Active RIS Versus Passive RIS: Which is Superior With the Same Power Budget?

Kangda Zhi, *Graduate Student Member, IEEE*, Cunhua Pan, *Member, IEEE*, Hong Ren, *Member, IEEE*, Kok Keong Chai, and Maged Elkashlan, *Senior Member, IEEE*

***Abstract*—This letter theoretically compares the active reconfigurable intelligent surface (RIS)-aided system with the passive RIS-aided system. For a fair comparison, we consider that these two systems have the same overall power budget that can be used at both the base station (BS) and the RIS. For active RIS, we first derive the optimal power splitting between the BS's transmit signal power and RIS's output signal power. We also analyze the impact of various system parameters on the optimal power splitting ratio. Then, we theoretically and numerically compare the performance between the active RIS and the passive RIS, which demonstrates that the active RIS would be superior if the power budget is not very small and the number of RIS elements is not very large.**

***Index Terms*—Reconfigurable intelligent surface (RIS), intelligent reflecting surface (IRS), active RIS, power budget.**

I. Introduction

PASSIVE reconfigurable intelligent surface (RIS)-aided systems have attracted extensive research attention recently [1]–[6]. By passively reflecting the impinging signal and intelligently adjusting the phase shifts, received signals from different paths can be constructively superimposed and enhanced. Besides, the passive RIS relies on low power tunable electronic circuits (e.g., PIN diodes or varactors) to shift the phase, and therefore nearly zero power is consumed.

However, the passive nature also has some drawbacks. The signal reflected by the RIS needs to pass through two paths, i.e., the base station (BS)-RIS and RIS-user paths. Without signal amplification, the received signal suffers from the product/double path loss attenuation and therefore becomes weak enough. This "double path loss" attenuation limits the potential of RIS to a large extent [7]. To tackle this challenge, the concept of active RIS has been proposed and investigated in [8]–[13]. The appealing feature of the active RIS is that it can not only adjust the phase shifts but also amplify the received signal attenuated from the first hop to a normal strength level. Accordingly, active RIS effectively circumvents the double path loss attenuation.

Manuscript received February 13, 2022; accepted March 5, 2022. Date of publication March 14, 2022; date of current version May 10, 2022. This work was supported in part by the National Key Research and Development Project (2019YFE0123600). The work of Kangda Zhi was supported by China Scholarship Council. The associate editor coordinating the review of this letter and approving it for publication was M. A. C. Gomes. *(Corresponding author: Cunhua Pan.)*

Kangda Zhi, Kok Keong Chai, and Maged Elkashlan are with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: k.zhi@qmul.ac.uk; michael.chai@qmul.ac.uk; maged.elkashlan@qmul.ac.uk).

Cunhua Pan and Hong Ren are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 211189, China (e-mail: cpan@seu.edu.cn; hren@seu.edu.cn).

Digital Object Identifier 10.1109/LCOMM.2022.3159525

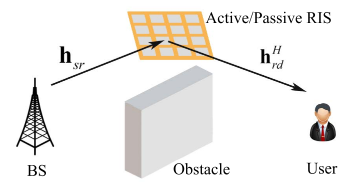

Fig. 1. Illustration of the considered system.

The hardware structure of active RIS is fundamentally different from that of passive RIS. To amplify the signal, the active RIS is equipped with phase shifts circuits and reflection-type amplifiers (e.g., current-inverting converters) [8]. Different from the low power passive RIS, similar power could be consumed by the BS and amplifiers of active RIS, which means that the power consumption of active RIS can no longer be neglected. Since the total power consumption in active RIS systems could be much higher than that in passive RIS-aided systems, it is fair to compare them under the same total power budget. Although the comparison of active and passive RISs under the same power budget has been considered in [8] via numerical simulations, this topic has not been studied from the perspective of theoretical analysis. To better understand the performance difference between active and passive RISs, explicit theoretical data rate expressions are necessary and some fundamental analytical insights are highly desirable.

In this letter, we perform a fair comparison between the active and passive RISs with the same overall power budget. Given total power, we first decide how much power should be split to the active RIS's amplifiers. Then, we theoretically analyze the impact of various system parameters on the derived optimal power splitting ratio. Finally, analytical and numerical results are provided to shed light on the performance difference between the active and passive RISs.

II. System Model

We consider a single-input single-output (SISO) system with the aid of an RIS equipped with N elements as illustrated in Fig. 1.[1] The BS-RIS and RIS-user channels are denoted by $\mathbf{h}_{sr} \in \mathbb{C}^{N\times 1}$ and $\mathbf{h}_{rd}^H \in \mathbb{C}^{1\times N}$, respectively. For brevity, the direct link is assumed to be entirely blocked [10].[2]

As in [9], [10], we assume that line-of-sight (LoS) paths exist in RIS-reflected channels. Based on the uniform rectangular array (URA) model, we denote that

[1]This tractable scenario enables us to provide some essential insights and gain a better understanding of the property of active RIS. The extension to more practical multi-user MIMO scenarios will be left for our future work.

[2]This assumption is reasonable for scenarios where buildings or walls exist between the BS and the user.

