Weighted Sum-Rate Maximization for Multi-Hop RIS-Aided Multi-User Communications: A Minorization-Maximization Approach

Zepeng Zhang and Ziping Zhao
School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China.
Email: {zhangzp1, zhaoziping}@shanghaitech.edu.cn

***Abstract*—The reconfigurable intelligent surface (RIS) has aroused much research attention recently due to its potential benefits in 5G and beyond wireless networks. This paper considers a general multi-hop RIS-aided multi-user communication system and the weighted sum-rate maximization problem is studied by jointly designing the active beamforming matrix at the base station and multiple phase-shift matrices at the RISs (considering both continuous and discrete phase constraints). To tackle the resulting highly nonconvex optimization problem, a problem-tailored low-complexity and globally convergent algorithm based on block minorization-maximization (BMM) is proposed. The effectiveness of the proposed BMM approach and the performance improvement gained with multi-hop RISs are both demonstrated through numerical simulations. The merits of the proposed algorithms are further illustrated by indicating their adaptivity in solving many other RIS-related system designs.**

***Index Terms*—Interference channel, power allocation, beamforming, inexact update, minorization-maximization, coordinate descent.**

I. Introduction

To meet the ever-increasing demand for high-speed and seamless data services in the next-generation wireless communication systems, which are expected to provide a 1000-fold increase in the network capacity, some techniques relying on massive multiple-input multiple-output, ultra-dense networks, and millimeter wave communications are expected to be the key enabling technologies [1]. However, such solutions will generally involve increased hardware costs and high power consumption. With the theoretical and experimental breakthroughs in micro electromechanical systems and metamaterials (e.g., metasurface), reconfigurable intelligent surface (RIS), as an emerging cost-effective technology, has recently been advocated as a powerful solution to enhance the spectrum efficiency and energy efficiency of wireless networks [2]–[4].

The RIS-aided wireless networks have aroused much research interest due to their potential benefits in 5G and beyond networks. Specifically, the joint beamforming and reflecting design with the assistance of RIS has received considerable attention in areas like multi-user multiple-input single-output (MISO) downlink communication systems [5], multigroup multicast MISO communication systems [6], and multicell multiple-input multiple-output (MIMO) communication systems [7], etc. While most of the existing works have focused on single-RIS assisted (i.e., single-hop) systems, the benefits of designing double-RIS assisted (i.e., two-hop) multi-user MIMO systems cooperatively have been explored recently in [8]. More generally, multi-hop RIS-aided systems can also be deployed as in the classical multi-hop relaying systems [9], [10] to combat the propagation distance problem and to improve the coverage range.

Consider an interference channel for multiuser downlink transmission, where the independent data streams are sent to some intended receivers simultaneously, one classical target of the system design is to achieve high total throughput (i.e., sum-rates) [11], [12]. In [5], the weighted sum-rate (WSR) maximization problem for a single-hop RIS-aided multi-user MISO systems was considered where the phase-shift matrix at the RIS and the active beamforming matrix at the base station (BS) are jointly designed. The WSR maximization problem is nonconvex due to the non-convexity of the objective and the constraints, for which two different approaches are proposed in [5], namely, an alternating minimization (AM) algorithm leveraging WMMSE [13] and Riemannian conjugate gradient [14], and the block coordinate descent (BCD) method aided by the fractional programming (FP) approach [15]. However, convergence of these approaches may not be favorable in practice. For the WMMSE-based AM method, it inherits a double-loop nature which may invoke many iterations to converge resulting in high per-iteration computational complexity. For the FP-based BCD method, it relies on the manifold structure of the continuous phase constraint making it become lame in the more practical discrete phase setting [16]. As for the multi-hop RIS-aided systems, to the best of our knowledge, there only exists a deep reinforcement learning-based approach [17] and its performance highly relies on the carefully chosen initializations which call for other iterative algorithms.

In this paper, to address the WSR maximization problem for the multi-hop RIS-aided multi-user systems, a problem-tailored low-complexity and globally convergent algorithm based on block minorization-maximization (BMM) [18]–[20] is proposed. The proposed BMM algorithm works for both the continuous-phase and the discrete-phase schemes for RISs and can be easily generalized to other RIS-related system design problems. Theoretically, we show that the sequence of iterates generated by the proposed algorithm converges to a KKT point of the WSR maximization problem. The superiority of the proposed BMM algorithm over existing algorithms under a single-hop setting and the performance improvement gained with multi-hop RISs are both demonstrated through numerical simulations, while the influence of the quantization levels of the discrete phases is also showcased.

This work was supported in part by the National Nature Science Foundation of China (NSFC) under Grant 62001295 and in part by the Shanghai Sailing Program under Grant 20YF1430800.

II. System Model and Problem Formulation

The multi-hop RIS-aided multi-user MISO downlink communication system is considered, where the BS equipped with M antennas communicates with K users with single antenna in a circular region. We assume the transmitted signal experiences I_k ($I_k \leq L$) hops on the reflecting RISs to arrive the k-th user, with the i-th ($i = 1, \ldots, L$) RIS containing N_i reflecting elements. Denote $\mathbf{G}_1 \in \mathbb{C}^{N_1 \times M}$, $\mathbf{G}_{i+1} \in \mathbb{C}^{N_{i+1} \times N_i}$, and $\boldsymbol{\Theta}_i = \mathsf{diag}(\boldsymbol{\theta}_i)$ with $\boldsymbol{\theta}_i = \left[[\boldsymbol{\theta}_i]_1, \ldots, [\boldsymbol{\theta}_i]_{N_i}\right]^T \in \mathbb{C}^{N_i}$ [1] as the channel matrix from the BS to the first RIS, the channel matrix from the i-th RIS to the $(i+1)$-th RIS, and the phase shift matrix of the i-th RIS, respectively. Denote $\mathbf{h}_{r,k} \in \mathbb{C}^{N_{I_k}}$ and $\mathbf{h}_{d,k} \in \mathbb{C}^M$ as the channel from the last RIS to the k-th user and the direct channel from the BS to the k-th user, respectively. Then the received signal at the k-th user is given by

