

Active RIS vs. Passive RIS: Which Will Prevail in 6G?

Zijian Zhang, *Student Member, IEEE*, Linglong Dai, *Fellow, IEEE*, Xibi Chen, *Student Member, IEEE*, Changhao Liu, *Student Member, IEEE*, Fan Yang, *Fellow, IEEE*, Robert Schober, *Fellow, IEEE*, and H. Vincent Poor, *Life Fellow, IEEE*

Abstract—As a revolutionary paradigm for controlling wireless channels, reconfigurable intelligent surfaces (RISs) have emerged as a candidate technology for future 6G networks. However, due to the “multiplicative fading” effect, the existing passive RISs only achieve limited capacity gains in many scenarios with strong direct links. In this paper, the concept of active RISs is proposed to overcome this fundamental limitation. Unlike passive RISs that reflect signals without amplification, active RISs can amplify the reflected signals via amplifiers integrated into their elements. To characterize the signal amplification and incorporate the noise introduced by the active components, we develop and verify the signal model of active RISs through the experimental measurements based on a fabricated active RIS element. Based on the verified signal model, we further analyze the asymptotic performance of active RISs to reveal the substantial capacity gain they provide for wireless communications. Finally, we formulate the sum-rate maximization problem for an active RIS aided multi-user multiple-input single-output (MU-MISO) system and a joint transmit beamforming and reflect precoding scheme is proposed to solve this problem. Simulation results show that, in a typical wireless system, passive RISs can realize only a limited sum-rate gain of 22%, while active RISs can achieve a significant sum-rate gain of 130%, thus overcoming the “multiplicative fading” effect.

Index Terms—Reconfigurable intelligent surface (RIS), beamforming, active RIS, signal model.

Manuscript received 6 September, 2022; revised 19 November, 2022; accepted 19 December, 2022. Date of publication XXX XX, 2022; date of current version XXX XX, 2022. This work was supported in part by the National Key Research and Development Program of China (Grant No. 2020YFB1807201), in part by the National Natural Science Foundation of China (Grant No. 62031019), in part by the European Commission through the H2020-MSCA-ITN META WIRELESS Research Project under Grant 956256, and in part by the U.S National Science Foundation under Grants CCF-1908308 and CNS-2128448. R. Schober’s work was partly supported by the Federal Ministry of Education and Research of Germany under the programme of “Souveran. Digital. Vernetzt.” joint project 6G-RIC (project identification number: PIN 16KISK023) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant SCHO 831/15-1. An earlier version of this paper was presented in part at the IEEE GLOBECOM’22, Rio de Janeiro, Brazil, December 2022 [1]. The associate editor coordinating the review of this article and approving it for publication was A. García Armada. (*Corresponding author: Linglong Dai*)

Zijian Zhang, Linglong Dai, Xibi Chen, Changhao Liu, and Fan Yang are with the Department of Electronic Engineering as well as the Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing 100084, China (e-mail: zhangzj20@mails.tsinghua.edu.cn, daill@mails.tsinghua.edu.cn, cxb17@mails.tsinghua.org.cn, liu-ch21@mails.tsinghua.edu.cn, fan_yang@mails.tsinghua.edu.cn).

Robert Schober is with the Institute for Digital Communications at Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91054 Erlangen, Germany (e-mail: robert.schober@fau.de).

H. Vincent Poor is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA (e-mail: poor@princeton.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCOMM.2022.3231893>.

Digital Object Identifier 10.1109/TCOMM.2022.3231893

I. INTRODUCTION

As wireless communications have advanced from the first generation (1G) to 5G, the system capacity has been significantly increased by improving the transceiver designs, while the wireless channel has been considered to be uncontrollable. Recently, due to the advances in meta-materials, reconfigurable intelligent surfaces (RISs) have been proposed [2]–[4] for the purpose of intelligently controlling wireless channels to achieve improved communication performance. Specifically, an RIS is an array composed of a very large number of passive elements that reflects electromagnetic signals in a desired manner so as to reconfigure the propagation properties of wireless environment [5]. Thanks to their high array gain, low cost, low power, and negligible noise [5]–[7], RISs promise to improve channel capacity [8], extend coverage [9], and save power [10] in future 6G networks. Additionally, RISs are also projected to have other applications such as in WiFi [11], precision measurement [12], and navigation [13].

As an important advantage of RISs, the negligible noise introduced by passive RISs enables a high array gain. Particularly, in a RIS aided single-user single-input single-output (SU-SISO) system, the achievable signal-to-noise ratio (SNR) gain enabled by an N -element RIS is proportional to N^2 [14]. Benefiting from this advantage, RISs are expected to introduce significant capacity gains in wireless systems [8]. However, in practice, these capacity gains are typically only observed in communication scenarios where the direct link between transmitter and receiver is completely blocked or very weak [8]–[10], [15]–[17]. By contrast, in many scenarios where the direct link is not weak, conventional RISs achieve limited capacity gains [18]. The reason behind this phenomenon is the “multiplicative fading” effect introduced by RISs, i.e., the equivalent path loss of the transmitter-RIS-receiver link is the product (instead of the sum) of the path losses of the transmitter-RIS and RIS-receiver links, which is usually thousands of times larger than that of the direct link [18]. As a result, the “multiplicative fading” effect makes it almost impossible for passive RISs to achieve noticeable capacity gains in many wireless environments. Many existing works on RISs have bypassed this effect by only considering scenarios with severely obstructed direct links [8]–[10], [15]–[17]. Therefore, to advance the practicability of RISs in future 6G wireless networks, a critical issue for RISs to be addressed is: *How to break the fundamental performance bottleneck caused by the “multiplicative fading” effect?*

To overcome the fundamental physical limitation, in this paper, a new RIS architecture called *active RISs* is proposed

for wireless communication systems. Specifically, different from *passive* RISs that passively reflect signals without amplification, the key feature of active RISs is their ability to actively reflect signals with amplification, which can be realized by integrating reflection-type amplifiers into their reflecting elements. At the expense of additional power consumption, active RIS can compensate for the large path loss of reflected links, which is promising to overcome the “multiplicative fading” effect. Some parallel works¹ have revealed the potential benefits of active RISs for enhancing wireless systems. For example, in [19], an active RIS was introduced into a single-user single-input multi-output (SU-SIMO) system to enhance the user’s SNR. In [20], the authors considered an active RIS aided SU-SISO system with limited RIS elements, and the impact of RIS location placement on the communication performance is analyzed.

In this paper², we propose the concept of active RISs and focuses on the signal model verification, asymptotic performance analysis, and multi-user beamforming design of active RISs aided communication systems. Specifically, our contributions are summarized as follows:

- We develop and verify the signal model of active RISs, which characterizes the amplification of the incident signal and accounts for the non-negligible thermal noise introduced by the active elements. Particularly, the verification is made via the experimental measurements based on a designed and fabricated active RIS element.
- Based on the verified signal model, we analyze the asymptotic performance of an active RIS with extremely large number of active elements, and we further compare it to that of the existing passive RISs, which reveals the notable capacity gain enabled by the use of active RISs.
- To evaluate the performance of active RISs in typical communication systems, we formulate a sum-rate maximization problem for an active RIS aided multi-user multiple-input single-output (MU-MISO) system. Then, by exploiting fractional programming (FP), a joint transmit beamforming and reflect precoding scheme is proposed to solve this problem.
- To account for the non-ideal factors in practical systems, we extend the studied beamforming design in the scenario with the self-interference of active RISs. We model the feedback-type self-interference of active RISs, which allows us to formulate an mean-squared error minimization problem to suppress the self-interference. Then, by utilizing alternating direction method of multipliers (ADMM) [22] and sequential unconstrained minimization techniques (SUMT) [23], an alternating optimization scheme is proposed to solve the formulated problem.

The rest of this paper is organized as follows. In Section II, the architectures as well as the signal models of passive

¹In October 2019, we started to design an active RIS element integrating a reflection-type amplifier [21]. The fabrication of this active RIS element was finished in August 2020. Subsequently, we set out to establish an experimental environment for signal measurements with this element, and all measurements were completed in February 2021. This paper first appeared on ArXiv in March 2021.

²Simulation codes are provided to reproduce the results presented in this article: <http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html>.

RISs and active RISs are introduced, respectively. In Section III, the asymptotic performance of active RISs is analyzed and compared to that of the passive RISs. In Section IV, a sum-rate maximization problem is formulated for an active RIS aided MU-MISO system, and a joint beamforming and precoding design is proposed to solve the formulated problem. In Section V, we extend the studied joint beamforming and precoding design to the practical case with self-interference. In Section VI, the convergence and complexity of the proposed schemes are analyzed. In Section VII, validation results are presented to validate the signal model of active RISs and evaluate the performance of active RISs in typical communication scenarios. Finally, conclusions are drawn and future works are discussed in Section VIII.

Notations: \mathbb{C} , \mathbb{R} , and \mathbb{R}_+ denote the sets of complex, real, and positive real numbers, respectively; $[\cdot]^{-1}$, $[\cdot]^*$, $[\cdot]^T$, and $[\cdot]^H$ denote the inverse, conjugate, transpose, and conjugate-transpose operations, respectively; $\|\cdot\|$ denotes the Euclidean norm of the argument; $\|\cdot\|_F$ denotes the Frobenius norm of the argument; $\text{diag}(\cdot)$ denotes the diagonalization operation; $\Re\{\cdot\}$ denotes the real part of the argument; \otimes denotes the Kronecker product; $\angle[\cdot]$ denotes the angle of the complex argument; $\ln(\cdot)$ denotes the natural logarithm of its argument; $\mathcal{CN}(\mu, \Sigma)$ denotes the complex multivariate Gaussian distribution with mean μ and variance Σ ; \mathbf{I}_L is an $L \times L$ identity matrix, and $\mathbf{0}_L$ is an $L \times 1$ zero vector.

II. PASSIVE RIS AND ACTIVE RIS

In this section, we introduce the architectures of RISs. First, in Subsection II-A, we review passive RISs and point out their physical limitation imposed by the “multiplicative fading” effect. Then, in Subsection II-B, to overcome this limitation, we propose the novel concept of active RISs along with their hardware structure and signal model. Finally, in Subsection II-C, we present the transmission model for an active RIS aided MU-MISO system.

A. Passive RIS

The RISs widely studied in most existing works are passive RISs [2]–[10]. Specifically, as shown in Fig. 1 (a), a passive RIS comprises a large number of passive elements each being able to reflect the incident signal with a controllable phase shift. In general, each passive RIS element consists of a reflective patch terminated with an impedance-adjustable circuit for phase shifting [24]. Thanks to the passive mode of operation without active radio-frequency (RF) components, a passive RIS element practically consumes zero direct-current power [24], and the introduced thermal noise is usually negligible [5]–[10]. Thereby, the signal model of an N -element passive RIS widely used in the literature is given as follows [6]:

$$\mathbf{y} = \Theta \mathbf{x}, \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^N$ denotes the incident signal, $\Theta := \text{diag}(e^{j\theta_1}, \dots, e^{j\theta_N}) \in \mathbb{C}^{N \times N}$ denotes the reflection coefficient matrix of the passive RIS with θ_n being the phase shift of the n -th passive element, and $\mathbf{y} \in \mathbb{C}^N$ denotes the signal reflected by the RIS. Note that the impact of noise is

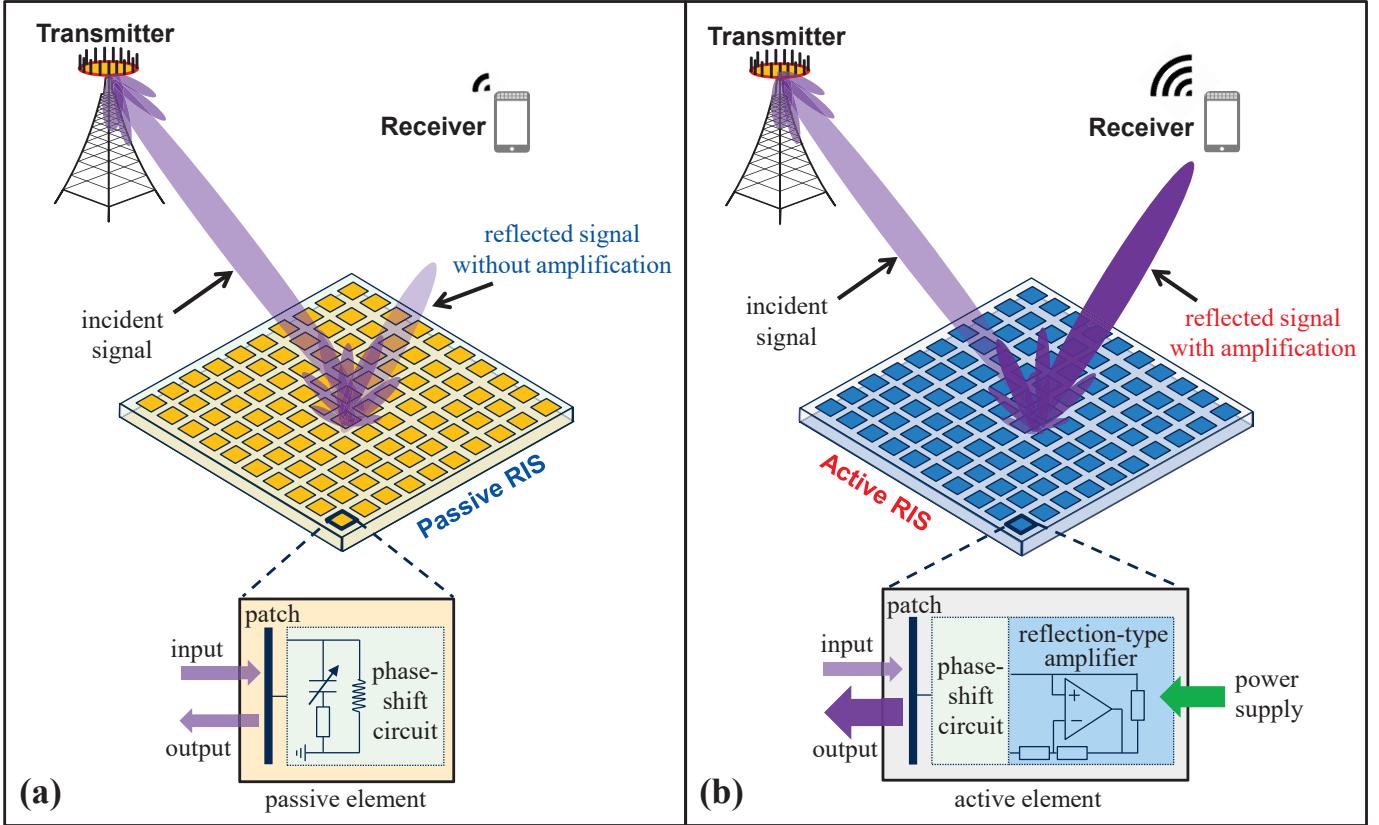


Fig. 1. An illustration of the hardware architectures of (a) a passive RIS and (b) an active RIS.

neglected in (1). As a consequence, by properly adjusting Θ to manipulate the N signals reflected by the N RIS elements to coherently add with the same phase at the receiver, a high array gain can be achieved. This is expected to significantly increase the receiver SNR [5]–[7], which is one of the key reasons for why RISs have attracted so much research interest recently [8]–[13], [15]–[17].

