

Joint Beamforming Algorithm for Multi-Stream MIMO Systems Assisted by Multiple Reconfigurable Intelligent Surfaces

NUNO SOUTO^{1,2} (Senior Member, IEEE), AND JOÃO C. SILVA^{1,2}

¹Department of Information Science and Technology, ISCTE-University Institute of Lisbon, 1649-026 Lisbon, Portugal

²Department of Information Science and Technology, Instituto de Telecomunicações, 1049-001 Lisbon, Portugal

CORRESPONDING AUTHOR: N. SOUTO (e-mail: nuno.souto@lx.it.pt)

This work was supported by FCT/MCTES through National Funds and when applicable co-funded by EU Funds under Project UIDB/50008/2020.

ABSTRACT In recent years there has been a growing interest in reconfigurable intelligent surfaces (RISs) as enablers for the realization of smart radio propagation environments which can provide performance improvements with low energy consumption in future wireless networks. However, to reap the potential gains of RIS it is crucial to jointly design both the transmit precoder and the phases of the RIS elements. Within this context, in this paper we study the use of multiple RIS panels in a parallel or multi-hop configuration with the aim of assisting a multi-stream multiple-input multiple-output (MIMO) communication. To solve the nonconvex joint optimization problem of the precoder and RIS elements targeted at maximizing the achievable rate, we propose a novel iterative algorithm based on the monotone accelerated proximal gradient (mAPG) method which includes an extrapolation step for improving the convergence speed and monitoring variables for ensuring sufficient descent of the algorithm. Based on the sufficient descent property we then present a detailed convergence analysis of the algorithm which includes expressions for the step size. Simulation results in different scenarios show that the use of multiple RIS panels combined with the proposed algorithm can be an effective solution for improving the achievable rates.

INDEX TERMS Reconfigurable intelligent surface (RIS), achievable rate, multiple-input multiple-output (MIMO), proximal minimization.

I. INTRODUCTION

OVER the last two decades we have been witnessing tremendous advances in wireless communication systems, which have enabled the support of a multitude of multimedia applications while also coping with the ever-growing number of mobile devices. Still, it is already clear that the full support of demanding applications such as immersive reality, holographic projection, autonomous driving, and tactile Internet, will only be feasible in future 6G and beyond networks [1]. Addressing this ever-increasing demand for higher data rates motivates further exploration of the mmWave bands that started to be integrated into 5G systems, and the move to even higher frequencies such as the Terahertz (THz) band [2]. While these bands can offer large available bandwidths they also pose several difficult

challenges for implementing viable communication systems. One of the main problems lies in the very high propagation loss, which can drastically limit the signal coverage range. To address this limitation, one can resort to antenna arrays comprising a massive number of elements since the short wavelengths at these bands enable the deployment of elements with small footprints. Combined with multi-input multi-output (MIMO) beamforming techniques, these massive antenna arrays are able to provide high directional gains. However, the transmitted signals are still very susceptible to obstacles [3] which has raised the interest in recent years for the adoption of reconfigurable intelligent surfaces (RISs).

A RIS is an artificial planar structure, typically composed of a large number of low-cost, quasi-passive elements which can be independently tuned to induce a specific phase

shift/amplitude on the reflected electromagnetic wave [4]. These elements can be reconfigured in real-time thus providing the capability to modify the wireless communication environment. Therefore, RISs have arisen as a potential solution for increasing coverage of communication systems as they have the capability to reflect and shape the incoming radio waves without the need for a power amplifier, RF chain, nor sophisticated signal processing [5], [6]. This allows the use of RIS for accomplishing passive beamforming which can significantly improve the spectral efficiency with low energy consumption [7], [8], [9]. Still, realizing RIS-based communication systems is extremely challenging as recent design formulations have shown [10]. The optimization problems tend to be large, the RIS matrices and precoding matrices are coupled, and they typically include non-convex constraints such as constant modulus or discrete-valued phase shifts. Therefore, finding global optimal solutions efficiently is non-trivial. A strategy to simplify the problem is to independently design the precoder and RIS phases. This approach is followed within the context of RIS-aided MIMO THz systems in [11], where an adaptive gradient method is applied for optimizing the RIS phases. Once the RIS matrix is computed, the precoder and combiner matrices are then obtained using a singular value decomposition (SVD) of the cascaded channel. While the approach achieves large gains over a RIS with random phase, the complexity tends to become very high when working with large antenna arrays and RISs with many elements. Also relying on the independent optimization of the precoder and RIS, [12] presented a low complexity proximal gradient algorithm which receives a previously computed SVD-based precoder and then optimizes the RIS matrix resorting to simple element-wise normalization. Due to its simplicity, it can cope with large problem settings while providing large gains. Still, to obtain improved performance it is necessary to jointly optimize both the precoder and RIS. However, since finding the global optimal solution is difficult, most works try to propose efficient approaches for finding local optimal solutions. For example, the work in [13] considered a joint beamforming scheme based on fractional programming optimization for RIS-assisted THz communications. Even though it corresponds to a limited scenario, it showed encouraging results. In [14] the authors targeted the minimization of the transmitted power constrained on the signal to interference plus noise ratio required at the receivers. To solve the problem, they proposed an alternating optimization (AO) algorithm that iteratively optimizes the RIS phase shifts and the transmit beamforming vector. The coverage and achievable rate performance are significantly improved when compared to conventional systems without RIS. However, it requires many iterations, especially when the number of RIS elements is large. Another iterative algorithm based on the AO approach was proposed in [15] but it only optimizes the transmit covariance matrix or one of the reflection coefficients with all the other variables being fixed. While this allows the implementation complexity to

be reduced, it can still be high when the number of reflecting elements is large. In [16], the authors addressed the joint optimization of the transmit beamformer and RIS phase shifts with the objective of maximizing the achievable rate in single-antenna single-user scenarios. Algorithms based on the fixed-point iteration method and manifold optimization were proposed for solving the problem. Also aiming at maximizing the achievable rate but in a multi-stream MIMO system, the authors in [17] proposed an iterative projected gradient method (PGM) for solving the joint optimization problem of the covariance matrix of the transmitted signal and the RIS elements. The PGM algorithm was shown to achieve similar achievable rates to the AO method from [14] in single-user scenarios but requiring with fewer iterations. However its complexity can become very high when using large transmitter and receiver arrays.

While all the previous works consider single RIS scenarios, the adoption of multiple RIS can potentially enable additional degrees of freedom and ensure uninterrupted connectivity inside a target area, as discussed in [18]. Therefore, a few recent works with multiple RIS have started to appear [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]. Regarding distributed RISs, two main approaches can be adopted. In the simpler approach, only one of the RIS panels selected according to some criteria participates in the transmission. In the second approach, all the RISs aid in establishing the communication which can potentially provide better performance at the cost of additional complexity [19]. Focusing on the second approach, in [20] the authors showed that unless the RIS is located near the base station (BS) or near the user, the adoption of multiple RISs is capable to achieve higher ergodic capacity than a single large panel RIS. Considering a single antenna BS and receiver, the outage and the rate performance of a multi-RIS aided system with limited backhaul capacity was analyzed in [21]. It was shown that, while rate performance degradation can occur due to backhaul limitation, the capacity can be improved when one of the RISs is deployed close to the line between the BS and the UE. Focusing on energy efficiency, in [22] the authors established a federated deep learning framework for the configuration optimization of multiple RISs in high-density networks with multiple single-antenna users. The work in [23] studied the joint active and passive precoding design, where multiple parallel RISs assist a single stream transmission between a BS and a single-antenna receiver. In this case, a semidefinite relaxation-based method was proposed for solving the problem. Also concerning multiple parallel RISs, the work in [24] considered a simpler approach where the RISs matrices are first optimized followed by the computation of a hybrid precoder/combiner using the SVD.

Regarding the use of multiple RISs in a multi-hop configuration, a few recent studies have started to appear such as in [25], which emphasized the need for innovative design approaches and optimization methods in this type of systems. The authors in [26] addressed the problem of optimizing the RIS deployment and proposed a hybrid strategy that includes

not only single reflection RIS-aided links but also double-reflection links. The approach can be exploited to establish additional line-of-sight (LOS) paths between the BS and users, especially when severe blocking exists in the direct and single-reflection links. Considering a single-antenna BS and user, [27] showed that given a sufficient total number of RIS elements, a LOS double reflection link can outperform a single RIS deployment, even if the former suffers a more severe pathloss due to the double signal reflection. Also addressing RIS-aided dual-hop links, the authors in [28] presented a joint active and passive beamforming design to maximize the secrecy rate in multiple-input single-output channels. To cope with the more general multi-hop scenario, the authors in [29] applied machine learning methods, in particular a deep reinforcement learning based framework, to find feasible solutions to the design problem of transmit beamforming at the BS and phase shift matrices at the RISs. This system model considered a multi-antenna BS transmitting to several single-antenna users. The work in [30], adopted a statistical approach to the multi-RIS beamforming problem in the absence of channel information. Similarly to other works, the system model is simplified as it assumes that the transmitter and receiver are equipped with one antenna each. Also focusing on multi-hop RIS-aided communications, [31] proposed the optimal active and cooperative passive beamforming for the downlink connection between a multi-antenna BS and a remote single-antenna user. For this setting, it also presented the optimal RIS selection and beam routing solution, which maximizes the cascaded LoS channel power. It was shown that there exists a fundamental trade-off between minimizing the end-to-end path loss and maximizing the cooperative passive beamforming gain.

Motivated by the work above, in this paper we study the realization of smart radio propagation environments where multiple passive RIS panels deployed in the vicinity of a BS assist a transmission towards a user. Whereas most of the previous works on active and passive beamforming designs restrict the system model to either parallel RIS or to dual-hop RIS links, and to simple configurations relying on a single stream, a single antenna user and sometimes a single-antenna BS, in this paper we consider a more general multi-stream MIMO setting where both the BS and user can be equipped with multiple antennas potentially arranged in large arrays, as envisioned in mmWave/THz communications. Targeting multiple RIS-aided links in parallel or multi-hop configurations, we address the joint design problem of the precoder and of all the RIS elements in the panels with the aim of maximizing the achievable rate. To efficiently tackle this nonconvex problem with large-scale settings, we apply the monotone accelerated proximal gradient (mAPG) method [32] in order to obtain a novel iterative algorithm that is simple to implement and is guaranteed to converge to a critical point of the problem. Although gradient-based methods have already been successfully applied for optimizing the achievable rate of RIS-aided systems such as the PGM in [17], they typically exhibit slow

convergence speed. To counter this problem, the adopted mAPG method includes an extrapolation step based on a linear combination of the previous and current solution estimates. It also includes the computation of additional monitoring variables for ensuring sufficient descent of the algorithm. This property facilitates the convergence analysis of the algorithm which, otherwise would be difficult to perform due to the nonconvex nature of the problem. The mAPG approach is first developed for multiple parallel RISs and then extended to multi-hop scenarios.

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

- We formulate the joint achievable rate maximization problem defined over the precoder and the phase shifts of all RIS elements, considering a multi-stream MIMO communication system assisted by multiple parallel RIS panels. To find a solution for this problem we derive an mAPG-based iterative algorithm, presenting exact expressions for the required gradients and proximal operators.
- We show that the objective function in the optimization problem has Lipschitz continuous gradients and provide a bound for the Lipschitz constants. Using the Lipschitz continuity of the gradients we then prove that the proposed joint precoding and RIS mAPG based (JPR-mAPG) algorithm is guaranteed to converge to a critical point.
- We extend the JPR-mAPG algorithm for multi-hop RIS-assisted scenarios, providing expressions for the required gradients. We establish the Lipschitz continuity of the gradients also for this scheme, presenting the respective bound for the Lipschitz constants. This ensures that the convergence of the algorithm is still valid for the multi-hop case.
- The achievable rate of the proposed approach is evaluated through numerical simulations and compared against the RIS-optimized APG algorithm (R-APG) algorithm that only optimizes the RIS phases shifts, the AO algorithm and the PGM algorithm which, while being a gradient-based algorithm, does not apply acceleration. Results show that, while the adoption of some type of optimization in the design of the RIS can offer clear gains over a static approach, the proposed JPR-mAPG approach is often able to provide higher rates. Furthermore, unlike the PGM and AO algorithms, JPR-mAPG does not have a computational complexity with a cubic or quadratic dependency on the transmitter and receiver array sizes, thus making it better suited for scenarios with very large antenna arrays which are common in the mmWave/THz bands. In the case of multi-hop, it was shown that the RIS panels optimized with the JPR-mAPG algorithm can overcome the multiplicative effect of the path losses and provide gains, enabling signal coverage when severe blockage between transmitter and receiver exists.

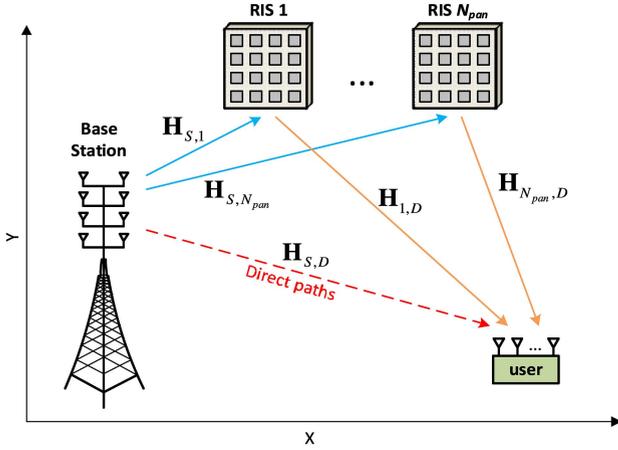


FIGURE 1. System model layout considered for the multiple parallel RIS-assisted MIMO communication system.

The rest of the paper is organized as follows. Section II defines the system model and optimization problem for the multiple parallel RIS-assisted scheme, presenting the proposed JPR-mAPG algorithm and providing the respective convergence analysis. Section III extends the work to a multi-hop RIS-assisted scheme. Section IV presents several simulation results in different scenarios, whereas the conclusions are provided in Section V.

