tuan, A. A. Nasir, E. Dutkiewicz, H. V. Poor, and L. Hanzo, "Active-RIS enhances the multi-user rate of multi-carrier communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 11, pp. 16948–16963, Nov. 2024.

[16] C. Pan et al., "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, Aug. 2020.

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

[18] W. Sun, S. Sun, T. Shi, X. Su, and R. Liu, "A new model of beyond diagonal reconfigurable intelligent surfaces (BD-RIS) for the corresponding quantization and optimization," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11521–11534, Sep. 2024.

[19] H. Yahya, H. Li, M. Nerini, B. Clerckx, and M. Debbah, "Beyond diagonal RIS: Passive maximum ratio transmission and interference nulling enabler," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 7613–7627, 2024.

[20] X. Zhou, T. Fang, and Y. Mao, "Joint active and passive beamforming optimization for beyond diagonal RIS-aided multi-user communications," *IEEE Commun. Lett.*, vol. 29, no. 3, pp. 517–521, Mar. 2025, doi: 10.1109/LCOMM.2025.3528968.

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

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

[23] H. Li, S. Shen, and B. Clerckx, "Synergizing beyond diagonal reconfigurable intelligent surface and rate-splitting multiple access," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8717–8729, Aug. 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10411856?casa_token=Xsi5L6kqfN0AAAAA:bQ602RFz8IHbhw_akwLBKQsQhXxACr9VzLk7XEfhk-BK9UcDPFToGHDYA3k53AJN6WDwX3CJX-I

[24] I. Santamaria, M. Soleymani, E. Jorswieck, and J. Gutiérrez, "Interference minimization in beyond-diagonal RIS-assisted MIMO interference channels," *IEEE Open J. Veh. Technol.*, vol. 6, pp. 1005–1017, 2025.

[25] H. Li and Z. Lin, "Accelerated proximal gradient methods for nonconvex programming," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 379–387.

[26] A. Beck, *First-Order Methods in Optimization*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2017. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9781611974997

[27] C. Feng, W. Shen, J. An, and L. Hanzo, "Joint hybrid and passive RIS-assisted beamforming for mmWave MIMO systems relying on dynamically configured subarrays," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 13913–13926, Aug. 2022.

[28] J. P. Pavia, V. Velez, N. Souto, M. Ribeiro, P. Sebastião, and A. Correia, "System-level assessment of low complexity hybrid precoding designs for massive MIMO downlink transmissions in beyond 5G networks," *Appl. Sci.*, vol. 12, no. 6, p. 2812, Mar. 2022.

[29] J. P. Pavia et al., "Low complexity hybrid precoding designs for multiuser mmWave/THz ultra massive MIMO systems," *Sensors*, vol. 21, no. 18, p. 6054, Sep. 2021.

[30] J. Praia, J. P. Pavia, N. Souto, and M. Ribeiro, "Phase shift optimization algorithm for achievable rate maximization in reconfigurable intelligent surface-assisted THz communications," *Electronics*, vol. 11, no. 1, p. 18, Dec. 2021.

[31] S. K. Jayaweera and H. V. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter channel state information," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2697–2709, Oct. 2003.

[32] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. JSAC-5, no. 2, pp. 128–137, Feb. 1987.

[33] F. Bohagen, P. Orten, and G. Oien, "On spherical vs. Plane wave modeling of line-of-sight MIMO channels," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 841–849, Mar. 2009.

[34] J. Ren and R. G. Vaughan, "Rice factor estimation from the channel phase," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 1976–1980, Jun. 2012.

[35] K. Dovelos, S. D. Assimonis, H. Q. Ngo, B. Bellalta, and M. Matthaiou, "Intelligent reflecting surfaces at terahertz bands: Channel modeling and analysis," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.

[36] E. Basar, I. Yildirim, and F. Kilinc, "Indoor and outdoor physical channel modeling and efficient positioning for reconfigurable intelligent surfaces in mmWave bands," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8600–8611, Dec. 2021.

[37] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[38] J. Huang et al., "Reconfigurable intelligent surfaces: Channel characterization and modeling," *Proc. IEEE*, vol. 110, no. 9, pp. 1290–1311, Sep. 2022.

[39] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

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

[41] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal-recovery problems," in *Convex Optimization in Signal Processing and Communications*. U.K.: Cambridge Univ. Press, 2009, pp. 42–88. [Online]. Available: https://api.semanticscholar.org/CorpusID:17437390

[42] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, Jan. 2014, doi: 10.1561/2400000003.

[43] R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson, "Active reconfigurable intelligent surface-aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4962–4975, Aug. 2021.

[44] K. Zhi, C. Pan, H. Ren, K. K. Chai, and M. Elkashlan, "Active RIS versus passive RIS: Which is superior with the same power budget?" *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1150–1154, May 2022.

[45] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel estimation with reconfigurable intelligent surfaces—A general framework," *Proc. IEEE*, vol. 110, no. 9, pp. 1312–1338, Sep. 2022.

[46] J. An, C. Xu, L. Gan, and L. Hanzo, "Low-complexity channel estimation and passive beamforming for RIS-assisted MIMO systems relying on discrete phase shifts," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1245–1260, Feb. 2022.