Unfortunately, in practice, this expected high capacity gain often cannot be realized, especially in communication scenarios where the direct link between the transmitter and the receiver is not weak. The reason for this negative result is the “multiplicative fading” effect introduced by passive RISs. Specifically, the equivalent path loss of the transmitter-RIS-receiver reflected link is the product (instead of the sum) of the path losses of the transmitter-RIS and RIS-receiver links, and therefore, it is thousands of times larger than that of the unobstructed direct link. Thereby, for an RIS to realize a noticeable capacity gain, thousands of RIS elements are required to compensate for this extremely large path loss.

Remark 1: To illustrate the above fact, let us consider an SU-SISO system aided by a passive RIS. Assume that the transceiver antennas are omnidirectional and RIS elements are tightly deployed with half-wavelength spacing [18]. Let $d = 200$ m, $d_t = 110$ m, and $d_r = 110$ m denote the distances between transmitter and receiver, transmitter and RIS, RIS and receiver, respectively. Assume that all channels are line-of-sight (LoS) and the RIS phase shift is optimally configured to maximize the channel gain of the transmitter-RIS-receiver

reflected link. Then, for carrier frequencies of 5/10/28 GHz, $N = \frac{4}{\lambda} d_t d_r = 4034/8067/22587$ RIS elements are required to make the reflected link as strong as the direct link [18]. The high signaling overhead introduced by the N pilots required for channel estimation [25] and the high complexity of $\mathcal{O}(N^2)$ for real-time beamforming [26] make the application of such a large number of passive RIS elements in practical wireless networks very challenging [18]. Consequently, many existing works have bypassed the “multiplicative fading” effect by only considering the scenario where the direct link is completely blocked or very weak [5]–[10], [15]–[17].

B. Active RISs

To overcome the fundamental performance bottleneck caused by the “multiplicative fading” effect of RISs, in this paper, we propose active RISs as a promising solution. As shown in Fig. 1 (b), similar to passive RISs, active RISs can also reflect the incident signals with reconfigurable phase shifts. Different from passive RISs that just reflect the incident signals without amplification, active RISs can further amplify the reflected signals. To achieve this goal, the key component of an active RIS element is the additionally integrated active reflection-type amplifier, which can be realized by different existing active components, such current-inverting converters [27], asymmetric current mirrors [28], or some integrated circuits [29].

With reflection-type amplifiers supported by a power supply, the reflected and amplified signal of an N -element active RIS can be modeled as follows:

$$\mathbf{y} = \underbrace{\Psi \mathbf{x}}_{\text{Desired signal}} + \underbrace{\Psi \mathbf{v}}_{\text{Dynamic noise}} + \underbrace{\mathbf{n}_s}_{\text{Static noise}}, \quad (2)$$

where $\Psi := \text{diag}(p_1 e^{j\theta_1}, \dots, p_N e^{j\theta_N}) \in \mathbb{C}^{N \times N}$ denotes the reflection coefficient matrix of the active RIS, wherein $p_n \in \mathbb{R}_+$ denotes the amplification factor of the n -th active element and p_n can be larger than one thanks to the integrated reflection-type amplifier. Due to the use of active components, active RISs consume additional power for amplifying the reflected signals, and the thermal noise introduced by active RIS elements cannot be neglected as is done for passive RISs. Particularly, as shown in (2), the noise introduced at active RISs can be classified into dynamic noise $\Psi \mathbf{v}$ and static noise \mathbf{n}_s , where $\Psi \mathbf{v}$ is the noise introduced and amplified by the reflection-type amplifier and \mathbf{n}_s is generated by the patch and the phase-shift circuit [28]. More specifically, \mathbf{v} is related to the input noise and the inherent device noise of the active RIS elements [28], while the static noise \mathbf{n}_s is unrelated to Ψ and is usually negligible compared to the dynamic noise $\Psi \mathbf{v}$, as will be verified by experimental results in Section VII-A. Thus, here we neglect \mathbf{n}_s and model \mathbf{v} as $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}_N, \sigma_v^2 \mathbf{I}_N)$.

Remark 2: Note that active RISs are fundamentally different from the relay-type RISs equipped with RF components [30]–[32] and relays [33]. Specifically, in [30]–[32], a subset of the passive RIS elements are connected to active RF chains, which are used for sending pilot signals and processing baseband signals. Thus, these relay-type RIS elements have signal processing capabilities [30]–[32]. On the contrary, active RISs do not have such capabilities but only reflect and amplify the incident signals to strengthen the reflected link. Besides, although active RISs can amplify the incident signals, similar to full-duplex amplify-and-forward (FD-AF) relays, their respective hardware architectures and transmission models are quite different. Specifically, an FD-AF relay is equipped with RF chains to receive the incident signal and then transmit it after amplification [33]. Due to the long delay inherent to this process, two timeslots are needed to complete the transmission of one symbol, and the received signal at the receiver in a timeslot actually depends on two different symbols, which were transmitted by the transmitter and the FD-AF relay, respectively [33]. As a consequence, in order to efficiently decode the symbols, the receiver in an FD-AF relay aided system has to combine the signals received in two successive timeslots to maximize the SNR. Thus, the transmission model for FD-AF relaying [33, Eq. (22), Eq. (25)] differs substantially from that for active RIS (3), which also leads to different achievable rates [33, Table I].

C. Active RIS Aided MU-MISO System

To characterize the performance gains enabled by our proposed active RISs in typical communication scenarios, we consider an active RIS aided downlink MU-MISO system as shown in Fig. 2, where an M -antenna base station (BS) serves K single-antenna users simultaneously with the aid of an N -element active RIS.

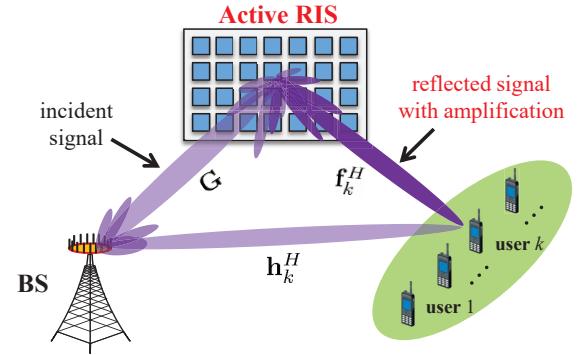


Fig. 2. An illustration of the downlink transmission in an active RIS aided MU-MISO system.

Let $\mathbf{s} := [s_1, \dots, s_K]^T \in \mathbb{C}^K$ denote the transmitted symbol vector for the K users with $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_K$. We assume that multi-user linear precoding is employed at the BS for downlink transmission. Then, according to (2), signal $r_k \in \mathbb{C}$ received at user k can be modeled as follows:

$$r_k = (\underbrace{\mathbf{h}_k^H}_{\text{Direct link}} + \underbrace{\mathbf{f}_k^H \Psi \mathbf{G}}_{\text{Reflected link}}) \sum_{j=1}^K \mathbf{w}_j s_j + \underbrace{\mathbf{f}_k^H \Psi \mathbf{v}}_{\text{Noise introduced by active RIS}} + \underbrace{z_k}_{\text{Noise introduced at user } k}, \quad (3)$$

where $\mathbf{G} \in \mathbb{C}^{N \times M}$, $\mathbf{h}_k^H \in \mathbb{C}^{1 \times M}$ and $\mathbf{f}_k^H \in \mathbb{C}^{1 \times N}$ denote the channel vector between the BS and the RIS, that between the BS and user k , and that between the RIS and user k , respectively; $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ denotes the BS beamforming vector for symbol s_k ; and z_k denotes the additive white Gaussian noise (AWGN) at user k with zero mean and variance σ^2 .

To analytically illustrate how active RISs can overcome the “multiplicative fading” effect, based on the signal model in (2), the performance gain enabled by the use of active RISs will be studied in the next section.

III. PERFORMANCE ANALYSIS

In this section, we analyze the asymptotic performance of active RISs to reveal their notable capacity gains. To this end, in order to make the problem analytically tractable and get insightful results, similar to [14], we consider a SU-SISO system with $M = 1$ BS antenna and $K = 1$ user, while the general MU-MISO case is studied in Section IV.

A. Asymptotic SNR for Passive RISs and Active RISs

To illustrate the capacity gain provided by passive/active RIS aided reflected links, for the moment, we ignore the direct link by setting \mathbf{h}_k to zero, as was done in, e.g., [14]. Furthermore, to obtain analytical results and find more insights, we assume that each active RIS element has the same amplification factor (i.e., $p_n := p$ for all $n \in \{1, \dots, N\}$), while the power allocation among active elements will be considered in Section IV. For a fair comparison with the asymptotic performance of passive RISs, similar to [14], we assume Rayleigh-fading channels.

For the above RIS aided SU-SISO system without direct link, we first redefine the BS-RIS channel matrix and the RIS-user channel vector as $\mathbf{G} := \mathbf{g} \in \mathbb{C}^{N \times 1}$ and $\mathbf{f}_k := \mathbf{f} \in \mathbb{C}^{N \times 1}$, respectively, to simplify the notations. Then, we recall the following lemma from [14] for the asymptotic SNR achieved by passive RISs.

Lemma 1 (Asymptotic SNR for passive RISs): Assuming $\mathbf{f} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_f^2 \mathbf{I}_N)$, $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_g^2 \mathbf{I}_N)$ and letting $N \rightarrow \infty$, the asymptotic SNR γ_{passive} of a passive RIS aided SU-SISO system is given by

$$\gamma_{\text{passive}} \rightarrow N^2 \frac{P_{\text{BS}}^{\max} \pi^2 \varrho_f^2 \varrho_g^2}{16\sigma^2}, \quad (4)$$

where P_{BS}^{\max} denotes the maximum transmit power at the BS.

Proof: The proof can be found in [14, Proposition 2]. ■

For comparison, under the same transmission conditions, we provide the asymptotic SNR of an active RIS aided SU-SISO system in the following lemma.

Lemma 2 (Asymptotic SNR for active RISs): Assuming $\mathbf{f} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_f^2 \mathbf{I}_N)$, $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_g^2 \mathbf{I}_N)$ and letting $N \rightarrow \infty$, the asymptotic SNR γ_{active} of an active RIS aided SU-SISO system is given by

$$\gamma_{\text{active}} \rightarrow N \frac{P_{\text{BS}}^{\max} P_{\text{A}}^{\max} \pi^2 \varrho_f^2 \varrho_g^2}{16 \left(P_{\text{A}}^{\max} \sigma_v^2 \varrho_f^2 + P_{\text{BS}}^{\max} \sigma^2 \varrho_g^2 + \sigma^2 \sigma_v^2 \right)}, \quad (5)$$

where P_{A}^{\max} denotes the maximum reflect power of the active RIS.

Proof: Please see Appendix A. ■

Remark 3: From (5) we observe that the asymptotic SNR of an active RIS aided SU-SISO system depends on both the BS transmit power P_{BS}^{\max} and the reflect power of the active RIS P_{A}^{\max} . When $P_{\text{BS}}^{\max} \rightarrow \infty$, it can be proved that the asymptotic SNR of the active RIS aided system will be upper-bounded by $\gamma_{\text{active}} \rightarrow N \frac{P_{\text{A}}^{\max} \pi^2 \varrho_f^2}{16\sigma^2}$, which only depends on the RIS-user channel gain ϱ_f^2 and the noise power at the user σ^2 . This indicates that, when the BS transmit power is high enough, the BS-RIS channel \mathbf{g} and the noise power at the active RIS have negligible impact on the user's SNR. Similarly, if $P_{\text{A}}^{\max} \rightarrow \infty$, the asymptotic SNR γ_{active} in (5) will be upper-bounded by $\gamma_{\text{active}} \rightarrow N \frac{P_{\text{BS}}^{\max} \pi^2 \varrho_g^2}{16\sigma_v^2}$. Compared with (5), this upper bound is independent of the RIS-user channel \mathbf{f} and the noise power at the user σ^2 . It indicates that, the negative impact of small \mathbf{f} and large σ^2 can be reduced by increasing the reflect power of the active RIS P_{A}^{\max} , which may provide guidance for the design of practical active RIS-aided systems.

Next, we compare the asymptotic SNRs for passive RISs in *Lemma 1* and active RISs in *Lemma 2* to reveal the superiority of active RISs in wireless communications.

B. Comparisons between Passive RISs and Active RISs

We can observe from *Lemma 1* and *Lemma 2* that, compared to the asymptotic SNR for passive RISs γ_{passive} in (4) which is proportional to N^2 , the asymptotic SNR for active RISs γ_{active} in (5) is proportional to N due to the noises additionally introduced by the use of active components. At first glance, it seems that the SNR proportional to N^2 achieved by passive

RISs γ_{passive} always exceeds the SNR proportional to N achieved by active RISs γ_{active} . However, this is actually not the case in many scenarios.

The reason behind this counterintuitive behavior is that, different from the denominator of (4) which depends on the noise power σ^2 , the denominator of (5) is determined by the much smaller multiplicative terms composed of path losses and noise power, i.e., $P_{\text{A}}^{\max} \sigma_v^2 \varrho_f^2$, $P_{\text{BS}}^{\max} \sigma^2 \varrho_g^2$, and $\sigma^2 \sigma_v^2$. In this case, the denominator of (5) is usually much smaller than that of (4). Thus, even if the numerator of (5) is smaller than that of (4) because of an N gain loss, the SNR gain of active RISs can still be higher than that of passive RISs in many scenarios.

Generally, due to the much smaller denominator of (5), only when N is unaffordably large can passive RISs outperform active RISs. To illustrate this claim, let us consider two different SU-SISO systems, which are aided by an active RIS and a passive RIS, respectively. Then, the following lemma specifies the condition that has to be met for passive RISs to outperform active RISs.