Notation: Matrices and vectors are represented by uppercase and lowercase boldface letters, respectively. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose of a matrix/vector, $\|\cdot\|_2$ is the 2-norm of a vector/matrix, $\|\cdot\|_F$ is the Frobenius norm, $\text{diag}(\cdot)$ represents a diagonal of a matrix if the argument is a matrix or represents a diagonal matrix if the argument is a vector, $\text{blkdiag}\{\mathbf{A}_i\}_{i=1,\dots,N}$ is a block diagonal matrix whose diagonal elements are matrices \mathbf{A}_i ($i = 1 \dots N$), $\partial g(\mathbf{x})$ is the sub-differential set of $g(\mathbf{x})$, \mathbf{I}_n is the $n \times n$ identity matrix and $I_{\mathcal{D}}(\mathbf{v})$ is the indicator function which returns 0 if $\mathbf{v} \in \mathcal{D}$ and $+\infty$ otherwise. We also consider that for matrices we have $\prod_{i=1}^n \mathbf{A}_i = \mathbf{A}_n \times \dots \times \mathbf{A}_1$.

II. MULTIPLE PARALLEL RIS-ASSISTED COMMUNICATION

A. SYSTEM MODEL AND PROBLEM FORMULATION

First, we address the case where, to help establish the communication, the receiver is served by N_{PAN} RIS panels simultaneously, as illustrated in Fig. 1. The transmitted signal comprises N_s data streams and is represented as $\mathbf{s} = [s_1 \dots s_{N_s}]^T$, where $s_i \in \mathbb{C}$ corresponds to an amplitude and phase modulated symbol with $E[|s|^2] = N_s$. The signal arriving at the receiver array, can be expressed as

$$\mathbf{r} = \sqrt{\rho} \mathbf{H} \mathbf{F} \mathbf{s} + \mathbf{n} \quad (1)$$

where $\sqrt{\rho}$ denotes the power per stream, $\mathbf{H} \in \mathbb{C}^{N_{rx} \times N_{tx}}$ is the RIS dependent channel matrix, N_{tx} is the number of transmit antennas, N_{rx} is the number of receive antennas, $\mathbf{F} \in \mathbb{C}^{N_{tx} \times N_s}$ is the BS precoder matrix and $\mathbf{n} \in \mathbb{C}^{N_{rx} \times 1}$ is the noise vector,

which contains independent zero-mean circularly symmetric Gaussian samples with covariance $\sigma_n^2 \mathbf{I}_{N_{rx}}$.

We consider the existence of blockages between multiple RIS which, combined with the high path-loss at high frequencies, will tend to result in negligible powers in the signals that are reflected by the RISs two or more times. Therefore, similarly to [24], we ignore the cross-reflections and approximate the channel matrix as

$$\mathbf{H} = \mathbf{H}_{S,D} + \sum_{i=1}^{N_{pan}} \mathbf{H}_{i,D} \Phi_i \mathbf{H}_{S,i}, \quad (2)$$

where $\mathbf{H}_{S,D} \in \mathbb{C}^{N_{rx} \times N_{tx}}$ denotes the direct channel between the transmitter and receiver, $\mathbf{H}_{S,i} \in \mathbb{C}^{N_{ris} \times N_{tx}}$ is the channel matrix between the transmitter and the i^{th} RIS panel, and $\mathbf{H}_{i,D} \in \mathbb{C}^{N_{rx} \times N_{ris}}$ is the channel matrix between the i^{th} RIS panel and the receiver. Matrix $\Phi_i \in \mathbb{C}^{N_{ris} \times N_{ris}}$ models the phase shifting effect of the i^{th} RIS panel. This matrix has a diagonal structure and can be written as $\Phi_i = \text{diag}(\varphi_i)$ with $\varphi_i = [ae^{j\theta_1^i}, \dots, ae^{j\theta_{N_{ris}}^i}]^T$, where $\theta_m^i \in [0, 2\pi)$ ($m = 1 \dots N_{pan}$) represents the phase shift of the m^{th} RIS element and a denotes the amplitude of the reflection coefficients. Along the paper we will assume perfect channel state information (CSI) at the BS. Even though, obtaining CSI in RIS aided schemes is challenging, several methods have been proposed in recent years [10], [33].

In this paper we design the precoder and RIS matrices to maximize the achievable rate of the system which, assuming Gaussian signaling, can be written as [8]

$$R = \log_2 \det \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right) \quad (3)$$

in bits/s/Hz, with P_n denoting the noise power (i.e., $P_n = \sigma_n^2$). It is important to remind that the channel matrix \mathbf{H} , and thus also the rate, depend on the RIS coefficient vectors φ_i , $i = 1 \dots N_{pan}$. The optimization problem can then be formulated as

$$\min_{\mathbf{F}, \varphi_i} f(\mathbf{F}, \varphi_1, \dots, \varphi_{N_{pan}}) = -\log \det \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right) \quad (4a)$$

$$\text{subject to } \|\mathbf{F}\|_F^2 \leq N_s, \quad (4b)$$

$$\varphi_i \in \mathcal{U}_{N_{RIS}}, \quad i = 1 \dots N_{pan}. \quad (4c)$$

where $\mathcal{U}_{N_{RIS}} = \{\mathbf{v} \in \mathbb{C}^{N_{RIS} \times 1} : |v_m| = a, m = 1 \dots N_{RIS}\}$ is the set of all complex valued vectors whose elements have a magnitude of a .

B. ACCELERATED PROXIMAL GRADIENT BASED SOLUTION

The coupling between optimization variables \mathbf{F} and φ_i in (4a), as well as constraint (4c) make the problem non-convex and difficult to solve efficiently. Furthermore, when considering a typical RIS assisted massive or ultra-massive MIMO scenario, formulation (4) will typically correspond to a large-scale optimization problem. This motivates the resort to first-order methods with low iteration cost for solving

the problem. In this case, we propose an algorithm based on the application of the monotone Accelerated Proximal Gradient (mAPG) method for nonconvex optimization problems described in [32]. First, we use the indicator function to encode constraints (4b) and (4c) and rewrite problem (4) as

$$\min_{\mathbf{F}, \boldsymbol{\varphi}_i} f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}}) + I_C(\mathbf{F}) + \sum_{i=1}^{N_{pan}} I_{\mathcal{U}_{NRIS}}(\boldsymbol{\varphi}_i) \quad (5)$$

where $\mathcal{C} = \{\mathbf{X} \in \mathbb{C}^{N_{tx} \times N_s} : \|\mathbf{X}\|_F^2 \leq N_s\}$. To find a solution for (5) we apply a gradient step consisting of $\mathbf{F}^{(q)} - \alpha \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)})$ and $\boldsymbol{\varphi}_i^{(q)} - \alpha \nabla_{\boldsymbol{\varphi}_i^*} f(\boldsymbol{\varphi}_i^{(q)})$, followed by a proximal mapping, whose operator for a function g is defined as

$$\text{prox}_g(\mathbf{a}) = \underset{\hat{\mathbf{x}}}{\text{argmin}} g(\hat{\mathbf{x}}) + \frac{1}{2} \|\hat{\mathbf{x}} - \mathbf{a}\|^2 \quad (6)$$

To improve the typical slow convergence rate of gradient-based methods, we can apply an extrapolation step as proposed in [34], which involves the computation of additional variables $(\mathbf{P}^{(q)}, \mathbf{y}_1^{(q)}, \dots, \mathbf{y}_{N_{pan}}^{(q)})$. However, this extrapolation may result in a bad update which, while not causing the objective function to increase, makes the convergence difficult to analyze, especially due to the nonconvex nature of the problem. Therefore, we add an extra set of variables $(\mathbf{U}^{(q)}, \mathbf{v}_1^{(q)}, \dots, \mathbf{v}_{N_{pan}}^{(q)})$ whose role is to monitor and correct the update of $(\mathbf{F}^{(q)}, \boldsymbol{\varphi}_1^{(q)}, \dots, \boldsymbol{\varphi}_{N_{pan}}^{(q)})$. Algorithm 1 summarizes all the steps of the proposed approach. Variables α , Q and t_k denote the step size, maximum number of iterations and extrapolation parameter, respectively. In step 5 and 6, the proximal operator of the indicator functions corresponds to the Euclidean projection over the respective sets and can be computed as

$$\text{prox}_{\alpha I_C}(\mathbf{X}) = \begin{cases} \mathbf{X}, & \|\mathbf{X}\|_F^2 \leq N_s \\ \frac{\sqrt{N_s}}{\|\mathbf{X}\|_F} \mathbf{X}, & \text{otherwise} \end{cases} \quad (7)$$

and

$$\text{prox}_{\alpha I_{\mathcal{U}_{NRIS}}}(\mathbf{x}) = a(\mathbf{x} \oslash |\mathbf{x}|), \quad (8)$$

where \oslash denotes the Hadamard division and $|\cdot|$ represents the absolute value applied elementwise. Note that these two projections enforce constraints (4b) and (4c). Regarding the later projection, we note that in the special case of $\text{prox}_{\alpha I_{\mathcal{U}_1}}(0)$, any point of the form $a e^{j\theta}$ with $\theta \in [0, 2\pi)$ is a valid solution. The complex-valued gradient of $f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}})$ required in the proposed algorithm can be computed according to the following Lemma.

Lemma 1: Let $f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}})$ be defined as in (4a) with the channel matrix represented as (2). Then, the corresponding complex-valued gradient with respect to \mathbf{F}^* and $\boldsymbol{\varphi}_i^*$ ($i = 1, \dots, N_{pan}$) can be computed as

$$\nabla_{\mathbf{F}^*} f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}}) = -\frac{\rho}{P_n} \mathbf{H}^H \mathbf{H} \mathbf{F} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \quad (9)$$

Algorithm 1 Joint Precoding and RIS - Monotone Accelerated Proximal Gradient Based Algorithm (JPR-mAPG)

- 1: **Input:** $\mathbf{r}, \mathbf{H}, \alpha, Q, \mathbf{F}^{(0)}, \boldsymbol{\varphi}_i^{(0)}$
 - 2: $\hat{\mathbf{x}}^{(0)} = \mathbf{z}^{(0)} = 0, t_0 = 0, t_1 = 1$
 - 3: **for** $q=1, 1, \dots, Q$ **do**
 - $\mathbf{P}^{(q)} = \mathbf{F}^{(q)} + \frac{t_{q-1}}{t_q} (\mathbf{W}^{(q)} - \mathbf{F}^{(q)}) + \frac{t_{q-1}-1}{t_q} (\mathbf{F}^{(q)} - \mathbf{F}^{(q-1)})$
 - $\mathbf{y}_i^{(q)} = \boldsymbol{\varphi}_i^{(q)} + \frac{t_{q-1}}{t_q} (\mathbf{z}^{(q)} - \boldsymbol{\varphi}_i^{(q)}) + \frac{t_{q-1}-1}{t_q} (\boldsymbol{\varphi}_i^{(q)} - \boldsymbol{\varphi}_i^{(q-1)}),$
 $i = 1, \dots, N_{pan}$
 - $\mathbf{W}^{(q+1)} = \text{prox}_{\alpha I_C}(\mathbf{P}^{(q)} - \alpha \nabla_{\mathbf{F}^*} f(\mathbf{P}^{(q)}))$
 $\mathbf{z}_i^{(q+1)} = \text{prox}_{\alpha I_{\mathcal{U}_{NRIS}}}(\mathbf{y}_i^{(q)} - \alpha \nabla_{\boldsymbol{\varphi}_i^*} f(\mathbf{y}_i^{(q)})),$
 $i = 1, \dots, N_{pan}$
 - $\mathbf{U}^{(q+1)} = \text{prox}_{\alpha I_C}(\mathbf{F}^{(q)} - \alpha \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)}))$
 $\mathbf{v}_i^{(q+1)} = \text{prox}_{\alpha I_{\mathcal{U}_{NRIS}}}(\boldsymbol{\varphi}_i^{(q)} - \alpha \nabla_{\boldsymbol{\varphi}_i^*} f(\boldsymbol{\varphi}_i^{(q)})),$
 $i = 1, \dots, N_{pan}$
 - if** $f(\mathbf{W}^{(q+1)}, \mathbf{z}_1^{(q+1)}, \dots, \mathbf{z}_{N_{pan}}^{(q+1)}) \leq f(\mathbf{U}^{(q+1)}, \mathbf{v}_1^{(q+1)}, \dots, \mathbf{v}_{N_{pan}}^{(q+1)})$ **then**
 - $\mathbf{F}^{(q+1)} = \mathbf{W}^{(q+1)}$
 $\boldsymbol{\varphi}_i^{(q+1)} = \mathbf{z}_i^{(q+1)}, i = 1, \dots, N_{pan}$
 - else**
 - $\mathbf{F}^{(q+1)} = \mathbf{U}^{(q+1)}$
 $\boldsymbol{\varphi}_i^{(q+1)} = \mathbf{v}_i^{(q+1)}, i = 1, \dots, N_{pan}$
 - end if**
 - $t_{q+1} = \frac{1}{2} \left(\sqrt{4(t_q)^2 + 1} + 1 \right)$
 - end for**
 - Output:** $\mathbf{F}^{(q+1)}, \boldsymbol{\varphi}_i^{(q+1)}, i = 1, \dots, N_{pan}.$
-

$$\nabla_{\boldsymbol{\varphi}_i^*} f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}}) = -\frac{\rho}{P_n} \text{diag} \left[\mathbf{H}_{i,D}^H \mathbf{H} \mathbf{F} \right] \times \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{H}_{S,i}^H \quad (10)$$

