Hybrid Active-Passive Reconfigurable Intelligent Surface-Assisted Multi-User MISO Systems

Nhan T. Nguyen*, V.-Dinh Nguyen†, Qingqing Wu‡, Antti Tölli*, Symeon Chatzinotas†, and Markku Juntti*
*Centre for Wireless Communications, University of Oulu, P.O.Box 4500, FI-90014, Finland
†Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, L-1855 Luxembourg
‡State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau 999078, China
{nhan.nguyen, antti.tolli, markku.juntti}@oulu.fi; {dinh.nguyen, symeon.chatzinotas}@uni.lu; qingqingwu@um.edu.mo

***Abstract*—We consider a multi-user multiple-input single-output (MISO) communications system which is assisted by a hybrid active-passive reconfigurable intelligent surface (RIS). Unlike conventional passive RISs, hybrid RIS is equipped with a few active elements with the ability to reflect and amplify incident signals to significantly improve the system performance. Towards a fairness design, we maximize the minimum rate among all users through jointly optimizing the transmit beamforming vectors and RIS reflecting/amplifying coefficients. Combining tools from block coordinate ascend and successive convex approximation, the challenging nonconvex problem is efficiently solved by a low-complexity iterative algorithm. The numerical results show that a hybrid RIS with 4 active elements out of a total of 50 elements with a power budget of -1 dBm offers an improvement of up to 80% to the considered system, while that achieved by a fully passive RIS is only 27%.**

***Index Terms*—Hybrid active-passive RIS, multi-user MISO, beamforming, RIS semi-passive beamforming, successive convex approximation.**

I. Introduction

Reconfigurable intelligent reflecting surfaces (RISs) have been advocated as a cost and energy-efficient solution to improve the performance of wireless communications systems [1], [2]. The reflecting elements of RISs can be configured to not only improve the received signal power but also mitigate interference in multi-user wireless systems [3]. The performance improvement of RISs in multi-user multiple-input multiple-output (MISO) systems are investigated in [4]–[13]. Through the joint optimization of the transmit beamforming/power at the base station (BS) and the reflecting coefficients of the RIS, it is shown in [4], [6], [9]–[12] that the system achievable sum-rate and/or fairness can be significantly improved thanks to the deployment of RISs. In [8], Li *et al.* exploited the concept of constructive interference and proposed an efficient design of RIS coefficients to minimize the symbol error rate. Whereas, the works in [9] and [10] focused on practical scenarios, where only statistical channel information and/or low-resolution phase shifts at the RIS are available for the joint design. To overcome the challenging nonconvexity and strongly coupled variables in the design, Huang *et al.* [11] developed a deep reinforcement learning-based algorithm to simultaneously obtain the transmit beamformers and RIS phase shifts. In particular, Zhang *et al.* [7] derived an expression of the system asymptotic capacity and determined the required number of RIS elements to meet a predefined quality of service (QoS).

All the above works considered the passive RIS, in which reflecting elements are unable to amplify incident signals, and thus, a large number of elements are required to compensate for the severe cascaded path loss [2]. Recently, hybrid active-passive RIS architectures have been introduced [14]–[18] to overcome the inherent limitation of passive RISs, especially in harsh transmission scenarios such as in low signal-to-noise ratio (SNR) regime and/or severe path loss. The key idea of the hybrid RIS is to add a few active elements to the conventional passive RIS, allowing them to reflect and amplify incident signals simultaneously. As a result, the hybrid RIS can reduce effects of the double path loss and significantly improve the system performance in terms of spectral efficiency [16], [19], [20], secrecy rate [21], harvested energy [22], and reliability [23]. These advantages are also reaped by fully active RISs [24]–[26], which, however, come at high cost of power consumption and hardware design, compared to the hybrid RIS with only a few active elements. Moreover, active elements with RF chains for processing of incident signals and channel estimation at the RIS are presented in [14], [15], [27].

In this work, we deploy the hybrid active-passive RIS to assist the multi-user MISO system. Thanks to the signal amplification, the hybrid RIS with a small-to-moderate size can efficiently compensate for the severe path loss and blockage in the communication links, especially in complex urban scenarios. Focusing on the fairness design, our goal is to maximize the minimum rate among all user equipments (UEs) by jointly optimizing the transmit beamformers and RIS amplifying/reflecting coefficients. To the best of the authors' knowledge, it has not been investigated in the literature. The problem is practically appealing but particularly more challenging, compared to those with conventional passive RISs, due to additional power constraints and amplified noise/interference caused by RIS active elements. To effectively solve the challenging optimization problem, we propose an efficient algorithm based on the block coordinate ascent (BCA) and successive approximation (SCA). Specifically, the original problem is first decomposed into two subproblems, which are then alternately solved by the SCA technique. Finally, the proposed design is evaluated by numerical results, which show that with the same total power budget, the hybrid RIS offers remarkable performance improvement compared to the systems without RIS and that with the conventional passive RIS, especially when the transmit power at the BS is limited.

II. System Model and Problem Formulation

A. System Model

We consider a downlink wireless network where a multiple-antenna BS serves K single-antenna UEs. The communication between BS and UEs is assisted by a hybrid active-passive RIS installed on the building facade. RIS is equipped with N elements, out of which, $N_{\rm a}$ ($N_{\rm a} \ll N$) elements are activated. The positions of active elements are predetermined in $\mathcal{A} \subset \{1, 2, \ldots, N\}$ with $|\mathcal{A}| = N_{\rm a}$. The RIS's active elements can potentially be realized by low-power reflection amplifiers [28]. We refer readers to [17], [24], [28] for more details on the reflection amplifier-based active RIS. It is seen that the fully passive RIS (i.e., with $N_{\rm a} = 0$) is just a special case of the hybrid RIS. Therefore, in this work, we will use the general term "*RIS*" for the discussion in the system model, while specific terms "*passive RIS*" or "*hybrid RIS*" are used in comparisons.