Lemma 3 (Case when passive RISs outperform active RISs): Assuming the number of RIS elements N is large, the required number of elements N for a passive RIS to outperform an active RIS has to satisfy

$$N \geq \frac{P_{\text{BS-A}}^{\max}}{P_{\text{BS-P}}^{\max}} \frac{P_{\text{A}}^{\max} \sigma^2}{\left(P_{\text{A}}^{\max} \sigma_v^2 \varrho_f^2 + P_{\text{BS-A}}^{\max} \sigma^2 \varrho_g^2 + \sigma^2 \sigma_v^2 \right)}, \quad (6)$$

where $P_{\text{BS-A}}^{\max}$ denotes the maximum BS transmit power for the active RIS aided system and $P_{\text{BS-P}}^{\max}$ denotes that for the passive RIS aided system.

Proof: Please see Appendix B. ■

Next, we consider a specific setup to compare the user's achievable SNRs in the above two systems. For a fair comparison, we constrain the total power consumption P^{\max} of the two systems to 2 W by setting $P_{\text{BS-P}}^{\max} = 2$ W for the passive RIS aided system and $P_{\text{BS-A}}^{\max} = P_{\text{A}}^{\max} = 1$ W for the active RIS aided system, respectively. Therefore, when $\sigma^2 = \sigma_v^2 = -100$ dBm and $\varrho_f^2 = \varrho_g^2 = -70$ dB, the required number of elements N for the passive RIS to outperform the active RIS is 2.5×10^6 according to (6), which is impractical to realize with current technology. Besides, the high overhead for channel estimation [25] and the high complexity for real-time beamforming [26] also make the application of such a large number of RIS elements impractical [18]. Conversely, for a more practical number of elements of $N = 256$, according to (5) and (4), the SNR achieved by the passive RIS is $\gamma_{\text{passive}} \approx 39.0$ dB, while the SNR achieved by the active RIS is $\gamma_{\text{active}} \approx 79.0$ dB, which is about 10,000 times higher than γ_{passive} .

Based on the above parameters, we show the asymptotic SNR versus the number of RIS elements N for both passive RISs and active RISs in Fig. 3, where N ranges from 10 to 1000 in Fig. 3 (a) and from 10^4 to 3×10^6 in Fig. 3 (b). From this figure we can observe that, when N ranges from 10 to 1000, the user's achievable SNR is about 40 dB higher in an active RIS aided system compared to a passive RIS aided system. Only when $N = 2.5 \times 10^6$ becomes the performance gain achieved by the passive RIS comparable to that achieved

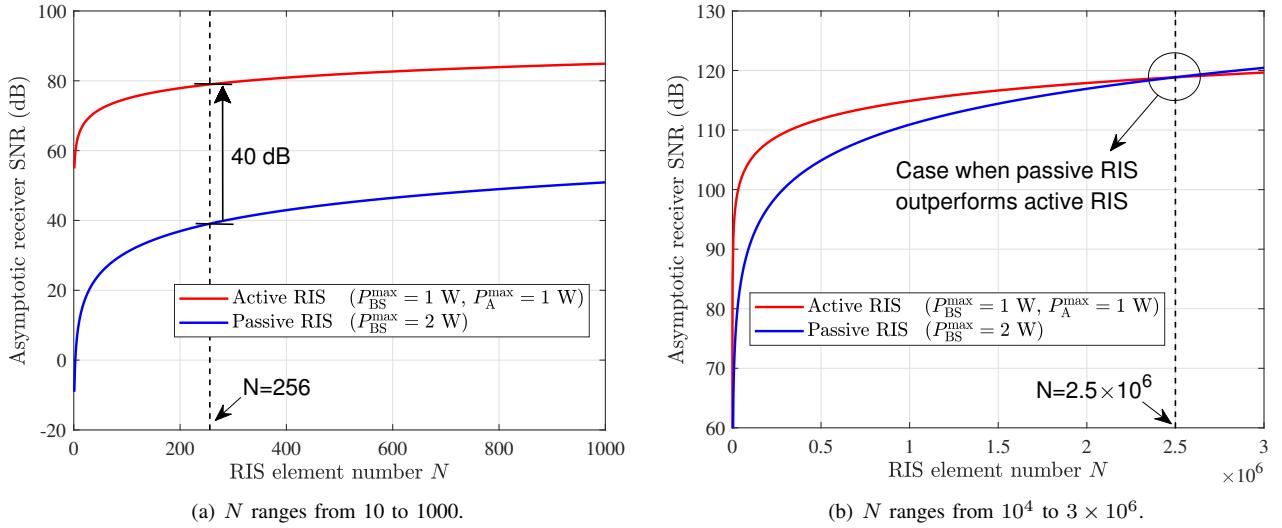


Fig. 3. Asymptotic SNR as a function of the number of RIS elements N for different ranges of N .

by the active RIS, which agrees well with our above analysis.

Remark 4: From the above comparisons we find that, although additional thermal noise is introduced by the active components, active RISs can still achieve a higher SNR gain than passive RISs. This is due to the fact that, the desired signals reflected by different active RIS elements are coherently added with the same phase at the user, while the introduced noises are not. Besides, when these introduced noises are received by the user, they have become much small due to the RIS-user path loss. In addition, different from the passive RIS aided system that all radiation power suffers from the multiplicative path loss of reflected links, the power radiated by active RISs only experiences the large fading of RIS-user link, thus the power attenuation is much less and the “multiplicative fading” effect can be well overcome.

C. Impact of Distances on RIS Performances

According to (4) and (5), the path losses of the wireless links are the key parameters influencing the RIS performances. Since the path losses highly depend on the inter-device distances, in this section, we analyze the impact of distances on the SNR gain of active RISs and passive RISs.

To characterize the relationship between distances and path losses, without loss of generality, we assume that the large-scaling fading of BS-RIS channel \mathbf{g} and RIS-user channel \mathbf{f} follow the far-field spherical-wave propagation model, which is widely used in existing works such as [14], [26]. Thus, the BS-RIS path loss ϱ_g^2 and the RIS-user path loss ϱ_f^2 can be rewritten as:

$$\varrho_g^2 = L_0 d_t^{-\alpha} \quad \text{and} \quad \varrho_f^2 = L_0 d_r^{-\beta}, \quad (7)$$

where L_0 is the path loss at the reference distance of 1 m, which is usually set to $L_0 = -30 \text{ dB}$ [14]; d_t and d_r denotes the BS-RIS distance and the RIS-user distance, respectively; α and β denote the path loss exponents of BS-RIS channel and RIS-user channel, respectively, whose values usually range from 2 to 4. To find more insights, here we assume that $\sigma^2 =$

σ_v^2 , $P_{BS-P}^{\max} = P^{\max}$, and $P_{BS-A}^{\max} = P_A^{\max} = P^{\max}/2$, wherein P^{\max} denotes the total radiation power. Then, we obtain the following lemma.

Lemma 4 (Scenario where passive RISs outperform active RISs): Given a large number of RIS elements N , the scenario where an active RIS can outperform a passive RIS should satisfy

$$\frac{1}{d_t^{-\alpha} + d_r^{-\beta}} \leq \frac{2NP^{\max}L_0}{P^{\max} - 4N\sigma^2}. \quad (8)$$

Proof: Substitute (7) into (42) in Appendix B and then solve $\gamma_{\text{passive}} \geq \gamma_{\text{active}}$ by exploiting $\frac{1}{N} \sum_{n=1}^N |f_n|^2 \approx \varrho_f^2$ and $\frac{1}{N} \sum_{n=1}^N |g_n|^2 \approx \varrho_g^2$. This completes the proof. ■

From (8) one can notice that, active RISs can outperform passive RISs in many scenarios. The reason is that, distances d_t and d_r are usually large, which makes the left part of (8) very large. By contrast, due to the large path loss $L_0 = -30 \text{ dB}$, the right part of (8) is usually small, which results in the fact that the inequality (8) follows in many practical scenarios. To see the above fact, here we fix the BS-RIS distance as $d_t = 20 \text{ m}$ and consider the following parameters: $L_0 = -30 \text{ dB}$, $\alpha = \beta = 2$, $P^{\max} = 2 \text{ W}$, $\sigma^2 = -100 \text{ dBm}$, and $N = 1024$. Then, we can calculate from (8) that, active RISs can outperform passive RISs as long as the RIS-user distance d_r satisfy $d_r \geq \left(\frac{P^{\max} - 4N\sigma^2}{2NP^{\max}L_0} - d_t^{-\alpha} \right)^{-\frac{1}{\beta}} = 1.43 \text{ m}$, which nearly covers the whole wireless communication region. In other words, to achieve the same performance, active RISs can be located much far away from terminals compared to passive RISs, which is one more advantage of using active RISs.

IV. JOINT TRANSMIT BEAMFORMING AND REFLECT PRECODING DESIGN

To investigate the capacity gain enabled by active RISs in typical communication scenarios, in this section, we consider a more general active RIS aided MU-MISO system. Specifically,

in Subsection IV-A, we formulate the problem of sum-rate maximization. Then, in Subsection IV-B, a joint transmit beamforming and reflect precoding scheme is proposed to solve the problem.

A. Sum-Rate Maximization Problem Formulation

According to the MU-MISO transmission model in (3), the signal-to-interference-plus-noise ratio (SINR) at user k can be obtained as

$$\gamma_k = \frac{|\bar{\mathbf{h}}_k^H \mathbf{w}_k|^2}{\sum_{j=1, j \neq k}^K |\bar{\mathbf{h}}_k^H \mathbf{w}_j|^2 + \|\mathbf{f}_k^H \Psi\|^2 \sigma_v^2 + \sigma^2}, \quad (9)$$

wherein $\bar{\mathbf{h}}_k^H = \mathbf{h}_k^H + \mathbf{f}_k^H \Psi \mathbf{G} \in \mathbb{C}^{1 \times M}$ is the equivalent channel from the BS to user k , which includes both the direct link and the reflected link. By solving the expectation of the squared Euclidean norm of the radiated signals, the BS transmit power, P_{BS} , and the reflect power of the active RIS, P_A , can be respectively derived as

$$P_{\text{BS}} = \mathbb{E} \left\{ \left\| \sum_{k=1}^K \mathbf{w}_k s_k \right\|^2 \right\} = \underbrace{\sum_{k=1}^K \|\mathbf{w}_k\|^2}_{\text{Desired signal power}}, \quad (10a)$$

$$\begin{aligned} P_A &= \mathbb{E} \left\{ \left\| \Psi \mathbf{G} \sum_{k=1}^K \mathbf{w}_k s_k + \Psi \mathbf{v} \right\|^2 \right\} \\ &= \underbrace{\sum_{k=1}^K \|\Psi \mathbf{G} \mathbf{w}_k\|^2}_{\text{Desired signal power}} + \underbrace{\|\Psi\|_{\text{F}}^2 \sigma_v^2}_{\text{Amplified noise power}}. \end{aligned} \quad (10b)$$

Note that, different from the BS transmit power P_{BS} which only includes the desired signal power, since the active RIS amplifies the noises as well, the additional power consumption due to the noise amplification should be taken into account in the reflect power of active RIS P_A .

Therefore, the original problem of sum-rate maximization, subject to the power constraints at the BS and the active RIS, can be formulated as follows:

$$\mathcal{P}_o : \max_{\mathbf{w}, \Psi} \quad R_{\text{sum}}(\mathbf{w}, \Psi) = \sum_{k=1}^K \log_2 (1 + \gamma_k), \quad (11a)$$

$$\text{s.t.} \quad C_1: \sum_{k=1}^K \|\mathbf{w}_k\|^2 \leq P_{\text{BS}}^{\max}, \quad (11b)$$

$$C_2: \sum_{k=1}^K \|\Psi \mathbf{G} \mathbf{w}_k\|^2 + \|\Psi\|_{\text{F}}^2 \sigma_v^2 \leq P_A^{\max}, \quad (11c)$$

where $\mathbf{w} := [\mathbf{w}_1^T, \dots, \mathbf{w}_K^T]^T$ is the overall transmit beamforming vector for the K users; C_1 and C_2 are the power constraints at the BS and active RIS, respectively.

Due to the non-convexity and highly coupled variables in problem \mathcal{P}_o in (11), the joint design of \mathbf{w} and Ψ is challenging. Specifically, the introduction of the active RIS brings many difficulties to the beamforming design, such as the additional power constraint, the power allocation among active elements, the cancellation of multi-user interference, and the amplified

Algorithm 1 Proposed joint transmit beamforming and reflect precoding scheme

Input: Channels \mathbf{G} , \mathbf{h}_k , and \mathbf{f}_k , $\forall k \in \{1, \dots, K\}$.
Output: Optimized BS beamforming vector \mathbf{w} , optimized RIS precoding matrix of active RIS Ψ , and optimized sum-rate R_{sum} .

- 1: Randomly initialize \mathbf{w} and Ψ ;
- 2: **while** no convergence of R_{sum} **do**
- 3: Update ρ by (14);
- 4: Update ϖ by (15);
- 5: Update \mathbf{w} by solving (17);
- 6: Update Ψ by solving (20);
- 7: Update R_{sum} by (11a);
- 8: **end while**
- 9: **return** Optimized \mathbf{w} , Ψ , and R_{sum} .

noise power. Therefore, to efficiently solve this problem, we develop a joint beamforming and precoding scheme based on alternating optimization and fractional programming (FP), as provided in the next subsection.

B. Proposed Joint Beamforming and Precoding Scheme

To solve the problem efficiently, we reformulate the problem first. For simplicity, here we refer to \mathbf{w} and Ψ as the BS beamforming vector and the RIS precoding matrix, respectively. In order to deal with the non-convex sum-of-logarithms and fractions in (11), we exploit FP methods proposed in [34] to decouple the variables in problem \mathcal{P}_o in (11), so that multiple variables can be optimized separately. This leads to the following lemma.