$\mathbf{h}_{sr} = h_{sr}\mathbf{a}_N(\vartheta^a, \vartheta^e)$ and $\mathbf{h}_{rd}^H = h_{rd}\mathbf{a}_N^T(\varsigma^a, \varsigma^e)$, where $\mathbf{a}_N(\omega^a, \omega^e) = \tilde{\mathbf{a}}_{N_x}(\sin\omega^a \sin\omega^e) \otimes \tilde{\mathbf{a}}_{N_y}(\cos\omega^e)$, $N = N_x N_y$, and $\tilde{\mathbf{a}}_L(\varpi) = [1, \ldots, e^{j2\pi\frac{q_{ris}}{\lambda}(L-1)\varpi}]^T$. ϑ^a, ϑ^e, ς^a and ς^e are the azimuth and elevation angles of arrival and departure, respectively. q_{ris} and λ denote the element spacing of the RIS and the wavelength, respectively. h_{sr} and h_{rd} represent the distance-dependent path-loss factors expressed as $h_{sr} = \sqrt{\beta_{sr} d_{sr}^{-\alpha_{sr}}}$ and $h_{rd} = \sqrt{\beta_{rd} d_{rd}^{-\alpha_{rd}}}$, where β_{sr} and β_{rd} represent the reference strength for the channel at a distance of 1 m, d_{sr} and d_{rd} are distances, and α_{sr} and α_{rd} denote the path-loss exponents. Note that if $\beta_{sr} = \beta_{rd} = -30$ dB and $\alpha_{sr} = \alpha_{rd} = 2$, the channel gains would be $h_{sr}^2 = -70$ dB and $h_{rd}^2 = -50$ dB, by assuming that the RIS is deployed close to the user so that $d_{sr} = 100$ m and $d_{rd} = 10$ m. Therefore, in practical scenarios, the channel gains h_{sr}^2 and h_{rd}^2 are very small values [7]. Without amplification, the received signal at the user could suffer from the product/double path loss $h_{sr}^2 \times h_{rd}^2 = -120$ dB and therefore becomes very weak. In the sequel of this letter, the order of magnitude of h_{sr}^2 and h_{rd}^2 could help us better understand the performance comparison between the active RIS and the passive RIS.

Define the reflection matrix of the RIS as $\mathbf{\Theta} = \mathrm{diag}\left\{\rho_1 e^{j\theta_1}, \ldots, \rho_N e^{j\theta_N}\right\}$, where θ_n denotes the phase shift of the n-th RIS reflecting element. For passive RIS, we have $\rho_n = 1, \forall n$. However, $\rho_n > 1, \forall n$ are feasible for active RIS due to its amplifiers. For simplicity, we assume that $\rho_n = \rho, \forall n$ and then define $\mathbf{\Theta} = \rho\,\mathrm{diag}\left\{e^{j\theta_1}, \ldots, e^{j\theta_N}\right\} = \rho\mathbf{\Phi}$. Let $x \sim \mathcal{CN}(0, 1)$ denote the symbol transmitted from the BS. Then, the signal reflected by the active RIS and finally received by the user is given by

$$y_{\text{act}} = \sqrt{P_{\text{BS}}^{\text{act}}}\rho\mathbf{h}_{rd}^H\mathbf{\Phi}\mathbf{h}_{sr}x + \rho\mathbf{h}_{rd}^H\mathbf{\Phi}\mathbf{n}_r + n, \tag{1}$$

where $P_{\text{BS}}^{\text{act}}$ is the transmit power of BS in active RIS systems, $\mathbf{n}_r \sim \mathcal{CN}\left(\mathbf{0}, \sigma_r^2\mathbf{I}_N\right)$ denotes the thermal noise introduced by active RIS components, and $n \sim \mathcal{CN}\left(0, \sigma^2\right)$ represents the thermal noise at the receiver. Letting $\rho = 1$ and $\mathbf{n}_r = \mathbf{0}$, we can obtain the received signal in the passive RIS-aided system as follows

$$y_{\text{pas}} = \sqrt{P_{\text{BS}}^{\text{pas}}}\mathbf{h}_{rd}^H\mathbf{\Phi}\mathbf{h}_{sr}x + n, \tag{2}$$

where $P_{\text{BS}}^{\text{pas}}$ is the transmit power of BS in passive RIS systems. Besides, the overall power consumption of active and passive RIS-aided systems are respectively given by [9]

$$\begin{aligned} Q_{\text{act}} &= P_{\text{BS}}^{\text{act}} + P_{\text{RIS}}^{\text{act}} + N\left(P_{\text{SW}} + P_{\text{DC}}\right), \\ Q_{\text{pas}} &= P_{\text{BS}}^{\text{pas}} + NP_{\text{SW}}, \end{aligned} \tag{3}$$

where $P_{\text{RIS}}^{\text{act}}$ is the power of amplified signal reflected by the active RIS, P_{SW} is the power consumed by the phase shift switch and control circuit in each RIS elements, P_{DC} is the direct current biasing power used by the amplifier in each active RIS element.

