Double-RIS Communication with DF Relaying for Coverage Extension: Is One Relay Enough?

Zaid Abdullah, Steven Kisseleff, Konstantinos Ntontin, Wallace Alves Martins,
Symeon Chatzinotas, and Björn Ottersten

Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg.
E-mails: {zaid.abdullah, steven.kisseleff, kostantinos.ntontin, wallace.alvesmartins,
symeon.chatzinotas, bjorn.ottersten}@uni.lu

Abstract—In this work, we investigate the decode-and-forward (DF) relay-aided double reconfigurable intelligent surface (RIS)-assisted networks, where the signal is subject to reflections from two RISs before reaching the destination. Different relay-aided network architectures are considered for maximum achievable rate under a total power constraint. Phase optimization for the double-RIS channels is tackled via the alternating optimization and majorization-minimization (MM) schemes. Moreover, closed-form solutions are obtained for each case. Numerical results indicate that the deployment of two relays, one near each RIS, achieves higher rates at low and medium signal-to-noise ratios (SNRs) compared to placing a single relay between the two RISs; while at high SNRs, the latter approach achieves higher rates only if the inter-relay interference for the former case is considerably high.

Index Terms—Reconfigurable Intelligent surface, decode-and-forward, multihop relaying.

I. Introduction

Conventional active relaying, such as decode-and-forward (DF), is a well known technology that is used to extend the coverage, and/or to enhance the quality-of-service (QoS) between a pair of transceiving nodes [1]. Ideally, the locations and/or number of relays should be optimized based on a certain cost function, such as to maximize the rate or to minimize the transmit power while satisfying a given QoS constraint [2], [3].

In contrast, the reconfigurable intelligent surface (RIS) technology is a new concept in wireless communications, where a large number of low-cost, nearly-passive reflecting elements are utilized to direct the impinging signal toward a desired destination, such that the multiple signal paths are constructively combined at the receiver [4]. One of the most attractive aspects about RISs is that they do not require power-demanding active radio-frequency chains. Another benefit compared to traditional active relaying is that RISs work on-the-fly, i.e. they do not introduce additional delays due to internal signal processing. Thanks to their low-cost and low power-consumption, it is highly anticipated that RISs will have a key role in future wireless networks [5]–[7].

However, due to the large degradation of the signal power with distance, which is caused by the absence of active amplification at the RIS and the double path-loss, few studies have shown that considerably large surfaces are required

to outperform a conventional single-antenna relay [8], [9]. Motivated by this fact, many researchers started adopting classical relays, such as DF or amplify-and-forward (AF), to enhance the performance of RIS-assisted networks [10]–[17]. However, in all these works, only a single relay was utilized to enhance the RIS-assisted transmission. Moreover, even though the work in [17] considers three RISs, only one of them was deployed near the single relay, whereas the other two were located within short distances of the source and destination.

However, in many real-world scenarios, the signal might go through multiple hops before reaching the destination. Therefore, we aim to find the optimal way of combining relays with RISs when there is more than a single RIS between the two transceiving nodes. In particular, we consider the double-RIS reflection case where the signal is subject to reflections from two RISs before reaching the destination, and we propose three different half-duplex (HD) relay-aided network architectures, and compare their effective rates under a total power constraint. The findings of this work can help understand how to perform optimal route optimization, relay placement, and RIS-relay pairing for future multihop RIS-relay assisted networks.

Notations: Matrices and vectors are denoted by boldface uppercase and lowercase letters, respectively. $\boldsymbol{x}^T$, $\boldsymbol{x}^*$, and $\|\boldsymbol{x}\|$ are the transpose, conjugate, and Euclidean norm of a vector $\boldsymbol{x}$, respectively, and $[\boldsymbol{x}]_i$ is the ith element of $\boldsymbol{x}$. $|x|$ and $\angle(x)$ are the absolute value and the phase of a complex number x, respectively. $\mathbb{E}\{a\}$ is the expected value of a, while $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. $\text{diag}\{\boldsymbol{x}\}$ is a diagonal matrix whose diagonal contains the elements of $\boldsymbol{x}$, while $\text{diag}\{\boldsymbol{X}\}$ is a vector whose elements are the diagonal of $\boldsymbol{X}$. Finally, $\Re\{x\}$ denotes the real part of a complex number x.

II. System Model and Phase Optimization

We consider a time division duplex scenario where there is a single-antenna source (S) aiming to transmit a signal to a single-antenna destination (D) with the help of two RISs (I_1 and I_2). Due to large distances, obstacles and path-loss, we assume that S has a direct link with only I_1, and similarly D has a direct link with only I_2.