Lemma 5 (Equivalent problem for sum-rate maximization): By introducing auxiliary variables $\rho := [\rho_1, \dots, \rho_K] \in \mathbb{R}_+^K$ and $\varpi := [\varpi_1, \dots, \varpi_K] \in \mathbb{C}^K$, the original problem \mathcal{P}_o in (11) can be equivalently reformulated as follows

$$\begin{aligned} \mathcal{P}_1: \max_{\mathbf{w}, \Psi, \rho, \varpi} \quad R'_{\text{sum}}(\mathbf{w}, \Psi, \rho, \varpi) &= \sum_{k=1}^K \ln(1 + \rho_k) - \\ &\quad \sum_{k=1}^K \rho_k + \sum_{k=1}^K g(\mathbf{w}, \Psi, \rho_k, \varpi_k), \end{aligned} \quad (12)$$

$$\text{s.t.} \quad C_1: \|\mathbf{w}\|^2 \leq P_{\text{BS}}^{\max},$$

$$C_2: \sum_{k=1}^K \|\Psi \mathbf{G} \mathbf{w}_k\|^2 + \|\Psi\|_{\text{F}}^2 \sigma_v^2 \leq P_A^{\max},$$

where function $g(\mathbf{w}, \Psi, \rho_k, \varpi_k)$ is defined as

$$g(\mathbf{w}, \Psi, \rho_k, \varpi_k) = 2\sqrt{(1 + \rho_k)} \Re \{ \varpi_k^* \bar{\mathbf{h}}_k^H \mathbf{w}_k \} - |\varpi_k|^2 \left(\sum_{j=1}^K |\bar{\mathbf{h}}_k^H \mathbf{w}_j|^2 + \|\mathbf{f}_k^H \Psi\|^2 \sigma_v^2 + \sigma^2 \right). \quad (13)$$

Proof: Constructive proof can be found in [34, Subsection III-C]. ■

Strong convergence of the FP methods was proved in [34]. Thus, if the updates in each iteration step of the BS beamforming vector \mathbf{w} , RIS precoding matrix Ψ , auxiliary variables ρ

and ϖ in (12) are all optimal, a locally optimal solution to (12) can be obtained by alternately optimizing these variables until R_{sum} converges. For clarity, we summarize the proposed joint beamforming and precoding scheme in **Algorithm 1**, and the specific optimal solutions for variables \mathbf{w} , Ψ , ρ , and ϖ are given in the following four steps, respectively.

1) *Fix $(\mathbf{w}, \Psi, \varpi)$ and then optimize ρ :* After fixing BS beamforming vector \mathbf{w} , RIS precoding matrix Ψ , and auxiliary variable ϖ , the optimal ρ can be obtained by solving $\frac{\partial R'_{\text{sum}}}{\partial \rho_k} = 0$ as

$$\rho_k^{\text{opt}} = \frac{\xi_k^2 + \xi_k \sqrt{\xi_k^2 + 4}}{2}, \quad \forall k \in \{1, \dots, K\}, \quad (14)$$

where $\xi_k = \Re\{\varpi_k^* \bar{\mathbf{h}}_k^H \mathbf{w}_k\}$.

2) *Fix (\mathbf{w}, Ψ, ρ) and then optimize ϖ :* After fixing the BS beamforming vector \mathbf{w} , RIS precoding matrix Ψ , and auxiliary variable ρ , the optimal ϖ can be derived by solving $\frac{\partial R'_{\text{sum}}}{\partial \varpi_k} = 0$ as

$$\varpi_k^{\text{opt}} = \frac{\sqrt{(1 + \rho_k)} \bar{\mathbf{h}}_k^H \mathbf{w}_k}{\sum_{j=1}^K |\bar{\mathbf{h}}_k^H \mathbf{w}_j|^2 + \|\mathbf{f}_k^H \Psi\|^2 \sigma_v^2 + \sigma^2}, \quad \forall k \in \{1, \dots, K\}. \quad (15)$$

3) *Fix (Ψ, ρ, ϖ) and then optimize \mathbf{w} :* To simplify the notations, we first introduce the following definitions:

$$\mathbf{b}_k^H = \sqrt{(1 + \rho_k)} \varpi_k^* \bar{\mathbf{h}}_k^H, \quad \mathbf{b} = [\mathbf{b}_1^T, \mathbf{b}_2^T, \dots, \mathbf{b}_N^T]^T, \quad (16a)$$

$$\mathbf{A} = \mathbf{I}_K \otimes \sum_{k=1}^K |\varpi_k|^2 \bar{\mathbf{h}}_k \bar{\mathbf{h}}_k^H, \quad \Xi = \mathbf{I}_K \otimes (\mathbf{G}^H \Psi^H \Psi \mathbf{G}), \quad (16b)$$

$$P_m^{\text{max}} = P_A^{\text{max}} - \|\Psi\|_F^2 \sigma_v^2. \quad (16c)$$

Then, for fixed RIS precoding matrix Ψ and auxiliary variables ρ and ϖ , problem \mathcal{P}_1 in (12) can be reformulated as follows

$$\begin{aligned} \mathcal{P}_2 : \quad & \max_{\mathbf{w}} \quad \Re\{2\mathbf{b}^H \mathbf{w}\} - \mathbf{w}^H \mathbf{A} \mathbf{w}, \\ & \text{s.t.} \quad C_1 : \|\mathbf{w}\|^2 \leq P_{\text{BS}}^{\text{max}}, \\ & \quad C_2 : \mathbf{w}^H \Xi \mathbf{w} \leq P_m^{\text{max}}. \end{aligned} \quad (17)$$

Since \mathcal{P}_2 in (17) is a standard quadratic constraint quadratic programming (QCQP) problem, by adopting the Lagrange multiplier method [22], the optimal solution \mathbf{w}^{opt} to \mathcal{P}_2 in (17) can be obtained as follows

$$\mathbf{w}^{\text{opt}} = (\mathbf{A} + \lambda_1 \mathbf{I}_{MK} + \lambda_2 \Xi)^{-1} \mathbf{b}, \quad (18)$$

where λ_1 and λ_2 are the Lagrange multipliers, which should be chosen such that the complementary slackness conditions of power constraints C_1 and C_2 are satisfied. The optimal Lagrange multipliers λ_1^{opt} and λ_2^{opt} can be obtained via a two-dimensional grid search [22].

4) *Fix $(\mathbf{w}, \rho, \varpi)$ and then optimize Ψ :* Define $\psi = [p_1 e^{j\theta_1}, \dots, p_N e^{j\theta_N}]^H$ as the vectorized RIS precoding matrix Ψ , i.e., $\text{diag}(\psi^H) := \Psi$. Thus, the equivalent channel $\bar{\mathbf{h}}_k^H$ can be rewritten as follows:

$$\bar{\mathbf{h}}_k^H = \mathbf{h}_k^H + \mathbf{f}_k^H \Psi \mathbf{G} = \mathbf{h}_k^H + \psi^H \text{diag}(\mathbf{f}_k^H) \mathbf{G}. \quad (19)$$

Utilizing (19), while fixing BS beamforming vector \mathbf{w} and auxiliary variables ρ and ϖ , problem \mathcal{P}_1 in (12) can be

reformulated as follows:

$$\begin{aligned} \mathcal{P}_3 : \quad & \max_{\psi} \quad \Re\{2\psi^H \mathbf{v}\} - \psi^H \Omega \psi, \\ & \text{s.t.} \quad C_2 : \psi^H \Pi \psi \leq P_A^{\text{max}}, \end{aligned} \quad (20)$$

wherein

$$\begin{aligned} \mathbf{v} = & \sum_{k=1}^K \sqrt{(1 + \rho_k)} \text{diag}(\varpi_k^* \mathbf{f}_k^H) \mathbf{G} \mathbf{w}_k - \\ & \sum_{k=1}^K |\varpi_k|^2 \text{diag}(\mathbf{f}_k^H) \mathbf{G} \sum_{j=1}^K \mathbf{w}_j \mathbf{w}_j^H \mathbf{h}_k, \end{aligned} \quad (21a)$$

$$\begin{aligned} \Omega = & \sum_{k=1}^K |\varpi_k|^2 \text{diag}(\mathbf{f}_k^H) \text{diag}(\mathbf{f}_k) \sigma_v^2 + \\ & \sum_{k=1}^K |\varpi_k|^2 \sum_{j=1}^K \text{diag}(\mathbf{f}_k^H) \mathbf{G} \mathbf{w}_j \mathbf{w}_j^H \mathbf{G}^H \text{diag}(\mathbf{f}_k), \end{aligned} \quad (21b)$$

$$\Pi = \sum_{k=1}^K \text{diag}(\mathbf{G} \mathbf{w}_k) (\text{diag}(\mathbf{G} \mathbf{w}_k))^H + \sigma_v^2 \mathbf{I}_N. \quad (21c)$$

Note that problem \mathcal{P}_3 in (20) is also a standard QCQP problem. Thus, the optimal solution ψ^{opt} can be obtained by adopting the Lagrange multiplier method and is given by

$$\psi^{\text{opt}} = (\Omega + \mu \Pi)^{-1} \mathbf{v}, \quad (22)$$

where μ is the Lagrange multiplier, which should be chosen such that the complementary slackness condition of power constrain C_2 is satisfied. Similarly, the optimal Lagrange multiplier μ^{opt} can be obtained via a binary search [22].

V. SELF-INTERFERENCE SUPPRESSION FOR ACTIVE RISs

Since active RISs work in full-duplex (FD) mode, the self-interference of active RISs occurs in practical systems. In this section, we extend the studied joint beamforming and precoding design to the practical system with the self-interference of active RISs. Specifically, in Subsection V-A, we first model the self-interference of active RISs, which allows us to account for the self-interference suppression in the beamforming design. In Subsection V-B, we formulate a mean-squared error minimization problem to suppress the self-interference of active RISs. In Subsection V-C, by utilizing ADMM [22] and SUMT [23], an alternating optimization scheme is proposed to solve the formulated problem.

A. Self-Interference Modeling

The self-interference of FD relays and that of active RISs are quite different. Specifically, due to the long processing delay at relays, the self-interference of FD relay originates from the different symbols that transmitted in the adjacent timeslot [35]–[37]. In this case, the self-interference at relays is usually viewed as colored Gaussian noise, which can be canceled by a zero-forcing suppression method [36]. Differently, since active RISs have nanosecond processing delay, the incident and reflected signals carry the same symbol in a timeslot. Due to the non-ideal inter-element isolation of practical arrays, part of the reflected signals may be received again by the active

RIS. In this case, the feedback-type self-interference occurs, which cannot be viewed as Gaussian noise anymore.

To distinguish the RIS precoding matrix in the ideal case Ψ , we denote the RIS precoding matrix in the non-ideal case with self-interference as $\Phi := \text{diag}(p_1 e^{j\theta_1}, \dots, p_N e^{j\theta_N})$. Recalling (2) and ignoring the negligible static noise for simplicity, the reflected signal of active RISs in the presence of self-interference can be modeled as follows:

$$\mathbf{y} = \underbrace{\Phi \mathbf{x}}_{\text{Desired signal}} + \underbrace{\Phi \mathbf{H} \mathbf{y}}_{\text{Self-interference}} + \underbrace{\Phi \mathbf{v}}_{\text{Dynamic noise}}, \quad (23)$$

where $\mathbf{H} \in \mathbb{C}^{N \times N}$ denotes the self-interference matrix [35]. In the general case without self-excitation (determinant of $(\mathbf{I}_N - \Phi \mathbf{H})$ is not zero), model (23) is a standard self-feedback loop circuit, of which the output \mathbf{y} naturally converges to the following steady state:

$$\mathbf{y} = \underbrace{(\mathbf{I}_N - \Phi \mathbf{H})^{-1} \Phi (\mathbf{x} + \mathbf{v})}_{\text{Equivalent RIS precoding matrix}}. \quad (24)$$

Comparing (24) and (2), one can observe that the difference is that the RIS precoding matrix Ψ in (2) is replaced by $(\mathbf{I}_N - \Phi \mathbf{H})^{-1} \Phi$. In particular, when all elements in \mathbf{H} are zero, the equivalent RIS precoding matrix $(\mathbf{I}_N - \Phi \mathbf{H})^{-1} \Phi$ is equal to diagonal matrix Φ .

B. Problem Formulation

To account for the self-interference of active RISs in the beamforming design, according to the new signal model (24), an intuitive way is to replace the RIS precoding matrix Ψ in problem \mathcal{P}_1 in (12) with the equivalent RIS precoding matrix $(\mathbf{I}_N - \Phi \mathbf{H})^{-1} \Phi$ and then solve \mathcal{P}_1 . Since this operation does not influence the optimizations of \mathbf{w} , ρ , and ϖ , here we focus on the optimization of Φ .

Consider replacing Ψ in (19) with $(\mathbf{I}_N - \Phi \mathbf{H})^{-1} \Phi$, thus the equivalent channel $\bar{\mathbf{h}}_k^H$ with self-interference can be written as:

$$\bar{\mathbf{h}}_k^H = \mathbf{h}_k^H + \mathbf{f}_k^H (\mathbf{I}_N - \Phi \mathbf{H})^{-1} \Phi \mathbf{G}. \quad (25)$$

However, due to the existence of self-interference matrix \mathbf{H} , Φ to be optimized exists in an inversion, thus the equivalent channel $\bar{\mathbf{h}}_k^H$ cannot be processed like (19), which makes Φ hard to be optimized. To address this challenge, we introduce the first-order Taylor expansion³ to approximate $(\mathbf{I}_N - \Phi \mathbf{H})^{-1} \approx \mathbf{I}_N + \Phi \mathbf{H}$, thus (25) can be rewritten as follows:

$$\begin{aligned} \bar{\mathbf{h}}_k^H &\approx \mathbf{h}_k^H + \mathbf{f}_k^H (\mathbf{I}_N + \Phi \mathbf{H}) \Phi \mathbf{G} \\ &\stackrel{(a)}{=} \mathbf{h}_k^H + (\mathbf{f}_k^H + \phi^H \text{diag}(\mathbf{f}_k^H) \mathbf{H}) \Phi \mathbf{G} \\ &\stackrel{(b)}{=} \mathbf{h}_k^H + \underbrace{(\phi^H + \phi^H \mathbf{H}_k \text{diag}(\phi^H))}_{\text{Equivalent precoding vector for user } k} \text{diag}(\mathbf{f}_k^H) \mathbf{G}, \end{aligned} \quad (26)$$

wherein RIS precoding vector ϕ satisfies $\Phi = \text{diag}(\phi^H)$; (a) holds since $\phi^H \text{diag}(\mathbf{f}_k^H) = \mathbf{f}_k^H \text{diag}(\phi^H)$; (b) holds by defining $\mathbf{H}_k = \text{diag}(\mathbf{f}_k^H) \mathbf{H} (\text{diag}(\mathbf{f}_k^H))^{-1}$.

³This approximation requires that the self-interference is not too strong (i.e., the values in \mathbf{H} are small), so that the high-order expansion items can be reasonably ignored.