III. Active RIS Versus Passive RIS

Based on (1), the achievable rate of active RIS systems is $R_{\text{act}} = \log_2(1 + \gamma_{\text{act}})$, in which the signal-to-noise-ratio (SNR) is expressed as

$$\begin{aligned} \gamma_{\text{act}} &= \frac{P_{\text{BS}}^{\text{act}}\rho^2\left|\mathbf{h}_{rd}^H\mathbf{\Phi}\mathbf{h}_{sr}\right|^2}{\rho^2\sigma_r^2\left\|\mathbf{h}_{rd}^H\mathbf{\Phi}\right\|^2 + \sigma^2} = \frac{P_{\text{BS}}^{\text{act}}\rho^2\left|\mathbf{h}_{rd}^H\mathbf{\Phi}\mathbf{h}_{sr}\right|^2}{\rho^2\sigma_r^2 N h_{rd}^2 + \sigma^2} \\ &\overset{(a)}{=} \frac{P_{\text{BS}}^{\text{act}}\rho^2 N^2 h_{sr}^2 h_{rd}^2}{\rho^2\sigma_r^2 N h_{rd}^2 + \sigma^2} = \frac{N^2 P_{\text{BS}}^{\text{act}} h_{sr}^2 h_{rd}^2}{N\sigma_r^2 h_{rd}^2 + \frac{\sigma^2}{\rho^2}}, \end{aligned} \tag{4}$$

where (a) utilizes the optimal design of $\mathbf{\Phi}$, i.e., $\theta_n = \arg\{[\mathbf{h}_{rd}]_n\} - \arg\{[\mathbf{h}_{sr}]_n\}, \forall n$.

Substituting (4) with $\rho = 1$ and $\sigma_r^2 = 0$, we obtain the achievable rate of passive RIS systems as $R_{\text{pas}} = \log_2(1 + \gamma_{\text{pas}})$, where

$$\gamma_{\text{pas}} = \frac{N^2 P_{\text{BS}}^{\text{pas}} h_{sr}^2 h_{rd}^2}{\sigma^2}. \tag{5}$$

Comparing (4) with (5), it is obvious that $\gamma_{\text{act}} > \gamma_{\text{pas}}$, if $P_{\text{BS}}^{\text{act}} = P_{\text{BS}}^{\text{pas}}$ (without the same power budget constraint), $Nh_{rd}^2 \ll 1$ (for $h_{rd}^2 \approx -50$ dB), $\sigma_r^2 \approx \sigma^2$, and $\rho^2 > 1$. Therefore, for a fair comparison, it is necessary to constrain that both two schemes have the same overall power budget $Q_{\text{act}} = Q_{\text{pas}}$, which means $P_{\text{BS}}^{\text{act}} < P_{\text{BS}}^{\text{pas}}$. In this context, the superiority of active RIS over passive RIS is non-trivial and needs to be re-examined.

A. Problem Formulation With the Same Power Budget

Assume that the total power budget is Q_{tot}, i.e., $Q_{\text{act}} = Q_{\text{pas}} = Q_{\text{tot}}$. From (3), we have

$$P_{\text{BS}}^{\text{act}} = Q_{\text{tot}} - N\left(P_{\text{SW}} + P_{\text{DC}}\right) - P_{\text{RIS}}^{\text{act}} \triangleq C - P_{\text{RIS}}^{\text{act}}, \tag{6}$$

$$P_{\text{BS}}^{\text{pas}} = Q_{\text{tot}} - NP_{\text{SW}} = C + NP_{\text{DC}}, \tag{7}$$

where $C = P_{\text{BS}}^{\text{act}} + P_{\text{RIS}}^{\text{act}}$ corresponds to the available power left for splitting to the BS and active RIS after supplying the hardware power consumption. Clearly, if $C \leq 0$, we have $P_{\text{BS}}^{\text{act}} = 0$ and $\gamma_{\text{act}} = 0$. Therefore, we only focus on the region of $C > 0$ in this section.

Different from the passive RIS, we need to additionally decide the optimal power splitting for $P_{\text{BS}}^{\text{act}}$ and $P_{\text{RIS}}^{\text{act}}$ given C. The optimization problem is formulated as

$$\max_{P_{\text{BS}}^{\text{act}},\,\rho} \quad \gamma_{\text{act}} \tag{8a}$$

$$\text{s.t.} \quad P_{\text{BS}}^{\text{act}}\rho^2\|\mathbf{\Phi}\mathbf{h}_{sr}\|^2 + \rho^2\sigma_r^2\|\mathbf{\Phi}\|^2 \leq P_{\text{RIS}}^{\text{act}}, \tag{8b}$$

$$P_{\text{BS}}^{\text{act}} + P_{\text{RIS}}^{\text{act}} = C. \tag{8c}$$

Since γ_{act} increases with ρ, substituting (8c) into (8b), the optimal ρ^2 should satisfy the following condition

$$\rho^2 = \frac{C - P_{\text{BS}}^{\text{act}}}{N\left(P_{\text{BS}}^{\text{act}} h_{sr}^2 + \sigma_r^2\right)}. \tag{9}$$

Then, the original problem is transformed to

$$\max_{P_{\text{BS}}^{\text{act}}} \quad \gamma_{\text{act}} = \frac{Nh_{sr}^2 h_{rd}^2\left(CP_{\text{BS}}^{\text{act}} - (P_{\text{BS}}^{\text{act}})^2\right)}{\sigma_r^2 h_{rd}^2\left(C - P_{\text{BS}}^{\text{act}}\right) + \sigma^2\left(P_{\text{BS}}^{\text{act}} h_{sr}^2 + \sigma_r^2\right)} \tag{10a}$$

$$\text{s.t.} \quad 0 \leq P_{\text{BS}}^{\text{act}} \leq C. \tag{10b}$$

It is readily to find $\gamma_{\text{act}}|_{P_{\text{BS}}^{\text{act}}=0} = 0$ and $\gamma_{\text{act}}|_{P_{\text{BS}}^{\text{act}}=C} = 0$. Therefore, it is necessary to decide the optimal power splitting between the BS and the active RIS.