Let α_n denote the coefficient associated with the nth element of the RIS. We can express α_n as $\alpha_n = |\alpha_n| e^{j\theta_n}$, where $\theta_n \in [0, 2\pi)$ represents the phase shift, $|\alpha_n| \in [0, 1]$ for $n \notin \mathcal{A}$, and $|\alpha_n| \leq a_{\max}$ for $n \in \mathcal{A}$. Here, $a_{\max}$ is the maximum power amplification gain that the active load can provide, which is up to 40 dB if active elements are realized by reflection amplifiers [24], [28]. We note here that to mitigate interference in multi-user systems, the reflection amplitude of passive RIS elements may not necessarily be unity [3]. Let $\boldsymbol{\Upsilon} \triangleq \text{diag}\{\alpha_1, \ldots, \alpha_N\} \in \mathbb{C}^{N \times N}$ be the diagonal matrix of the RIS coefficients. For ease of exposition in the following analysis, we define an additive decomposition $\boldsymbol{\Upsilon} = \boldsymbol{\Phi} + \boldsymbol{\Psi}$, where $\boldsymbol{\Psi} = \mathbb{1}_N^{\mathcal{A}} \circ \boldsymbol{\Upsilon}$ and $\boldsymbol{\Phi} = \left(\boldsymbol{I}_N - \mathbb{1}_N^{\mathcal{A}}\right) \circ \boldsymbol{\Upsilon}$ contain the active and passive coefficients, respectively. Here, $\mathbb{1}_N^{\mathcal{A}}$ is an $N \times N$ diagonal matrix whose non-zero elements are all unity and have positions determined by $\mathcal{A}$, and $\circ$ represents a Hadamard product.

Let $\boldsymbol{h}_{0,k}^H \in \mathbb{C}^{1 \times N_t}$, $\boldsymbol{H}_1 \in \mathbb{C}^{N \times N_t}$, and $\boldsymbol{h}_{2,k}^H \in \mathbb{C}^{1 \times N}$ denote the channels between BS and UE k, between BS and RIS, and between RIS and UE k, respectively. The effective channel between BS and UE k can be expressed as $\boldsymbol{h}_k^H = \boldsymbol{h}_{0,k}^H + \boldsymbol{h}_{2,k}^H \boldsymbol{\Upsilon} \boldsymbol{H}_1$. Denote by s_k with $\mathbb{E}\{|s_k|^2\} = 1$ and $\boldsymbol{w}_k \in \mathbb{C}^{N_t \times 1}$ the transmitted symbol and the beamforming vector intended for UE k, respectively. The transmitted signal from the BS can be given as $\boldsymbol{x} = \sum_{k=1}^K \boldsymbol{w}_k s_k \in \mathbb{C}^{N_t \times 1}$. Thus, the total transmit power at the BS is $p_{\rm BS} = \sum_{k=1}^K \|\boldsymbol{w}_k\|^2 \leq p_{\rm BS}^{\max}$, where $p_{\rm BS}^{\max}$ is the maximum transmit power of the BS. The received signal at UE k can be given as

$$y_k = \boldsymbol{h}_k^H \boldsymbol{w}_k s_k + \sum\nolimits_{j \neq k} \boldsymbol{h}_k^H \boldsymbol{w}_j s_j + n_k, \tag{1}$$

where $n_k = \boldsymbol{h}_{2,k}^H \boldsymbol{\Psi} \boldsymbol{n}_{\rm r} + n_{\rm u}$ is the aggregated noise at UE k, with $n_{\rm u} \sim \mathcal{CN}(0, \sigma_{\rm u}^2)$ being the additive white Gaussian noise (AWGN) at UE k; and $\boldsymbol{n}_{\rm r} \sim \mathcal{CN}(\boldsymbol{0}, \mathbb{1}_N^{\mathcal{A}} \circ \sigma_{\rm r}^2 \boldsymbol{I}_N)$ is the total effective noise including self-interference and AWGN noise caused by RIS active elements operating in full-duplex mode [17].

B. Problem Formulation

From (1), the achievable rate of UE k (in nats/s/Hz) can be expressed as

$$R_k = \log\left(1 + \frac{|\boldsymbol{h}_k^H \boldsymbol{w}_k|^2}{\sum_{j \neq k} |\boldsymbol{h}_k^H \boldsymbol{w}_j|^2 + \sigma_{\rm r}^2 \|\boldsymbol{h}_{2,k}^H \boldsymbol{\Psi}\|^2 + \sigma_{\rm u}^2}\right). \tag{2}$$

Let $p_{\rm RIS}$ denote the transmit power of active elements of the RIS. It can be expressed as $p_{\rm RIS} = \mathbb{E}\{\|\boldsymbol{\Psi}(\boldsymbol{H}_1 \boldsymbol{x} + \boldsymbol{n}_{\rm r})\|^2\} \overset{(a)}{=} \sum_{n \in \mathcal{A}} |\alpha_n|^2 \xi_n$, where $\xi_n \triangleq \sigma_{\rm r}^2 + \|\boldsymbol{h}_{1,n}\|^2 \sum_{k=1}^K \|\boldsymbol{w}_k\|^2$; $\boldsymbol{h}_{1,n}$ denotes the nth row of $\boldsymbol{H}_1$, and equality (a) follows the diagonal structure of $\boldsymbol{\Psi}$ whose non-zero elements are in $\mathcal{A}$. The total transmit power at the RIS is constrained as $p_{\rm RIS} \leq p_{\rm RIS}^{\max}$, where $p_{\rm RIS}^{\max}$ is the power budget.