Algorithm 2 Proposed self-interference suppression scheme

Input: Ideally optimized active RIS precoding vector ψ^{opt} , self-interference matrix \mathbf{H} , and channel \mathbf{f}_k , $\forall k \in \{1, \dots, K\}$.
Output: Active RIS precoding matrix Φ in the non-ideal case with self-interference.
1: Initialization: $\phi \leftarrow \psi^{\text{opt}}$, $\phi' \leftarrow \psi^{\text{opt}}$, and $\zeta \leftarrow 10^{-3}$;
2: **while** no convergence of $q(\phi, \phi')$ **do**
3: Update ϕ by (30);
4: Update ϕ' by (31);
5: Update ζ by $\zeta \leftarrow 2 \times \zeta$;
6: **end while**
7: $\Phi \leftarrow \text{diag}(\phi^H)$;
8: **return** Optimized active RIS precoding matrix Φ .

Comparing (26) and (19), the difference is that the RIS precoding vector ψ in (19) is replaced by the equivalent precoding vector $\phi + \text{diag}(\phi) \mathbf{H}_k^H \phi$ for user k . Therefore, an efficient way to eliminate the impact of self-interference is to design a ϕ to make all $\phi + \text{diag}(\phi) \mathbf{H}_k^H \phi$ approach the ideally optimized RIS precoding vector ψ^{opt} as close as possible. To achieve this, we temporarily omit the power constraint of active RISs in (11c) and formulate the following mean-squared error minimization problem:

$$\mathcal{P}_4 : \min_{\phi} f(\phi) = \frac{1}{K} \sum_{k=1}^K \|(\phi + \text{diag}(\phi) \mathbf{H}_k^H \phi) - \psi^{\text{opt}}\|^2, \quad (27)$$

where objective $f(\phi)$ is the cost function, defined as the mean of the squared approximation errors.

C. Proposed Self-Interference Suppression Scheme

To ensure the communication performance of active RIS aided systems, in this subsection, we propose a self-interference suppression scheme to solve problem \mathcal{P}_4 in (27).

Obviously, in the ideal case without self-interference (i.e., self-interference matrix \mathbf{H} is zero matrix), the optimal solution to problem \mathcal{P}_4 in (27) is $\phi = \psi^{\text{opt}}$ and satisfies $f(\phi) = 0$. Here we focus on the non-ideal case with a non-zero \mathbf{H} . In this case, problem \mathcal{P}_4 is challenging to solve due to the three reasons. Firstly, the objective $f(\phi)$ is usually non-convex since \mathbf{H}_k is asymmetric and indefinite. Secondly, $f(\phi)$ is in quartic form with respect to ϕ thus \mathcal{P}_4 has generally no closed-form solution. Finally, the coupled term $\text{diag}(\phi) \mathbf{H}_k^H \phi$ is a non-standard quadratic thus is hard to be preprocessed and optimized like (20).

To tackle this issue, inspired by ADMM [22] and SUMT [23], we turn to find a feasible solution to problem \mathcal{P}_4 by alternating optimization, as shown in **Algorithm 2**. The key idea of this algorithm includes two aspects: i) ADMM: Fix some variables and then optimize the others, so that $f(\phi)$ becomes temporarily convex thus can be minimized by alternating optimizations. ii) SUMT: Introduce an initially small but gradually increasing penalty term into the objective, so that the variables to be optimized can converge as an achievable solution to the original problem.

Following this idea, \mathcal{P}_4 in (27) can be reformulated as

$$\mathcal{P}_5 : \min_{\phi, \phi'} q(\phi, \phi') = f(\phi, \phi') + \underbrace{\zeta \|\phi' - \phi\|^2}_{\text{Penalty term}}, \quad (28)$$

wherein $f(\phi, \phi')$ is defined as

$$f(\phi, \phi') = \frac{1}{K} \sum_{k=1}^K \|(\phi + \text{diag}(\phi') \mathbf{H}_k^H \phi) - \psi^{\text{opt}}\|^2 \quad (29)$$

and $\zeta > 0$ is the penalty coefficient that increases in each iteration. For simplicity, here we assume ζ doubles in each update. In particular, when $\zeta \rightarrow \infty$, problem \mathcal{P}_5 in (28) is equivalent to \mathcal{P}_4 in (27).

Observing (29), we note that $q(\phi, \phi') = f(\phi, \phi') = f(\phi)$ when $\phi' = \phi$. Particularly, when ϕ (or ϕ') is fixed, objective $q(\phi, \phi')$ becomes a convex quadratic as a function of ϕ' (or ϕ). Therefore, for a given ζ , $q(\phi, \phi')$ in (28) can be minimized by optimizing ϕ and ϕ' alternatively. By solving $\partial q(\phi, \phi')/\partial\phi = 0$ and $\partial q(\phi, \phi')/\partial\phi' = 0$, we obtain the updating formulas of ϕ and ϕ' respectively, given by

$$\phi = \left(\frac{1}{K} \sum_{k=1}^K \mathbf{B}_k^H \mathbf{B}_k + \zeta \mathbf{I}_N \right)^{-1} \left(\zeta \phi' + \frac{1}{K} \sum_{k=1}^K \mathbf{B}_k^H \psi^{\text{opt}} \right), \quad (30)$$

$$\phi' = \left(\frac{1}{K} \sum_{k=1}^K \mathbf{D}_k^H \mathbf{D}_k + \zeta \mathbf{I}_N \right)^{-1} \left(\zeta \phi + \frac{1}{K} \sum_{k=1}^K \mathbf{D}_k^H (\psi^{\text{opt}} - \phi) \right), \quad (31)$$

where $\mathbf{B}_k = \mathbf{I}_N + \text{diag}(\phi') \mathbf{H}_k^H$ and $\mathbf{D}_k = \text{diag}(\mathbf{H}_k^H \phi)$. Besides, due to the existence of penalty term $\zeta \|\phi' - \phi\|^2$, as ζ increases, the converged solution to \mathcal{P}_5 in (28) tends to satisfy $\phi' = \phi$. After several alternating updates, ϕ and ϕ' will converge to the same value ($\phi = \phi'$), thus we obtain the desired RIS precoding matrix $\Phi = \text{diag}(\phi^H)$, which is exactly the output of **Algorithm 2**.

Recall that we temporarily omitted the power constraint in (11c) while optimizing ϕ . Here we introduce a scaling factor $\tau > 0$ for Φ to satisfy (11c), leading to the final solution Φ^{opt} , i.e.,

$$\Phi^{\text{opt}} = \tau \Phi. \quad (32)$$

According to (24) and (10), τ can be obtained by replacing Ψ in P_A in (10) with $(\mathbf{I}_N - \Phi^{\text{opt}} \mathbf{H})^{-1} \Phi^{\text{opt}}$ and then doing a binary search to find a proper τ that satisfies $P_A = P_A^{\text{max}}$. This completes the proposed self-interference suppression scheme.

VI. CONVERGENCE AND COMPLEXITY

A. Convergence Analysis

Algorithm 1 converges to a local optimal point after several iterations, since the updates in each iteration step of the algorithm are all optimal solutions to the respective subproblems. To prove this, here we introduce superscript t as the iteration index, e.g., \mathbf{w}^t refers to the transmit beamforming vector at

the end of the t -th iteration. Then, **Algorithm 1** converges as

$$\begin{aligned} R'_{\text{sum}}(\mathbf{w}^{t+1}, \Psi^{t+1}, \rho^{t+1}, \varpi^{t+1}) &\stackrel{(a)}{\geq} \\ R'_{\text{sum}}(\mathbf{w}^{t+1}, \Psi^t, \rho^{t+1}, \varpi^{t+1}) &\stackrel{(b)}{\geq} R'_{\text{sum}}(\mathbf{w}^t, \Psi^t, \rho^{t+1}, \varpi^{t+1}) \\ &\stackrel{(c)}{\geq} R'_{\text{sum}}(\mathbf{w}^t, \Psi^t, \rho^{t+1}, \varpi^t) \stackrel{(d)}{\geq} R'_{\text{sum}}(\mathbf{w}^t, \Psi^t, \rho^t, \varpi^t), \end{aligned} \quad (33)$$

where (a) and (b) follow since the updates of Ψ and \mathbf{w} are the optimal solutions to subproblems \mathcal{P}_3 in (20) and \mathcal{P}_2 in (17), respectively; (c) and (d) follow because the updates of ϖ and ρ maximize R'_{sum} when the other variables are fixed, respectively. Therefore, the objective R'_{sum} is monotonically non-decreasing in each iteration. Since the value of R'_{sum} is upper-bounded due to power constraints C_1 and C_2 , **Algorithm 1** will converge to a local optimum.

As an exterior point method, **Algorithm 2** meets two standard convergence conditions [23], which determines that it converges to a local optimal point where $\phi = \phi'$ and $q(\phi, \phi') = f(\phi, \phi') = f(\phi)$. Firstly, for a given penalty coefficient ζ in each iteration, the value of $q(\phi, \phi')$ in (28) is lower-bounded by zero and experiences the following monotonically non-increasing update:

$$q(\phi^{t+1}, (\phi')^{t+1}) \stackrel{(a)}{\leq} q(\phi^{t+1}, (\phi')^t) \stackrel{(b)}{\leq} q(\phi^t, (\phi')^t), \quad (34)$$

where (a) follows because the update of ϕ' minimizes $q(\phi, \phi')$ in \mathcal{P}_5 in (28) when ϕ is fixed and (b) follows since the update of ϕ minimizes $q(\phi, \phi')$ when ϕ' is fixed. Secondly, as penalty coefficient ζ increases to be sufficiently large ($\zeta \rightarrow \infty$), $q(\phi, \phi')$ in (28) is dominated by the penalty term $\zeta \|\phi' - \phi\|^2$. The updating formulas (30) becomes $\phi = \phi'$ and (31) becomes $\phi' = \phi$. It indicates that, ϕ and ϕ' do not update anymore and $\phi = \phi'$ always holds. As a result, penalty term $\zeta \|\phi' - \phi\|^2$ is equal to zero and the converged objective $q(\phi, \phi')$ finally satisfies $q(\phi, \phi') = f(\phi, \phi') = f(\phi)$.

B. Computational Complexity Analysis

The computational complexity of **Algorithm 1** is mainly determined by the updates of the four variables ρ , ϖ , \mathbf{w} , and Ψ via (14), (15), (17), and (20), respectively. Specifically, the computational complexity of updating ρ is $\mathcal{O}(KM)$. The complexity of updating ϖ is $\mathcal{O}(K^2 M + KN)$. Considering the complexity of solving standard convex QCQP problem, for a given accuracy tolerance ε , the computational complexity of updating \mathbf{w} is $\mathcal{O}(\log_2(1/\varepsilon) \sqrt{MK+2} (1+MK) M^3 K^3)$. Similarly, the computational complexity of updating Ψ is $\mathcal{O}(\log_2(1/\varepsilon) \sqrt{N+1} (1+2N) N^3)$. Thus, the overall computational complexity of **Algorithm 1** is given by $\mathcal{O}(\log_2(1/\varepsilon) I_o (M^{4.5} K^{4.5} + N^{4.5}))$, wherein I_o denotes the number of iterations required by **Algorithm 1** for convergence.

Similarly, the computational complexity of **Algorithm 2** is mainly determined by updating ϕ and ϕ' via (30) and (31), respectively. As closed-form updating formulas, their computational complexity are both $\mathcal{O}((K+1) N^3 + (K+1) N^2)$, which are mainly caused by matrix inversions. Thus, the overall computational complexity of **Algorithm 2** is $\mathcal{O}(I_s K N^3)$, wherein I_s is the number of iterations required by **Algorithm 2** for convergence.

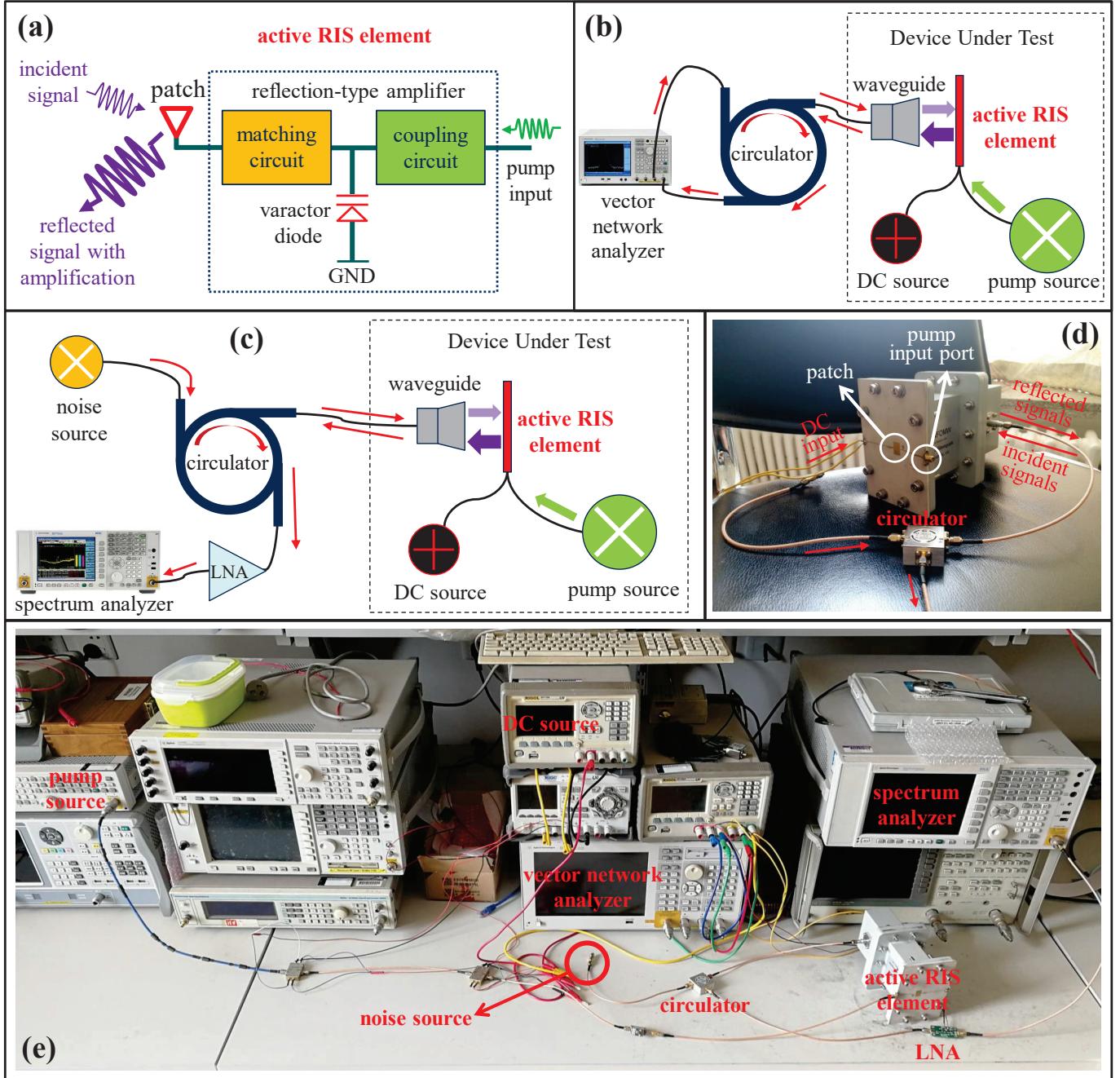


Fig. 4. The experimental devices and environment for validating the signal model (2) of the active RIS.