$$y_k = \big(\mathbf{h}_{r,k}^H \prod_{i=1,\cdots,I_k} \boldsymbol{\Theta}_i \mathbf{G}_i + \mathbf{h}_{d,k}^H\big)\mathbf{w}_k s_k + \sum_{j,j\neq k}^{K} \big(\mathbf{h}_{r,k}^H \prod_{i=1,\cdots,I_k} \boldsymbol{\Theta}_i \mathbf{G}_i + \mathbf{h}_{d,k}^H\big)\mathbf{w}_j s_j + e_k,$$

where $\mathbf{w}_k \in \mathbb{C}^M$ and $s_k \in \mathcal{CN}(0,1)$ are the beamforming vector and the independent user symbols under the assumption of Gaussian signals, and $e_k \in \mathcal{CN}(0, \sigma^2)$ represents the noise. For simplicity, only the direct transmission paths from the BS to users and the transmission paths through I_k RISs to users have been considered. Then the SINR at the k-th user can be expressed as

$$\mathsf{SINR}_k = \frac{\left|(\mathbf{h}_{r,k}^H \prod_{i=1,\cdots,I_k} \boldsymbol{\Theta}_i \mathbf{G}_i + \mathbf{h}_{d,k}^H)\mathbf{w}_k\right|^2}{\sum_{j,j\neq k}^{K}\left|(\mathbf{h}_{r,k}^H \prod_{i=1,\cdots,I_k} \boldsymbol{\Theta}_i \mathbf{G}_i + \mathbf{h}_{d,k}^H)\mathbf{w}_j\right|^2 + \sigma^2}.$$

Assume that all the channels $\{\mathbf{h}_{r,i}\}$, $\{\mathbf{h}_{d,i}\}$, and $\{\mathbf{G}_i\}$ are perfectly known by both the BS and the users, this paper aims at maximizing the WSR of the system by jointly designing the beamforming vectors $\mathbf{w}_i$'s and the phase shift matrices $\boldsymbol{\Theta}_i$'s. Denote the data rate at the k-th user as $\mathsf{R}_k(\{\mathbf{w}_i\}, \{\boldsymbol{\Theta}_i\}) = \log(1 + \mathsf{SINR}_k)$,[2] the WSR is computed as follows:

$$\mathsf{WSR}(\{\mathbf{w}_i\}, \{\boldsymbol{\Theta}_i\}) = \sum_{k=1}^{K} \omega_k \mathsf{R}_k(\{\mathbf{w}_i\}, \{\boldsymbol{\Theta}_i\}). \quad \text{(WSR)}$$

where $\omega_k \geq 0$ is the predefined weight for the data rate of k-th user. Finally, the WSR maximization problem is given as follows:

$$\begin{array}{ll} \underset{\{\mathbf{w}_i\},\{\boldsymbol{\Theta}_i\}}{\mathsf{maximize}} & \mathsf{WSR}(\{\mathbf{w}_i\}, \{\boldsymbol{\Theta}_i\}) \\ \mathsf{subject\ to} & \sum_{k=1}^{K} \|\mathbf{w}_k\|_2^2 \leq P_{\mathsf{max}} \\ & \boldsymbol{\Theta}_l \in \mathcal{R}_l,\ \forall l = 1, \ldots, L, \end{array} \quad \text{(WSRMax)}$$

where P_{max} is the transmit power limit of the BS, and $\mathcal{R}_i$ represents the constant modulus constraint indicating that there is no energy loss of the signal by going through the RISs. In this paper, $\mathcal{R}_i$ can take either a continuous-phase scheme as

$$\mathcal{R}_i^{\mathsf{C}} = \left\{\boldsymbol{\Theta}_i = \mathsf{diag}(\boldsymbol{\theta}_i) \,|\, \boldsymbol{\theta}_i \in \mathbb{C}^{N_i}, \left|[\boldsymbol{\theta}_i]_j\right| = 1, \forall j = 1, \ldots, N_i\right\},$$

indicating phases of the RIS elements can take any angles or a discrete-phase scheme (more practical for hardware implementation) which is given as follows:

$$\mathcal{R}_i^{\mathsf{D}} = \big\{\boldsymbol{\Theta}_i = \mathsf{diag}(\boldsymbol{\theta}_i) \mid \boldsymbol{\theta}_i \in \mathbb{C}^{N_i}, \left|[\boldsymbol{\theta}_i]_j\right| = 1, \mathsf{arg}([\boldsymbol{\theta}_i]_j) \in \Phi_i, \forall j = 1, \ldots, N_i\big\},$$

where Φ_i denotes the set of fixed angles that achievable for the i-th RIS. In general, different RISs can take different phase schemes. Problem (WSRMax) is non-convex and hence NP-hard. In the following, a low-complexity and globally convergent algorithm will be developed for problem resolution.

III. Solving the WSR Maximization Problem via Block Minorization-Maximization

The BMM [18] method can be regarded as a combination of the BCD methods [21] and the minorization-maximization (for maximization problem) methods [19]. BCD, a.k.a. Gauss-Seidel, is an optimization method aiming at finding a local optimum of the problem by optimizing along one variable block at a time and then for different blocks repeatedly. The direct variable update in BCD can be difficult to attain, especially when the problem is non-convex and/or when a unique global optimal solution is required for convergence claim. Instead of solving for an exact update, BMM solves a series of simple surrogate problems w.r.t. one variable block each time via carrying out an inexact variable update (refer to [18], [19] for details). Surrogate functions in BMM can be chosen in a flexible way [19] while a properly chosen one can make the updates easy and hence can lead to a fast convergence over iterations. In practice, the surrogate subproblems are applaudable if they are convex and efficiently solvable or they have closed-form solutions. In the following, we will derive a BMM algorithm for Problem (WSRMax).