To enhance the performance of the double-RIS channel, we deploy a single-antenna HD-DF relay(s) between the two

ends. In particular, we investigate three different relay-aided scenarios. The first one corresponds to the case of a single relay (R) present between I_1 and I_2, and the communication takes place over two time-slots. In the second scenario, we assume that there are two relays R_1 and R_2, placed near I_1 and I_2, respectively, and only a single node in the network can transmit at any given time-instant. In the last scenario, we aim to enhance the second scenario by allowing the second relay R_2 to transmit to D while S transmits its data to R_1. Note that throughout this work, we assume perfect channel estimation for all links.[1] We also assume centralized processing for the different double-RIS communication schemes. Fig. 1 shows the different relay-aided double-RIS network configurations, where we compare with the no-relay scenario as a benchmark scheme. We next start formulating the received signals and corresponding achievable rates for each scenario.

A. Transmission through only the RISs

In this scenario, we assume that the transmission is realized through the two RISs only. Therefore, the received signal at the destination can be written as

$$y_D^{(1)}(n) = \sqrt{p}\left(\boldsymbol{h}_{I_2D}^T \boldsymbol{\Phi G \Theta h}_{I_1S}\right)x_s(n) + w_D(n), \qquad (1)$$

where the superscript in $y_D^{(1)}$ indicates that this is a single-hop transmission, n is the time index, p is the total transmit power at any given time-instant n, $\boldsymbol{h}_{I_1S} \in \mathbb{C}^M$ and $\boldsymbol{h}_{I_2D} \in \mathbb{C}^M$ are the channels between $S \rightarrow I_1$, and $I_2 \rightarrow D$, respectively,[2] and M is the number of reflecting elements at each RIS. $\boldsymbol{G} \in \mathbb{C}^{M \times M}$ is the channel between the two RISs; while $\boldsymbol{\Theta} = \text{diag}\{\boldsymbol{\theta}\} \in \mathbb{C}^{M \times M}$ and $\boldsymbol{\Phi} = \text{diag}\{\boldsymbol{\phi}\} \in \mathbb{C}^{M \times M}$ are the reflection matrices for I_1 and I_2, respectively. x_s is the information symbol transmitted from S with $\mathbb{E}\{|x_s|^2\} = 1$, and $w_D \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN) at the destination. Therefore, the received signal-to-noise ratio (SNR) at the destination is given as

$$\gamma_D^{(1)} = \rho \left|\boldsymbol{h}_{I_2D}^T \boldsymbol{\Phi G \Theta h}_{I_1S}\right|^2, \qquad (2)$$

where $\rho = p/\sigma^2$. The achievable rate in this case is

$$\mathcal{R}^{(1)} = \log_2\left(1 + \gamma_D^{(1)}\right). \qquad (3)$$

Clearly, the achievable rate depends on both $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$. However, to optimize the achievable rate, we first need to reformulate the cascaded channel as follows:

$$\left|\boldsymbol{h}_{I_2D}^T \boldsymbol{\Phi G \Theta h}_{I_1S}\right|^2 = \left|\boldsymbol{\phi}^T \text{diag}\{\boldsymbol{h}_{I_2D}\}\boldsymbol{G}\, \text{diag}\{\boldsymbol{h}_{I_1S}\}\boldsymbol{\theta}\right|^2$$
$$= \left|\boldsymbol{\phi}^T \boldsymbol{F \theta}\right|^2. \qquad (4)$$

Now we can present the following optimization problem

$$\underset{\boldsymbol{\theta}, \boldsymbol{\phi}}{\text{maximize}} \quad \rho \left|\boldsymbol{\phi}^T \boldsymbol{F \theta}\right|^2 \qquad (5)$$

subject to

$$\left|[\boldsymbol{\theta}]_m\right| = 1, \qquad \left|[\boldsymbol{\phi}]_m\right| = 1, \quad \forall m \in \mathcal{M}, \qquad (5a)$$

[1] Note that the authors in [18], [19] have proposed channel estimation schemes for the double-RIS channels with satisfactory estimation accuracy.