Proof: Proof is shown in Appendix A. ■

C. COMPUTATIONAL COMPLEXITY

Regarding the complexity of the proposed JPR-mAPG algorithm, it is associated mainly to steps 5, 6 and 7. These steps require the computation of the channel matrix using (2) which has a complexity cost of $\mathcal{O}(N_{tx} N_{ris} N_{rx} N_{pan})$. Steps 5 and 6 are dominated by the computation of the gradients $\nabla_{\mathbf{F}^*} f$ and $\nabla_{\boldsymbol{\varphi}_i^*} f$ ($i = 1 \dots N_{pan}$), which entail several matrix multiplications and a small matrix inversion, requiring $\mathcal{O}(N_s N_{tx} N_{rx})$ and $\mathcal{O}(N_{tx} N_{ris} N_{rx} N_{pan})$ operations, respectively. As for the projections (7) and (8), they are relatively cheap having a complexity of $\mathcal{O}(N_s N_{tx})$ and $\mathcal{O}(N_{ris})$, respectively. Step 7 computes the objective function $f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}})$ which also comprises several matrix multiplications and the determinant of a small matrix. In

this case the complexity is $\mathcal{O}(N_{tx}N_{ris}N_{rx}N_{pan} + N_sN_{tx}N_{rx})$. Steps 5-7 are repeated for Q iterations which, keeping only the dominant terms, results in an overall computational complexity order of $\mathcal{O}(Q(N_{tx}N_{ris}N_{rx}N_{pan} + N_sN_{tx}N_{rx}))$. As a comparison, the complexity of the R-APG algorithm [12] (directly extended to the parallel multi-RIS case) is $\mathcal{O}(Q(N_{tx}N_{ris}N_{rx}N_{pan} + N_sN_{tx}N_{rx} + N_{tx}N_{rx}^2 + N_{rx}^3))$. While the complexity of both JPR-mAPG and R-APG grows with the product of the transmitter, receiver and RIS array sizes and has a linear dependency on the number of RIS panels, the R-APG algorithm has a stronger dependency on the receiver array size (grows with $\mathcal{O}(QN_{tx}N_{rx}^2)$ and $\mathcal{O}(QN_{rx}^3)$). Bearing in mind that they were designed for a single panel RIS (i.e., $N_{pan} = 1$), the AO method [15] has a complexity of $\mathcal{O}((L+1)N_{tx}N_{ris}N_{rx} + L \min(N_{tx}, N_{rx})^3 + \frac{L}{2}N_{tx}^2 \min(N_{tx}, N_{rx}) + Q_{AM}(N_{tx}^3 + N_{tx}^2N_{ris} + 2N_{tx}N_{ris}N_{rx} + (2N_{tx}N_{rx}^2 + 2N_{rx}^3)N_{ris} + N_{tx}^2N_{ris} + 2N_{tx}N_{ris}N_{rx} + (2N_{tx}N_{rx}^2 + 2N_{rx}^3)N_{ris} + \min(N_{tx}, N_{rx})^3 + \frac{1}{2}N_{tx}^2 \min(N_{tx}, N_{rx})))$ (L is the number of initial random phase shifts realizations and Q_{AM} is the number of outer iterations) and the PGM approach [17] has a complexity of $\mathcal{O}(Q_{PGM}(N_{rx}^2N_{tx} + N_{rx}N_{tx}^2 + N_{rx}N_{tx}N_{ris} + N_{rx}^3 + N_{tx}^3))$ (Q_{PGM} is the number of iterations of PGM). The PGM algorithm has a dependency on the product of the transmitter, receiver and RIS array sizes, as happens with JPR-mAPG. However, its complexity also has a cubic dependency on the transmitter and receiver array sizes (grows with $\mathcal{O}(Q_{PGM}N_{tx}^3)$ and $\mathcal{O}(Q_{PGM}N_{rx}^3)$) which may hinder the application to very large arrays. The AO algorithm also suffers from the same problem as it also grows with the cubic size of the transmitter and receiver arrays.

D. CONVERGENCE ANALYSIS

In this subsection we show that with the proper selection of the step size, the algorithm can be guaranteed to converge to a critical point. To help prove the convergence of the algorithm we first present the following theorem.

Theorem 1: Function $f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}})$ defined as in (4a) with the channel matrix represented as (2) has Lipschitz continuous gradients over feasible set (4b)–(4c), i.e.,

$$\begin{aligned} & \left(\left\| \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_2)}, \boldsymbol{\varphi}_1^{(q_2)}, \dots, \boldsymbol{\varphi}_{N_{pan}}^{(q_2)}) - \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_1)}, \boldsymbol{\varphi}_1^{(q_1)}, \dots, \boldsymbol{\varphi}_{N_{pan}}^{(q_1)}) \right\|_F^2 \right. \\ & + \sum_{i=1}^{N_{pan}} \left\| \nabla_{\boldsymbol{\varphi}_i^*} f(\mathbf{F}^{(q_2)}, \boldsymbol{\varphi}_1^{(q_2)}, \dots, \boldsymbol{\varphi}_{N_{pan}}^{(q_2)}) \right. \\ & \left. - \nabla_{\boldsymbol{\varphi}_i^*} f(\mathbf{F}^{(q_1)}, \boldsymbol{\varphi}_1^{(q_1)}, \dots, \boldsymbol{\varphi}_{N_{pan}}^{(q_1)}) \right\|_2^2 \Big)^{\frac{1}{2}} \\ & \leq L \left(\left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F^2 + \sum_{i=1}^{N_{pan}} \left\| \boldsymbol{\varphi}_i^{(q_2)} - \boldsymbol{\varphi}_i^{(q_1)} \right\|_2^2 \right)^{\frac{1}{2}}, \end{aligned} \quad (11)$$

where constant L is given by

$$L = \sqrt{\max(b^2 + bc + d^2 + de, c^2 + bc + e^2 + de)}, \quad (12)$$

with

$$b = \frac{\rho}{P_n} \zeta^2 \left(1 + 2 \frac{N_s \rho}{P_n} \zeta^2 \right) \quad (13)$$

$$c = 2\sqrt{N_s} \frac{\rho}{P_n} \zeta \lambda_{\max}(\tilde{\mathbf{H}}_{R,D}) \lambda_{\max}(\tilde{\mathbf{H}}_{S,R}) \left(1 + \frac{N_s \rho}{P_n} \zeta^2 \right) \quad (14)$$

$$d = 2\sqrt{N_s} \frac{\rho}{P_n} \zeta \left(1 + \frac{N_s \rho}{P_n} \zeta^2 \right) \sum_{i=1}^{N_{pan}} \lambda_{\max}(\mathbf{H}_{i,D}) \lambda_{\max}(\mathbf{H}_{S,i}) \quad (15)$$

$$\begin{aligned} e &= \frac{N_s \rho}{P_n} \left(1 + 2 \frac{N_s \rho}{P_n} \zeta^2 \right) \lambda_{\max}(\tilde{\mathbf{H}}_{R,D}) \lambda_{\max}(\tilde{\mathbf{H}}_{S,R}) \\ & \times \sum_{i=1}^{N_{pan}} \lambda_{\max}(\mathbf{H}_{i,D}) \lambda_{\max}(\mathbf{H}_{S,i}) \end{aligned} \quad (16)$$

$$\zeta = \lambda_{\max}(\mathbf{H}_{S,D}) + a \sum_{i=1}^{N_{pan}} \lambda_{\max}(\mathbf{H}_{i,D}) \lambda_{\max}(\mathbf{H}_{S,i}) \quad (17)$$

and $\lambda_{\max}(\mathbf{X})$ denoting the largest singular value of matrix \mathbf{X} .

Proof: Proof is shown in Appendix B. ■

After establishing the Lipschitz continuity of the gradients, the convergence of the algorithm can be stated in the following theorem.

Theorem 2: Let the step size in Algorithm 1 be selected as $\alpha < \frac{1}{L}$. Then both sequence sets $\{(\mathbf{F}^{(q)}, \boldsymbol{\varphi}_1^{(q)}, \dots, \boldsymbol{\varphi}_{N_{pan}}^{(q)})\}$ and $\{(\mathbf{U}^{(q)}, \mathbf{v}_1^{(q)}, \dots, \mathbf{v}_{N_{pan}}^{(q)})\}$ generated by the algorithm are bounded. Furthermore, let $\{(\mathbf{F}', \boldsymbol{\varphi}'_1, \dots, \boldsymbol{\varphi}'_{N_{pan}})\}$ be any accumulation point of set $\{(\mathbf{F}^{(q)}, \boldsymbol{\varphi}_1^{(q)}, \dots, \boldsymbol{\varphi}_{N_{pan}}^{(q)})\}$, then it must be a critical point of (4).

Proof: Proof is shown in Appendix C. ■

III. MULTI-HOP RIS-ASSISTED COMMUNICATION

A. SYSTEM MODEL

The second case that we address in this paper corresponds to a multi-hop scheme where, to help circumvent existing obstructions, the signal transmitted by the BS arrives at the receiver after being sequentially reflected by N_{PAN} aiding RIS panels, as illustrated in Fig. 2. For this type of scenario, we still use the received signal model (1), but with the channel matrix defined as

$$\mathbf{H} = \mathbf{H}_{S,D} + \mathbf{H}_{N_{pan}+1} \prod_{i=1}^{N_{pan}} \boldsymbol{\Phi}_i \mathbf{H}_i. \quad (18)$$

In this expression $\mathbf{H}_1 \in \mathbb{C}^{N_{ris} \times N_{tx}}$ is the channel matrix between the transmitter and the first RIS panel, $\mathbf{H}_i \in \mathbb{C}^{N_{ris} \times N_{ris}}$ ($i=2, \dots, N_{pan}$) is the channel matrix between RIS panels i and $i+1$ and $\mathbf{H}_{N_{pan}+1} \in \mathbb{C}^{N_{rx} \times N_{ris}}$ is the channel matrix between the N_{pan}^{th} RIS panel and the receiver. It is important to note that, with the exception of the first and last RIS, in this model we are neglecting the direct links between the BS and a RIS and also between a RIS and the user, as these are assumed to be all NLOS and suffer a high propagations loss (considering the envisioned mmWave/THz

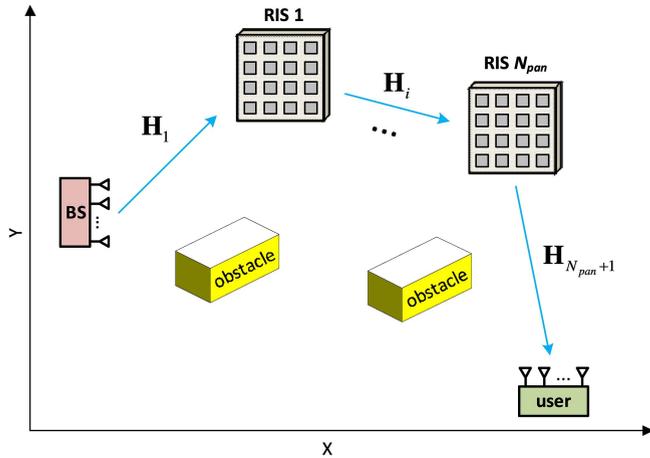


FIGURE 2. System model layout considered for the multi-hop RIS-assisted MIMO communication system.

frequencies). The same happens to the multiple reflections between non-consecutive RIS panels. These assumptions are often used in multi-hop RIS aided works [25], [29]. If a more accurate channel model that includes for example the direct links between the middle RISs and the BS/user is required, the proposed algorithm can still be directly applied (only the gradient expression with respect to φ_i^* will need to be re-computed).

B. PRECODER AND RIS OPTIMIZATION

Since the received signal model in the multi-hop case is essentially the same as for the multiple parallel RIS scheme, we can also formulate the precoder and RIS design problem as (4a)–(4c) and solve it using Algorithm 1. However, due to the different channel matrix representation, the complex-valued gradient of $f(\mathbf{F}, \varphi_1, \dots, \varphi_{N_{pan}})$ will be different, as stated in the following lemma.

Lemma 2: Let $f(\mathbf{F}, \varphi_1, \dots, \varphi_{N_{pan}})$ be defined as in (4a) with the channel matrix represented as (18). Then, the corresponding complex-valued gradient with respect to \mathbf{F}^* can be computed as (9) whereas the gradient with respect to φ_i^* ($i = 1, \dots, N_{pan}$) can be computed as

$$\nabla_{\varphi_i^*} f(\mathbf{F}, \varphi_1, \dots, \varphi_{N_{pan}}) = -\frac{\rho}{P_n} \text{diag} \left[\bar{\mathbf{H}}_{i,D}^H \mathbf{H} \mathbf{F} \right. \\ \left. \times \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \mathbf{F}^H \bar{\mathbf{H}}_{S,i}^H \right] \quad (19)$$

with

$$\bar{\mathbf{H}}_{S,i} = \mathbf{H}_i \prod_{j=1}^{i-1} \Phi_j \mathbf{H}_j \quad (20)$$

and

$$\bar{\mathbf{H}}_{i,D} = \mathbf{H}_{N_{pan}+1} \prod_{l=i+1}^{N_{pan}} \Phi_l \mathbf{H}_l. \quad (21)$$

Proof: The derivation is straightforward by directly extending the approach presented in Appendix A for the multiple parallel RIS-assisted channel model.

Convergence of the JPR-mAPG algorithm can also be established for the multi-hop case through Theorem 2, as long the gradients are still Lipschitz continuous. The following theorem states that this condition is still satisfied in this case.