VII. VALIDATION RESULTS

In this section, we present validation results. To validate the signal model (2), in Subsection VII-A, we present experimental results based on a fabricated active RIS element. Then, in Subsection VII-B, simulation results are provided to evaluate the sum-rate of active RIS aided MU-MISO systems. Finally, in Subsection VII-C, the impact of active RIS self-interference on system performance is discussed.

A. Validation Results for Signal Model

To validate the signal model (2), we designed and fabricated an active RIS element with an integrated reflection-type ampli-

fier for experimental measurements in [21]. Particularly, since the phase-shifting ability of RISs has been widely verified [24], we focus on studying the reflection gain and the noise introduced by an active RIS element. Thus, the validation of signal model (2) is equivalent to validating

$$P_y = \underbrace{GP_x}_{\text{Desired-signal power}} + \underbrace{G\sigma_v^2 + \sigma_s^2}_{\text{noise power}}, \quad (35)$$

where P_y is the power of the signals reflected by the active RIS element; P_x is the power of the incident signal; $G := p^2$ is the reflection gain of the active RIS element; $G\sigma_v^2$ and σ_s^2 are the powers of the dynamic noise and static noise introduced

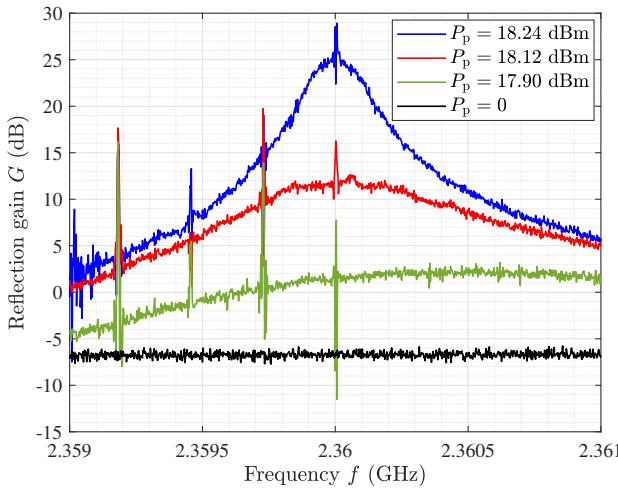


Fig. 5. Experimental measurement result for reflection gain G versus signal frequency f .

by the active RIS element, respectively.

1) *Hardware platform*: To validate the model in (35), we first establish the hardware platform used for our experimental measurements, see Fig. 4. Specifically, we show the following aspects:

- Fig. 4 (a) illustrates the structure of the fabricated active RIS element operating at a frequency of 2.36 GHz [21]. A pump input at a frequency of 4.72 GHz is used to supply the power required by the active RIS element. The incident signal and the pump input are coupled in a varactor-diode-based reflection-type amplifier to generate the reflected signal with amplification.
- Fig. 4 (b) illustrates the system used for measuring the reflection gain G of the active RIS element. A direct-current (DC) source is used to provide a bias voltage of 7.25 V for driving the active RIS element, and a controllable pump source is used to reconfigure the reflection gain G . A circulator is used to separate the incident signal and the reflected signal, and the reflection gain is directly measured by a vector network analyzer.
- Fig. 4 (c) illustrates the system for measuring the noises introduced at the active RIS element, where a spectrum analyzer is used to measure the noise power. The noise source is a 50Ω impedance, which aims to simulate a natural input noise of -174 dBm/Hz at each patch. The reflected signal is amplified by a low-noise amplifier (LNA) so that the spectrum analyzer can detect it.
- Fig. 4 (d) shows a photo of the fabricated active RIS element under test, which is connected by a waveguide for incident/reflected signal exchanges.
- Fig. 4 (e) shows a photo of the experimental environment with the required equipment for device driving and signal measurement.

2) *Reflection gain measurement*: Using the measurement system for the reflection gain depicted in Fig. 4 (b), we first investigate the reflection gain G of the active RIS element. The reflection gain G can be reconfigured by the input power of

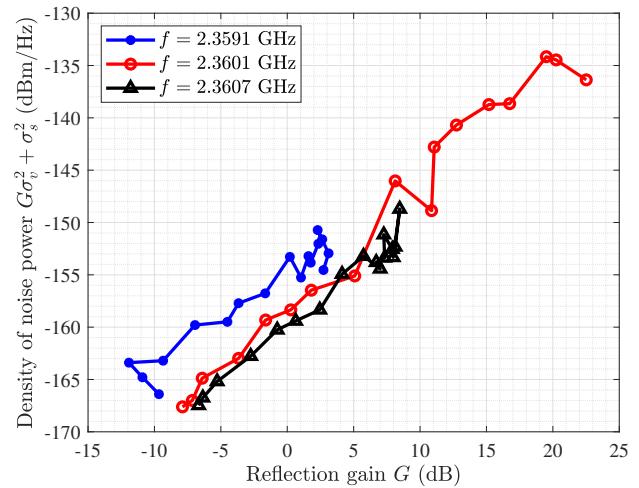


Fig. 6. Experimental measurement result for the density of noise power $G\sigma_v^2 + \sigma_s^2$ versus reflection gain G .

the pump source P_p . By setting the input power of the vector network analyzer as $P_x = -50$ dBm, the reflection gain G as a function of the signal frequency can be directly measured via the vector network analyzer. Then, in Fig. 5, we show the measurement results for reflection gain G as a function of signal frequency f for different input powers of the pump source P_p . We observe that the active RIS element can achieve a reflection gain G of more than 25 dB, when $P_p = 18.24$ dBm, which confirms the significant reflection gains enabled by active RISs. On the other hand, when $P_p = 0$, we observe that G falls to -6 dB, which is lower than the expected 0 dB. This loss is mainly caused by the inherent power losses of the circulator and transmission lines used for measurement.

3) *Noise power measurement*: We further study the noise power introduced and amplified by the active RIS element, i.e., $G\sigma_v^2 + \sigma_s^2$ in (35), where $G\sigma_v^2$ and σ_s^2 are the powers of the dynamic noise and static noise introduced at the active RIS element, respectively. Using the noise measurement system in Fig. 4 (c), we show the measurement results for the spectral density of noise power $G\sigma_v^2 + \sigma_s^2$ as a function of G for different operating frequencies in Fig. 6. We can observe that the noise power increases nearly linearly with G , which verifies the noise model $G\sigma_v^2 + \sigma_s^2$ in (35). Particularly, for $f = 2.3601$ GHz, the spectral density of σ_s^2 is about -174 dBm/Hz, while that of σ_v^2 is about -160 dBm/Hz, which is about 15 dB higher. The reason for this is that the input noise is amplified by the noise factor [28], and additional noises are also introduced by the other active components the measurement equipment, such as the leakage noise from the DC source.

B. Simulation Results for Joint Beamforming and Precoding Design

To evaluate the effectiveness of the proposed joint beamforming and precoding design, in this subsection, we present simulation results for passive RIS and active RIS aided MU-MISO systems, respectively.

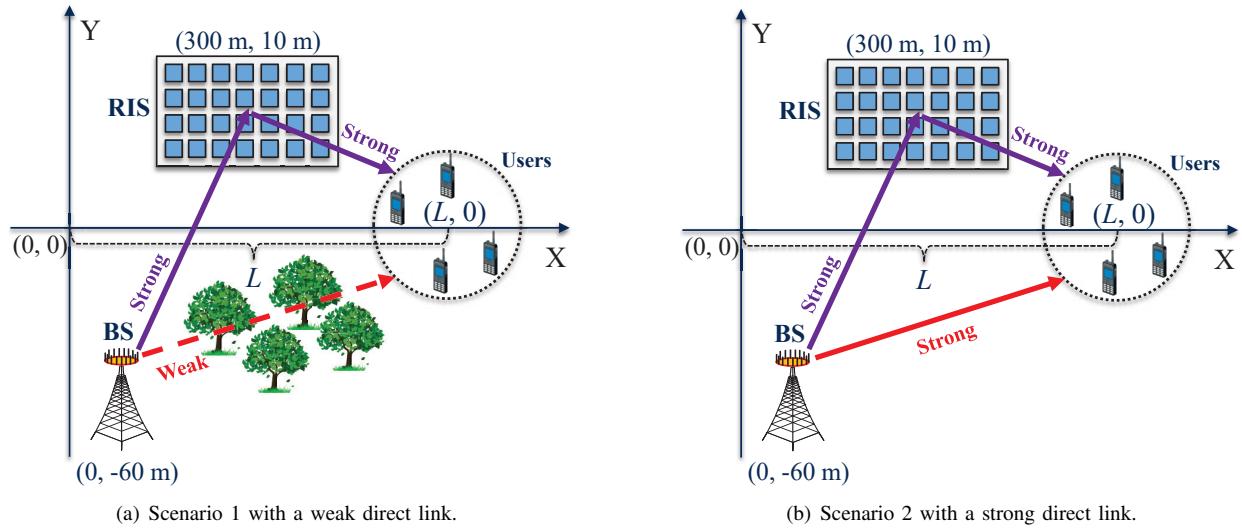


Fig. 7. Two simulation scenarios with different channel conditions, where a BS aided by an active RIS serves four users.

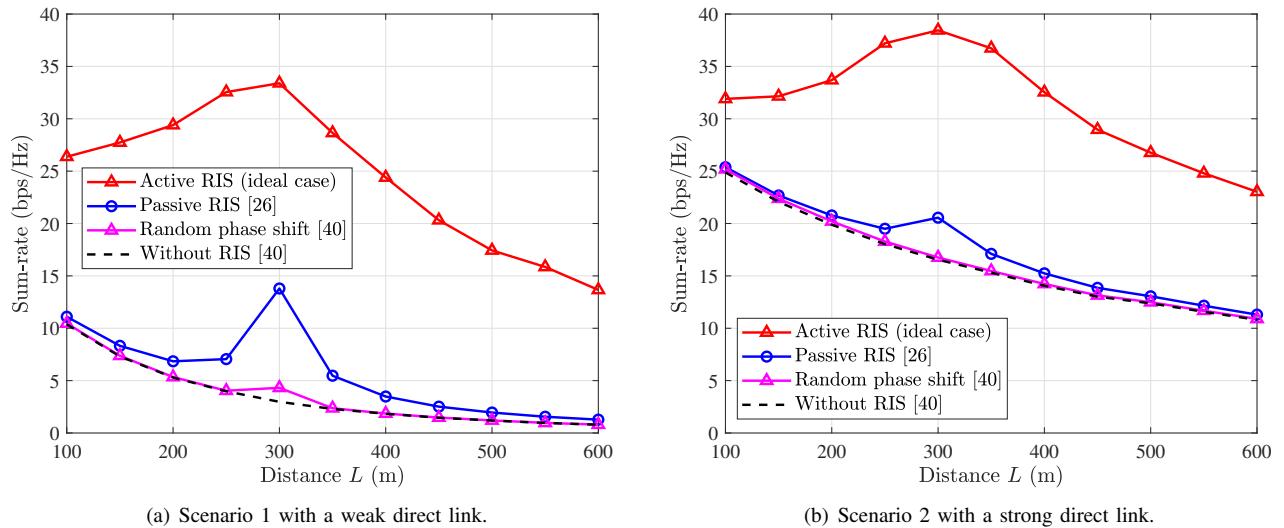


Fig. 8. Simulation results for the sum-rate as a function of distance L in an RIS-aided MU-MISO system.

1) *Simulation setup*: For the simulation setup, we consider an active/passive RIS aided MU-MISO system operating at a frequency of 5 GHz as shown in Fig. 7. Particularly, we consider two scenarios with different channel conditions. In Fig. 7 (a), the direct link is weak due to severe obstruction, while the direct link is strong in Fig. 7 (b). To be specific, two different path loss models from the 3GPP standard [38, B.1.2.1] are utilized to characterize the large-scale fading of the channels:

$$\begin{aligned} PL_s &= 37.3 + 22.0 \log d, \\ PL_w &= 41.2 + 28.7 \log d, \end{aligned} \quad (36)$$

where d is the distance between two devices. Path loss model PL_w is used to generate the weak BS-user link in scenario 1, while PL_s is used to generate the strong BS-user link in scenario 2. For both scenarios in Fig. 7, PL_s is used to generate the BS-RIS and the RIS-user channels. To account for small-scale fading, following [39], we adopt the Ricean

fading channel model for all channels involved. In this way, an arbitrary channel matrix \mathbf{H} is generated by

$$\mathbf{H} = \sqrt{PL} \left(\sqrt{\frac{\kappa}{\kappa+1}} \mathbf{H}_{\text{LoS}} + \sqrt{\frac{1}{\kappa+1}} \mathbf{H}_{\text{NLoS}} \right), \quad (37)$$

where PL is the corresponding path loss of \mathbf{H} ; κ is the Ricean factor; and \mathbf{H}_{LoS} and \mathbf{H}_{NLoS} represent the deterministic LoS and Rayleigh fading components, respectively. In particular, here we assume $\kappa = 1$.