[2] The different narrowband fading channels adopted in this work will be explained in detail in Section III.
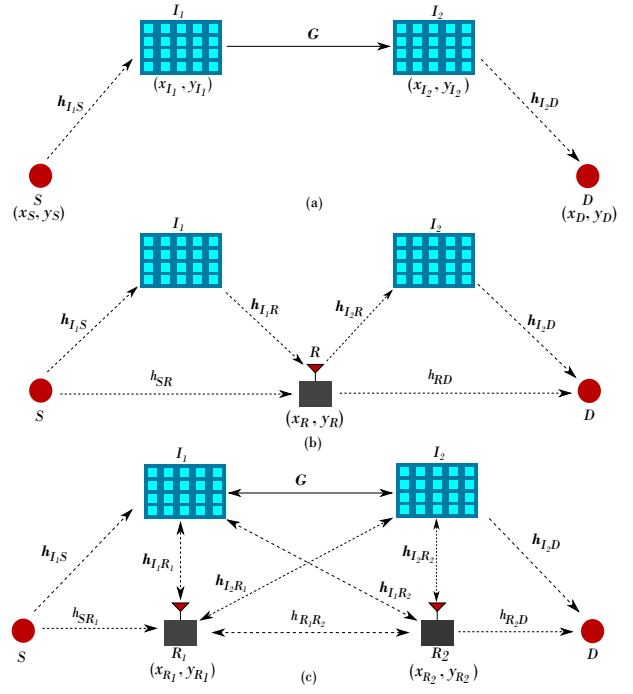

Fig. 1. The considered relay-aided network architectures: (a) Communication through two RISs only, (b) Communication through two RISs and a single relay, and (c) Communication through two RISs and two relays.

where $\mathcal{M} = \{1, 2, ..., M\}$ is the set of all reflecting elements at each RIS. The optimization problem in (5) is non-convex, due to the unit-modulus constraint and the coupled optimization variables. Therefore, we adopt an alternating approach where we fix $\boldsymbol{\phi}$ to solve for $\boldsymbol{\theta}$, and vice-versa.

1) Optimizing $\boldsymbol{\theta}$ for a given $\boldsymbol{\phi}$: Let $\boldsymbol{r}^T = \boldsymbol{\phi}^T \boldsymbol{F} \in \mathbb{C}^{1 \times M}$, now we can formulate the following optimization problem

$$\underset{\boldsymbol{\theta}}{\text{maximize}} \quad \rho \left|\boldsymbol{r}^T \boldsymbol{\theta}\right|^2 \qquad (6)$$

$$\text{subject to} \quad \left|[\boldsymbol{\theta}]_m\right| = 1, \quad \forall m \in \mathcal{M}. \qquad (6a)$$

The solution to the above optimization problem can be given in a closed-form as follows:

$$[\boldsymbol{\theta}^\star]_m = \exp\left(j\angle\left([\boldsymbol{r}]_m\right)\right)^*. \qquad (7)$$

2) Optimizing $\boldsymbol{\phi}$ for a given $\boldsymbol{\theta}$: Similar to the previous approach, and after defining $\boldsymbol{v} = \boldsymbol{F \theta} \in \mathbb{C}^M$, we can write $\left|\boldsymbol{\phi}^T \boldsymbol{F \theta}\right| = \left|\boldsymbol{\phi}^T \boldsymbol{v}\right|$. Therefore, the solution to $\boldsymbol{\phi}$ is

$$[\boldsymbol{\phi}^\star]_m = \exp\left(j\angle\left([\boldsymbol{v}]_m\right)\right)^*. \qquad (8)$$

We alternate the optimization process between $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ until the increment in the achievable rate between two successive iterations falls below a certain threshold, or we reach a maximum number of optimization iterations.

B. Communication through RISs and a single relay

In this case, we assume that there is one HD-DF relay (R), placed in the middle between I_1 and I_2,[3] and the transmission

[3] The choice of placing the single relay between I_1 and I_2 comes intuitively to balance the SNRs of first and second hops.

is carried out through two time-slots.

1) First-Hop: In the first time-slot, S transmits its data to R through the direct link and the reflected signal from I_1. Therefore, the received signal at R is given as

$$y_R^{(2)}(n) = \sqrt{p}\Big(h_{SR} + \boldsymbol{h}_{I_1R}^T\boldsymbol{\Theta}\boldsymbol{h}_{I_1S}\Big)x_s(n) + w_R(n), \quad (9)$$

where the superscript in $y_R^{(2)}$ indicates that there are two hops in this case, $\boldsymbol{h}_{I_1R} \in \mathbb{C}^M$ is the channel vector between I_1 and R, $h_{SR} \in \mathbb{C}$ is the channel between S and R, and $w_R \sim \mathcal{CN}(0, \sigma^2)$ is the AWGN at R. To maximize the received SNR at R,[4] the phase-shifts of $\boldsymbol{\Theta}$ should be selected as follows

$$[\boldsymbol{\Theta}^\star]_{m,m} = \exp\Big(j\Big(\angle(h_{SR}) - \angle([\boldsymbol{h}_{I_1R}]_m[\boldsymbol{h}_{I_1S}]_m)\Big)\Big), \quad (10)$$