We aim to maximize the minimum rate among all UEs through jointly optimizing the transmit beamformers and RIS coefficients, which can be mathematically formulated as

$$\begin{aligned}
\underset{\{\boldsymbol{w}_k\}, \{\alpha_n\}}{\text{maximize}} \quad & \min_k \{R_k\} && \text{(3a)}\\
\text{subject to} \quad & 0 \leq \sum\nolimits_{k=1}^K \|\boldsymbol{w}_k\|^2 \leq p_{\rm BS}^{\max}, && \text{(3b)}\\
& 0 \leq \theta_n \leq 2\pi, \ \forall n, && \text{(3c)}\\
& |\alpha_n| \leq 1, \ \forall n \notin \mathcal{A}, && \text{(3d)}\\
& |\alpha_n| \leq a_{\max}, \ \forall n \in \mathcal{A}, && \text{(3e)}\\
& \sum\nolimits_{n \in \mathcal{A}} |\alpha_n|^2 \xi_n \leq p_{\rm RIS}^{\max}, && \text{(3f)}
\end{aligned}$$

where (3c)–(3f) are constrains of the hybrid RIS. Note in (3e) that is only active elements ($n \in \mathcal{A}$) can amplify the signals with amplification gains restricted by $a_{\max}$ [24]. It is clear that the objective function is non-concave and non-smooth, resulting in a non-convexity of problem (3).

III. Proposed Design

The strong coupling between $\{\boldsymbol{w}_k\}$ and $\{\alpha_n\}$ in the rate function makes problem (3) more difficult to solve. A direct application of SCA comes at a cost of high computational complexity. In what follows, we first transform problem (3) into a more tractable form as (3) as

$$\begin{aligned}
\underset{\tau, \{\boldsymbol{w}_k\}, \{\alpha_n\}}{\text{maximize}} \quad & \tau && \text{(4a)}\\
\text{subject to} \quad & R_k \geq \tau, \ \forall k, && \text{(4b)}\\
& \text{(3b)} - \text{(3f)}, && \text{(4c)}
\end{aligned}$$

where τ is a newly introduced variable. By utilizing the BCA approach, we decouple (4) into two sub-problems with respect to $\{\boldsymbol{w}_k\}$ and $\{\alpha_n\}$, each of which is efficiently solved by the SCA method.

A. Transmit Beamforming Design

For given $\{\alpha_n\}$, the optimal beamformers $\{\boldsymbol{w}_k\}$ at the BS can be found by solving the following problem:

$$\underset{\tau, \{\boldsymbol{w}_k\}}{\text{maximize}} \ \tau, \ \text{subject to (4b), (3b), (3f).} \tag{5}$$

where (3b) and (3f) are convex. To convexify constraint (4b), we introduce slack variables $\{\gamma_k\}$ to express it equivalently as

$$\log(1+\gamma_k) \geq \tau,\ \forall k, \tag{6a}$$

$$\frac{|\boldsymbol{h}_k^H \boldsymbol{w}_k|^2}{\sum_{j\neq k}|\boldsymbol{h}_k^H \boldsymbol{w}_j|^2 + \sigma_k^2} \geq \gamma_k,\ \forall k, \tag{6b}$$

where $\sigma_k^2 \triangleq \sigma_{\mathrm{r}}^2 \|\boldsymbol{h}_{2,k}^H \boldsymbol{\Psi}\|^2 + \sigma_{\mathrm{u}}^2$ is a constant with respect to $\{\boldsymbol{w}_k\}$. Let us define $\boldsymbol{w} \triangleq [\boldsymbol{w}_1^T, \ldots, \boldsymbol{w}_K^T]^T \in \mathbb{C}^{KN_t \times 1}$, $\tilde{\boldsymbol{H}}_k \triangleq \boldsymbol{h}_k \boldsymbol{h}_k^H \in \mathbb{C}^{N_t \times N_t}$, and

$$\hat{\boldsymbol{H}}_k \triangleq \mathrm{blockdiag}\{\boldsymbol{0}, \ldots, \boldsymbol{0}, \tilde{\boldsymbol{H}}_k,\ \boldsymbol{0}, \ldots, \boldsymbol{0}\} \in \mathbb{C}^{KN_t \times KN_t},$$

$$\bar{\boldsymbol{H}}_k \triangleq \mathrm{blockdiag}\{\tilde{\boldsymbol{H}}_k, \ldots, \tilde{\boldsymbol{H}}_k,\ \boldsymbol{0},\ \tilde{\boldsymbol{H}}_k, \ldots, \tilde{\boldsymbol{H}}_k\} \in \mathbb{C}^{KN_t \times KN_t}.$$

As a result, we can further rewrite constraint (6b) as $(\boldsymbol{w}^H \hat{\boldsymbol{H}}_k \boldsymbol{w})/(\boldsymbol{w}^H \bar{\boldsymbol{H}}_k \boldsymbol{w} + \sigma_k^2) \geq \gamma_k, \forall k$, which is equivalent to

$$\boldsymbol{w}^H \bar{\boldsymbol{H}}_k \boldsymbol{w} + \sigma_k^2 - \frac{\boldsymbol{w}^H \hat{\boldsymbol{H}}_k \boldsymbol{w}}{\gamma_k} \leq 0,\ \forall k. \tag{7}$$