A. The Update Step of $\{\mathbf{w}_i\}$

Since variables $\{\mathbf{w}_i\}$ are coupled in constraints, we take it as one block. At iterate $\{\{\underline{\mathbf{w}_i}\}, \{\underline{\boldsymbol{\Theta}_i}\}\}$[3], WSR w.r.t. $\{\mathbf{w}_i\}$ is

$$\mathsf{WSR}_{\{\mathbf{w}_i\}}(\{\mathbf{w}_i\} | \{\underline{\mathbf{w}_i}\}, \{\underline{\boldsymbol{\Theta}_i}\}) = \sum_{k=1}^{K} \omega_k \log\big(1 + \frac{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{\sum_{j,j\neq k}^{K}\left|\mathbf{h}_k^H \underline{\mathbf{w}_j}\right|^2 + \sigma^2}\big),$$

where $\mathbf{h}_k = \prod_{i=I_k,\cdots,1} \mathbf{G}_i^H \underline{\boldsymbol{\Theta}_i}^H \mathbf{h}_{r,k} + \mathbf{h}_{d,k}$. Then we introduce the following two useful results.

Lemma 1. *The* $\log(1+\frac{x^2}{y})$ *with* $x \geq 0$ *and* $y > 0$ *is minorized at* $(\underline{x}, \underline{y})$ *as follows:*

$$\log(1+\frac{x^2}{y}) \geq \log(1+\frac{\underline{x}^2}{\underline{y}}) - \frac{\underline{x}^2}{\underline{y}} + \frac{2\underline{x}}{\underline{y}}x - \frac{\underline{x}^2}{\underline{y}(\underline{y}+\underline{x}^2)}(y+x^2),$$

where the equality is attained at $(x, y) = (\underline{x}, \underline{y})$.

Proof: It can be proved by the convexity of $\log(\frac{1}{x})$ for $x > 0$ and the convexity of $\frac{x^2}{y}$ for $x \geq 0$ and $y > 0$. ■

Lemma 2. *The* $|x|$ *with* $x \in \mathbb{C}$ *is minorized at* $\underline{x}$ *as follows:*

$$|x| \geq |\underline{x}|^{-1}\mathsf{Re}(\underline{x}^* x),$$

where $|\underline{x}| \neq 0$ *and the equality is attained at* $x = \underline{x}$.

[1] $[\mathbf{x}]_i$ denotes the i-th element in $\mathbf{x}$.

[2] Note the natural logarithm is used since optimal solutions of the WSR maximization problem given later is irrelevant to bases of the log-functions.

[3] In this paper, underlined variables denote those whose values are given.

Based on Lemma 1 and Lemma 2, a minorizing function for $\mathsf{WSR}_{\{\mathbf{w}_i\}}$ w.r.t. $\{\mathbf{w}_i\}$ can be constructed as follows:

$$\underline{\mathsf{WSR}_{\{\mathbf{w}_i\}}}(\{\mathbf{w}_i\}|\{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}) = \sum_{k=1}^{K}\omega_k\big(\mathsf{Re}(\eta_{k,1}\mathbf{h}_k^H\mathbf{w}_k) - \eta_{k,2}\sum_{j=1}^{K}\big|\mathbf{h}_k^H\mathbf{w}_j\big|^2 + const_{w,k}\big).$$

where $\eta_{k,1} = \frac{2\alpha_k^*}{\beta_k}$, $\eta_{k,2} = \frac{|\alpha_k|^2}{\beta_k(\beta_k+|\alpha_k|^2)}$ with $\alpha_k = \mathbf{h}_k^H\underline{\mathbf{w}_k}$ and $\beta_k = \sum_{j,j\neq k}^{K}\big|\mathbf{h}_k^H\underline{\mathbf{w}_j}\big|^2 + \sigma^2$, and the constant term $const_{w,k} = \log(1+\frac{|\alpha_k|^2}{\beta_k}) - \frac{|\alpha_k|^2}{\beta_k}$. With the quadratic minorizing function $\underline{\mathsf{WSR}_{\{\mathbf{w}_i\}}}$ and ignoring the constant terms, the resulting subproblem for $\{\mathbf{w}_i\}$ which is convex is given by

$$\begin{array}{ll}\underset{\{\mathbf{w}_i\}}{\mathsf{minimize}} & \sum_{k=1}^{K}\omega_k\eta_{k,2}\big(\big\|\mathbf{h}_k^H\mathbf{w}_k - \frac{\eta_{k,1}^*}{2\eta_{k,2}}\big\|_2^2 + \sum_{j,j\neq k}^{K}\big\|\mathbf{h}_k^H\mathbf{w}_j\big\|_2^2\big)\\ \mathsf{subject\ to} & \sum_{k=1}^{K}\|\mathbf{w}_k\|_2^2 \le P_{\mathsf{max}}.\end{array} \quad (1)$$

Proposition 3. *By solving the KKT system, the optimal solution to Problem* (1) *can be obtained in the following way*

$$\mathbf{w}_k^\star = \begin{cases}\mathbf{R}^{-1}\mathbf{p}_k & \sum_{k=1}^{K}\|\mathbf{R}^{-1}\mathbf{p}_k\|_2^2 \le P_{\mathsf{max}}\\ (\mathbf{R}+2\lambda^\star\mathbf{I})^{-1}\mathbf{p}_k & \textit{otherwise}\end{cases}$$
$$\textit{for } k = 1,\ldots,K,$$

where $\mathbf{R} = \sum_{k=1}^{K}2\omega_k\eta_{k,2}\mathbf{h}_k\mathbf{h}_k^H$, $\mathbf{p}_k = \omega_k\eta_{1,k}^*\mathbf{h}_k$, *and the optimal dual variable* $\lambda^\star > 0$ *satisfies* $\sum_{k=1}^{K}\|(\mathbf{R}+2\lambda^\star\mathbf{I})^{-1}\mathbf{p}_k\|_2^2 = P_{\mathsf{max}}$ *and can be readily found via one-dimensional line search methods [21] to meet*[4]

$$\sum_{n=1}^{N}\frac{\sum_{k=1}^{K}\big([\mathbf{V}^H\mathbf{p}_k]_n\big)^2}{([\boldsymbol{\Lambda}]_{nn}+2\lambda^\star)^2} = P_{\mathsf{max}}$$

with $\mathbf{V}$ *and* $\boldsymbol{\Lambda}$ *from the eigendecompositon* $\mathbf{R} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^H$.