B. Optimal Power Splitting

Theorem 1: If $\sigma_r^2 h_{rd}^2 = \sigma^2 h_{sr}^2$, the optimal power split to the BS is $(P_{\mathrm{BS}}^{\mathrm{act}})^\star = C/2$. Otherwise, we have

$$\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^\star = \frac{1}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2} \times \Big\{ C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2 - \sqrt{\left(C\sigma^2 h_{sr}^2 + \sigma^2\sigma_r^2\right)\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right)} \Big\}. \tag{11}$$

Meanwhile, the optimal power split to the RIS is $(P_{\mathrm{RIS}}^{\mathrm{act}})^\star = C - (P_{\mathrm{BS}}^{\mathrm{act}})^\star$.

Proof: Please refer to Appendix A. ■

If $\sigma_r^2 = \sigma^2$, $\beta_{sr} = \beta_{rd}$, and $\alpha_{sr} = \alpha_{rd}$, condition $\sigma_r^2 h_{rd}^2 = \sigma^2 h_{sr}^2$ can be satisfied when the RIS is located in the middle between the BS and the user. In this special case, it is optimal to equally split the power to the BS and the RIS. Unless otherwise stated, we focus on the case $\sigma_r^2 h_{rd}^2 \neq \sigma^2 h_{sr}^2$ in the following.

Corollary 1: When $\sigma_r^2 \to 0$, $(P_{\mathrm{BS}}^{\mathrm{act}})^\star \to 0$ and $(P_{\mathrm{RIS}}^{\mathrm{act}})^\star \to C$. When $\sigma_r^2 \to \infty$, $(P_{\mathrm{BS}}^{\mathrm{act}})^\star \to C + \frac{\sigma^2}{h_{rd}^2} - \sqrt{\frac{\sigma^2}{h_{rd}^2}(C + \frac{\sigma^2}{h_{rd}^2})}$, which further tends to C if $\frac{\sigma^2}{h_{rd}^2} \ll C$.

As $\sigma_r^2 \to 0$, the receiver noise σ^2 becomes dominant, and therefore larger $P_{\mathrm{RIS}}^{\mathrm{act}}$ is preferred, which achieves larger ρ and then reduces the term $\frac{\sigma^2}{\rho^2}$ in (4). As $\sigma_r^2 \to \infty$, the RIS noise becomes dominant, and it is useless to increase ρ since it also amplifies the RIS noise term $\rho^2\sigma_r^2 N h_{rd}^2$ in (4). In this case, increasing $P_{\mathrm{BS}}^{\mathrm{act}}$ can effectively improve the SNR.

Corollary 2: Both $(P_{\mathrm{BS}}^{\mathrm{act}})^\star$ and $(P_{\mathrm{RIS}}^{\mathrm{act}})^\star$ are increasing functions of Q_{tot} but decreasing functions of N.

Proof: Please refer to Appendix B. ■

Corollary 2 shows that as power budget Q_{tot} grows, it is optimal to simultaneously increase the power of the BS and the RIS. Fully splitting the increased power to the BS/RIS unfairly will sacrifice the performance. Meanwhile, when N increases, less power is left for the BS and the RIS, and it is optimal to simultaneously cut down their power.

Corollary 3: When $\sigma_r^2 h_{rd}^2 > \sigma^2 h_{sr}^2$, $(P_{\mathrm{BS}}^{\mathrm{act}})^\star - (P_{\mathrm{RIS}}^{\mathrm{act}})^\star$ is an increasing function of Q_{tot} but decreasing function of N. On the contrary, when $\sigma_r^2 h_{rd}^2 < \sigma^2 h_{sr}^2$, $(P_{\mathrm{BS}}^{\mathrm{act}})^\star - (P_{\mathrm{RIS}}^{\mathrm{act}})^\star$ is a decreasing function of Q_{tot} but increasing function of N.

Proof: Please refer to Appendix C. ■

Although the power should be split fairly, Corollary 3 unveils that there do exist some splitting priorities for increased Q_{tot}. With larger $\sigma_r^2 h_{rd}^2$, more power should be split to the BS, since $P_{\mathrm{BS}}^{\mathrm{act}}$ reduces the impact of RIS noise. By contrast, with larger $\sigma^2 h_{sr}^2$, more power should be split to the active RIS, which leads to larger ρ and therefore decreases the impact of receiver noise. Besides, by moving the RIS closer to the user, h_{rd}^2 increases while h_{sr}^2 decreases. Hence, when the RIS is located near the BS (user), more power should be split to the RIS (BS). This is because the received signal at the RIS is stronger with larger h_{sr}^2, and larger $P_{\mathrm{RIS}}^{\mathrm{act}}$ is needed to amplify a stronger signal for a certain multiple ρ as shown in (9).

C. SNR Comparison

Using $(P_{\mathrm{BS}}^{\mathrm{act}})^\star$, (4), (5), and (9), by solving $\gamma_{\mathrm{act}} > \gamma_{\mathrm{pas}}$, it is readily to obtain the following results.