By applying the first-order Taylor approximation around the point $[\boldsymbol{w}^{(i)}, \gamma_k^{(i)}]$ found at iteration i, the concave function $f_{\mathtt{qol}}(\boldsymbol{w}, \gamma_k) \triangleq -\boldsymbol{w}^H \hat{\boldsymbol{H}}_k \boldsymbol{w}/\gamma_k$ is linearized as

$$\begin{aligned} f_{\mathtt{qol}}(\boldsymbol{w}, \gamma_k) &\leq F_{\mathtt{qol}}(\boldsymbol{w}, \gamma_k; \boldsymbol{w}^{(i)}, \gamma_k^{(i)}) \\ &\triangleq \frac{\boldsymbol{w}^{(i)H} \hat{\boldsymbol{H}}_k \boldsymbol{w}^{(i)}}{\gamma_k^{(i)2}} \gamma_k - \frac{2\Re\left(\boldsymbol{w}^{(i)H} \hat{\boldsymbol{H}}_k \boldsymbol{w}\right)}{\gamma_k^{(i)}}. \end{aligned} \tag{8}$$

As a result, (7) can be iteratively replaced by the convex constraint

$$\boldsymbol{w}^H \bar{\boldsymbol{H}}_k \boldsymbol{w} + \sigma_k^2 + F_{\mathtt{qol}}(\boldsymbol{w}, \gamma_k; \boldsymbol{w}^{(i)}, \gamma_k^{(i)}) \leq 0,\ \forall k. \tag{9}$$

Given $\sum_{k=1}^K \|\boldsymbol{w}_k\|^2 = \|\boldsymbol{w}\|^2$, (3f) becomes

$$\sum_{n\in\mathcal{A}} |\alpha_n|^2 \left(\sigma_{\mathrm{r}}^2 + \|\boldsymbol{h}_{1,n}\|^2 \|\boldsymbol{w}\|^2\right) \leq p_{\mathrm{RIS}}^{\max}. \tag{10}$$

Summary, we solve the following convex program of (5) at iteration i:

$$\underset{\tau, \boldsymbol{w}}{\text{maximize}}\ \tau,\ \text{subject to (3b), (6a), (9), (10).} \tag{11}$$

B. Optimization of RIS Coefficients

Given $\{\boldsymbol{w}_k\}$, the RIS coefficients ($\{\alpha_n\}$) can be optimized by solving the following problem

$$\underset{\tau, \{\alpha_n\}}{\text{maximize}}\ \tau,\ \text{subject to (4b), (3c) – (3f),} \tag{12}$$

where constraints (3c)-(3f) are convex with respect to $\{\alpha_n\}$. The optimization variables $\{\alpha_n\}$ have not been exposed in the current form of the nonconvex constraint (4b). To address this issue, we denote $\bar{h}_{0,kj} \triangleq \boldsymbol{h}_{0,k}^H \boldsymbol{w}_j$ and $\bar{\boldsymbol{h}}_{1,k} \triangleq \boldsymbol{H}_1 \boldsymbol{w}_k, \forall k, j$, yielding $\boldsymbol{h}_k^H \boldsymbol{w}_k = \bar{h}_{0,kk} + \boldsymbol{h}_{2,k}^H \boldsymbol{\Upsilon} \bar{\boldsymbol{h}}_{1,k}$ and $\boldsymbol{h}_k^H \boldsymbol{w}_j = \bar{h}_{0,kj} + \boldsymbol{h}_{2,k}^H \boldsymbol{\Upsilon} \bar{\boldsymbol{h}}_{1,j}$. Then, we can write the SINR term in (2) as

$$\mathrm{SINR} = \frac{|\bar{h}_{0,kk} + \boldsymbol{h}_{2,k}^H \boldsymbol{\Upsilon} \bar{\boldsymbol{h}}_{1,k}|^2}{\sum_{j\neq k} |\bar{h}_{0,kj} + \boldsymbol{h}_{2,k}^H \boldsymbol{\Upsilon} \bar{\boldsymbol{h}}_{1,j}|^2 + \sigma_{\mathrm{r}}^2 \|\boldsymbol{h}_{2,k}^H \boldsymbol{\Psi}\|^2 + \sigma_{\mathrm{u}}^2}.$$

By defining $\boldsymbol{\alpha} \triangleq [\alpha_1, \ldots, \alpha_N]^T \in \mathbb{C}^{N\times 1}$, $\tilde{\boldsymbol{H}}_{2,k} \triangleq \mathrm{diag}\{\boldsymbol{h}_{2,k}^H\} \in \mathbb{C}^{N\times N}$, and $\tilde{\boldsymbol{h}}_{12,kj} \triangleq \tilde{\boldsymbol{H}}_{2,k} \bar{\boldsymbol{h}}_{1,j} \in \mathbb{C}^{N\times 1}$, we have $\bar{h}_{0,kj} + \boldsymbol{h}_{2,k}^H \boldsymbol{\Upsilon} \bar{\boldsymbol{h}}_{1,j} = \bar{h}_{0,kj} + \boldsymbol{\alpha}^T \tilde{\boldsymbol{h}}_{12,kj}$ and $\boldsymbol{h}_{2,k}^H \boldsymbol{\Psi} = \boldsymbol{\alpha}^T \mathbb{1}_N^{\mathcal{A}} \tilde{\boldsymbol{H}}_{2,k}, \forall k, j$. After straightforward algebraic manipulations, the numerator and denominator of the SINR can be expressed as $\boldsymbol{\alpha}^H \boldsymbol{Q}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \boldsymbol{t}_k\big) + e_k$ and $\boldsymbol{\alpha}^H \tilde{\boldsymbol{Q}}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \tilde{\boldsymbol{t}}_k\big) + \tilde{e}_k$, respectively; Here, $\boldsymbol{Q}_k = \tilde{\boldsymbol{h}}_{12,kk}^* \tilde{\boldsymbol{h}}_{12,kk}^T$, $\boldsymbol{t}_k = \tilde{\boldsymbol{h}}_{12,kk}^* \bar{h}_{0,kk}$, and $e_k = |\bar{h}_{0,kk}|^2$; furthermore, $\tilde{\boldsymbol{Q}}_k = \sigma_{\mathrm{r}}^2 \mathbb{1}_N^{\mathcal{A}} \tilde{\boldsymbol{H}}_{2,k}^* \tilde{\boldsymbol{H}}_{2,k}^T \mathbb{1}_N^{\mathcal{A}} + \sum_{j\neq k} \tilde{\boldsymbol{h}}_{12,kj}^* \tilde{\boldsymbol{h}}_{12,kj}^T$, $\tilde{\boldsymbol{t}}_k = \sum_{j\neq k} \tilde{\boldsymbol{h}}_{12,kj}^* \bar{h}_{0,kj}$, $\tilde{e}_k = \sigma_{\mathrm{u}}^2 + \sum_{j\neq k} |\bar{h}_{0,kj}|^2$, and $\Re(\cdot)$ denotes the real part of a complex number. Thus, the rate function of UE k is rewritten as