As common settings, the BS and the active/passive RIS are located at $(0, -60 \text{ m})$ and $(300 \text{ m}, 10 \text{ m})$, respectively. The locations of the four users will be specified later. Unless specified otherwise, the numbers of BS antennas and RIS elements are set as $M = 4$ and $N = 512$, respectively. The noise power is set as $\sigma^2 = \sigma_v^2 = -100 \text{ dBm}$. Let P_{BS}^{\max} denote the maximum transmit power at the BS and P_A^{\max} denote the maximum reflect power of the active RIS, which don't include the

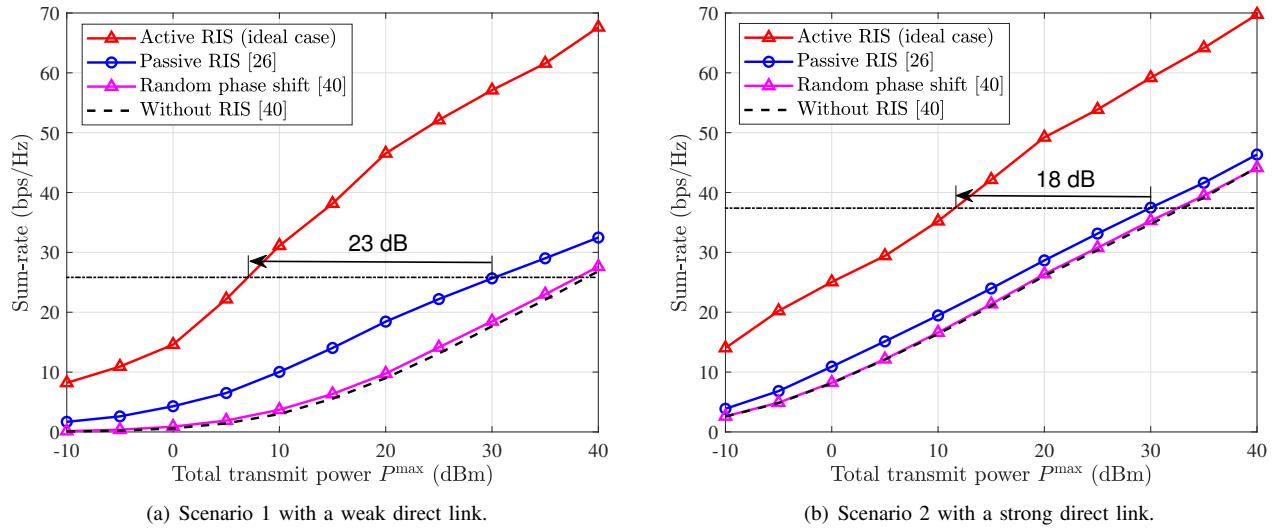


Fig. 9. Simulation results for the sum-rate versus total power consumption P^{\max} in an RIS-aided MU-MISO system.

hardware static power. For fair comparison, we constrain the total power consumption $P^{\max} := P_{\text{BS}}^{\max} + P_{\text{A}}^{\max}$ to 10 dBm by setting $P_{\text{BS}}^{\max} = 0.99 \times P^{\max}$ and $P_{\text{A}}^{\max} = 0.01 \times P^{\max}$ for the active RIS aided system, and $P_{\text{BS}}^{\max} = 10$ dBm for the other benchmark systems. To show the effectiveness of beamforming designs, here we consider the following four schemes for simulations:

- **Active RIS (ideal case):** In an ideal active RIS-aided MU-MISO system without self-interference, the proposed **Algorithm 1** is employed to jointly optimize the BS beamforming and the precoding at the active RIS.
- **Passive RIS [26]:** In a passive RIS-aided MU-MISO system, the algorithm proposed in [26] is adopted to jointly optimize the BS beamforming and the precoding at the passive RIS.
- **Random phase shift [40]:** In a passive RIS-aided MU-MISO system, the phase shifts of all passive RIS elements are randomly set. Then, relying on the equivalent channels from the BS to users, the weighted mean-squared error minimization (WMMSE) algorithm from [40] is used to optimize the BS beamforming.
- **Without RIS [40]:** In an MU-MISO system without RIS, the WMMSE algorithm from [40] is adopted to optimize the BS beamforming.

2) *Coverage performance of active RISs:* To observe the coverage performance of active RISs, we assume that the four users are randomly located in a circle with a radius of 5 m from the center ($L, 0$). In Fig. 8 (a) and (b), we plot the sum-rate versus distance L for the two considered scenarios, where the direct link is weak and strong, respectively. Based on these results, we have two observations. Firstly, in scenario 1 with a weak direct link, the passive RIS can indeed achieve an obvious performance improvement, while the active RIS achieves a much higher sum-rate gain. Secondly, in scenario 2 with a strong direct link, the passive RIS only achieves a limited sum-rate gain, while the active RIS still realizes a noticeable sum-rate gain. For example, when $L = 300$ m,

the capacities without RIS, with passive RIS, and with active RIS in scenario 1 are 2.98 bps/Hz, 13.80 bps/Hz, and 33.39 bps/Hz respectively, while in scenario 2, these values are 16.75 bps/Hz, 20.56 bps/Hz, and 38.45 bps/Hz, respectively. For this position, the passive RIS provides a 363% gain in scenario 1 and a 22% gain in scenario 2. By contrast, the active RIS achieves noticeable sum-rate gains of 1020% in scenario 1 and 130% in scenario 2, which are much higher than those achieved by the passive RIS in the corresponding scenarios. These results demonstrate that, compared with the passive RIS, the active RIS can overcome the “multiplicative fading” effect and achieve noticeable sum-rate gains even when direct link is strong.

3) *Sum-rate versus total power consumption P^{\max} :* To evaluate the averaged performance in the coverage of active/passive RIS, we assume that all users are randomly distributed in a large circle with a radius of 50 m from the center (300 m, 0). We show the users' sum-rate versus the total power consumption P^{\max} in Fig. 9. From Fig. 9 we observe that the passive RIS achieves visible performance gains in scenario 1 where the direct link is weak, while the passive RIS only achieves limited sum-rate gains in scenario 2 where the direct link is strong. By contrast, in both scenarios, the active RIS realizes a high performance gain. Particularly, to achieve the same performance as the passive RIS aided system, the required power consumption for the active RIS aided system is much lower. For example, when the total power consumption of the passive RIS aided system is $P^{\max} = 30$ dBm, to achieve the same sum-rate, the active RIS aided system only requires 7 dBm in scenario 1 and 12 dBm in scenario 2, which correspond to power savings of 23 dB and 18 dB, respectively. The reason for this result is that, for the passive RIS, the total power is only allocated to BS. Thus, all transmit power is affected by the large path loss of the full BS-RIS-user link. However, for the active RIS, part of the transmit power is allocated to the active RIS, and this part of the power is only affected by the path loss of the RIS-user link. Thus, the

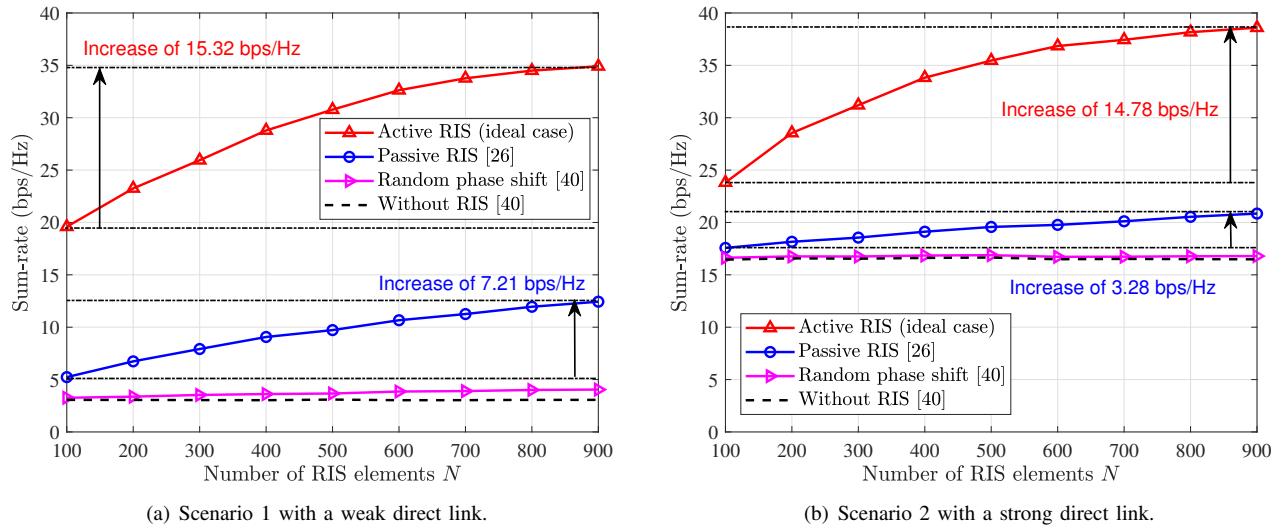


Fig. 10. Simulation results for the sum-rate versus the number of RIS elements N in an RIS-aided MU-MISO system.

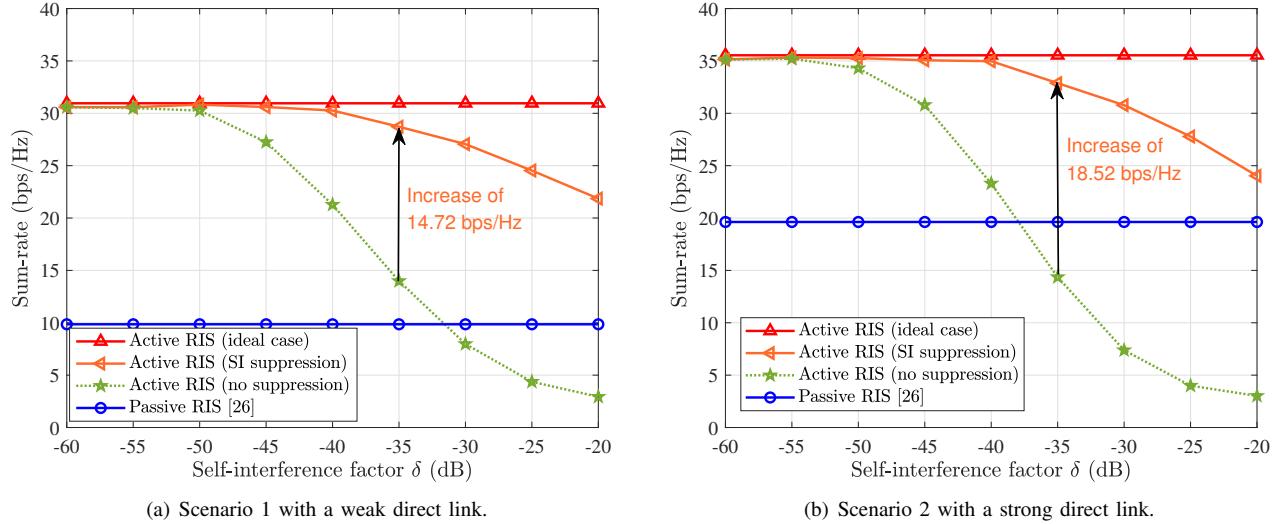


Fig. 11. Simulation results for the sum-rate versus the self-interference factor δ of the active RIS.

active RIS is promising for reducing the power consumption of communication systems.

4) *Sum-rate versus number of RIS elements N :* For the same setup as in Fig. 9, we plot the users' sum-rate versus the number of RIS elements N in Fig. 10. We observe that, as the number of RIS elements N increases, both the passive RIS and the active RIS achieve higher sum-rate gains, while the performance improvement for the active RIS aided system is much larger than that for the passive RIS aided system. For example, when N increases from 100 to 900, the sum-rate of the passive RIS aided system increases from 5.23 bps/Hz to 12.44 bps/Hz in scenario 1 (increase of 7.21 bps/Hz) and from 17.57 bps/Hz to 20.85 bps/Hz in scenario 2 (increase of 3.28 bps/Hz), respectively. By contrast, the sum-rate of the active RIS aided system increases from 19.59 bps/Hz to 34.91 bps/Hz in scenario 1 (increase of 15.32 bps/Hz) and from 23.81 bps/Hz to 38.59 bps/Hz in scenario 2 (increase of 14.78

bps/Hz), respectively. These results show that the sum-rate increase of the active RIS aided system is much higher than that of the passive RIS aided system. This indicates that, as long as the number of RIS elements N is not exceedingly large (such as millions of elements), compared with the passive RIS, increasing the number of elements of the active RIS is much more efficient for improving the communication performance, which is in agreement with the performance analysis in Section III.

C. Simulation Results for Self-Interference Suppression

In this subsection, we present simulation results to verify the effectiveness of the proposed self-interference suppression scheme for active RISs.

1) *Simulation setup:* To avoid the impact of other factors, we adopt the same setup in Subsection VII-B, which is used in Fig. 9 and Fig. 10. Without loss of generality, we assume

that each element in self-interference matrix \mathbf{H} is distributed as $\sim \mathcal{CN}(0, \delta^2)$ [35]–[37], where we name δ as the self-interference factor, which is inversely proportional to the inter-element isolation of practical arrays [41]–[43]. To evaluate the impact of self-interference on sum-rate, we add two new benchmarks for simulations:

- **Active RIS (SI suppression):** In a non-ideal active RIS-aided MU-MISO system with self-interference, **Algorithm 1** is employed to optimize the BS beamforming and the active RIS precoding, and **Algorithm 2** is employed to suppress the self-interference. Then, the performance is evaluated under the condition of self-interference.
- **Active RIS (no suppression):** In a non-ideal active RIS-aided MU-MISO system with self-interference, only **Algorithm 1** is employed to design the BS beamforming and the active RIS precoding and the self-interference is ignored. Then, the performance is evaluated under the condition of self-interference.

2) *Impact of self-interference on sum-rate:* We plot the users' sum-rate versus the self-interference factor δ in Fig. 11. We observe that, when $\delta < -50$ dB, the self-interference has almost no impact on the sum-rate. However, as the self-interference strengthens, the active RIS aided system without self-interference suppression suffers an increasingly high performance loss. Particularly, when $\delta = -35$ dB, the active RIS without self-interference suppression does not even perform as well as the passive RIS in scenario 2. The reason is that, the existence of self-interference matrix \mathbf{H} makes the reflected signals unable to focus on the users, or even worse, cancel the desired signals of the direct link. Fortunately, thanks to our proposed **Algorithm 2**, the active RIS aided system with self-interference suppression can still hold a considerable performance. For example, when $\delta = -35$ dB, compared with the active RIS aided system without self-interference suppression, the system with self-interference suppression can compensate for the sum-rate loss of 14.72 bps/Hz in scenario 1 and that of 18.52 bps/Hz in scenario 2.