Lemma 1: Passive RIS performs better if

$$N h_{rd}^2 \frac{\sigma_r^2}{\sigma^2} > \frac{(P_{\mathrm{BS}}^{\mathrm{act}})^\star}{P_{\mathrm{BS}}^{\mathrm{pas}}}. \tag{12}$$

Otherwise, active RIS performs better when

$$\frac{C - (P_{\mathrm{BS}}^{\mathrm{act}})^\star}{N\left(\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^\star h_{sr}^2 + \sigma_r^2\right)} = \left(\rho^\star\right)^2 > \frac{1}{\frac{(P_{\mathrm{BS}}^{\mathrm{act}})^\star}{P_{\mathrm{BS}}^{\mathrm{pas}}} - N h_{rd}^2 \frac{\sigma_r^2}{\sigma^2}}. \tag{13}$$

We firstly focus on condition (12). Since $\frac{(P_{\mathrm{BS}}^{\mathrm{act}})^\star}{P_{\mathrm{BS}}^{\mathrm{pas}}} < 1$, passive RIS must be better if $N > \frac{\sigma^2}{h_{rd}^2\sigma_r^2}$, i.e., for sufficiently large N, since in this case active RIS suffers from severe noise $N\sigma_r^2 h_{rd}^2$. (12) is easier to hold for small σ^2 or large σ_r^2. This is because active RIS can effectively mitigate the impact of σ^2 but it is impaired by σ_r^2. If $\sigma_r^2 = \sigma^2$, condition (12) holds if $N h_{rd}^2 > 1$, which means that the attenuation from path loss h_{rd}^2 is compensated by RIS's gain N and therefore double path loss attenuation no longer exists. However, we emphasize that the N satisfying condition (12) may be very large due to the small value of h_{rd}^2.

Lemma 2: Define $g(C) = \frac{(P_{\mathrm{BS}}^{\mathrm{act}})^\star}{P_{\mathrm{BS}}^{\mathrm{pas}}}$. If $\sigma_r^2 h_{rd}^2 > \sigma_r h_{sr}^2$, $g(C)$ increases with C. If $\sigma_r^2 h_{rd}^2 < \sigma_r h_{sr}^2$, depending on the values of N, $g(C)$ could be an increasing function or a function which firstly increases but then decreases with C. In addition, $g(0) \to 0$ and $g(\infty) \to \frac{\sigma_r^2 h_{rd}^2 - \sqrt{\sigma^2 h_{sr}^2 \sigma_r^2 h_{rd}^2}}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2} \in (0,1)$ which approaches 1 as $\sigma_r^2 h_{rd}^2 \to \infty$.

Proof: Please refer to Appendix D. ■

Corollary 4: Passive RIS outperforms active RIS for small C, i.e., for low power budget Q_{tot}.

Proof: Define $C^* = \arg\max_C g(C)$. If $g(C^*) < N h_{rd}^2 \frac{\sigma_r^2}{\sigma^2}$, condition (12) holds for all C (also for small C). If not, from Lemma 2, there exists an intersection point $\widetilde{C} \in (0, C^*)$ so that $g(\widetilde{C}) = N h_{rd}^2 \frac{\sigma_r^2}{\sigma^2}$. Then, condition (12) holds for all $C \in (0, \widetilde{C})$. ■

The above discussions demonstrate that passive RIS is better for very large N and very small Q_{tot}. We next use condition (13) to demonstrate the superiority of active RIS when N is not very large and Q_{tot} is not very small.

Corollary 5: Active RIS outperforms passive RIS if $\sigma^2 \approx \sigma_r^2 \ll P_{\mathrm{DC}}$, $N h_{sr}^2, N h_{rd}^2 \ll 1$, and $C \geq N P_{\mathrm{DC}}$.

Proof: We prove $\gamma_{\mathrm{act}}\left((P_{\mathrm{BS}}^{\mathrm{act}})^\star\right) > \gamma_{\mathrm{pas}}$ by proving $\gamma_{\mathrm{act}}\left(C/2\right) > \gamma_{\mathrm{pas}}$ due to $\gamma_{\mathrm{act}}\left((P_{\mathrm{BS}}^{\mathrm{act}})^\star\right) \geq \gamma_{\mathrm{act}}\left(C/2\right)$. By replacing $(P_{\mathrm{BS}}^{\mathrm{act}})^\star$ in (13) with $C/2$, the left-hand side of (13) is lower bounded by

$$\rho_{\mathrm{sub}}^2 = \frac{C}{C N h_{sr}^2 + 2N\sigma_r^2} \geq \min\left\{\frac{1}{2N h_{sr}^2}, \frac{C}{4N\sigma_r^2}\right\} \geq \min\left\{\frac{1}{2N h_{sr}^2}, \frac{P_{\mathrm{DC}}}{4\sigma_r^2}\right\} \gg 1. \tag{14}$$

The right-hand side of (13) is approximately upper bounded by

$$\frac{1}{\frac{C}{2(C+NP_{\mathrm{DC}})} - N h_{rd}^2 \frac{\sigma_r^2}{\sigma^2}} \leq \frac{1}{\frac{1}{4} - N h_{rd}^2 \frac{\sigma_r^2}{\sigma^2}} \approx 4 < \rho_{\mathrm{sub}}^2, \tag{15}$$

which completes the proof. ■

The reason behind Corollary 5 is that the signal received at the RIS has been attenuated by path loss h_{sr}^2 and therefore

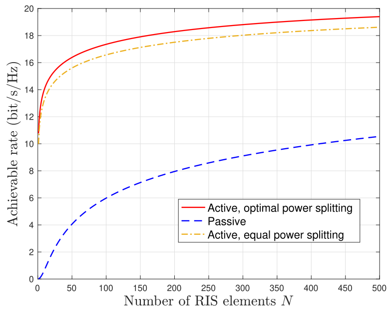

Fig. 2. Rate comparison versus N.