$$R_k = \log\left(1 + \frac{\boldsymbol{\alpha}^H \boldsymbol{Q}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \boldsymbol{t}_k\big) + e_k}{\boldsymbol{\alpha}^H \tilde{\boldsymbol{Q}}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \tilde{\boldsymbol{t}}_k\big) + \tilde{e}_k}\right), \tag{13}$$

where $\boldsymbol{\alpha}$ is clearly exposed. By introducing new variables $\{\tilde{N}_k, \tilde{\gamma}_k\}$, constraint (4b) is equivalently rewritten as

$$\log(1+\tilde{\gamma}_k) \geq \tau,\ \forall k, \tag{14a}$$

$$\frac{\tilde{N}_k^2}{\boldsymbol{\alpha}^H \tilde{\boldsymbol{Q}}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \tilde{\boldsymbol{t}}_k\big) + \tilde{e}_k} \geq \tilde{\gamma}_k,\ \forall k, \tag{14b}$$

$$\boldsymbol{\alpha}^H \boldsymbol{Q}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \boldsymbol{t}_k\big) + e_k \geq \tilde{N}_k^2,\ \forall k. \tag{14c}$$

The nonconvex constraints include (14b) and (14c), which are expressed as

$$\boldsymbol{\alpha}^H \tilde{\boldsymbol{Q}}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \tilde{\boldsymbol{t}}_k\big) + \tilde{e}_k - \frac{\tilde{N}_k^2}{\tilde{\gamma}_k} \leq 0,\ \forall k, \tag{15a}$$

$$\tilde{N}_k^2 - \|\bar{\boldsymbol{Q}}_k \boldsymbol{\alpha}\|^2 - 2\Re\big(\boldsymbol{\alpha}^H \boldsymbol{t}_k\big) - e_k \leq 0,\ \forall k, \tag{15b}$$

where $\bar{\boldsymbol{Q}} = \boldsymbol{Q}_k^{1/2}$. To address the nonconvexity of (15a) and (15b), we use the following approximations

$$\begin{aligned} -\frac{\tilde{N}_k^2}{\tilde{\gamma}_k} &\leq F_{\mathtt{qol}}(\tilde{N}_k, \tilde{\gamma}_k; \tilde{N}_k^{(i)}, \tilde{\gamma}_k^{(i)}) \\ -\|\bar{\boldsymbol{Q}}_k \boldsymbol{\alpha}\|^2 &\leq F_{\mathtt{qua}}(\bar{\boldsymbol{Q}}_k \boldsymbol{\alpha}; \bar{\boldsymbol{Q}}_k \boldsymbol{\alpha}^{(i)}),\ \forall k, \end{aligned} \tag{16}$$

where $F_{\mathtt{qol}}(\cdot;\cdot)$ is defined in (8), and $F_{\mathtt{qua}}(\boldsymbol{x}; \boldsymbol{x}_0) \triangleq 2\boldsymbol{x}_0^H(\boldsymbol{x}_0 - \boldsymbol{x}) - \|\boldsymbol{x}_0\|^2$ is a convex approximation of the concave function $-\|\boldsymbol{x}\|^2$ around $\boldsymbol{x}_0$. As a result, (15a) and (15b) are iteratively replaced by the following convex constraints

$$\begin{aligned} &\boldsymbol{\alpha}^H \tilde{\boldsymbol{Q}}_k \boldsymbol{\alpha} + 2\Re\big(\boldsymbol{\alpha}^H \tilde{\boldsymbol{t}}_k\big) + \tilde{e}_k \\ &\quad + F_{\mathtt{qol}}(\tilde{N}_k, \tilde{\gamma}_k; \tilde{N}_k^{(i)}, \tilde{\gamma}_k^{(i)}) \leq 0, \forall k, \end{aligned} \tag{17a}$$

$$\tilde{N}_k^2 + F_{\mathtt{qua}}(\bar{\boldsymbol{Q}}_k \boldsymbol{\alpha}; \bar{\boldsymbol{Q}}_k \boldsymbol{\alpha}^{(i)}) - 2\Re\big(\boldsymbol{\alpha}^H \boldsymbol{t}_k\big) - e_k \leq 0. \tag{17b}$$

Furthermore, (3f) is transformed to a more compact form as

$$\boldsymbol{\alpha}^H \boldsymbol{\Xi} \boldsymbol{\alpha} \leq p_{\mathrm{RIS}}^{\max}, \tag{18}$$

where $\boldsymbol{\Xi} = \mathrm{diag}\{\tilde{\xi}_1, \ldots, \tilde{\xi}_N\}$ with $\tilde{\xi}_n = \xi_n$ for $n \in \mathcal{A}$, and $\tilde{\xi}_n = 0$, otherwise.