$\forall m \in \mathcal{M}$. Then, the received SNR at R with optimal phase-shifts can be expressed as

$$\gamma_R^{(2)} = \rho\Big(|h_{SR}| + \sum_{m\in\mathcal{M}}\big|[\boldsymbol{h}_{I_1R}]_m[\boldsymbol{h}_{I_1S}]_m\big|\Big)^2. \quad (11)$$

2) Second-Hop: During the second time-slot, the relay re-transmits the signal, with power p, to the destination through the direct link and I_2. Assuming successful decoding of x_s at the relay, the received signal at D can be given as

$$y_D^{(2)}(n+1) = \sqrt{p}\Big(h_{RD} + \boldsymbol{h}_{I_2D}^T\boldsymbol{\Phi}\boldsymbol{h}_{I_2R}\Big)x_s(n) + w_D(n+1), \quad (12)$$

where $h_{RD} \in \mathbb{C}$ and $\boldsymbol{h}_{I_2R} \in \mathbb{C}^M$ are the channels between $R \to D$ and $R \to I_2$, respectively, and $w_D \sim \mathcal{CN}(0, \sigma^2)$ is the AWGN at D. Assuming perfect phase-shifts at I_2 for $\boldsymbol{\Phi}$, the received SNR at D is

$$\gamma_D^{(2)} = \rho\Big(|h_{RD}| + \sum_{m\in\mathcal{M}}\big|[\boldsymbol{h}_{I_2D}]_m[\boldsymbol{h}_{I_2R}]_m\big|\Big)^2, \quad (13)$$

and the corresponding achievable rate is

$$\mathcal{R}^{(2)} = \frac{1}{2}\log_2\Big(1 + \min\{\gamma_R^{(2)},\ \gamma_D^{(2)}\}\Big), \quad (14)$$

and the $(\frac{1}{2})$ pre-log factor is due to the two-hop transmission.

C. Communication through RISs and two relays

In this case, we assume that there are two HD-DF relays, R_1 and R_2, to assist the communication between S and D. In particular, we assume that R_1 is placed in close proximity to I_1, while R_2 is placed near I_2. Furthermore, the transmission takes place over three time-slots, since we assume that only one node can transmit at any given time-instant.

1) First-Hop: In the first-hop, S transmits its signal to R_1 through the direct link and I_1. Therefore, the received signal at R_1 is

$$y_{R_1}^{(3,3)}(n) = \sqrt{p}\Big(h_{SR_1} + \boldsymbol{h}_{I_1R_1}^T\boldsymbol{\Theta}\boldsymbol{h}_{I_1S}\Big)x_s(n) + w_{R_1}(n), \quad (15)$$

where the superscript in $y_{R_1}^{(3,3)}$ indicates that there are three hops and S transmits a new block of data every three transmission time-slots, $\boldsymbol{h}_{I_1R_1} \in \mathbb{C}^M$ is the channel vector

[4]Note that there exists another path from $S \to I_1 \to I_2 \to R$, however, this path is neglected here as the received signal through $S \to I_1 \to R$ will be dominant due to shorter travel distance and less reflections.

between I_1 and R_1, and $w_{R_1} \sim \mathcal{CN}(0, \sigma^2)$ is the AWGN at R_1. Assuming perfect phase-optimization for $\boldsymbol{\Theta}$ at I_1, the received SNR at R_1 is

$$\gamma_{R_1}^{(3,3)} = \rho\Big(|h_{SR_1}| + \sum_{m\in\mathcal{M}}\big|[\boldsymbol{h}_{I_1R_1}]_m[\boldsymbol{h}_{I_1S}]_m\big|\Big)^2. \quad (16)$$

2) Second-Hop: During the second-hop, R_1 re-transmits the signal, assuming successful decoding, to R_2 through the direct link, direct reflection links from both I_1 and I_2, as well as double-reflection link. Therefore, and assuming perfect decoding of x_s at R_1, the received signal at R_2 can be given as shown in (17) at the top of the next page, where we have $\boldsymbol{h}_{I_1R_2} \in \mathbb{C}^M$, $\boldsymbol{h}_{I_2R_1} \in \mathbb{C}^M$, $\boldsymbol{h}_{I_2R_2} \in \mathbb{C}^M$ and $h_{R_1R_2} \in \mathbb{C}$ are the channels between $I_1 \to R_2$, $I_2 \to R_1$, $I_2 \to R_2$ and $R_1 \to R_2$, respectively. $\boldsymbol{\Psi}_1 \in \mathbb{C}^{M\times M}$ and $\boldsymbol{\Psi}_2 \in \mathbb{C}^{M\times M}$ are the reflection matrices for I_1 and I_2, respectively, during the second-hop, and $w_{R_2} \sim \mathcal{CN}(0, \sigma^2)$ is the AWGN at R_2.