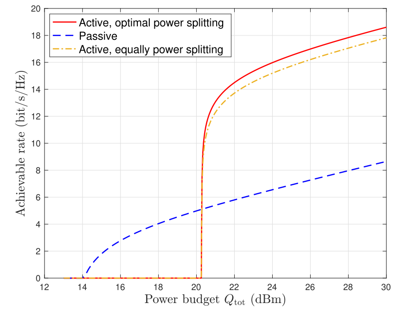

Fig. 3. Rate comparison versus power budget Q_{tot}.

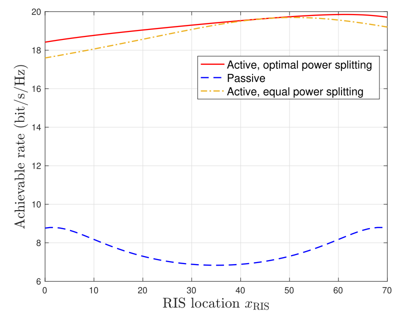

Fig. 4. Rate comparison versus RIS location x_{RIS}.

becomes very weak. Accordingly, large ρ^2 is feasible to amplify this weak signal to a normal strength while satisfying the RIS power constraint (8b).

IV. Simulation Results

If not specified otherwise, we set $P_{\text{DC}} = -5$ dBm, $P_{\text{SW}} = -10$ dBm [9], $\sigma^2 = \sigma_r^2 = -80$ dBm, $Q_{\text{tot}} = 30$ dBm, and $q_{ris} = \lambda/4$. Based on 3GPP UMi LoS scenarios with 5 GHz carrier frequency [14] and assuming 5 dBi antenna at the BS and the RIS [7], the path-loss parameters are given by $\beta_{sr} = -31.9$ dB, $\beta_{rd} = -36.9$ dB, and $\alpha_{sr} = \alpha_{rd} = 2.2$. The BS, the RIS, and the user are located at $(0,0,0)$, $(x_{\text{RIS}}, 10\text{ m}, 2\text{ m})$, and $(70\text{ m}, 0, 0)$, respectively. $x_{\text{RIS}} = 5$ m and $N = 256$ are adopted by default.[3] Besides, an equal power splitting scheme, i.e., $P_{\text{BS}}^{\text{act}} = P_{\text{RIS}}^{\text{act}} = C/2$, is considered as the baseline.

We first validate our conclusion in Corollary 5. Based on the above simulation setup and use the sub-optimal solution $P_{\text{BS}}^{\text{act}} = C/2$, we can calculate the left- and right-hand side of (13) as $\rho_{\text{sub}}^2 = 1.3 \times 10^3$ and 2.18, respectively, which verifies the correctness of Corollary 5.

Fig. 2 illustrates the superiority of active RIS when N is not very large. This is because active RIS can exploit a small amount of power to amplify the signal attenuated after the transmission in the first hop, and therefore overcome the double path loss attenuation and significantly improve the strength of the signal received at the user. Besides, note that the active RIS requires higher hardware complexity than the passive RIS given the same N. However, it can be observed from Fig. 2 that compared with the rate achieved by passive RIS with hundreds of elements, the active RIS could achieve a higher achievable rate with only tens of elements. Therefore, the higher hardware complexity would not limit the potential of applying the active RIS.

Being consistent with Corollaries 4 and 5, Fig. 3 unveils that the passive RIS performs better for small power budgets while the active RIS is superior when the power budget is sufficient. Meanwhile, it can be seen that the equal splitting strategy begins to show its defects as Q_{tot} increases. This is because as the power budget grows, the optimal splitting ratio between the BS and active RIS needs to be adjusted accordingly, as discussed in Corollary 3.

Fig. 4 unveils the superiority of active RIS in all deployment locations. The passive RIS should be deployed near the BS or the user in order to alleviate the attenuation from double path loss. By contrast, the active RIS can achieve promising performance in all locations since the double-fading effect is effectively circumvented thanks to the integration of amplifiers. Meanwhile, as expected in Theorem 1, the equal splitting scheme is optimal only in the case of $\sigma_r^2 h_{rd}^2 = \sigma^2 h_{sr}^2$.

Finally, it is worth noting that Fig. 2 - 4 all show that the equal power splitting scheme is a high-quality sub-optimal solution. This is because this scheme could satisfy the fairness requirement in Corollary 2. Since the equal power splitting scheme has low implementation complexity, our results validate the feasibility of applying active RIS in practical communication systems.

[3]The far-field channel models are valid when the communication distance is smaller than Fraunhofer distance $2D^2/\lambda$, where $D = \sqrt{2N}q_{ris}$ is the maximal aperture of the RIS with a square array [15, (134)]. Accordingly, the far-field assumption holds for $N < 757$.

V. Conclusion

This letter theoretically compared the active RIS with the passive RIS under the same power budget. We firstly derived the optimal power splitting ratio between the BS and the active RIS. We then provided some insights based on the derived splitting ratio and discussed the conditions when active or passive RIS performs better.