Finally, problem (12) can be approximated by the following convex program at iteration i

$$\underset{\tau, \boldsymbol{\alpha}, \{\tilde{N}_k\}, \{\tilde{\gamma}_k\}}{\text{maximize}}\ \tau,\ \text{s. t. (3c) – (3e), (14a), (17a) – (18).} \tag{19}$$

Algorithm 1 Iterative Algorithm to Solve Problem (3)

1: **Initialize** $\{\boldsymbol{w}_k^{(0)}, \alpha_n^{(0)}\}$. The feasible points for $\{\tilde{N}_k^{(0)}, \tilde{\gamma}_k^{(0)}\}$ are set to hold the equalities in (14b) and (14c). Set $i = 0$.
2: **repeat**
3: Solve problem (11) for given $\{\alpha_n^{(i)}\}$ to obtain the solution $\{\boldsymbol{w}_k^\star\}$. Update $\{\boldsymbol{w}_k^{(i+1)}\} := \{\boldsymbol{w}_k^\star\}$.
4: Solve problem (19) for given $\{\boldsymbol{w}_k^{(i+1)}\}$ to obtain the solutions $\boldsymbol{\alpha}^\star, \{\tilde{N}_k^\star\}, \{\tilde{\gamma}_k^\star\}$. Update $\alpha_n^{(i+1)} := \alpha_n^\star, \forall n$, $\tilde{N}_k^{(i+1)} := \tilde{N}_k^\star$, and $\tilde{\gamma}_k^{(i+1)} := \tilde{\gamma}_k^\star, \forall k$.
5: Set $i = i + 1$.
6: **until** convergence.

We summarize the proposed iterative algorithm based on the BCA and SCA methods for solving (3) in Algorithm 1. In step 1, the initial points for $\{\boldsymbol{w}_k^{(0)}\}$, $\{\alpha_n^{(0)}\}$, $\{\tilde{N}_k^{(0)}\}$, and $\{\tilde{\gamma}_k^{(0)}\}$ are generated to guarantee that Algorithm 1 is successfully solved in the first iteration. In steps 2–6, subproblems (11) and (19) are alternatively solved and $\{\boldsymbol{w}_k^{(i)}\}$, $\{\alpha_n^{(i)}\}$, $\{\tilde{N}_k^{(i)}\}$, and $\{\tilde{\gamma}_k^{(i)}\}$ are updated after each iteration until the objective value τ converges. Note that convex problems (11) and (19) can be solved with standard optimization toolbox, such as CVX or YALMIP-MOSEK. The computational complexities required to solve subproblems (11) and (19) are $\mathcal{O}\left(\sqrt{2K+2}(KN_t+1)^3\right)$ and $\mathcal{O}\left(\sqrt{3K+N+1}(2K+N+1)^3\right)$, respectively. Thus, the overall complexity of Algorithm 1 is $\mathcal{O}\left(I(\sqrt{2K+2}(KN_t+1)^3+\sqrt{3K+N+1}(2K+N+1)^3)\right)$, where I is the number of iterations until convergence.

IV. Numerical Results

In this section, numerical results are provided to evaluate the effectiveness of Algorithm 1. We assume that the BS and RIS are deployed in a two-dimensional coordinate system at $(0, 0)$ and $(20, 0)$ m, respectively, while the UEs are randomly and uniformly distributed in a square area of 200×200 m^2. The Rayleigh fading model is considered for the direct BS-UEs channels, while those for the BS-RIS and RIS-UEs reflecting channels are Rician fading models with Rician factors of 100 and 10, respectively. The path loss for link distance d is given by $\beta(d) = \beta_0(d/1\text{m})^{-\epsilon}$, where β_0 is the path loss at the reference distance of 1 m, and $\epsilon \in \{\epsilon_0, \epsilon_1, \epsilon_2\}$ represents the path loss exponents of BS-UEs, BS-RIS, and RIS-UEs channels, respectively. We set $\beta_0 = -30$ dB and $\sigma_{\text{u}}^2 = -80$ dBm, $\{\epsilon_0, \epsilon_1, \epsilon_2\} = \{3.2, 2.2, 2.5\}$. The total power of noise and residual self-interference of the RIS is computed as $\sigma_{\text{r}}^2 = (\eta + 1)\sigma_{\text{u}}^2$, with $\eta = 1$ dB reflecting the possible residual self-interference caused by active elements operating in full-duplex mode [17], [29]. The positions of the RIS active elements are fixed to $\mathcal{A} = \{1, \ldots, N_{\text{a}}\}$. Towards a fair comparison, the power budget at the BS in the hybrid RIS-aided system is reduced by $p_{\text{RIS}}^{\max}$, so that all the compared schemes have the same total power budget as $p_{\text{BS}}^{\max}$.