B. The Update Step of $\{\boldsymbol{\Theta}_l\}$

In this section, we choose to update the phase shift matrices $\boldsymbol{\Theta}_l,\ l = 1,\ldots,L$ successively. Since elements in each phase shift matrix $\boldsymbol{\Theta}_l$ are separable over constraints, they can be updated either in series or in parallel. Given the iterate $\{\{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}\}$, the data rate at the k-th user w.r.t. $\boldsymbol{\Theta}_l$ is

$$\mathsf{R}_{l,k}(\boldsymbol{\Theta}_l|\{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}) = \log\big(1 + \frac{\big|\boldsymbol{\theta}_l^T\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_k} + \mathbf{h}_{d,k}^H\underline{\mathbf{w}_k}\big|^2}{\sum_{j,j\neq k}^{K}\big|\boldsymbol{\theta}_l^T\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_j} + \mathbf{h}_{d,k}^H\underline{\mathbf{w}_j}\big|^2 + \sigma^2}\big),$$

where $\tilde{\mathbf{G}}_{r,l,k} = \mathsf{diag}\big(\mathbf{G}_{l+1}^H\underline{\boldsymbol{\Theta}}_{l+1}^H\cdots\mathbf{G}_{I_k}^H\underline{\boldsymbol{\Theta}}_{I_k}^H\mathbf{h}_{r,k}\big)\mathbf{G}_l\cdots\underline{\boldsymbol{\Theta}_1}\mathbf{G}_1$.[5] Then a minorizing function for $\mathsf{WSR}_{\boldsymbol{\Theta}_l}$ w.r.t. $\boldsymbol{\Theta}_l$ can be constructed based on Lemma 1 and Lemma 2 as follows:

$$\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}(\boldsymbol{\Theta}_l \mid \{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}) = \sum_{k=1}^{K}\omega_k\Big(\mathsf{Re}\left(\eta_{k,1}\boldsymbol{\theta}_l^T\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_k}\right) - \eta_{k,2}\sum_{j=1}^{K}\big|\boldsymbol{\theta}_l^T\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_j} + \mathbf{h}_{d,k}^H\underline{\mathbf{w}_j}\big|^2 + const_{l,k}\Big),$$

[4]A further majorization can be applied to the objective in (1) for a closed-form solution without line search, which is not detailed due to space limitation.

[5]Note that w.l.o.g. we have assumed $I_k \ge l$ in this section.

where $const_{l,k} = \log(1+\frac{|\alpha_k|^2}{\beta_k}) - \frac{|\alpha_k|^2}{\beta_k} + \eta_{k,1}\frac{\alpha_k^*}{|\alpha_k|}\mathbf{h}_{d,k}^H\underline{\mathbf{w}_k} - \eta_{k,2}K\sigma^2$. The $\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}$ can be written compactly as

$$\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}(\boldsymbol{\Theta}_l \mid \{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}) = -\boldsymbol{\theta}_l^T\mathbf{L}_l\boldsymbol{\theta}_l + \mathsf{Re}(\boldsymbol{\theta}_l^T\mathbf{d}_l) + const_l,$$

where $\mathbf{L}_l = \sum_{k=1}^{K}\omega_k\eta_{k,2}\sum_{j=1}^{K}\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_j}\underline{\mathbf{w}_j}^H\tilde{\mathbf{G}}_{r,l,k}^H$, $\mathbf{d}_l = \sum_{k=1}^{K}\omega_k\eta_{k,1}\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_k} - \sum_{k=1}^{K}\omega_k\eta_{k,2}\sum_{j=1}^{K}2\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_j}(\mathbf{h}_{d,k}^H\underline{\mathbf{w}_j})^*$, and $const_l = \sum_{k=1}^{K}\omega_k const_{l,k} - \sum_{k=1}^{K}\omega_k\eta_{k,2}\sum_{j=1}^{K}\big|\mathbf{h}_{d,k}^H\underline{\mathbf{w}_j}\big|^2$.

Remark 4. Although the quadratic monorizing function $\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}$ is derived in a way for the whole phase shift matrix $\boldsymbol{\Theta}_l$, it is also a minorizing function for each element of it, i.e., $[\boldsymbol{\theta}_l]_i,\ i = 1,\ldots,N_l$ with values of the other elements fixed.

Based on the results in Remark 4, two ways for updating the phase shift matrices $\{\boldsymbol{\Theta}_l\}$ are presented in the following.

1) Serial Update of $\boldsymbol{\Theta}_l$*:* Consider variable $[\boldsymbol{\theta}_l]_j$ and define $\underline{[\boldsymbol{\theta}_l]_{-j}} = \big[\underline{[\boldsymbol{\theta}_l]_1},\ldots\underline{[\boldsymbol{\theta}_l]_{j-1}},0,\underline{[\boldsymbol{\theta}_l]_{j+1}},\ldots,\underline{[\boldsymbol{\theta}_l]_{N_l}}\big]^T$, we can obtain a minorizing function for $\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}$ w.r.t. $[\boldsymbol{\theta}_l]_j$ as follows:

$$\underline{\mathsf{WSR}_{[\boldsymbol{\theta}_l]_j}}([\boldsymbol{\theta}_l]_j \mid \{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}) = \mathsf{Re}\big((2\underline{[\boldsymbol{\theta}_l]_{-j}^T}[\mathbf{L}_l]_{:,j} + [\mathbf{d}_l]_j)[\boldsymbol{\theta}_l]_j\big) + [\mathbf{L}_l]_{jj}|[\boldsymbol{\theta}_l]_j|^2 - \underline{[\boldsymbol{\theta}_l]_{-j}^T}\mathbf{L}_l\underline{[\boldsymbol{\theta}_l]_{-j}} + \mathsf{Re}(\underline{[\boldsymbol{\theta}_l]_{-j}^T}\mathbf{d}_l),$$