Theorem 3: Function $f(\mathbf{F}, \varphi_1, \dots, \varphi_{N_{pan}})$ defined as in (4a), with the channel matrix represented as (18), has Lipschitz continuous gradients over feasible set (4b)–(4c), thus satisfying (11) with constant L given by

$$L = \sqrt{(N_{pan} + 1) \max(b^2 + N_{pan}c^2, c^2 + N_{pan}d^2)}, \quad (22)$$

with

$$b = \frac{\rho}{P_n} \zeta^2 \left(1 + 2 \frac{N_s \rho}{P_n} \zeta^2 \right) \quad (23)$$

$$c = 2\sqrt{N_s} \frac{\rho}{P_n} \zeta a^{N_{pan}-1} \left(\prod_{i=1}^{N_{pan}+1} \lambda_{\max}(\mathbf{H}_i) \right) \left(1 + \frac{N_s \rho}{P_n} \zeta^2 \right) \quad (24)$$

$$d = \frac{N_s \rho}{P_n} a^{2N_{pan}-2} \left(\prod_{i=1}^{N_{pan}+1} \lambda_{\max}^2(\mathbf{H}_i) \right) \left(1 + 2 \frac{N_s \rho}{P_n} \zeta^2 \right) \quad (25)$$

$$\zeta = \lambda_{\max}(\mathbf{H}_{S,D}) + a^{N_{pan}} \prod_{i=1}^{N_{pan}+1} \lambda_{\max}(\mathbf{H}_i). \quad (26)$$

Proof: Proof is similar to the one in Appendix B for Theorem 1. Therefore, the details are omitted here. ■

C. COMPUTATIONAL COMPLEXITY

In the case of the multi-hop RIS-assisted communication, the main differences lie on the computation of the channel matrix (18) which has complexity of $\mathcal{O}(N_{tx} N_{ris}^2 (N_{pan} - 1) + N_{tx} N_{ris} N_{rx})$, on the gradients $\nabla_{\varphi_i^*} f$ ($i = 1 \dots N_{pan}$) which require $\mathcal{O}(N_{ris}^2 N_{rx} (N_{pan} - 1) + N_s N_{tx} N_{rx} + N_{tx} N_{ris} N_{rx} N_{pan})$ operations and also on the computation of the objective function $f(\mathbf{F}, \varphi_1, \dots, \varphi_{N_{pan}})$ in step 7, which takes $\mathcal{O}(N_{tx} N_{ris}^2 (N_{pan} - 1) + N_{tx} N_{ris} N_{rx} + N_s N_{tx} N_{rx})$ operations. Keeping only the dominant terms, the overall computational complexity order of the JPR-mAPG algorithm becomes $\mathcal{O}(Q((N_{tx} N_{ris}^2 + 16 N_{ris}^2 N_{rx})(N_{pan} - 1) + N_{tx} N_{ris} N_{rx} N_{pan} + N_s N_{tx} N_{rx}))$. Similarly to the parallel RIS configuration, the complexity grows with the product of the transmitter, receiver and RIS array sizes. However, in this case it also reveals a quadratic dependency on the panel size as a result of the cascaded channels. This dependency also exists in the R-APG algorithm [12] (extended to the multi-hop RIS-assisted configuration), since its computational complexity is $\mathcal{O}(N_{rx}^3 + Q((N_{tx} N_{ris}^2 + N_{ris}^2 N_{rx})(N_{pan} - 1) + N_{rx}^3 + N_{tx} N_{rx}^2 + N_{tx} N_{ris} N_{rx} N_{pan}))$.

IV. NUMERICAL RESULTS

In this section we evaluate the achievable rate of the proposed approach with the aid of Monte Carlo simulations. Two main

scenarios are considered according to the two versions of the JPR-mAPG algorithm described previously namely, one based on multiple parallel RIS and one based on multi-hop.

A. CHANNEL MODEL

Even though the proposed approach is independent of a specific channel model, in this evaluation we adopt a clustered geometric channel [35], that is commonly assumed in mmWave and sub-THz literature [36], [37]. In the case of the channels between the transmitter, RIS and receiver, we consider that the RIS panels were properly located so that they consist of a LOS and a non-line-of-sight (NLOS) component, i.e., $\mathbf{H}_{S,i} = \mathbf{H}_{S,i}^{LOS} + \mathbf{H}_{S,i}^{NLOS}$ and $\mathbf{H}_{i,D} = \mathbf{H}_{i,D}^{LOS} + \mathbf{H}_{i,D}^{NLOS}$. We consider a spherical wave model for distances between transmitter, RIS and receiver smaller than the Fraunhofer distance defined as $D_F \triangleq 2 \frac{L_{\text{array}}^2}{\lambda}$ [38], where L_{array} is the maximum overall dimension of the array. Assuming unit normalized power radiation patterns along the directions of interest for both the antennas and RIS elements we can express the components of $\mathbf{H}_{S,i}^{LOS}$ as [39]

$$H_{S,i}^{LOS}(n, m) = \sqrt{\frac{G_{\text{tx}} A_{\text{RIS}}}{4\pi (d_{n,m}^{LOS,S,i})^2} e^{-k_{\text{abs}}(f) d_{n,m}^{LOS,S,i}} e^{-j2\pi \frac{d_{n,m}^{LOS,S,i}}{\lambda}}}, \quad (27)$$

where $d_{n,m}^{LOS,S,i}$ denotes the distance of the LOS path between the m^{th} transmit antenna element and the n^{th} RIS element of the i^{th} panel, G_{tx} is the transmit antenna gain, A_{RIS} is the area of a RIS element, and $k_{\text{abs}}(f)$ represents the molecular absorption coefficient at frequency f . Regarding the components of $\mathbf{H}_{S,i}^{NLOS}$, assuming N_{ray} paths, they can be depicted as

$$H_{S,i}^{NLOS}(n, m) = \frac{1}{\sqrt{K_{\text{Rice}}}} \sum_{l=1}^{N_{\text{ray}}} \sqrt{\frac{G_{\text{tx}} A_{\text{RIS}}}{4\pi (d_{l,n,m}^{S,i})^2} e^{-k_{\text{abs}}(f) d_{l,n,m}^{S,i}}} \times |\alpha_{l,n,m}^{S,i}| e^{-j2\pi \frac{d_{l,n,m}^{S,i}}{\lambda}}, \quad (28)$$

where K_{Rice} defines the energy ratio between the LOS and NLOS components, $\alpha_{l,n,m}^{S,i}$ is the complex gain of the l^{th} path between element m and n of the BS and i^{th} RIS panel (normalized as $\sum_{l=1}^{N_{\text{ray}}} E[|\alpha_{l,n,m}^{\text{RIS},S}|^2] = 1$) and $d_{l,n,m}^{S,i}$ denotes the respective distance. The components of $\mathbf{H}_{i,D}^{LOS}$ and $\mathbf{H}_{i,D}^{NLOS}$ can be written using expressions similar to (27) and (28).

When the distances between transmitter, RIS and receiver are larger than the Fraunhofer distance, we adopt a planar wave model. In this case, the respective NLOS channel matrix is approximated as

$$\mathbf{H}_{S,i}^{NLOS} = \sqrt{\frac{\beta_{\text{NLOS}}^{S,i}}{K_{\text{Rice}}}} \sum_{l=1}^{N_{\text{ray}}} \alpha_{l,n,m}^{S,i} \mathbf{a}_{\text{RIS}}(\phi_l^{i \leftarrow S}, \theta_l^{i \leftarrow S}) \mathbf{a}_S(\phi_l^{S \rightarrow i}, \theta_l^{S \rightarrow i})^H \quad (29)$$

where $\mathbf{a}_{\text{RIS}}(\phi_l^{i \leftarrow S}, \theta_l^{i \leftarrow S})$ and $\mathbf{a}_S(\phi_l^{S \rightarrow i}, \theta_l^{S \rightarrow i})$ denote the array responses of the BS and RIS panel at the respective angles of departure (AoD) and angles of arrival (AoA)

$(\phi_l^{i \leftarrow S}, \theta_l^{i \leftarrow S})$ and $(\phi_l^{S \rightarrow i}, \theta_l^{S \rightarrow i})$ (ϕ denotes the azimuth and θ the elevation). $\beta_{\text{NLOS}}^{S,i}$ represents the path loss, which we express as

$$\beta_{\text{NLOS}}^{S,i} = \frac{G_{\text{tx}} A_{\text{RIS}}}{4\pi (d_{S \leftrightarrow \text{RIS}}^i)^\gamma} e^{-k_{\text{abs}}(f) d_{S \leftrightarrow \text{RIS}}^i}, \quad (30)$$

where $d_{S \leftrightarrow \text{RIS}}^i$ denotes the distance between the center of the transmitter array and the center of the i^{th} RIS panel and γ is the path loss exponent. A similar expression is used for the LOS matrix (in this case with only one path). For the direct channel between the BS and the user, we consider that due to the presence of many obstructing obstacles it will typically consist solely of NLOS components and, therefore, can be expressed similarly to (28) or (29).

In all simulations uniform planar arrays (UPAs) with $\sqrt{N_{\text{tx}}} \times \sqrt{N_{\text{tx}}}$, $\sqrt{N_{\text{rx}}} \times \sqrt{N_{\text{rx}}}$ and $\sqrt{N_{\text{ris}}} \times \sqrt{N_{\text{ris}}}$ elements are adopted at the transmitter, receiver and RIS, respectively. In this case, the steering vectors for the transmitter array are given by

$$\mathbf{a}_S(\phi_l^{S \rightarrow i}, \theta_l^{S \rightarrow i}) = \left[1, \dots, e^{j\frac{2\pi}{\lambda} d_S (p \sin \phi_l^{S \rightarrow i} \sin \theta_l^{S \rightarrow i} + q \cos \theta_l^{S \rightarrow i})}, \dots, e^{j\frac{2\pi}{\lambda} d_S ((\sqrt{N_{\text{tx}}}-1) \sin \phi_l^{S \rightarrow i} \sin \theta_l^{S \rightarrow i} + (\sqrt{N_{\text{tx}}}-1) \cos \theta_l^{S \rightarrow i})} \right]^T, \quad (31)$$

where p, q are the antenna indices, and d_S is the inter-element spacing. A similar expression is adopted for the RIS and receiver arrays. In the simulations, the inter-element spacing of the transmitter and receiver arrays, as well as the distance between the centers of adjacent elements of the RIS are always set as $\lambda/2$, whereas the sizes of the RIS elements are $\lambda/2 \times \lambda/2$.

B. MULTIPLE PARALLEL RIS SCENARIO

In the first scenario we consider a large indoor environment (office/shopping mall) where one or more RISs are deployed in the vicinity of the BS. The transmission is configured with $N_s = 3$, $N_{\text{tx}} = 64$ and $N_{\text{rx}} = 16$. The operating frequency is set to 28 GHz with an occupied bandwidth of 800 MHz. We assume that the amplitude of the reflection coefficients is $a = 1$. Based on [40] we apply a path loss exponent of 1.90 for the LOS propagation and 4.39 for the NLOS propagation. The number of propagation paths is set to $N_{\text{ray}} = 10$, with the azimuth and elevation AoD/AoA uniformly random distributed in $(-\pi, \pi)$ and $(-\pi/2, \pi/2)$, respectively [41]. For the indirect link through the RIS we apply $K_{\text{Rice}} = 10$, whereas the direct link has no LOS component ($K_{\text{Rice}} = 0$). All simulated results are averaged over 500 random channel realizations. To simplify the setups, we work in a two-dimensional coordinate system (the third coordinate is assumed to be the same for the BS, RIS panels and receiver) with the BS located at (0,0) m and the receiver at (50,0) m.

In Fig. 3 we evaluate the convergence behavior of the JPR-mAPG algorithm. The RIS panels have $N_{\text{ris}} = 256$ elements each and are placed at positions (10,10), (10,-10),

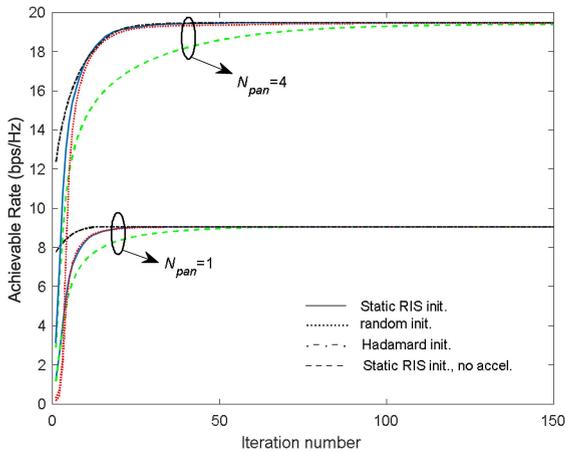


FIGURE 3. Convergence behavior of the JPR-mAPG algorithm with multiple parallel RISs. Scenario with: $N_s = 3$, $N_{tx} = 64$, $N_{rx} = 16$, $N_{ris} = 256$ and $P_{tx} = 30$ dBm.

(7,12) and (7,-12) m. For the $N_{pan} = 1$ curves, the panel is located at the first position. Different initializations of the algorithm are considered in the figure, namely a static RIS with all the elements set as $\phi_i^{(0)} = 1$ (simple reflector), a random initialization of the phases and a Hadamard based codebook scheme [42]. It can be observed that even if the initialization corresponds to a higher initial achievable rate, it does not seem to have a relevant impact on the quality of the final solution provided by the algorithm. The curves referred in the legend as “no accel.” represent the case where the JPR-mAPG algorithm is operating with $\mathbf{P}^{(q)} = \mathbf{F}^{(q)}$ and $\mathbf{y}_i^{(q)} = \phi_i^{(q)}$ in step 4. This results in $\mathbf{F}^{(q+1)} = \mathbf{W}^{(q+1)} = \mathbf{U}^{(q+1)}$ and $\phi_i^{(q+1)} = \mathbf{z}_i^{(q+1)} = \mathbf{v}_i^{(q+1)}$, which means that no extrapolated and monitoring variables are being used and the algorithm reduces to a conventional proximal gradient approach (only steps 6 and 9 of Algorithm 1 are applied). It can be observed that while the normal JPR-mAPG algorithm requires less than 25 iterations to converge, the version without acceleration can take from 50 to over 100 iterations, depending on the number of panels.

Fig. 4 shows the achievable rate as a function of the transmitted power with $N_{ris} = 256$ and only one panel. The evaluation is performed with the RIS in two different positions: (10,10) m and (25,10) m. As benchmarks, we include curves for a static RIS (acting as a simple mirror), the R-APG from [12], the AO method from [15] and the PGM approach from [17]. To limit the computational complexity, the maximum number of iterations in all these benchmark algorithms is set to 300. As a reference we also include the case of no RIS. It can be seen that the adoption of a RIS can substantially improve the achievable rate over a transmission without RIS, with the proposed JPR-mAPG algorithm providing a 5-fold improvement at $P_{tx} = 30$ dBm. We can also observe that this gain is higher when the RIS is closer to the BS (or closer to the user, even though these curves were not included). Comparing the different RIS approaches, we can

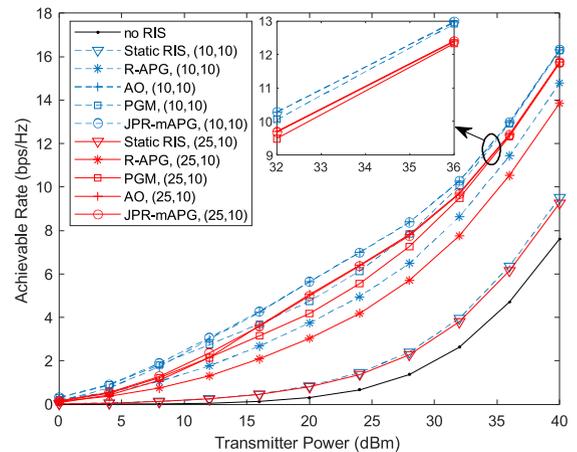


FIGURE 4. Achievable rate versus transmitted power with different algorithms and different locations of a single RIS panel. Scenario with: $N_s = 3$, $N_{tx} = 64$, $N_{rx} = 16$ and $N_{ris} = 256$.