[47] L. Zhang, C. Pan, Y. Wang, H. Ren, and K. Wang, "Robust beamforming design for intelligent reflecting surface aided cognitive radio systems with imperfect cascaded CSI," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 1, pp. 186–201, Mar. 2022.

[48] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 210–214, Feb. 2020.

[49] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1931–1945, Jul. 2021.

[50] H. Li, S. Shen, Y. Zhang, and B. Clerckx, "Channel estimation and beamforming for beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 72, pp. 3318–3332, 2024.

[51] A. L. F. de Almeida, B. Sokal, H. Li, and B. Clerckx, "Channel estimation for beyond diagonal RIS via tensor decomposition," *IEEE Trans. Signal Process.*, vol. 73, pp. 4764–4779, 2025.

[52] Y. Liu, W. Mei, H. Sun, D. Wang, and Z. Chen, "Power-measurement-based channel estimation for beyond diagonal RIS," *IEEE Commun. Lett.*, vol. 29, no. 11, pp. 2666–2670, Nov. 2025.

[53] N. Ginige, A. S. de Sena, N. H. Mahmood, N. Rajatheva, and M. Latva-Aho, "Efficient channel prediction for beyond diagonal RIS-assisted MIMO systems with channel aging," *IEEE Trans. Veh. Technol.*, vol. 74, no. 8, pp. 12658–12672, Aug. 2025.

[54] Y. Xing, T. S. Rappaport, and A. Ghosh, "Millimeter wave and sub-THz indoor radio propagation channel measurements, models, and comparisons in an office environment," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3151–3155, Oct. 2021.

[55] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[56] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[57] I. Santamaria, M. Soleymani, E. Jorswieck, and J. Gutiérrez, "MIMO capacity maximization with beyond-diagonal RIS," in *Proc. IEEE 25th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2024, pp. 936–940.

[58] Y. Zhao, H. Li, B. Clerckx, and M. Franceschetti, "MIMO channel shaping and rate maximization using beyond-diagonal RIS," *IEEE Trans. Signal Process.*, vol. 73, pp. 4397–4414, 2025.
[59] E. Björnson and Ö. T. Demir, "Capacity maximization for MIMO channels assisted by beyond-diagonal RIS," in *Proc. 19th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2025, pp. 1–5.
[60] A. Hjørungnes, *Complex-valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

DIOGO MENDES (Graduate Student Member, IEEE) received the Graduate and master's degrees in telecommunications and computer engineering from the Instituto Universitário de Lisboa (ISCTE), Lisbon, Portugal, in 2019, where he is currently pursuing the Ph.D. degree in information science and technology in the areas of wireless communications and signal processing for telecommunications. He is currently a Student Researcher with the Instituto de Telecomunicações (IT) and ISCTE. His research interests include physical layer algorithms for massive MIMO and reconfigurable intelligent surfaces, wireless transceivers, channel estimation, modulation and coding schemes, and signal processing for communications. He is a member of the IEEE Communications Society and the Chair of the IEEE ISCTE Student Branch.

NUNO SOUTO (Senior Member, IEEE) received the Graduate degree in aerospace engineering-avionics from the Instituto Superior Técnico, Lisbon, Portugal, in 2000, and the Ph.D. degree in 2006. From November 2000 to January 2002, he was a Researcher in automatic speech recognition with the Instituto de Engenharia e Sistemas de Computadores, Lisbon. He joined the University Institute of Lisbon (ISCTE) as an Assistant Professor in 2006. He has been a Researcher with the Instituto de Telecomunicações (IT), Portugal, since 2002, and has been involved in several international research projects and many national projects. His research interests include wireless networks, signal processing for communications, multicarrier communications, frequency domain equalization, channel coding, modulation, channel estimation, synchronization, MIMO schemes, wireless sensor networks, uncrewed aerial vehicles, mmWave/terahertz communications, machine learning for wireless communications, and molecular communications. He is a member of the IEEE Communications Society.

JOÃO PEDRO PAVIA (Member, IEEE) received the bachelor's degree in digital technologies and information security from the ISCTE—University Institute of Lisbon and the Ph.D. degree in information sciences and technologies in 2022. He is currently an Assistant Professor with the Digital Technologies Department, School of Applied Digital Technologies, ISCTE. He is a member of the COST—European Cooperation in Science and Technology, where he collaborates with professionals from various countries to develop solutions focused on the field of intelligent radio communications for seamless, inclusive interactions. In addition, another of his collaborations is related to the development of physical layer security solutions for reliable and resilient 6G systems. He has been involved as a local organizer and a reviewer at several conferences and symposia. He is also a reviewer of several journals in the field of information sciences and technologies. His research interests include topics related to cybersecurity, wireless communication networks, artificial intelligence, and big data.

JOÃO SILVA received the B.S. degree in aerospace engineering from the Instituto Superior Técnico, Lisbon Technical University, in 2000, and the M.B.A. degree from ISEG in 2015. From 2000 to 2002, he worked as a Business Consultant with McKinsey & Company. He finished his Ph.D. thesis with IST in 2006, focusing on spread spectrum techniques, multiuser detection schemes, and MIMO systems. Since 2006, he has been working on computer networks as an Assistant Professor with ISCTE-IUL.