We first show in Fig. 1 the convergence of Algorithm 1. For initialization, we set $\boldsymbol{w}_k^{(0)} = \sqrt{\frac{p_{\text{BS}}^{\max}}{K}}\frac{\boldsymbol{h}_k}{\|\boldsymbol{h}_k\|}, \forall k$, implying the conjugate beamforming with equal power. Furthermore,

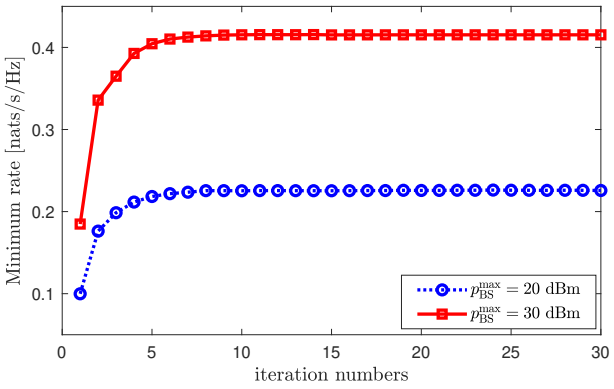

Fig. 1. Convergence of Algorithm 1 with $N_t = 2$, $K = 5$, $N = 50$, $p_{\text{BS}}^{\max} = \{20, 30\}$ dBm, $N_{\text{a}} = 4$ and $p_{\text{RIS}}^{\max} = 0$ dBm.

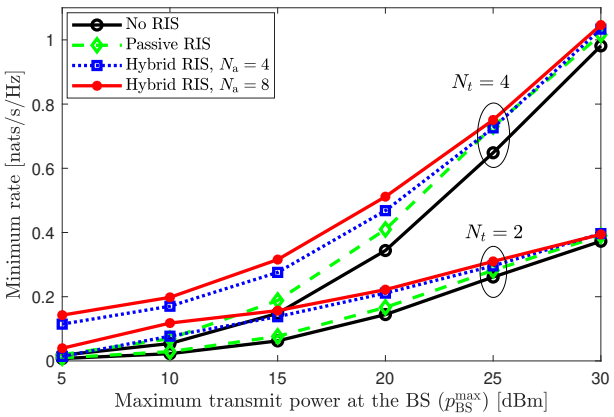

Fig. 2. Minimum rates versus $p_{\text{BS}}^{\max}$ with $N_t = \{2, 4\}$, $K = 5$, $N = 50$, $N_{\text{a}} = \{4, 8\}$, $p_{\text{BS}}^{\max} = [0, 30]$ dBm, and $p_{\text{RIS}}^{\max} = -1$ dBm.

$\{\alpha_n^{(0)}\}$ are initialized as $\{e^{j\theta_n^{(0)}}\}$ with initial phase shifts $\{\theta_n^{(0)}\}$ being randomly generated on $[0, 2\pi)$. These initial values belong to the feasible region of problem (3). For both cases $p_{\text{BS}}^{\max} = \{20, 30\}$ dBm, it is observed that Algorithm 1 converges after only a few iterations. With $p_{\text{BS}}^{\max} = 20$ dBm, the algorithm converges slightly faster, but obviously to a lower rate, compared with the case $p_{\text{BS}}^{\max} = 30$ dBm.

In Fig. 2, we plot the minimum rate versus the maximum transmit power of the BS with various deployments of RISs. As expected, the minimum rate performance of both passive and hybrid RISs increases significantly when $p_{\text{BS}}^{\max}$ increases for both $N_t = \{2, 4\}$. However, the performance gain of the former is marginal, especially at low SNRs. In contrast, the hybrid RIS with only $N_{\text{a}} = 4$ active elements can provide significant performance improvement to the system, and the gain is more remarkable at a low SNR regime. For example, with $N_t = 4$ and $p_{\text{BS}}^{\max} = 20$ dBm, the passive RIS achieves 27% improvement, while that attained by the hybrid RIS is up to 80% with the requirement of only $N_{\text{a}} = 4$ active elements and a power budget of $p_{\text{RIS}}^{\max} = -1$ dBm. Furthermore, the hybrid RIS with $N_{\text{a}} = 8$ provides only slightly better performance compared to the case $N_{\text{a}} = 4$, as will be further discussed below.

In Fig. 3, we show the minimum rate versus the transmit power budget at RISs, i.e., $p_{\text{RIS}}^{\max}$. Unsurprisingly, the perfor-

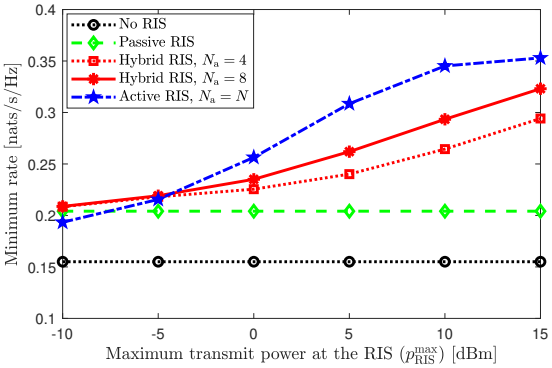

Fig. 3. Minimum rates versus $p_{\text{RIS}}^{\max}$ with $N_t = 2$, $K = 5$, $N = 50$, $N_{\text{a}} = \{4, 8, N\}$, $p_{\text{BS}}^{\max} = 20$ dBm, and $p_{\text{RIS}}^{\max} = [-10, 15]$ dBm.

mance improvement attained by the hybrid RIS significantly increases with $p_{\text{RIS}}^{\max}$. Furthermore, as $p_{\text{RIS}}^{\max}$ increases, the hybrid RIS with more active elements provides significantly larger performance gains. However, this is not valid for a low $p_{\text{RIS}}^{\max}$. It is observed that hybrid RISs with $N_{\text{a}} = 4$ and $N_{\text{a}} = 8$ offer almost the same performance for $p_{\text{RIS}}^{\max} \leq -5$ dBm, while the fully active RIS (with $N_{\text{a}} = N = 50$) performs even worse than the passive RIS. This is because when a limited power budget needs to be shared among a large number of active elements, the amplitudes of these elements become very small, causing signal attenuation on reflecting channels. Note that a RIS with a larger number of active elements also requires a higher power consumption [24]. These further explain the motivations of hybrid RISs, especially when the power budget allocated to RISs is limited.