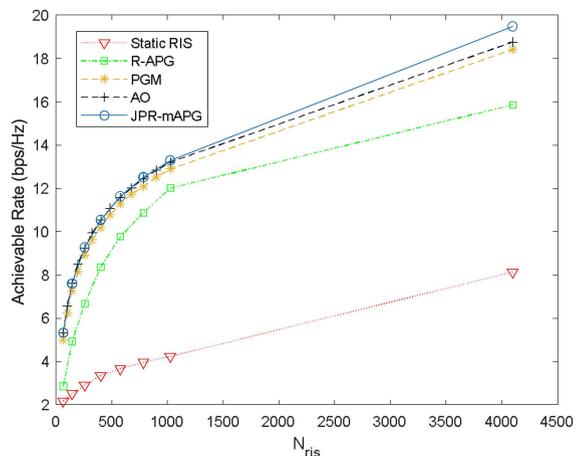


FIGURE 5. Achievable rate versus N_{ris} with different algorithms and a single RIS panel. Scenario with: $N_s = 3$, $N_{tx} = 64$, $N_{rx} = 16$ and $P_{tx} = 30$ dBm.

see that both JPR-mAPG and AO achieve similar results, with PGM also achieving close results at high transmission powers. R-APG displays a somewhat inferior performance as it only tries to optimize the RIS matrix after the precoder is computed.

Keeping basically the same setup but with $P_{tx} = 30$ dBm and a single RIS positioned at (10,10) m, Fig. 5 evaluates the impact of the number of RIS elements using different algorithms. We can observe that the achievable rate clearly improves with the number of RIS elements, with the curves showing a higher slope for low N_{ris} values, which gradually decreases as the number of elements grows. Similarly to the previous figure, all the evaluated RIS optimization algorithms provide substantial gains over a static RIS, with the proposed JPR-mAPG approach achieving the best results, especially with larger RISs.

Considering the same scenario, Fig. 6 shows the achievable rate versus transmitter-receiver distance, depicting the effect of using multiple RIS panels combined with the

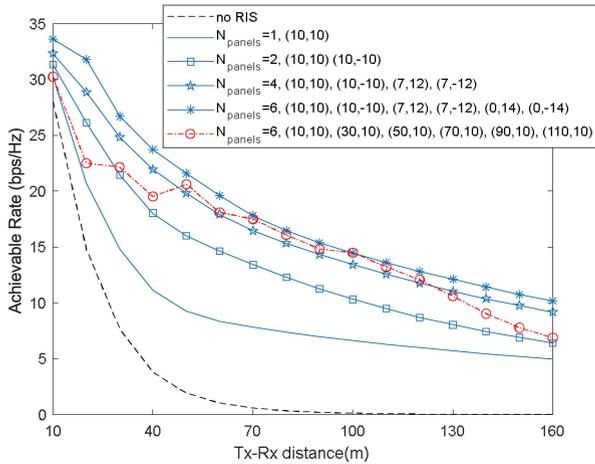


FIGURE 6. Achievable rate versus distance of the proposed JPR-mAPG algorithm with different multi-RIS assisted configurations. Scenario with $N_s = 3$, $N_{tx} = 64$, $N_{rx} = 16$, $N_{ris} = 256$ and $P_{tx} = 30$ dBm.

proposed algorithm. As expected, the adoption of multiple RIS panels can improve the rates even further. However, similarly to what was observed in the single panel case, the location of the various panels can have a significant impact on the results. Considering the case $N_{pan} = 6$, it can be seen that spreading the panels along a straight path starting from the transmitter instead of locating them all in the vicinity of the BS can result in a reduced number of panels being effective in aiding the communication depending on the position of the receiver. It is important to note that while the optimal deployment of multiple RIS is outside the scope of this paper, this behavior is in line with previous studies that stated that placing RISs near the BS and user can not only minimize the product-distance path-loss but also increase the path diversity and increase the resilience of the network against signal blockage [25], [26]. Still, finding optimal RIS deployments as well as RIS-user association and multi-hop RIS-aided routing are still active topics of research [25], [30].

While the proposed JPR-mAPG algorithm assumes continuous phase shifts, in practical implementations of RIS-assisted systems only a finite number of phase shifts can be supported in the RIS elements [43]. In this case we adopt the simplest approach that is commonly followed with continuous based algorithms which relies on directly quantizing the obtained final solutions to the nearest discrete values [17]. Using this approach, Fig. 7 evaluates the impact of using discrete phase shifts at the RIS, considering the same setup of the previous figures. Curves for different numbers of multiple RIS panels are presented, where the locations are assumed to be the same as those used in Fig. 6 (the case $N_{pan} = 6$ corresponds to the deployment where the panels are all in the vicinity of the BS). As expected, the quantization of the phase shifts results in some performance loss which is worse when using one bit quantization. In this case the degradation ranges between 16%-22%, depending on the number of panels. Still, increasing the number of quantization bits

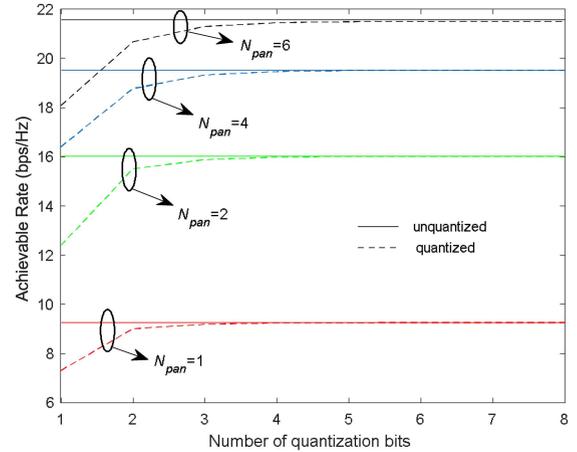


FIGURE 7. Achievable rate versus number of quantization bits when using the proposed JPR-mAPG algorithm with different multi-RIS assisted configurations with discrete phase shifts. Scenario with $N_s = 3$, $N_{tx} = 64$, $N_{rx} = 16$, $N_{ris} = 256$ and $P_{tx} = 30$ dBm.

strongly reduces the degradation, and with 3 bits the loss to the ideal unquantized curves is under 1% for all the different sets of panels.

C. MULTI-HOP RIS-ASSISTED SCENARIO

In the second scenario we consider the transmission between an access point (AP) and user which is supported with the aid of two RIS panels, establishing a two-hop communication scheme. The layout with the positions of AP, user and RIS panels is presented in Fig. 8. In this case the system is operating at the sub-THz band with the frequency set to 142 GHz and the occupied bandwidth kept at 800 MHz. The number of propagation paths is set to $N_{ray} = 3$ and $K_{Rice} = 10$. For the AP→RIS1, RIS1→RIS2 and RIS2→user links we consider LOS propagation with a path loss exponent of 2.05 [40]. For the direct link AP→user (as well as for the RIS1→user link considered for the $N_{pan} = 1$ results) we assume NLOS propagation with a path loss exponent of 4.6 [40]. The other parameters are configured as $N_s = 1$, $N_{tx} = 256$, $N_{rx} = 36$.

In Fig. 9 we study the convergence behavior of the JPR-mAPG algorithm in this multi-hop scenario ($N_{pan} = 2$). Both panels are set with $N_{ris} = 2500$ elements. As previously observed in the parallel RIS scenario, the normal algorithm with the extrapolation steps converges faster than the “no accel.” version which relies on a direct projected gradient step. The former requires around 50 iterations whereas the latter needs almost 150 iterations to converge. It can also be observed that in this multi-hop configuration, there seems to be a small impact on the convergence speed when initializing both RIS panels as static reflectors instead of adopting random phases.

Keeping the same sizes of the RIS panels as in the previous figure, Fig. 10 shows the achievable rate versus transmitted power of the proposed JPR-mAPG approach as well as of the other benchmark algorithms, considering both two-hop and one-hop links. The one-hop case corresponds to having RIS1

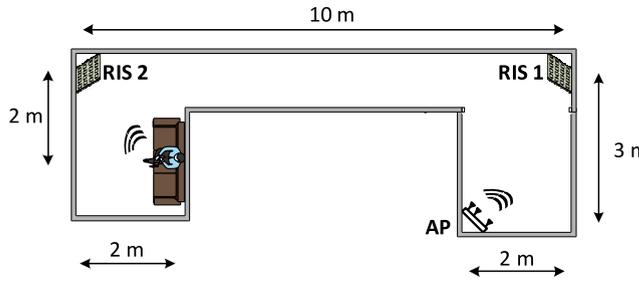


FIGURE 8. Overhead layout of the simulated multi-hop scenario.

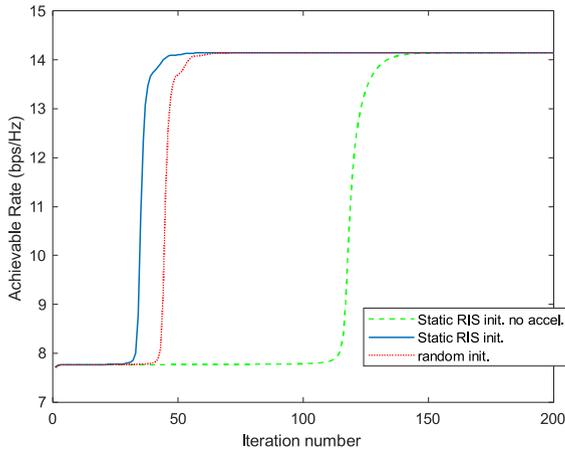


FIGURE 9. Convergence behavior of the JPR-mAPG algorithm in a two-hop RIS-aided link. Scenario with: $N_s = 1$, $N_{tx} = 256$, $N_{rx} = 36$, $N_{ris} = 2500$, $N_{pan} = 2$ and $P_{tx} = 30$ dBm.

in operation and RIS2 removed. Curves with $N_{pan} = 2$ are not included for PGM and AM since these algorithms were designed for one-hop links. With only one RIS panel, the gain provided by any of the algorithms over the case without a RIS is very small. This small gain is a consequence of the RIS1→user link being NLOS with a high path loss exponent. So, even if the AP→RIS1 is LOS and has a low path loss exponent, the path loss of the indirect link depends on the product of these partial path losses and therefore, it would be necessary to adopt a very large RIS to obtain a strong indirect link which could result in substantial gains. In this case, the two-hop link is more effective since by using two similarly sized panels that enable establishing an indirect link where all segments have LOS propagation with a low path loss exponent, the optimized RISs are able to provide substantial gains even under the multiplicative path loss effect. In fact, for a transmitted power of $P_{tx} = 30$ dBm, the two-hop JPR-mAPG algorithm is able to increase the achievable rate from 7.4 bps/Hz to 13.4 bps/Hz. Similarly to what was observed in the multiple parallel RIS scenario, adopting a joint design of the precoder and RIS (JPR-mAPG), achieves better performances than applying an algorithm that only optimizes the RIS panels (R-APG).

In order to get a better insight on the impact of the RIS panels, especially at these high frequency bands, Fig. 11 presents the achievable rate for the same settings of Fig. 10, but considering the existence of the indirect link only. Curves

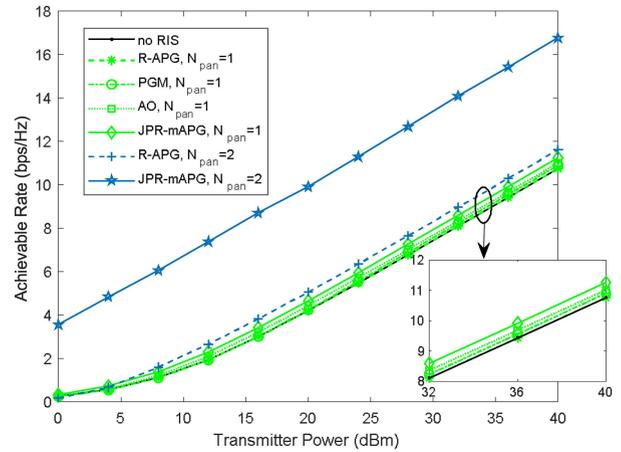


FIGURE 10. Achievable rate versus transmitter power in the multi-hop RIS-assisted setting (direct link and indirect link). Scenario with $N_s = 1$, $N_{tx} = 256$, $N_{rx} = 36$, $N_{ris} = 2500$.