where $[\mathbf{L}_l]_{:,j}$ is the j-th column of $\mathbf{L}_l$. By ignoring the constant terms in $\underline{\mathsf{WSR}_{[\boldsymbol{\theta}_l]_j}}$, the subproblem w.r.t. $[\boldsymbol{\theta}_l]_j$ is given by

$$\mathsf{maximize}_{[\boldsymbol{\theta}_l]_j}\ \ \mathsf{Re}\big([\boldsymbol{\theta}_l]_j[\mathbf{b}_l^{\mathsf{S}}]_j\big)\ \ \mathsf{subject\ to}\ \ [\boldsymbol{\theta}_l]_j \in \mathcal{R}_l \quad (2)$$

with $\mathbf{b}_l^{\mathsf{S}} = \big[\underline{[\boldsymbol{\theta}_l]_{-1}^T}[\mathbf{L}_l]_{:,1} + [\mathbf{d}_l]_1,\ldots,\underline{[\boldsymbol{\theta}_l]_{-N_l}^T}[\mathbf{L}_l]_{:,N_l} + [\mathbf{d}_l]_{N_l}\big]^T$.

Proposition 5. *Optimal solutions to Problem* (2) *can be obtained in closed-forms as follows:*

- *if* $\mathcal{R}_l^{\mathsf{C}}$ *is chosen, the solution is* $[\boldsymbol{\theta}_l]_j^\star = e^{\jmath\arg\left([\mathbf{b}_l^{\mathsf{S}}]_j^*\right)}$;
- *if* $\mathcal{R}_l^{\mathsf{D}}$ *is chosen, the solution is given by* $[\boldsymbol{\theta}_l]_j^\star = e^{\jmath\arg\min_{\phi_l\in\Phi_l}\left\|\phi_l - \arg\left([\mathbf{b}_l^{\mathsf{S}}]_j^*\right)\right\|_2}$, *which can be efficiently implemented on hardwares leveraging a look-up table.*

Finally, all the phase shift matrices $\{\boldsymbol{\Theta}_l\}$ will be updated successively with the elements in each $\boldsymbol{\theta}_l$ updated in series.

2) Parallel Update of $\boldsymbol{\Theta}_l$*:* Besides updating the elements in $\boldsymbol{\theta}_l$ serially, we can also update $\boldsymbol{\theta}_l$ as a whole (i.e., in parallel) by constructing a linear minorizing function of $\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}$.

Lemma 6. *[22] Let* $\mathbf{L},\mathbf{M} \in \mathbb{H}^n$ *such that* $\mathbf{M} \succeq \mathbf{L}$*. The function* $\mathbf{x}^H\mathbf{M}\mathbf{x}$ *with* $\mathbf{x} \in \mathbb{C}^n$ *is minorized at* $\underline{\mathbf{x}}$ *as follows:*

$$\mathbf{x}^H\mathbf{M}\mathbf{x} \ge \mathbf{x}^H\mathbf{L}\mathbf{x} + 2\mathsf{Re}(\mathbf{x}^H(\mathbf{M}-\mathbf{L})\underline{\mathbf{x}}) + \underline{\mathbf{x}}^H(\mathbf{L}-\mathbf{M})\underline{\mathbf{x}}$$

where the equality is attained at $\mathbf{x} = \underline{\mathbf{x}}$.

Lemma 7. *Given* $\mathbf{M} = \mathsf{diag}(\mathbf{a}^H)\mathsf{diag}(\mathbf{a})$ *and* $\mathbf{L} = \mathbf{a}\mathbf{a}^H$ *with* $\mathbf{a} \in \mathbb{C}^N$*, it follows that* $\mathbf{M} \succeq \mathbf{L}$.

Proof: For any $\mathbf{x} \in \mathbb{C}^N$, we can obtain $\mathbf{x}^H(\mathbf{M}-\mathbf{L})\mathbf{x} = \mathbf{x}^H\mathsf{diag}(\mathbf{a}^H)\mathsf{diag}(\mathbf{a})\mathbf{x} - \mathbf{x}^H\mathbf{a}\mathbf{a}^H\mathbf{x} = \sum_{i=1}^{N}\big|[\mathbf{a}]_i^*x_i\big|^2 - \big|\mathbf{a}^H\mathbf{x}\big|^2 \ge 0$, which completes the proof. ∎

Based on Lemma 6 and Lemma 7, a linear minorizing function for $\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}$ can be obtained as in Eq. (3) with $\mathbf{M}_l = \sum_{k=1}^{K}\omega_k\eta_{k,2}\sum_{j=1}^{K}\mathsf{diag}\big(\underline{\mathbf{w}_j}^H\tilde{\mathbf{G}}_{r,l,k}^H\big)\mathsf{diag}\big(\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_j}\big)$. Eq. (3) can be written compactly as follows:

$$\underline{\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}}(\boldsymbol{\Theta}_l \mid \{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}) = \mathsf{Re}(\boldsymbol{\theta}_l^T\mathbf{b}_l^{\mathsf{P}}) + const_l',$$