V. Conclusion

We proposed the deployment of the hybrid active-passive RIS architecture to enhance the performance of multi-user MISO systems. In this hybrid RIS, a very few active elements are employed to enhance reflecting and amplifying gains of the RISs' elements. We proposed an iterative algorithm based on the BCA and SCA approaches to effectively solve the formulated max-min rate problem. Extensive numerical results were provided to verify the merits of the proposed algorithm. They revealed that the hybrid RIS offers remarkable performance improvement compared to existing schemes (i.e. without RIS and conventional passive RISs).

References

[1] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, 2019.

[2] 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, 2019.

[3] ——, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2019.

[4] J. Hu, Y.-C. Liang, and Y. Pei, "Reconfigurable intelligent surface enhanced multi-user MISO symbiotic radio system," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2359–2371, 2020.

[5] 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 Commun.*, vol. 38, no. 8, pp. 1809–1822, 2020.

[6] A. Kammoun, A. Chaaban, M. Debbah, M.-S. Alouini *et al.*, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7748–7764, 2020.

[7] H. Zhang, B. Di, Z. Han, H. V. Poor, and L. Song, "Reconfigurable intelligent surface assisted multi-user communications: How many reflective elements do we need?" *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1098–1102, 2021.

[8] A. Li, L. Song, B. Vucetic, and Y. Li, "Interference exploitation precoding for reconfigurable intelligent surface aided multi-user communications with direct links," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1937–1941, 2020.

[9] X. Gan, C. Zhong, C. Huang, and Z. Zhang, "Ris-assisted multi-user MISO communications exploiting statistical CSI," *IEEE Trans. Wireless Commun.*, vol. 69, no. 10, pp. 6781–6792, 2021.

[10] B. Di, H. Zhang, L. Li, L. Song, Y. Li, and Z. Han, "Practical hybrid beamforming with finite-resolution phase shifters for reconfigurable intelligent surface based multi-user communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4565–4570, 2020.

[11] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, 2020.

[12] X. Ma, S. Guo, H. Zhang, Y. Fang, and D. Yuan, "Joint beamforming and reflecting design in reconfigurable intelligent surface-aided multi-user communication systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3269–3283, 2021.

[13] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, "Energy efficient multi-user MISO communication using low resolution large intelligent surfaces," in *IEEE Global Commun. Conf. (GLOBECOM) Workshop*, 2018, pp. 1–6.

[14] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Deep learning for large intelligent surfaces in millimeter wave and massive MIMO systems," in *IEEE Global Commun. Conf. (GLOBECOM)*, 2019, pp. 1–6.

[15] G. C. Alexandropoulos and E. Vlachos, "A hardware architecture for reconfigurable intelligent surfaces with minimal active elements for explicit channel estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 2020, pp. 9175–9179.

[16] N. T. Nguyen, Q.-D. Vu, K. Lee, and M. Juntti, "Spectral efficiency optimization for hybrid relay-reflecting intelligent surface," *Proc. IEEE Int. Conf. Commun. Workshop*, 2021.

[17] ——, "Hybrid relay-reflecting intelligent surface-assisted wireless communication," *to appear in IEEE Trans. Veh. Technol.*, 2022.

[18] N. T. Nguyen *et al.*, "Hybrid relay-reflecting intelligent surface-aided wireless communications: Opportunities, challenges, and future perspectives," *arXiv preprint arXiv:2104.02039*, 2021.

[19] ——, "Downlink throughput of cell-free massive MIMO systems assisted by hybrid relay-reflecting intelligent surfaces," in *Proc. IEEE Int. Conf. Commun.*, 2022.

[20] P. Zeng, D. Qiao, Q. Wu, and Y. Wu, "Active IRS aided WPCNs: A new paradigm towards higher efficiency and wider coverage," *arXiv preprint arXiv:2111.11600*, 2021.

[21] K.-H. Ngo *et al.*, "Low-latency and secure computation offloading assisted by hybrid relay-reflecting intelligent surface," in *IEEE Int. Conf. Advanced Tech. Commun. (ATC)*, 2021, pp. 306–311.

[22] S. Ahmed, A. E. Kamal, and M. Y. Selim, "Adding active elements to reconfigurable intelligent surfaces to enhance energy harvesting for IoT devices," in *IEEE Military Commun. Conf.*, 2021, pp. 297–302.

[23] Z. Yigit, E. Basar, M. Wen, and I. Altunbas, "Hybrid reflection modulation," *CoRR*, vol. abs/2111.08355, 2021.

[24] R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson, "Active reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Wireless Commun.*, 2021.

[25] M. H. Khoshafa, T. M. Ngatched, M. H. Ahmed, and A. R. Ndjiongue, "Active reconfigurable intelligent surfaces-aided wireless communication system," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3699–3703, 2021.

[26] G. Chen, Q. Wu, C. He, W. Chen, J. Tang, and S. Jin, "Active IRS aided multiple access for energy-constrained IoT systems," *arXiv preprint arXiv:2201.12565*, 2022.

[27] R. Schroeder, J. He, and M. Juntti, "Passive RIS vs. Hybrid RIS: A comparative study on channel estimation," in *Proc. IEEE Veh. Technol. Conf.*, June 2021, pp. 1–7.

[28] N. Landsberg and E. Socher, "A low-power 28-nm CMOS FD-SOI reflection amplifier for an active F-band reflectarray," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 10, pp. 3910–3921, 2017.

[29] R. Malik and M. Vu, "Optimal transmission using a self-sustained relay in a full-duplex MIMO system," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 2, pp. 374–390, 2018.