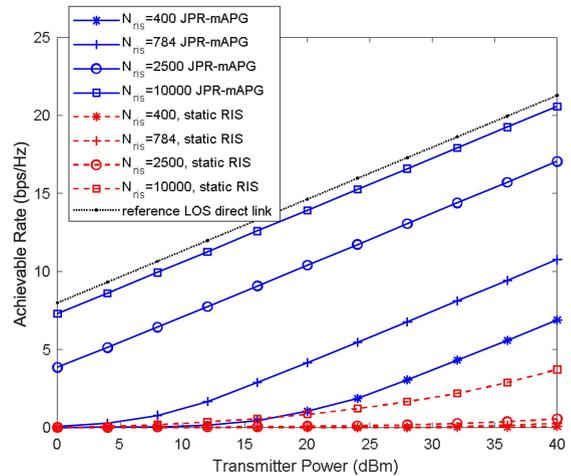


FIGURE 11. Achievable rate versus transmitter power in a two-hop RIS-assisted setting (only indirect link). Scenario with $N_s = 1$, $N_{tx} = 256$, $N_{rx} = 36$, $N_{pan} = 2$.

with different numbers of RIS elements per panel, N_{ris} , are shown. Besides the results with the JPR-mAPG algorithm, we include results with static RISs (simple reflectors) and also a curve representing a reference direct transmission with LOS propagation, a path loss exponent of 2.05 and a distance equal to the overall length of the dual-hop RIS aided link. We can see that working in a higher frequency band and with a multi-hop scheme requires the adoption of properly optimized RIS panels with a higher number of elements (even though not necessarily larger due to the smaller wavelength) in order to achieve acceptable rates. In fact, adopting simple reflectors requires large panels in order to obtain some noticeable improvements on the achievable rate. In contrast, if the RISs are optimized using the JPR-mAPG algorithm, they can overcome the multiplicative effect of the path losses of the segments comprising the indirect link and large improvements can be observed in the achievable rate. This capability to overcome the higher path loss caused by multiple signal reflections and provide substantial beam-forming gains becomes more evident when the number of

RIS elements is large, which is in line with the conclusions in [31]. In fact, with $N_{ris} = 10000$, the behavior of the indirect link is becoming close to the reference LOS direct link whose path loss depends solely on the sum of the lengths of the different segments. This type of behavior was previously discussed for single RIS in [44].

V. CONCLUSION

In this paper we addressed the joint design of the transmit precoder and RIS phases within the context of multi-stream MIMO communications assisted by multiple RIS panels. Aiming at solving the nonconvex problem of maximizing the achievable rate, we derived an algorithm based on the mAPG method that can be tailored for the case of parallel RISs as well as for multi-hop RIS-assisted scenarios. The mAPG approach relies on an extrapolation step for improving the convergence speed and a set of monitoring variables that ensure the sufficient descent of the algorithm. Detailed convergence analysis was presented which included expressions for the step size that guarantee convergence to a critical point. Simulation results showed that the proposed JPR-mAPG algorithm can often achieve higher rates than other benchmarked approaches. Furthermore, it was observed that the use of properly located multiple RIS panels operating in parallel can have a substantial impact on the achieved rates, whereas the adoption of several RIS panels in a multi-hop configuration can be an effective approach for providing coverage in the case of severe blockage effects.

APPENDIX A PROOF OF LEMMA 1

The following derivations for the complex-valued gradient of $f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}})$ with respect to \mathbf{F}^* and $\boldsymbol{\varphi}_i^*$ ($i = 1, \dots, N_{pan}$) are based on the procedure described in [45]. First, we derive the complex differential of $f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}})$ with respect to \mathbf{F}^* as

$$\begin{aligned} df &= -\text{Tr} \left\{ \frac{\rho}{P_n} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \right. \\ &\quad \left. \times \left(\mathbf{F}^H \mathbf{H}^H \mathbf{H} d\mathbf{F} + d\mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right) \right\} \\ &= -\text{Tr} \left\{ \frac{\rho}{P_n} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{H}^H \mathbf{H} d\mathbf{F} \right. \\ &\quad \left. + \frac{\rho}{P_n} \left(\mathbf{H}^H \mathbf{H} \mathbf{F} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \right)^T d\mathbf{F}^* \right\}. \end{aligned} \quad (32)$$

where we used the relation [45, Table 3.1]

$$d(\text{Indet}(\mathbf{X})) = \text{Tr} \left\{ \mathbf{X}^{-1} d\mathbf{X} \right\}. \quad (33)$$

Then, using [45, Table 3.2] we obtain (9). Regarding the complex differential of $f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}})$ with respect to

$\boldsymbol{\varphi}_i^*$, it can be written as

$$\begin{aligned} df &= -\text{Tr} \left\{ \frac{\rho}{P_n} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \right. \\ &\quad \left. \times \left(\mathbf{F}^H \mathbf{H}^H d\mathbf{H} \mathbf{F} + \mathbf{F}^H d\mathbf{H}^H \mathbf{H} \mathbf{F} \right) \right\}. \end{aligned} \quad (34)$$

Noting that

$$d\mathbf{H} = \mathbf{H}_{i,D} \text{diag}(d\boldsymbol{\varphi}_i) \mathbf{H}_{S,i} \quad (35)$$

and

$$\text{Tr} \{ \mathbf{Z} \text{diag}(d\boldsymbol{\varphi}_i) \} = \text{diag}(\mathbf{Z})^T d\boldsymbol{\varphi}_i \quad (36)$$

then we obtain

$$\begin{aligned} df &= -\frac{\rho}{P_n} \text{diag} \left[\mathbf{H}_{S,i} \mathbf{F} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{H}^H \mathbf{H}_{i,D} \right]^T d\boldsymbol{\varphi}_i \\ &\quad - \frac{\rho}{P_n} \text{diag} \left[\mathbf{H}_{i,D}^H \mathbf{H} \mathbf{F} \left(\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{H}_{S,i} \right]^T d\boldsymbol{\varphi}_i^*. \end{aligned} \quad (37)$$

Again, using [45, Table 3.2] gives the gradient expression in (10).

APPENDIX B PROOF OF THEOREM 1

The following inequalities will be useful for the proof. For matrices \mathbf{A} and \mathbf{B} , we have

$$\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\| \|\mathbf{B}\| \quad (38)$$

for the Frobenius and the 2-norm. Also

$$\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F \quad (39)$$

and

$$\|\mathbf{A}\mathbf{B}\|_F \leq \|\mathbf{A}\|_2 \|\mathbf{B}\|_F. \quad (40)$$

Before proving Theorem 1, we first introduce the following Lemma which will be applied in several steps of the proof.

Lemma 2: Let $\{\boldsymbol{\Psi}_k^{(i)}\}$ with $i=1,2$, and $k=1,\dots,N$ ($N \geq 2$) be a set of matrices. Then we have the following inequality

$$\begin{aligned} \left\| \prod_{k=1}^N \boldsymbol{\Psi}_k^{(2)} - \prod_{k=1}^N \boldsymbol{\Psi}_k^{(1)} \right\|_F &\leq \sum_{l=1}^N \left(\prod_{m=l+1}^N \|\boldsymbol{\Psi}_m^{(1)}\|_2 \right) \|\boldsymbol{\Psi}_l^{(2)} - \boldsymbol{\Psi}_l^{(1)}\|_F \\ &\quad \times \left(\prod_{k=1}^{l-1} \|\boldsymbol{\Psi}_k^{(2)}\|_2 \right) \end{aligned} \quad (41)$$

Proof: We prove by induction on N . For $N=2$ we can write:

$$\begin{aligned} &\left\| \boldsymbol{\Psi}_2^{(2)} \boldsymbol{\Psi}_1^{(2)} - \boldsymbol{\Psi}_2^{(1)} \boldsymbol{\Psi}_1^{(1)} \right\|_F \\ &\leq \left\| \boldsymbol{\Psi}_2^{(2)} \boldsymbol{\Psi}_1^{(2)} - \boldsymbol{\Psi}_2^{(1)} \boldsymbol{\Psi}_1^{(2)} \right\|_F + \left\| \boldsymbol{\Psi}_2^{(1)} \boldsymbol{\Psi}_1^{(2)} - \boldsymbol{\Psi}_2^{(1)} \boldsymbol{\Psi}_1^{(1)} \right\|_F \\ &\leq \left\| \boldsymbol{\Psi}_2^{(2)} - \boldsymbol{\Psi}_2^{(1)} \right\|_F \|\boldsymbol{\Psi}_1^{(2)}\|_2 + \left\| \boldsymbol{\Psi}_2^{(1)} \right\|_2 \|\boldsymbol{\Psi}_1^{(2)} - \boldsymbol{\Psi}_1^{(1)}\|_F \end{aligned}$$

$$\leq \sum_{l=1}^2 \left(\prod_{m=l+1}^2 \left\| \Psi_m^{(1)} \right\|_2 \right) \left\| \Psi_l^{(2)} - \Psi_l^{(1)} \right\|_F \left(\prod_{k=1}^{l-1} \left\| \Psi_k^{(2)} \right\|_2 \right) \quad (42)$$

thus showing that the inequality is valid. For $N+1$ we can write

$$\begin{aligned} \left\| \prod_{k=1}^{N+1} \Psi_k^{(2)} - \prod_{k=1}^{N+1} \Psi_k^{(1)} \right\|_F &= \left\| \Psi_{N+1}^{(2)} \prod_{k=1}^N \Psi_k^{(2)} - \Psi_{N+1}^{(1)} \prod_{k=1}^N \Psi_k^{(1)} \right\|_F \\ &\leq \left\| \Psi_{N+1}^{(2)} \prod_{k=1}^N \Psi_k^{(2)} - \Psi_{N+1}^{(1)} \prod_{k=1}^N \Psi_k^{(2)} \right\|_F \\ &\quad + \left\| \Psi_{N+1}^{(1)} \prod_{k=1}^N \Psi_k^{(2)} - \Psi_{N+1}^{(1)} \prod_{k=1}^N \Psi_k^{(1)} \right\|_F \\ &\leq \left\| \Psi_{N+1}^{(2)} - \Psi_{N+1}^{(1)} \right\|_F \prod_{k=1}^N \left\| \Psi_k^{(2)} \right\|_2 \\ &\quad + \left\| \Psi_{N+1}^{(1)} \right\|_2 \left\| \prod_{k=1}^N \Psi_k^{(2)} - \prod_{k=1}^N \Psi_k^{(1)} \right\|_F. \end{aligned} \quad (43)$$

Assume that inequality (41) holds for N . Then, we can rewrite (43) as

$$\begin{aligned} \left\| \prod_{k=1}^{N+1} \Psi_k^{(2)} - \prod_{k=1}^{N+1} \Psi_k^{(1)} \right\|_F &\leq \left\| \Psi_{N+1}^{(2)} - \Psi_{N+1}^{(1)} \right\|_F \prod_{k=1}^N \left\| \Psi_k^{(2)} \right\|_2 \\ &\quad + \left\| \Psi_{N+1}^{(1)} \right\|_2 \sum_{l=1}^N \left(\prod_{m=l+1}^N \left\| \Psi_m^{(1)} \right\|_2 \right) \\ &\quad \times \left\| \Psi_l^{(2)} - \Psi_l^{(1)} \right\|_F \left(\prod_{k=1}^{l-1} \left\| \Psi_k^{(2)} \right\|_2 \right) \\ &\leq \left\| \Psi_{N+1}^{(2)} - \Psi_{N+1}^{(1)} \right\|_F \prod_{k=1}^N \left\| \Psi_k^{(2)} \right\|_2 \\ &\quad + \sum_{l=1}^N \left(\prod_{m=l+1}^{N+1} \left\| \Psi_m^{(1)} \right\|_2 \right) \left\| \Psi_l^{(2)} - \Psi_l^{(1)} \right\|_F \left(\prod_{k=1}^{l-1} \left\| \Psi_k^{(2)} \right\|_2 \right). \end{aligned} \quad (44)$$

Noting that

$$\begin{aligned} \left\| \Psi_{N+1}^{(2)} - \Psi_{N+1}^{(1)} \right\|_F \prod_{k=1}^N \left\| \Psi_k^{(2)} \right\|_2 &= \left(\prod_{m=N+1+1}^{N+1} \left\| \Psi_m^{(1)} \right\|_2 \right) \left\| \Psi_{N+1}^{(2)} - \Psi_{N+1}^{(1)} \right\|_F \prod_{k=1}^{N+1-1} \left\| \Psi_k^{(2)} \right\|_2 \end{aligned}$$

we can finally write

$$\left\| \prod_{k=1}^{N+1} \Psi_k^{(2)} - \prod_{k=1}^{N+1} \Psi_k^{(1)} \right\|_F$$

$$\leq \sum_{l=1}^{N+1} \left(\prod_{m=l+1}^{N+1} \left\| \Psi_m^{(1)} \right\|_2 \right) \left\| \Psi_l^{(2)} - \Psi_l^{(1)} \right\|_F \left(\prod_{k=1}^{l-1} \left\| \Psi_k^{(2)} \right\|_2 \right) \quad (45)$$

which shows that the inequality is also valid. ■

We note that inequality (41) can also be written using the spectral norm since $\|\mathbf{X}\|_2 \leq \|\mathbf{X}\|_F$.