$$\underline{\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}}(\boldsymbol{\Theta}_l \mid \{\underline{\mathbf{w}_i}\},\{\underline{\boldsymbol{\Theta}_i}\}) = \mathsf{Re}\Big(\sum_{k=1}^{K}\omega_k\big(\eta_{k,1}\boldsymbol{\theta}_l^T\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_k} - 2\eta_{k,2}\sum_{j=1}^{K}\boldsymbol{\theta}_l^T\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_j}(\mathbf{h}_{d,k}^H\underline{\mathbf{w}_j})^*\big) - 2\boldsymbol{\theta}_l^T(\mathbf{L}_l-\mathbf{M}_l)\underline{\boldsymbol{\theta}_l}\Big) + \sum_{k=1}^{K}\omega_k const_{l,k} + \mathsf{Re}(\eta_{k,1}\mathbf{h}_{d,k}^H\underline{\mathbf{w}_k}) - \eta_{k,2}\sum_{j=1}^{K}\left|\mathbf{h}_{d,k}^H\underline{\mathbf{w}_j}\right|^2 - \mathsf{tr}(\mathbf{M}_l) - \underline{\boldsymbol{\theta}_l^T}(\mathbf{M}_l-\mathbf{L}_l)\underline{\boldsymbol{\theta}_l} \qquad (3)$$

where $\mathbf{b}_l^{\mathsf{P}} = \sum_{k=1}^{K}\omega_k\eta_{k,1}\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_k} - 2(\mathbf{L}_l-\mathbf{M}_l)\underline{\boldsymbol{\theta}_l} - 2\eta_{k,2}\sum_{j=1}^{K}\tilde{\mathbf{G}}_{r,l,k}\underline{\mathbf{w}_j}(\mathbf{h}_{d,k}^H\underline{\mathbf{w}_j})^*$ and $const_l' = \sum_{k=1}^{K}\omega_k const_{l,k} + \mathsf{Re}(\eta_{k,1}\mathbf{h}_{d,k}^H\underline{\mathbf{w}_k}) - \eta_{k,2}\sum_{j=1}^{K}\left|\mathbf{h}_{d,k}^H\underline{\mathbf{w}_j}\right|^2 - \mathsf{tr}(\mathbf{M}_l) - \underline{\boldsymbol{\theta}_l^T}(\mathbf{M}_l-\mathbf{L}_l)\underline{\boldsymbol{\theta}_l}$.

Lemma 8. *Given* $\mathbf{M}_1 = \mathsf{diag}(\mathbf{a}^H)\mathsf{diag}(\mathbf{a})$ *and* $\mathbf{M}_2 = \mathbf{a}^H\mathbf{a}\mathbf{I}$ *with* $\mathbf{a} \in \mathbb{C}^N$*, it follows that* $\mathbf{M}_2 \succeq \mathbf{M}_1$.

Proof: For $\forall \mathbf{x} \in \mathbb{C}^N$, we can obtain $\mathbf{x}^H(\mathbf{M}_2-\mathbf{M}_1)\mathbf{x} = \mathbf{x}^H\mathbf{a}^H\mathbf{a}\mathbf{x} - \mathbf{x}^H\mathsf{diag}(\mathbf{a}^H)\mathsf{diag}(\mathbf{a})\mathbf{x} = \sum_{i=1}^{N}\left|[\mathbf{a}]_i\right|^2\sum_{i=1}^{N}\left|x_i\right|^2 - \sum_{i=1}^{N}\left|[\mathbf{a}]_i\right|^2\left|x_i\right|^2 \geq 0$, which completes the proof. ■

Lemma 8 shows that the chosen minorizing function of this paper is tighter than the widely used one in the literature, where the largest eigenvalue scaled identity is used to majorize the sum-of-rank-1 matrix [5], [22], [23]. This can partially explain the faster convergence speed of the proposed algortihm.

By discarding the constant in $\underline{\underline{\mathsf{WSR}_{\boldsymbol{\Theta}_l}}}$, the subproblem for $\boldsymbol{\Theta}_l$ is

$$\mathsf{maximize}_{\boldsymbol{\theta}_l} \quad \mathsf{Re}(\boldsymbol{\theta}_l^T\mathbf{b}_l^{\mathsf{P}}) \quad \mathsf{subject\ to} \quad \boldsymbol{\Theta}_l \in \mathcal{R}_l, \qquad (4)$$

which is separable over different elements in $\boldsymbol{\theta}_l$ and can be solved in parallel with optimal solutions given in Proposition 5. Finally, all the phase shift matrices $\{\boldsymbol{\Theta}_l\}$ will be updated successively with elements of each $\boldsymbol{\theta}_l$ updated in parallel.

Remark 9. Based on Lemma 7, $\mathbf{b}_l^{\mathsf{P}}$ shares intrinsically the same expression as $\mathbf{b}_l^{\mathsf{S}}$, only differing in the choice of values of $\underline{\boldsymbol{\theta}_l}$'s.

In summary, via BMM the variable blocks $\{\mathbf{w}_i\}$ and $\{\boldsymbol{\Theta}_i\}$ are updated cyclically in closed-forms until some convergence criterion is met. The SQUAREM [24] can be used as an off-the-shelf accelerator to solve the nonlinear fixed-point update in BMM. The overall algorithm is summarized in Algorithm 1 with convergence properties established in Theorem 10.

Theorem 10. *Every limit point, denoted by* $\{\{\mathbf{w}_i^{(\infty)}\},\{\boldsymbol{\Theta}_i^{(\infty)}\}\}$*, of the sequence* $\{\{\mathbf{w}_i^{(t)}\},\{\boldsymbol{\Theta}_i^{(t)}\}\}_{t\in\mathbb{N}}$ *generated by Algorithm 1 is a KKT point (stationary point) of Problem* (WSRMax).