Appendix A

The first-order derivative of γ_{act} in (10a) with respect to $P_{\text{BS}}^{\text{act}}$ is given by

$$\frac{\partial \gamma_{\text{act}}}{\partial P_{\text{BS}}^{\text{act}}} = \frac{N h_{sr}^2 h_{rd}^2 \times f_1\left(P_{\text{BS}}^{\text{act}}\right)}{\left(\sigma_r^2 h_{rd}^2 \left(C - P_{\text{BS}}^{\text{act}}\right) + \sigma^2 \left(P_{\text{BS}}^{\text{act}} h_{sr}^2 + \sigma_r^2\right)\right)^2}, \quad (16)$$

where

$$\begin{aligned} & f_1\left(P_{\text{BS}}^{\text{act}}\right) \\ & \quad = \left(\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2\right)\left(P_{\text{BS}}^{\text{act}}\right)^2 - 2\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right) P_{\text{BS}}^{\text{act}} \\ & \qquad + C\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right). \end{aligned} \quad (17)$$

If $\sigma_r^2 h_{rd}^2 = \sigma^2 h_{sr}^2$, $f_1\left(P_{\text{BS}}^{\text{act}}\right)$ is a linear function with $f_1(C/2) = 0$. Clearly, the optimal solution is $\left(P_{\text{BS}}^{\text{act}}\right)^\star = C/2$.

When $\sigma_r^2 h_{rd}^2 \neq \sigma^2 h_{sr}^2$, $f_1\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)$ is a quadratic function with $f_1(0) = C\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right) > 0$ and $f_1(C) = -C\left(C\sigma^2 h_{sr}^2 + \sigma^2\sigma_r^2\right) < 0$. Therefore, there must exist one and only one root $\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\mathrm{rt}}$ for $f_1(P_{\mathrm{BS}}^{\mathrm{act}})$ within $(0, C)$. When $0 \leq P_{\mathrm{BS}}^{\mathrm{act}} < \left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\mathrm{rt}}$, $f_1\left(P_{\mathrm{BS}}^{\mathrm{act}}\right) > 0$. When $\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\mathrm{rt}} < P_{\mathrm{BS}}^{\mathrm{act}} \leq C$, $f_1\left(P_{\mathrm{BS}}^{\mathrm{act}}\right) < 0$. Accordingly, based on (16), γ_{act} is maximized at $\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star} = \left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\mathrm{rt}}$. We next derive the root $\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\mathrm{rt}}$ which locates in $(0, C)$. After some simplifications, two roots of $f_1\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)$ are given by

$$\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\mathrm{RTs}} = \frac{1}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2} \times \Big\{C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2 \pm\sqrt{\left(C\sigma^2 h_{sr}^2 + \sigma^2\sigma_r^2\right)\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right)}\Big\}. \tag{18}$$

If $\sigma_r^2 h_{rd}^2 > \sigma^2 h_{sr}^2$, we have $C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2 = \sqrt{(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2)(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2)} > \sqrt{\left(C\sigma^2 h_{sr}^2 + \sigma^2\sigma_r^2\right)\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right)}$. Thus, both two roots are positive, and the left root is the smaller one located in $(0, C)$, as given in (11). If $\sigma_r^2 h_{rd}^2 < \sigma^2 h_{sr}^2$, we have $C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2 < \sqrt{\left(C\sigma^2 h_{sr}^2 + \sigma^2\sigma_r^2\right)\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right)}$. In this case, the left root is negative but the right root is positive. Hence, the right root, written in (11), is the solution.

APPENDIX B

The first-order derivative of $\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}$ with respect to C is

$$\frac{\partial\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{\partial C} = \frac{\sigma_r^2 h_{rd}^2 - \sigma^2\sigma_r^2\sqrt{f_2(C)}}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2}, \tag{19}$$

where $f_2(C) = \frac{\left(2Ch_{sr}^2 h_{rd}^2 + \sigma^2 h_{sr}^2 + \sigma_r^2 h_{rd}^2\right)^2}{4\left(C\sigma^2 h_{sr}^2 + \sigma^2\sigma_r^2\right)\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right)}$ and

$$\frac{\partial f_2(C)}{\partial C} = \frac{4\left(2Ch_{sr}^2 h_{rd}^2 + \sigma^2 h_{sr}^2 + \sigma_r^2 h_{rd}^2\right)\sigma^2\sigma_r^2}{\left\{4\left(C\sigma^2 h_{sr}^2 + \sigma^2\sigma_r^2\right)\left(C\sigma_r^2 h_{rd}^2 + \sigma^2\sigma_r^2\right)\right\}^2} \times\left\{2\sigma^2\sigma_r^2 h_{sr}^2 h_{rd}^2 - \left(\left(\sigma^2 h_{sr}^2\right)^2 + \left(\sigma_r^2 h_{rd}^2\right)^2\right)\right\} \overset{(b)}{<} 0, \tag{20}$$

where (b) utilizes the inequality $x^2 + y^2 \geq 2xy$ and $\sigma_r^2 h_{rd}^2 \neq \sigma^2 h_{sr}^2$ as assumed before. Therefore, the function in (19) monotonously increases (decreases) with C if $\sigma_r^2 h_{rd}^2 > \sigma^2 h_{sr}^2$ ($\sigma_r^2 h_{rd}^2 < \sigma^2 h_{sr}^2$). Besides, we have $\lim\limits_{C\to 0}\frac{\partial\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{\partial C} \to \frac{1}{2} > 0$ and $\lim\limits_{C\to\infty}\frac{\partial\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{\partial C} \to \frac{\sigma_r^2 h_{rd}^2 - \sqrt{\sigma^2\sigma_r^2 h_{sr}^2 h_{rd}^2}}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2} > 0$. Due to the monotonicity, there must be $\frac{\partial\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{\partial C} > 0, \forall C > 0$. Similarly, using $\left(P_{\mathrm{RIS}}^{\mathrm{act}}\right)^{\star} = C - \left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}$, we can prove $\frac{\partial\left(P_{\mathrm{RIS}}^{\mathrm{act}}\right)^{\star}}{\partial C} > 0$. Since C increases with Q_{tot} but decreases with N, the proof is completed.