Proof of Theorem 1: First we define the following concatenated matrices and vectors

$$\tilde{\boldsymbol{\varphi}} = \left[\boldsymbol{\varphi}_1^T, \dots, \boldsymbol{\varphi}_{N_{pan}}^T \right]^T, \quad (46)$$

$$\tilde{\mathbf{H}}_{R,D} = \left[\mathbf{H}_{1,D} \cdots \mathbf{H}_{N_{pan},D} \right], \quad (47)$$

$$\tilde{\mathbf{H}}_{S,R} = \left[\mathbf{H}_{S,1}^T \cdots \mathbf{H}_{S,N_{pan}}^T \right]^T \quad (48)$$

and $\tilde{\Phi} = \text{diag}(\tilde{\boldsymbol{\varphi}})$ which allow us to rewrite (2) in a more compact form

$$\mathbf{H} = \mathbf{H}_{S,D} + \tilde{\mathbf{H}}_{R,D} \tilde{\Phi} \tilde{\mathbf{H}}_{S,R}. \quad (49)$$

This allows us to also work with the concatenated gradient vector

$$\begin{aligned} \nabla_{\tilde{\boldsymbol{\varphi}}}^* f(\mathbf{F}, \tilde{\boldsymbol{\varphi}}) &= \left[\nabla_{\boldsymbol{\varphi}_1^*} f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}}) \right]^T, \dots, \\ &\quad \left[\nabla_{\boldsymbol{\varphi}_{N_{pan}}^*} f(\mathbf{F}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{N_{pan}}) \right]^T \end{aligned} \quad (50)$$

and write

$$\begin{aligned} &\left\| \nabla_{\tilde{\boldsymbol{\varphi}}}^* f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\tilde{\boldsymbol{\varphi}}}^* f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_2 \\ &= \frac{\rho}{P_n} \times \left\| \text{diag} \left(\text{blkdiag} \left\{ (\mathbf{H}_{i,D})^H \mathbf{H}^{(q_2)} \right. \right. \right. \\ &\quad \left. \left. \left. \mathbf{F}^{(q_2)} \mathbf{K}(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) \mathbf{F}^{(q_2)H} (\mathbf{H}_{S,i})^H, \dots \right. \right. \right. \\ &\quad \left. \left. \left. - (\mathbf{H}_{i,D})^H \mathbf{H}^{(q_1)} \mathbf{F}^{(q_1)} \mathbf{K}(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \mathbf{F}^{(q_1)H} (\mathbf{H}_{S,i})^H \right\} \right) \right\|_2. \end{aligned} \quad (51)$$

with $\mathbf{K}(\mathbf{F}, \tilde{\boldsymbol{\varphi}}) = (\mathbf{I}_{N_s} + \frac{\rho}{P_n} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F})^{-1}$. We can rewrite this expression as

$$\begin{aligned} &\left\| \nabla_{\tilde{\boldsymbol{\varphi}}}^* f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\tilde{\boldsymbol{\varphi}}}^* f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_2 \\ &\leq \frac{\rho}{P_n} \sum_{i=1}^{N_{pan}} \left\| (\mathbf{H}_{i,D})^H \mathbf{H}^{(q_2)} \mathbf{F}^{(q_2)} \mathbf{K}(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) \mathbf{F}^{(q_2)H} (\mathbf{H}_{S,i})^H \right. \\ &\quad \left. - (\mathbf{H}_{i,D})^H \mathbf{H}^{(q_1)} \mathbf{F}^{(q_1)} \mathbf{K}(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \mathbf{F}^{(q_1)H} (\mathbf{H}_{S,i})^H \right\|_F. \end{aligned} \quad (52)$$

It can be easily shown that $\|\mathbf{K}(\mathbf{F}, \tilde{\boldsymbol{\varphi}})\|_2 \leq 1$ and $\|\Phi_i\|_2 = a$. Using Lemma 2 we can write

$$\left\| \mathbf{H}^{(q_2)} - \mathbf{H}^{(q_1)} \right\|_F \leq \lambda_{\max}(\tilde{\mathbf{H}}_{R,D}) \lambda_{\max}(\tilde{\mathbf{H}}_{S,R}) \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2 \quad (53)$$

$$\begin{aligned} &\left\| \mathbf{K}(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \mathbf{K}(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_F \\ &\leq \left\| \mathbf{K}(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)})^{-1} - \mathbf{K}(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)})^{-1} \right\|_F \end{aligned}$$

$$\begin{aligned} &\leq \frac{2\rho}{P_n} \left(\sqrt{N_s} \zeta^2 \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F \right. \\ &\quad \left. + N_s \zeta \lambda_{\max}(\tilde{\mathbf{H}}_{R,D}) \lambda_{\max}(\tilde{\mathbf{H}}_{S,R}) \right. \\ &\quad \left. \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2 \right) \end{aligned} \quad (54)$$

Applying these inequalities and Lemma 2 to (52) results

$$\begin{aligned} &\left\| \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_2 \\ &\leq \frac{\rho}{P_n} \left(\sum_{i=1}^{N_{pan}} \lambda_{\max}(\mathbf{H}_{i,D}) \lambda_{\max}(\mathbf{H}_{S,i}) \right) \\ &\quad \times \left(2\sqrt{N_s} \zeta \left(1 + \frac{N_s \rho}{P_n} \zeta^2 \right) \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F + N_s \lambda_{\max}(\tilde{\mathbf{H}}_{R,D}) \right. \\ &\quad \left. \times \lambda_{\max}(\tilde{\mathbf{H}}_{S,R}) \left(1 + 2\frac{N_s \rho}{P_n} \zeta^2 \right) \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2 \right) \\ &= d \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F + e \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2. \end{aligned} \quad (55)$$

Similarly, we can also use Lemma 2 and write the following inequality

$$\begin{aligned} &\left\| \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_F \\ &= \frac{\rho}{P_n} \left\| \mathbf{H}^{(q_2)H} \mathbf{H}^{(q_2)} \mathbf{F}^{(q_2)} \mathbf{K}(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) \right. \\ &\quad \left. - \mathbf{H}^{(q_1)H} \mathbf{H}^{(q_1)} \mathbf{F}^{(q_1)} \mathbf{K}(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_2 \\ &\leq \frac{\rho}{P_n} \zeta^2 \left(1 + 2\frac{N_s \rho}{P_n} \zeta^2 \right) \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F + 2\sqrt{N_s} \frac{\rho}{P_n} \zeta \\ &\quad \times \lambda_{\max}(\tilde{\mathbf{H}}_{R,D}) \lambda_{\max}(\tilde{\mathbf{H}}_{S,R}) \left(1 + \frac{N_s \rho}{P_n} \zeta^2 \right) \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2 \\ &= b \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F + c \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2. \end{aligned} \quad (56)$$

If we square both sides of the expressions in (55) and (56) we can rewrite the inequalities as

$$\begin{aligned} &\left\| \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_2^2 \\ &\leq (d^2 + de) \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F^2 + (e^2 + de) \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2^2 \\ &\left\| \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_F^2 \\ &\leq (b^2 + bc) \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F^2 + (c^2 + bc) \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2^2, \end{aligned} \quad (57)$$

where we applied the relation

$$\begin{aligned} &2 \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2 \\ &\leq \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F^2 + \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2^2. \end{aligned}$$

Combining both expressions and applying the square root to both sides we finally obtain

$$\begin{aligned} &\left(\left\| \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_F^2 \right. \\ &\quad \left. + \left\| \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q_2)}, \tilde{\boldsymbol{\varphi}}^{(q_2)}) - \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q_1)}, \tilde{\boldsymbol{\varphi}}^{(q_1)}) \right\|_2^2 \right)^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned} &\leq \left((b^2 + bc + d^2 + de) \left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F^2 \right. \\ &\quad \left. + (c^2 + bc + e^2 + de) \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2^2 \right)^{\frac{1}{2}} \\ &\leq L \left(\left\| \mathbf{F}^{(q_2)} - \mathbf{F}^{(q_1)} \right\|_F^2 + \left\| \tilde{\boldsymbol{\varphi}}^{(q_2)} - \tilde{\boldsymbol{\varphi}}^{(q_1)} \right\|_2^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (59)$$

This completes the proof. \blacksquare

APPENDIX C PROOF OF THEOREM 2

The following proof is based on the background provided in [32]. First, using definition (6) and noting that $\text{Tr}\{\mathbf{A} + \mathbf{A}^H\} = 2\text{Re}\{\text{Tr}\{\mathbf{A}\}\}$ for a square matrix \mathbf{A} , we can write the update of the monitoring variables (step 6 of Algorithm 1) as

$$\begin{aligned} &(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) = \underset{(\mathbf{X}, \mathbf{x})}{\text{argmin}} I_C(\mathbf{X}) + I_{\mathcal{U}_{N_{pan}N_{RIS}}}(\mathbf{x}) \\ &\quad + \text{Re} \left\{ \text{Tr} \left\{ \left(\nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \right)^H (\mathbf{X} - \mathbf{F}^{(q)}) \right\} \right\} + \frac{1}{2\alpha} \left\| \mathbf{X} - \mathbf{F}^{(q)} \right\|_F^2 \\ &\quad + \text{Re} \left\{ \left(\nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \right)^H (\mathbf{x} - \tilde{\boldsymbol{\varphi}}^{(q)}) \right\} + \frac{1}{2\alpha} \left\| \mathbf{x} - \tilde{\boldsymbol{\varphi}}^{(q)} \right\|_2^2, \end{aligned} \quad (60)$$

where we work with the concatenated vectors (46) and $\tilde{\mathbf{v}} = [\mathbf{v}_1^T, \dots, \mathbf{v}_{N_{pan}}^T]^T$. Considering that $(\mathbf{F}^{(0)}, \tilde{\boldsymbol{\varphi}}^{(0)})$ are initialized over the feasible set then we have $I_C(\mathbf{F}^{(q)}) = 0$, $I_{\mathcal{U}_{N_{pan}N_{RIS}}}(\tilde{\boldsymbol{\varphi}}^{(q)}) = 0$, $I_C(\mathbf{U}^{(q+1)}) = 0$ and $I_{\mathcal{U}_{N_{pan}N_{RIS}}}(\tilde{\mathbf{v}}^{(q+1)}) = 0$ for any q . Furthermore, since $(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)})$ minimizes (60) we can write

$$\begin{aligned} &\text{Re} \left\{ \text{Tr} \left\{ \left(\nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \right)^H (\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}) \right\} \right. \\ &\quad \left. + \left(\nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \right)^H (\tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}) \right\} \\ &\leq -\frac{1}{2\alpha} \left(\left\| \mathbf{U}^{(q+1)} - \mathbf{F}^{(q)} \right\|_F^2 + \left\| \tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)} \right\|_2^2 \right), \end{aligned} \quad (61)$$

According to Theorem 1, $f(\mathbf{F}, \tilde{\boldsymbol{\varphi}})$ has Lipschitz continuous gradients over the feasible set with constant L . Therefore, according to the Descent Lemma ([46], chapter 5) we can write the following inequality

$$\begin{aligned} &f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \leq f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \\ &\quad + \text{Re} \left\{ \text{Tr} \left\{ \left(\nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \right)^H (\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}) \right\} \right\} \\ &\quad + \frac{L}{2} \left\| \mathbf{U}^{(q+1)} - \mathbf{F}^{(q)} \right\|_F^2 \\ &\quad + \text{Re} \left\{ \left(\nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \right)^H (\tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}) \right\} \\ &\quad + \frac{L}{2} \left\| \tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)} \right\|_2^2. \end{aligned} \quad (62)$$

Combining this expression with (61) allows us to write

$$\begin{aligned} &f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \leq f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) - \left(\frac{1}{2\alpha} - \frac{L}{2} \right) \\ &\quad \left(\left\| \mathbf{U}^{(q+1)} - \mathbf{F}^{(q)} \right\|_F^2 + \left\| \tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)} \right\|_2^2 \right). \end{aligned} \quad (63)$$

Since we have $\alpha < \frac{1}{L}$ then we obtain $f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \leq f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)})$. Noting that the update of $(\mathbf{F}^{(q+1)}, \tilde{\boldsymbol{\varphi}}^{(q+1)})$ is performed according to steps 7-11 of Algorithm 1, then we can conclude that

$$f(\mathbf{F}^{(q+1)}, \tilde{\boldsymbol{\varphi}}^{(q+1)}) \leq f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \leq f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}), \quad (64)$$

i.e., the algorithm is nonincreasing over the sequence $\{(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)})\}$. Furthermore, sequences $\{(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)})\}$ and $\{(\mathbf{U}^{(q)}, \tilde{\mathbf{v}}^{(q)})\}$ are bounded since for finite valued channel matrices $f(\mathbf{F}, \tilde{\boldsymbol{\varphi}})$ is bounded from below and the constraint sets $\mathcal{U}_{N_{\text{pan}}N_{\text{RIS}}}$ and \mathcal{C} are bounded and closed. Therefore, $\{(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)})\}$ has accumulation points, with $f(\mathbf{F}, \tilde{\boldsymbol{\varphi}})$ taking the same value at those points as it is nonincreasing. We denote this value as \hat{f} . Combining (63) and (64) we have

$$\left(\frac{1}{2\alpha} - \frac{L}{2}\right) \left(\|\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}\|_F^2 + \|\tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}\|_2^2\right) \leq f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) - f(\mathbf{F}^{(q+1)}, \tilde{\boldsymbol{\varphi}}^{(q+1)}). \quad (65)$$

Applying a summation over $q = 0, \dots, +\infty$ on both sides yields

$$\left(\frac{1}{2\alpha} - \frac{L}{2}\right) \sum_{q=0}^{+\infty} \left(\|\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}\|_F^2 + \|\tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}\|_2^2\right) \leq f(\mathbf{F}^{(0)}, \tilde{\boldsymbol{\varphi}}^{(0)}) - \hat{f} < +\infty, \quad (66)$$

which implies

$$\|\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}\|_F^2 \rightarrow 0, \quad \|\tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}\|_2^2 \rightarrow 0 \quad (67)$$

when $q \rightarrow +\infty$. This means that if we denote any accumulation point of $\{(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)})\}$ as $\{(\mathbf{F}', \tilde{\boldsymbol{\varphi}}')\}$, we can always build a subsequence $\{(\mathbf{U}^{(q_j)}, \tilde{\mathbf{v}}^{(q_j)})\}$ that converges to that point when $j \rightarrow +\infty$. Furthermore, $f(\mathbf{F}, \tilde{\boldsymbol{\varphi}})$ has a continuous gradient and, thus, $f(\mathbf{U}^{(q_j+1)}, \tilde{\mathbf{v}}^{(q_j+1)}) \rightarrow f(\mathbf{F}', \tilde{\boldsymbol{\varphi}}')$. $(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)})$ is computed as in step 6 of Algorithm 1 which, according to (6) allows us to write

$$\begin{cases} 0 \in \partial I_{\mathcal{C}}(\mathbf{U}^{(q+1)}) + \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) + \frac{1}{2\alpha}(\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}) \\ 0 \in \partial I_{\mathcal{U}_{N_{\text{pan}}N_{\text{RIS}}}}(\tilde{\mathbf{v}}^{(q+1)}) + \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) + \frac{1}{2\alpha}(\tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}). \end{cases} \quad (68)$$

This condition can be rewritten as

$$\begin{aligned} 0 \in \partial G(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \\ - \left(\nabla_{\mathbf{F}^*} f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) - \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)})\right), \\ \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) - \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \\ + \frac{1}{2\alpha}(\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}, \tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}) \end{aligned} \quad (69)$$

with

$$\begin{aligned} G(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) = I_{\mathcal{C}}(\mathbf{U}^{(q+1)}) + \partial I_{\mathcal{U}_{N_{\text{pan}}N_{\text{RIS}}}}(\tilde{\mathbf{v}}^{(q+1)}) \\ + f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \end{aligned} \quad (70)$$

and

$$\begin{aligned} \partial G(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) = \left(\partial I_{\mathcal{C}}(\mathbf{U}^{(q+1)}) + \nabla_{\mathbf{F}^*} f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}\right), \\ \partial I_{\mathcal{U}_{N_{\text{pan}}N_{\text{RIS}}}}(\tilde{\mathbf{v}}^{(q+1)}) + \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \end{aligned} \quad (71)$$