VIII. CONCLUSIONS AND FUTURE WORKS

In this paper, we have proposed the concept of active RISs to overcome the fundamental limitation of the “multiplicative fading” effect. Specifically, we have developed and verified a signal model for active RISs by a fabricated active RIS element through experimental measurements. Based on the verified signal model, we have analyzed the asymptotic performance of active RISs and then formulated an optimization problem to maximize the sum-rate in an active RIS aided MU-MISO system. Subsequently, we have proposed a joint beamforming and precoding scheme to solve this problem. Finally, experimental and simulation results have shown that, compared with the benchmark scheme without RIS, the passive RIS can realize only a limited sum-rate gain of about 22% in a typical application scenario, while the proposed active RIS can achieve a substantial sum-rate gain of about 130%, thus indeed overcoming the fundamental limitation of the “multiplicative fading” effect. In the future, many research directions for active RISs are worth pursuing, including hardware design

[27], prototype development [7], channel estimation [25], and energy efficiency analysis [10].

APPENDIX A PROOF OF LEMMA 2

For notational simplicity, we rewrite some matrices and vectors in (3) as $\mathbf{g} = [g_1, \dots, g_N]^T$, $\mathbf{f} = [f_1, \dots, f_N]^T$, and $\mathbf{w}_k := w$. Thus, the downlink transmission model in (3) can be rewritten as

$$r = \underbrace{\mathbf{p}\mathbf{f}^H\Theta\mathbf{g}}_{\text{Reflected link}} \underbrace{ws}_{\text{Noise introduced by active RIS}} + \underbrace{z}_{\text{Noise introduced at user}}, \quad (38)$$

where $r \in \mathbb{C}$ is the signal received by the user. Based on the transmission model in (38), the maximization of the user's SNR γ , subject to the power constraints at the BS and the active RIS, can be formulated as follows:

$$\begin{aligned} \max_{w, p, \Theta} \quad & \gamma = \frac{|p\mathbf{f}^H\Theta\mathbf{g}w|^2}{p^2\|\mathbf{f}^H\Theta\|^2\sigma_v^2 + \sigma^2}, \\ \text{s.t.} \quad & C_1 : |w|^2 \leq P_{\text{BS}}^{\max}, \\ & C_2 : p^2\|\mathbf{f}^H\Theta\mathbf{g}w\|^2 + p^2N\sigma_v^2 \leq P_{\text{A}}^{\max}, \end{aligned} \quad (39)$$

where P_{BS}^{\max} and P_{A}^{\max} denote the maximum transmit power and the maximum reflect power at the BS and the active RIS, respectively. Then, the optimal solution of problem (39) can be obtained by the Lagrange multiplier method as follows:

$$w^{\text{opt}} = \sqrt{P_{\text{BS}}^{\max}}, \quad (40a)$$

$$\theta_n^{\text{opt}} = \angle f_n - \angle g_n, \quad \forall n \in \{1, \dots, N\}, \quad (40b)$$

$$p^{\text{opt}} = \sqrt{\frac{P_{\text{A}}^{\max}}{P_{\text{BS}}^{\max} \sum_{n=1}^N |g_n|^2 + N\sigma_v^2}}. \quad (40c)$$

By substituting (40) into (39), the user's maximum achievable SNR for active RISs can be obtained as

$$\gamma_{\text{active}} = \frac{P_{\text{BS}}^{\max} P_{\text{A}}^{\max} \left| \sum_{n=1}^N |f_n| |g_n| \right|^2}{P_{\text{A}}^{\max} \sigma_v^2 \sum_{n=1}^N |f_n|^2 + \sigma^2 \left(P_{\text{BS}}^{\max} \sum_{n=1}^N |g_n|^2 + N\sigma_v^2 \right)}. \quad (41)$$

Note that we assume $\mathbf{f} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_f^2 \mathbf{I}_N)$ and $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_g^2 \mathbf{I}_N)$. Thus, by letting $N \rightarrow \infty$ in (41), according to the law of large numbers, we have $\sum_{n=1}^N |f_n| |g_n| \rightarrow N \frac{\pi \varrho_f \varrho_g}{4}$, $\sum_{n=1}^N |g_n|^2 \rightarrow N \varrho_g^2$, and $\sum_{n=1}^N |f_n|^2 \rightarrow N \varrho_f^2$. After substituting these asymptotic equations into (41), we obtain for the asymptotic SNR for active RISs the expression in (5), which completes the proof.

APPENDIX B PROOF OF LEMMA 3

According to the related analysis in [14] and Appendix A, the user's achievable SNR for an SU-SISO system aided by a passive RIS and that aided by an active RIS can be respectively written as follows

$$\gamma_{\text{passive}} = \frac{P_{\text{BS-P}}^{\max} \left| \sum_{n=1}^N |f_n| |g_n| \right|^2}{\sigma^2}, \quad (42a)$$

$$\gamma_{\text{active}} = \frac{P_{\text{BS-A}}^{\max} P_{\text{A}}^{\max} \left| \sum_{n=1}^N |f_n| |g_n| \right|^2}{P_{\text{A}}^{\max} \sigma_v^2 \sum_{n=1}^N |f_n|^2 + \sigma^2 \left(P_{\text{BS-A}}^{\max} \sum_{n=1}^N |g_n|^2 + N \sigma_v^2 \right)}, \quad (42b)$$

where $P_{\text{BS-A}}^{\max}$ denotes the maximum BS transmit power for the active RIS aided system and $P_{\text{BS-P}}^{\max}$ denotes that for the passive RIS aided system. By solving $\gamma_{\text{passive}} \geq \gamma_{\text{active}}$ according to (42), we have

$$N \geq \frac{P_{\text{BS-A}}^{\max}}{P_{\text{BS-P}}^{\max}} \frac{P_{\text{A}}^{\max} \sigma^2}{P_{\text{A}}^{\max} \sigma_v^2 \frac{1}{N} \sum_{n=1}^N |f_n|^2 + \sigma^2 \left(P_{\text{BS}}^{\max} \frac{1}{N} \sum_{n=1}^N |g_n|^2 + \sigma_v^2 \right)} \\ \approx \frac{P_{\text{BS-A}}^{\max}}{P_{\text{BS-P}}^{\max}} \frac{P_{\text{A}}^{\max} \sigma^2}{\left(P_{\text{A}}^{\max} \sigma_v^2 \varrho_f^2 + P_{\text{BS-A}}^{\max} \sigma^2 \varrho_g^2 + \sigma^2 \sigma_v^2 \right)}, \quad (43)$$

where we assume again that $\mathbf{f} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_f^2 \mathbf{I}_N)$ and $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}_N, \varrho_g^2 \mathbf{I}_N)$. Since the number of RIS elements N is usually large, the components $\frac{1}{N} \sum_{n=1}^N |f_n|^2$ and $\frac{1}{N} \sum_{n=1}^N |g_n|^2$ in (43) were approximated by ϱ_f^2 and ϱ_g^2 . This completes the proof.

REFERENCES

- [1] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RISs: Signal modeling, asymptotic analysis, and beamforming design," in *Proc. 2022 IEEE Global Commun. Conf. (IEEE GLOBECOM'22)*, Rio de Janeiro, Brazil, Dec. 2022, pp. 1–7.
- [2] L. Zhang, X. Q. Chen, S. Liu, Q. Zhang, J. Zhao, J. Y. Dai, G. D. Bai, X. Wan, Q. Cheng, G. Castaldi, V. Galdi, and T. J. Cui, "Space-time-coding digital metasurfaces," *Nat. Commun.*, vol. 9, no. 4338, Oct. 2018.
- [3] H. Ren, "A light-programmable metasurface," *Nat. Electr.*, vol. 3, pp. 137–138, Mar. 2020.
- [4] S. Venkatesh, X. Lu, H. Saeidi, and K. Sengupta, "A high-speed programmable and scalable terahertz holographic metasurface based on tiled CMOS chips," *Nat. Electr.*, vol. 3, pp. 785–793, Dec. 2020.
- [5] M. Di Renzo, A. Zappone, M. Debbah, M. S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "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.
- [6] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Aug. 2019.
- [7] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. Di Renzo, C. B. Chae, and L. Hanzo, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45 913–45 923, Mar. 2020.
- [8] 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, Aug. 2020.
- [9] 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. 14 960–14 973, Dec. 2020.
- [10] 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, Aug. 2019.
- [11] H. Zhao, Y. Shuang, M. Wei, T. J. Cui, P. Hougne, and L. Li, "Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals," *Nat. Commun.*, vol. 11, no. 3926, Aug. 2020.
- [12] M. Faraji-Dana, E. Arbab, A. Arbab, S. M. Kamali, H. Kwon, and A. Faraon, "Compact folded metasurface spectrometer," *Nat. Commun.*, vol. 9, no. 4196, Oct. 2013.
- [13] J. Park, B. G. Jeong, S. I. Kim, D. Lee, J. Kim, C. Shin, C. B. Lee, T. Otsuka, J. Kyoungh, S. Kim, K. Yang, Y. Park, J. Lee, I. Hwang, J. Jang, S. H. Song, M. L. Brongersma, K. Ha, S. Hwang, H. Choo, and B. L. Choi, "All-solid-state spatial light modulator with independent phase and amplitude control for three-dimensional LiDAR applications," *Nat. Nanotechnol.*, vol. 16, p. 69–76, Oct. 2020.
- [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.
- [15] W. Zhao, G. Wang, S. Atapattu, T. A. Tsiftsis, and C. Tellambura, "Is backscatter link stronger than direct link in reconfigurable intelligent surface-assisted system?" *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1342–1346, Jun. 2020.
- [16] T. Hou, Y. Liu, Z. Song, X. Sun, and Y. Chen, "MIMO-NOMA networks relying on reconfigurable intelligent surface: A signal cancellation-based design," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6932–6944, Nov. 2020.
- [17] Z. Zhang and L. Dai, "A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network," *IEEE Trans. Signal Process.*, vol. 69, pp. 4085–4101, Aug. 2021.
- [18] M. Najafi, V. Jamali, R. Schober, and H. V. Poor, "Physics-based modeling and scalable optimization of large intelligent reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2673–2691, Apr. 2021.
- [19] R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson, "Active reconfigurable intelligent surface-aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4962–4975, Aug. 2021.
- [20] C. You and R. Zhang, "Wireless communication aided by intelligent reflecting surface: Active or passive?" *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2659–2663, Dec. 2021.
- [21] X. Chen and F. Yang, "Nonlinear electromagnetic surfaces: Theory, design and application," *Master Thesis in Tsinghua University*, May 2020, [Online] Available: <http://etds.lib.tsinghua.edu.cn/Thesis>.
- [22] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Nov. 2014, [Online] Available: https://stanford.edu/~boyd/papers/pdf/admm_distr_stats.pdf.
- [23] A. V. Fiacco and G. P. McCormick, *Nonlinear programming: Sequential unconstrained minimization techniques*. SIAM, 1990.
- [24] H. Yang, F. Yang, X. Cao, S. Xu, J. Gao, X. Chen, M. Li, and T. Li, "A 1600-element dual-frequency electronically reconfigurable reflectarray at X/Ku-band," *IEEE Trans. Antennas Propag.*, vol. 65, no. 6, pp. 3024–3032, Jun. 2017.
- [25] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021.
- [26] 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, Aug. 2020.
- [27] J. Lončar, Z. Šipuš, and S. Hrabar, "Ultrathin active polarization-selective metasurface at X-band frequencies," *Physical Review B*, vol. 100, no. 7, p. 075131, Oct. 2019.
- [28] J. Bousquet, S. Magierowski, and G. G. Messier, "A 4-GHz active scatterer in 130-nm CMOS for phase sweep amplify-and-forward," *IEEE Trans. Circuits Syst. I*, vol. 59, no. 3, pp. 529–540, Mar. 2012.
- [29] K. K. Kishor and S. V. Hum, "An amplifying reconfigurable reflectarray antenna," *IEEE Trans. Antennas Propag.*, vol. 60, no. 1, pp. 197–205, Jan. 2012.
- [30] J. He, N. T. Nguyen, R. Schroeder, V. Tapiola, J. Kokkonieti, and M. Juntti, "Channel estimation and hybrid architectures for RIS-assisted communications," in *Proc. 2021 Joint European Conf. Netw. Commun. 6G Summit (EuCNC/6G Summit'21)*, Jun. 2021, pp. 60–65.
- [31] N. T. Nguyen, Q.-D. Vu, K. Lee, and M. Juntti, "Hybrid relay-reflecting intelligent surface-assisted wireless communication," *arXiv preprint arXiv:2103.03900*, Mar. 2021.
- [32] E. Basar, "Transmission through large intelligent surfaces: A new frontier in wireless communications," in *Proc. 2019 European Conf. Netw. Commun. (EuCNC'19)*, Jun. 2019, pp. 1–6.
- [33] K. Ntontin, J. Song, and M. D. Renzo, "Multi-antenna relaying and reconfigurable intelligent surfaces: End-to-end SNR and achievable rate," *arXiv preprint arXiv:1908.07967*, Aug. 2019.
- [34] K. Shen and W. Yu, "Fractional programming for communication systems—part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.
- [35] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 913–927, Jan. 2014.

- [36] P. Lioliou, M. Viberg, M. Coldrey, and F. Athley, "Self-interference suppression in full-duplex mimo relays," in *Proc. 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ACSSC'10)*, Nov. 2010, pp. 658–662.
- [37] P. Xing, J. Liu, C. Zhai, X. Wang, and L. Zheng, "Self-interference suppression for the full-duplex wireless communication with large-scale antenna," *Trans. Emerging Tel. Tech.*, vol. 27, no. 6, pp. 764–774, Feb. 2016.
- [38] "Further advancements for E-UTRA physical layer aspects (release 9)," *3GPP TS 36.814*, Mar. 2010.
- [39] H. Guo, Y. 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, May 2020.
- [40] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
- [41] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 91–101, May 2015.
- [42] M. Heino, S. N. Venkatasubramanian, C. Icheln, and K. Haneda, "Design of wavetraps for isolation improvement in compact in-band full-duplex relay antennas," *IEEE Trans. Antennas Propag.*, vol. 64, no. 3, pp. 1061–1070, Dec. 2016.
- [43] D. Korpi, M. Heino, C. Icheln, K. Haneda, and M. Valkama, "Compact inband full-duplex relays with beyond 100 dB self-interference suppression: Enabling techniques and field measurements," *IEEE Trans. Antennas Propag.*, vol. 65, no. 2, pp. 960–965, Nov. 2017.