APPENDIX C

Define $\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star} - \left(P_{\mathrm{RIS}}^{\mathrm{act}}\right)^{\star} \triangleq f_3(C)$. Then, we have

$$\frac{\partial f_3(C)}{\partial C} = \frac{\sigma_r^2 h_{rd}^2 + \sigma^2 h_{sr}^2 - 2\sigma^2\sigma_r^2\sqrt{f_2(C)}}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2}, \tag{21}$$

and $\lim\limits_{C\to 0}\frac{\partial f_3(C)}{\partial C} \to 0$. Note that we have proved that $f_2(C)$ is a decreasing function. When $\sigma_r^2 h_{rd}^2 > \sigma^2 h_{sr}^2$, $\frac{\partial f_3(C)}{\partial C}$ is an increasing function leading to $\frac{\partial f_3(C)}{\partial C} > 0$ for $C > 0$. When $\sigma_r^2 h_{rd}^2 < \sigma^2 h_{sr}^2$, $\frac{\partial f_3(C)}{\partial C}$ is a decreasing function leading to $\frac{\partial f_3(C)}{\partial C} < 0$ for $C > 0$.

APPENDIX D

Using $P_{\mathrm{BS}}^{\mathrm{pas}} = C + NP_{\mathrm{DC}}$, we have $\frac{\partial}{\partial C}\frac{\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{P_{\mathrm{BS}}^{\mathrm{pas}}} = \frac{f_4(C)}{(C + NP_{\mathrm{DC}})^2}$ where $f_4(C) = (C + NP_{\mathrm{DC}})\left\{\frac{\partial}{\partial C}\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}\right\} - \left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}$ with $f_4(0) = \frac{NP_{\mathrm{DC}}}{2}$. Besides, we have $\frac{\partial}{\partial C}f_4(C) = (C + NP_{\mathrm{DC}})\left\{\frac{\partial^2}{\partial C^2}\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}\right\}$. As proved in Appendix B, $\frac{\partial^2}{\partial C^2}\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star} > 0$ if $\sigma_r^2 h_{rd}^2 > \sigma^2 h_{sr}^2$, which leads to $f_4(C) > f_4(0) > 0$ and $\frac{\partial}{\partial C}\frac{\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{P_{\mathrm{BS}}^{\mathrm{pas}}} > 0$. By contrast, if $\sigma_r^2 h_{rd}^2 < \sigma^2 h_{sr}^2$, we have $\frac{\partial^2}{\partial C^2}\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star} < 0$ and then $f_4(C)$ is a decreasing function. As $C \to \infty$, we have

$$f_4(\infty) \to \frac{\sigma^2\sigma_r^2\left(\frac{\sigma^2 h_{sr}^2 + \sigma_r^2 h_{rd}^2}{2\sqrt{\sigma^2 h_{sr}^2 \sigma_r^2 h_{rd}^2}} - 1\right)}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2} + NP_{\mathrm{DC}}\frac{\sigma_r^2 h_{rd}^2 - \sqrt{\sigma^2\sigma_r^2 h_{sr}^2 h_{rd}^2}}{\sigma_r^2 h_{rd}^2 - \sigma^2 h_{sr}^2}, \tag{22}$$

where the first term is negative while the second term is positive. Thus, when N is larger than a threshold, $f_4(C) > f_4(\infty) > 0$ and then $\frac{\partial}{\partial C}\frac{\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{P_{\mathrm{BS}}^{\mathrm{pas}}} > 0$. Otherwise, $f_4(C)$ decreases from positive value $f_4(0)$ to negative value $f_4(\infty)$ which means that $\frac{\left(P_{\mathrm{BS}}^{\mathrm{act}}\right)^{\star}}{P_{\mathrm{BS}}^{\mathrm{pas}}}$ firstly increases but then decreases.

REFERENCES

[1] M. D. Renzo *et al.*, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[2] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[3] M. Najafi *et al.*, "Physics-based modeling and scalable optimization of large intelligent reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2673–2691, Apr. 2021.

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

[5] C. Huang *et al.*, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[6] C. Pan *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[7] E. Björnson, O. Özdogan, and E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

[8] Z. Zhang *et al.*, "Active RIS vs. passive RIS: Which will prevail in 6G?" 2021, *arXiv:2103.15154*.

[9] R. Long *et al.*, "Active reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4962–4975, Aug. 2021.

[10] 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.

[11] M. H. Khoshafa *et al.*, "Active reconfigurable intelligent surfaces-aided wireless communication system," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3699–3703, Nov. 2021.

[12] K. Liu, Z. Zhang, L. Dai, S. Xu, and F. Yang, "Active reconfigurable intelligent surface: Fully-connected or sub-connected?" *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 167–171, Jan. 2022.

[13] G. Chen *et al.*, "Active IRS aided multiple access for energy-constrained IoT systems," 2022, *arXiv:2201.12565*.

[14] *Further Advancements for E-UTRA Physical Layer Aspects (Release 9)*, 3GPP, document TR36.814, Aug. 2010.

[15] C. Pan *et al.*, "An overview of signal processing techniques for RIS/IRS-aided wireless systems," 2021, *arXiv:2112.05989*.