To satisfy condition (69) we must have

$$\begin{aligned} \left(\nabla_{\mathbf{F}^*} f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) - \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)})\right), \\ \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) - \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q)}, \tilde{\boldsymbol{\varphi}}^{(q)}) \\ - \frac{1}{2\alpha}(\mathbf{U}^{(q+1)} - \mathbf{F}^{(q)}, \tilde{\mathbf{v}}^{(q+1)} - \tilde{\boldsymbol{\varphi}}^{(q)}) \in \partial G(\mathbf{U}^{(q+1)}, \tilde{\mathbf{v}}^{(q+1)}) \end{aligned} \quad (72)$$

Working with the subsequence $(\mathbf{U}^{(q_j+1)}, \tilde{\mathbf{v}}^{(q_j+1)}) \rightarrow (\mathbf{F}', \tilde{\boldsymbol{\varphi}}')$, and since $f(\mathbf{U}^{(q_j+1)}, \tilde{\mathbf{v}}^{(q_j+1)}) \rightarrow f(\mathbf{F}', \tilde{\boldsymbol{\varphi}}')$, then the subgradient in the left side of (72) must tend to a subgradient in $\partial G(\mathbf{F}', \tilde{\boldsymbol{\varphi}}')$ when $j \rightarrow +\infty$. Furthermore, we can write

$$\begin{aligned} \left(\left\|\nabla_{\mathbf{F}^*} f(\mathbf{U}^{(q_j+1)}, \tilde{\mathbf{v}}^{(q_j+1)}) - \nabla_{\mathbf{F}^*} f(\mathbf{F}^{(q_j)}, \tilde{\boldsymbol{\varphi}}^{(q_j)})\right\| \right. \\ \left. - \frac{1}{2\alpha}(\mathbf{U}^{(q_j+1)} - \mathbf{F}^{(q_j)})\right)\|_F^2 \\ + \left\|\nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{U}^{(q_j+1)}, \tilde{\mathbf{v}}^{(q_j+1)}) - \nabla_{\tilde{\boldsymbol{\varphi}}^*} f(\mathbf{F}^{(q_j)}, \tilde{\boldsymbol{\varphi}}^{(q_j)})\right. \\ \left. - \frac{1}{2\alpha}(\tilde{\mathbf{v}}^{(q_j+1)} - \tilde{\boldsymbol{\varphi}}^{(q_j)})\right\|_2^2 \Big)^{\frac{1}{2}} \\ \leq \left(L + \frac{1}{2\alpha}\right) \sqrt{\|\mathbf{U}^{(q_j+1)} - \mathbf{F}^{(q_j)}\|_F^2 + \|\tilde{\mathbf{v}}^{(q_j+1)} - \tilde{\boldsymbol{\varphi}}^{(q_j)}\|_2^2}. \end{aligned} \quad (73)$$

Using (67) we can then conclude that this norm tends to 0 and thus $0 \in \partial G(\mathbf{F}', \tilde{\boldsymbol{\varphi}}')$, i.e., $(\mathbf{F}', \tilde{\boldsymbol{\varphi}}')$ is a critical point of (4). ■

REFERENCES

- [1] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021, doi: [10.1109/JPROC.2021.3061701](https://doi.org/10.1109/JPROC.2021.3061701).
- [2] M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, "The road to 6G: Ten physical layer challenges for communications engineers," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 64–69, Jan. 2021, doi: [10.1109/MCOM.001.2000208](https://doi.org/10.1109/MCOM.001.2000208).
- [3] P. Wang, J. Fang, L. Dai, and H. Li, "Joint transceiver and large intelligent surface design for massive MIMO mmWave systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1052–1064, Feb. 2021, doi: [10.1109/TWC.2020.3030570](https://doi.org/10.1109/TWC.2020.3030570).
- [4] Y. Liu et al., "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd Quart., 2021, doi: [10.1109/COMST.2021.3077737](https://doi.org/10.1109/COMST.2021.3077737).
- [5] M. Jian et al., "Reconfigurable intelligent surfaces for wireless communications: Overview of hardware designs, channel models, and estimation techniques," 2022, *arXiv:2203.03176*.
- [6] E. Björnson, Ö. Özdogan, and E. G. Larsson, "Reconfigurable intelligent surfaces: Three myths and two critical questions," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 90–96, Dec. 2020, doi: [10.1109/MCOM.001.2000407](https://doi.org/10.1109/MCOM.001.2000407).

- [7] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020, doi: [10.1109/JSAC.2020.3007211](https://doi.org/10.1109/JSAC.2020.3007211).
- [8] B. Ning, Z. Tian, Z. Chen, C. Han, J. Yuan, and S. Li, "Prospective beamforming technologies for ultra-massive MIMO in terahertz communications: A tutorial," 2021, *arXiv:2107.03032*.
- [9] Z. Chen, X. Ma, C. Han, and Q. Wen, "Towards intelligent reflecting surface empowered 6G terahertz communications: A survey," *China Commun.*, vol. 18, no. 5, pp. 93–119, May 2021, doi: [10.23919/JCC.2021.05.007](https://doi.org/10.23919/JCC.2021.05.007).
- [10] C. Pan et al., "An overview of signal processing techniques for RIS/IRS-aided wireless systems," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 883–917, Aug. 2022, doi: [10.1109/JSTSP.2022.3195671](https://doi.org/10.1109/JSTSP.2022.3195671).
- [11] X. Ma, Z. Chen, L. Yan, C. Han, and Q. Wen, "Joint hardware design and capacity analysis for intelligent reflecting surface enabled terahertz MIMO communications," 2021, *arXiv:cs.IT/2012.06993*.
- [12] 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, 2022, doi: [10.3390/electronics11010018](https://doi.org/10.3390/electronics11010018).
- [13] S. Nie and I. F. Akyildiz, "Beamforming in intelligent environments based on ultra-massive MIMO platforms in millimeter wave and terahertz bands," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Barcelona, Spain, May 2020, pp. 8683–8687, doi: [10.1109/ICASSP40776.2020.9053786](https://doi.org/10.1109/ICASSP40776.2020.9053786).
- [14] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019, doi: [10.1109/TWC.2019.2936025](https://doi.org/10.1109/TWC.2019.2936025).
- [15] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020, doi: [10.1109/JSAC.2020.3000814](https://doi.org/10.1109/JSAC.2020.3000814).
- [16] X. Yu, D. Xu, and R. Schober, "MISO wireless communication systems via intelligent reflecting surfaces: (invited paper)," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2019, pp. 735–740, doi: [10.1109/ICCCChina.2019.8855810](https://doi.org/10.1109/ICCCChina.2019.8855810).
- [17] N. S. Perović, L.-N. Tran, M. Di Renzo, and M. F. Flanagan, "Achievable rate optimization for MIMO systems with reconfigurable intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3865–3882, Jun. 2021, doi: [10.1109/TWC.2021.3054121](https://doi.org/10.1109/TWC.2021.3054121).
- [18] C. Liaskos, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "Using any surface to realize a new paradigm for wireless communications," *Commun. ACM*, vol. 61, no. 11, pp. 30–33, Oct. 2018, doi: [10.1145/3192336](https://doi.org/10.1145/3192336).
- [19] T. N. Do, G. Kaddoum, T. L. Nguyen, D. B. da Costa, and Z. J. Haas, "Multi-RIS-aided wireless systems: Statistical characterization and performance analysis," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8641–8658, Dec. 2021, doi: [10.1109/TCOMM.2021.3117599](https://doi.org/10.1109/TCOMM.2021.3117599).
- [20] Y. Zhang et al., "Reconfigurable intelligent surfaces with outdated channel state information: Centralized vs. distributed deployments," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2742–2756, Apr. 2022.
- [21] Z. Xie, W. Yi, X. Wu, Y. Liu, and A. Nallanathan, "Downlink multi-RIS aided transmission in backhaul limited networks," *IEEE Wireless Commun. Lett.*, vol. 11, no. 7, pp. 1458–1462, Jul. 2022, doi: [10.1109/LWC.2022.3174448](https://doi.org/10.1109/LWC.2022.3174448).
- [22] L. Li, D. Ma, H. Ren, P. Wang, W. Lin, and Z. Han, "Toward energy-efficient multiple IRSS: Federated learning-based configuration optimization," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 2, pp. 755–765, Jun. 2022, doi: [10.1109/TGCN.2021.3136306](https://doi.org/10.1109/TGCN.2021.3136306).
- [23] P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14960–14973, Dec. 2020, doi: [10.1109/TVT.2020.3031657](https://doi.org/10.1109/TVT.2020.3031657).
- [24] J. Choi, G. Kwon, and H. Park, "Joint beamforming design for LOS MIMO systems with multiple intelligent reflecting surfaces," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, San Diego, CA, USA, Nov. 2021, pp. 267–272, doi: [10.1109/MILCOM52596.2021.9653118](https://doi.org/10.1109/MILCOM52596.2021.9653118).
- [25] W. Mei, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless networks: From single-reflection to multireflection design and optimization," *Proc. IEEE*, vol. 110, no. 9, pp. 1380–1400, 2022, doi: [10.1109/JPROC.2022.3170656](https://doi.org/10.1109/JPROC.2022.3170656).
- [26] C. You, B. Zheng, W. Mei, and R. Zhang, "How to deploy intelligent reflecting surfaces in wireless network: BS-side, user-side, or both sides?" *J. Commun. Inf. Netw.*, vol. 7, no. 1, pp. 1–10, Mar. 2022, doi: [10.23919/JCIN.2022.9745477](https://doi.org/10.23919/JCIN.2022.9745477).
- [27] Y. Han, S. Zhang, L. Duan, and R. Zhang, "Cooperative double-IRS aided communication: Beamforming design and power scaling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1206–1210, Aug. 2020, doi: [10.1109/LWC.2020.2986290](https://doi.org/10.1109/LWC.2020.2986290).
- [28] J. Shao and J. Zhu, "Joint beamforming and phase shifts design in double intelligent reflect surface aided secrecy MISO channel," *Progr. Elect. Res. C*, vol. 108, pp. 89–101, Jan. 2021.
- [29] C. Huang et al., "Multi-hop RIS-empowered terahertz communications: A DRL-based hybrid beamforming design," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1663–1677, Jun. 2021, doi: [10.1109/JSAC.2021.3071836](https://doi.org/10.1109/JSAC.2021.3071836).
- [30] F. Xu et al., "Coordinating multiple intelligent reflecting surfaces without channel information," 2023, *arXiv:2302.09717v1*.
- [31] W. Mei and R. Zhang, "Cooperative beam routing for multi-IRS aided communication," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 426–430, Feb. 2021.
- [32] H. Li and Z. Lin, "Accelerated proximal gradient methods for non-convex programming," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 379–387.
- [33] 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, doi: [10.1109/JPROC.2022.3170358](https://doi.org/10.1109/JPROC.2022.3170358).
- [34] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009, doi: [10.1109/TIP.2009.2028250](https://doi.org/10.1109/TIP.2009.2028250).
- [35] A. A. M. Saleh and R. A. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987, doi: [10.1109/JSAC.1987.1146527](https://doi.org/10.1109/JSAC.1987.1146527).
- [36] G. Zhou, C. Pan, H. Ren, K. Wang, M. Elkashlan, and M. D. Renzo, "Stochastic learning-based robust beamforming design for ris-aided millimeter-wave systems in the presence of random blockages," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 1057–1061, Jan. 2021, doi: [10.1109/TVT.2021.3049257](https://doi.org/10.1109/TVT.2021.3049257).
- [37] X. Wang, Z. Lin, F. Lin, and L. Hanzo, "Joint hybrid 3D beamforming relying on sensor-based training for reconfigurable intelligent surface aided terahertz-based multi-user massive MIMO systems," *IEEE Sensors J.*, vol. 22, no. 14, pp. 14540–14552, Jul. 2022, doi: [10.1109/JSEN.2022.3182881](https://doi.org/10.1109/JSEN.2022.3182881).
- [38] F. Bohagen, P. Orten, and G. E. 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, doi: [10.1109/TCOMM.2009.03.070062](https://doi.org/10.1109/TCOMM.2009.03.070062).
- [39] 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)*, Montreal, QC, Canada, Jun. 2021, pp. 1–6, doi: [10.1109/ICCWorkshops50388.2021.9473890](https://doi.org/10.1109/ICCWorkshops50388.2021.9473890).
- [40] 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, doi: [10.1109/LCOMM.2021.3088264](https://doi.org/10.1109/LCOMM.2021.3088264).
- [41] 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," 2022, *arXiv:2201.04328*.
- [42] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.
- [43] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020, doi: [10.1109/MCOM.001.1900107](https://doi.org/10.1109/MCOM.001.1900107).
- [44] S. W. Ellingson, "Path loss in reconfigurable intelligent surface-enabled channels," in *Proc. IEEE Ann. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Oct. 2021, pp. 829–835.
- [45] A. Hjörungnes, *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [46] A. Beck, *First-Order Methods in Optimization* (MOS-SIAM Series on Optimization). Philadelphia, PA, USA: SIAM, 2017.



NUNO SOUTO (Senior Member, IEEE) graduated in aerospace engineering (avionics) from the Instituto Superior Técnico, Lisbon, Portugal, in 2000. He received the Ph.D. degree in 2006, and the Habilitation degree in 2021. From November 2000 to January 2002 he worked as a Researcher of Automatic Speech Recognition with the Instituto de Engenharia e Sistemas de Computadores, Lisbon. He joined the ISCTE-University Institute of Lisbon as an Assistant Professor in 2006. He has been a Researcher

with the Instituto de Telecomunicações, 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, OFDM, single carrier transmission with frequency domain equalization, channel coding, modulation, channel estimation, synchronization, MIMO schemes, wireless sensor networks, unmanned aerial vehicles, mmWave, and terahertz communications. He is a member of the IEEE Communications Society.



JOÃO C. 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.