Proof: Due to space limitation, an outline of the proof is given. First note the sub-problem of $\mathbf{W}$ is convex. For the sub-problem of $\boldsymbol{\Theta}_l$ in (2) or (4) with nonconvex constraint $\mathcal{R}_l$, it can be verified that the regularity condition liner independence constraint qualification (LICQ) [25] holds everywhere on set $\mathcal{R}_l$. The continuous phase constraint $\mathcal{R}_l^{\mathsf{C}}$ and the discrete phase constraint $\mathcal{R}_l^{\mathsf{D}}$ w.r.t. $\boldsymbol{\theta}_l$ can be rewritten as equality constraints $[\mathbf{h}_l^{\mathsf{C}}(\boldsymbol{\theta}_l)]_i = ([\boldsymbol{\theta}_l]_i)^2 - 1 = 0$, $i = 1,\ldots,N_l$ and $[\mathbf{h}_l^{\mathsf{D}}]_i(\boldsymbol{\theta}_l) = \prod_{p=1,\cdots,P_l}([\boldsymbol{\theta}_l]_i - e^{j\phi_{l_p}}) = 0$, $i = 1,\ldots,N_l$, respectively. The gradient $\nabla[\mathbf{h}_l^{\mathsf{C}}(\boldsymbol{\theta}_l)]_1,\ldots,\nabla[\mathbf{h}_l^{\mathsf{C}}(\boldsymbol{\theta}_l)]_{N_l}$ and $\nabla[\mathbf{h}_l^{\mathsf{D}}(\boldsymbol{\theta}_l)]_1,\ldots,\nabla[\mathbf{h}_l^{\mathsf{D}}(\boldsymbol{\theta}_l)]_{N_l}$ meet the LICQ evidently since $[\boldsymbol{\theta}_l]_j$ appears only at the j-the entry of vector $\nabla[\mathbf{h}_l^{\mathsf{C}}(\boldsymbol{\theta}_l)]_i$ and $\nabla[\mathbf{h}_l^{\mathsf{D}}(\boldsymbol{\theta}_l)]_i$. Therefore, the updated $\boldsymbol{\theta}_l$ via Lemma 5 is a KKT point of problem (2) or problem (4). Considering the properties of minorizing functions [18], it can be shown $\{\{\mathbf{w}_i^{(\infty)}\},\{\boldsymbol{\Theta}_i^{(\infty)}\}\}$ is a KKT point of Problem (WSRMax). ■

Algorithm 1 The BMM Approach to Problem (WSRMax).

Input: $\{\mathbf{h}_{r,i}\}$, $\{\mathbf{h}_{d,i}\}$, $\{\mathbf{G}_i\}$, $P_{\max}$, $\{\omega_k\}$, σ^2, initial feasible values of $\{\mathbf{w}_i\}$ and $\{\boldsymbol{\Theta}_i\}$.

Repeat

1. Update $\{\mathbf{w}_i\}$ in closed-form via Proposition 3
2. Update $\boldsymbol{\Theta}_1,\ldots,\boldsymbol{\Theta}_L$ successively in closed-forms (either in series or in parallel for each $\boldsymbol{\Theta}_l$) via Proposition 5

Until the value of the objective function converges.

Remark 11. The proposed BMM algorithm is general enough to be applied to many other systems like the single-hop system [5], the two-hop system with two cooperative RISs [8], RIS-aided MIMO interfering broadcast channels [13], etc.

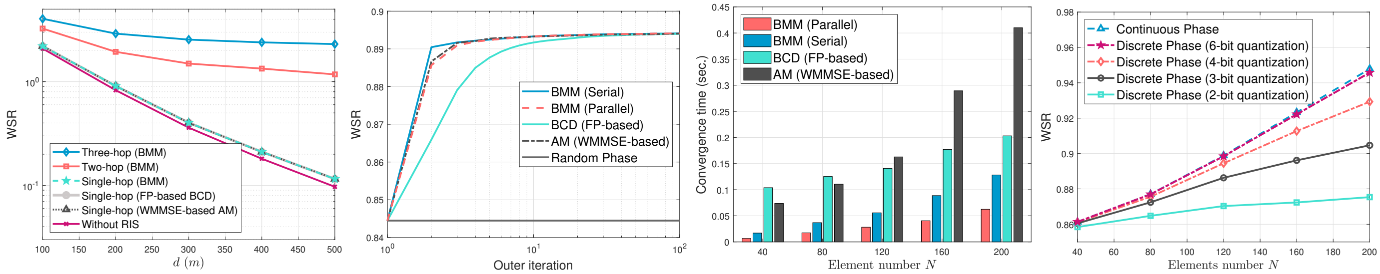
IV. Numerical Simulations

In this section, numerical simulations are conducted to demonstrate the effectiveness of the proposed BMM algorithm and the benefits of multi-hop RIS-aided systems. The simulation settings are given as follows. For the single-hop RIS system, locations of the BS and the RIS are set as $(0,0,10)m$ and $(d,0,10)m$, and the K users are randomly distributed in a circle centered at $(d,30,0)m$ with radius of $10m$. For the two-hop RIS system, an extra RIS located at $(\frac{d}{2},0,10)m$ is added, while for the three-hop RIS system, the third RIS located at $(\frac{d}{4},0,10)m$ is included. We assume that the BS is equipped with a uniform linear array with antenna spacing of $\frac{\lambda}{2}$ and the RISs are equipped with a uniform planar array with its element spacing of $\frac{\lambda}{8}$, where λ is the wavelength. We further assume that the channel fading is frequency flat and adopt the Rician fading model for all channels. For example, the channel from the BS to the RIS at the single-hop RIS system is modeled as

$$\mathbf{G}_1 = \sqrt{\frac{\kappa_G(d)}{K_G+1}}(\sqrt{K_G}\mathbf{G}_{1,\mathsf{LoS}} + \mathbf{G}_{1,\mathsf{NLoS}}),$$

where $\kappa_G(d)$ is the distance-dependent path loss, $K_G \in [0,\infty)$ is the Rician factor, $\mathbf{G}_{1,\mathsf{LoS}}$ is the LoS component that equals the product of the array responses at the two sides, and $[\mathbf{G}_{1,\mathsf{NLoS}}]_{ij} \sim \mathcal{CN}(0,1)$ with $i = 1,\ldots N_1$ and $j = 1,\ldots,M$ is the NLoS component following the Rayleigh fading model. Specifically, we have chosen $\kappa_G(d) = T_0(\frac{d}{d_0})^{-\varrho_G}$ with $T_0 = -30$dB, the reference distance $d_0 = 1m$, $\varrho_G = 2.2$, and $K_G = 3$. Similarly, the other channels are modeled as the $\mathbf{G}_1$ and the path loss exponents for the RIS-RIS link(s), the BS-user link, and the RIS-user link are set as $\varrho_G = 2.2$, $\varrho_d = 3.5$, and $\varrho_r = 2.8$, respectively. The Rician factors for $\{\mathbf{G}_i\}$, $\{\mathbf{h}_{d,i}\}$, and $\{\mathbf{h}_{r,i}\}$ are set as $K_G = 3$, $K_d = 0$, and $K_r = 3$, respectively. Besides, we have considered the noise power spectrum density of -169dBm/Hz and the transmission bandwidth of 240kHz. All the simulation curves are averaged over 100 independent channel realizations.

In the single-hop RIS scenario, four benchmarks are considered: i) AM (WMMSE-based) method [5]; ii) BCD (FP-based) method [5]; iii) Random Phase where the phase shift



(a) Achievable WSR. (b) Convergence behavior. (c) Convergence time. (d) Ach. WSR in discrete phase.

Fig. 1: Performance comparisons ($P_{\mathsf{max}} = 0$dBm, $K = 4$, and $M = 4$).

matrices $\{\boldsymbol{\Theta}_i\}$ are randomly initialized and $\{\mathbf{w}_i\}$ is optimized via minorization-maximization (MM); iv) Without RIS, i.e., no RIS is used and $\{\mathbf{w}_i\}$ is optimized via MM. The comparison in terms of achievable WSR is depicted in Fig. 1a where we have set $N = 100$, $\sigma^2 = 1$, and $I_1 = \cdots = I_k = L$, from which we can see that AM, BCD, and the proposed BMM approach converge to the same WSR as expected, and the benefits of introducing multiple RIS to improve WSR is obvious. The convergence behaviors of different approaches with $d = 200m$ under single-hop scenario are shown in Fig. 1b, from which we can see that the serial BMM acquires the least number of iterations to converge.[6] The convergence speed in terms of CPU time is further presented in Fig. 1c, where the parallel BMM achieves the lowest time consumption among all cases. Besides, the influence of the quantization level of the discrete phase is also investigated. Fig. 1d showcases the approximation performance of WSR under discrete schemes compared to the continuous scheme.

V. Conclusions

In this paper, the weighted sum-rate maximization problem for a multi-hop RIS-aided multi-user communication system has been studied. Algorithms based on the block minorization-maximization technique has been proposed with its superior performance demonstrated through numerical simulations.

References

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, 2014.

[2] 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, 2019.

[3] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. D. Renzo, and M. Debbah, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.

[4] X. Yuan, Y.-J. A. Zhang, Y. Shi, W. Yan, and H. Liu, "Reconfigurable-intelligent-surface empowered wireless communications: Challenges and opportunities," *IEEE Trans. Wireless Commun.*, 2021.

[5] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, 2020.

[6] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3236–3251, 2020.

[6]To make it consistant with all the methods (especailly the double-loop methods) under comparison, "one iteration" is defined after all the variable blocks are updated once. The convergence speed can be better validated with the complement of the convergence plot in terms of CPU time (i.e., Fig. 1c).

[7] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, 2020.

[8] B. Zheng, C. You, and R. Zhang, "Double-IRS assisted multi-user MIMO: Cooperative passive beamforming design," *IEEE Trans. Wireless Commun.*, 2021.

[9] G. Shen, J. Liu, D. Wang, J. Wang, and S. Jin, "Multi-hop relay for next-generation wireless access networks," *Bell Labs Tech. J.*, vol. 13, no. 4, pp. 175–193, 2009.

[10] D. Gündüz, M. A. Khojastepour, A. Goldsmith, and H. V. Poor, "Multi-hop MIMO relay networks: diversity-multiplexing trade-off analysis," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1738–1747, 2010.

[11] C. W. Tan, M. Chiang, and R. Srikant, "Maximizing sum rate and minimizing mse on multiuser downlink: Optimality, fast algorithms and equivalence via max-min sinr," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6127–6143, 2011.

[12] P. C. Weeraddana, M. Codreanu, M. Latva-aho, A. Ephremides, C. Fischione *et al.*, "Weighted sum-rate maximization in wireless networks: A review," *Found. and Trends® in Netw.*, vol. 6, no. 1–2, 2012.

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

[14] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*. Princeton Univ. Press, 2009.

[15] K. Shen and W. Yu, "Fractional programming for communication system–Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, 2018.

[16] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas in Commun.*, vol. 38, no. 8, pp. 1809–1822, 2020.

[17] C. Huang, Z. Yang, G. C. Alexandropoulos, K. Xiong, L. Wei, C. Yuen, Z. Zhang, and M. Debbah, "Multi-hop RIS-empowered terahertz communications: A DRL-based hybrid beamforming design," *arXiv preprint arXiv:2101.09137*, 2021.

[18] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.

[19] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, 2016.

[20] Z. Zhao, "Joint transmit waveforms and receive filters design for large-scale MIMO beampattern synthesis," in *2020 IEEE 11th Sens. Array Multichannel Signal Process. Workshop (SAM)*. IEEE, 2020, pp. 1–5.

[21] D. P. Bertsekas, *Nonlinear programming*. Belmont, MA, USA: Athena Sci., 1999.

[22] Z. Zhao and D. P. Palomar, "MIMO transmit beampattern matching under waveform constraints," in *2018 IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2018, pp. 3281–3285.

[23] R. Zhou, Z. Zhao, and D. P. Palomar, "Unified framework for minimax MIMO transmit beampattern matching under waveform constraints," in *2019 IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2019, pp. 4150–4154.

[24] R. Varadhan and C. Roland, "Simple and globally convergent methods for accelerating the convergence of any EM algorithm," *Scandinavian J. of Statist.*, vol. 35, no. 2, pp. 335–353, 2008.

[25] D. W. Peterson, "A review of constraint qualifications in finite-dimensional spaces," *SIAM Review*, vol. 15, no. 3, pp. 639–654, 1973.