Let $\boldsymbol{Q} = \mathrm{diag}\{\boldsymbol{h}_{I_2R_2}\}\boldsymbol{G}\mathrm{diag}\{\boldsymbol{h}_{I_1R_1}\} \in \mathbb{C}^{M\times M}$, $\boldsymbol{u}_{I_1} = \mathrm{diag}\{\boldsymbol{h}_{I_1R_2}\}\boldsymbol{h}_{I_1R_1} \in \mathbb{C}^M$, $\boldsymbol{u}_{I_2} = \mathrm{diag}\{\boldsymbol{h}_{I_2R_2}\}\boldsymbol{h}_{I_2R_1} \in \mathbb{C}^M$, $\boldsymbol{\psi}_1 = \mathrm{diag}\{\boldsymbol{\Psi}_1\} \in \mathbb{C}^M$, and $\boldsymbol{\psi}_2 = \mathrm{diag}\{\boldsymbol{\Psi}_2\} \in \mathbb{C}^M$. Then, the received SNR at R_2 can be written as follows:

$$\gamma_{R_2}^{(3,3)} = \rho\big|h_{R_1R_2} + \boldsymbol{\psi}_1^T\boldsymbol{u}_{I_1} + \boldsymbol{\psi}_2^T\boldsymbol{u}_{I_2} + \boldsymbol{\psi}_2^T\boldsymbol{Q}\boldsymbol{\psi}_1\big|^2. \quad (18)$$

Clearly the SNR depends on both $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_2$, therefore, we can formulate the following optimization problem

$$\underset{\boldsymbol{\psi}_1,\ \boldsymbol{\psi}_2}{\text{maximize}} \quad \rho\big|h_{R_1R_2} + \boldsymbol{\psi}_1^T\boldsymbol{u}_{I_1} + \boldsymbol{\psi}_2^T\boldsymbol{u}_{I_2} + \boldsymbol{\psi}_2^T\boldsymbol{Q}\boldsymbol{\psi}_1\big|^2 \quad (19)$$

subject to

$$\big|[\boldsymbol{\psi}_i]_m\big| = 1, \quad \forall m \in \mathcal{M}, \quad i \in \{1, 2\}, \quad (19a)$$

the above optimization problem is also non-convex. Accordingly, we adopt an alternating optimization scheme where we fix one of the optimization variables and solve for the other one. In particular, and for a given $\boldsymbol{\psi}_2$, we have $\gamma_{R_2}^{(3,3)} = \rho\big|\boldsymbol{z}^T\boldsymbol{\psi}_1 + c\big|^2$, where $\boldsymbol{z}^T = \big(\boldsymbol{u}_{I_1}^T + \boldsymbol{\psi}_2^T\boldsymbol{Q}\big)$, and $c = \big(\boldsymbol{\psi}_2^T\boldsymbol{u}_{I_2} + h_{R_1R_2}\big)$. Therefore, it is straightforward to see that the optimal phase-shift for the mth element of $\boldsymbol{\psi}_1$ is

$$[\boldsymbol{\psi}_1^\star]_m = \exp\Big(j\Big(\angle(c) - \angle([\boldsymbol{z}]_m)\Big)\Big). \quad (20)$$

Similarly, to optimize the phase-shifts of $\boldsymbol{\psi}_2$, we can fix $\boldsymbol{\psi}_1$ to obtain $\gamma_{R_2}^{(3,3)} = \rho\big|\boldsymbol{v}^T\boldsymbol{\psi}_2 + r\big|^2$, where $\boldsymbol{v} = \big(\boldsymbol{u}_{I_2} + \boldsymbol{Q}\boldsymbol{\psi}_1\big)$, and $r = \big(h_{R_1R_2} + \boldsymbol{\psi}_1^T\boldsymbol{u}_{I_1}\big)$. It follows that the optimal phase-shift for the mth element of $\boldsymbol{\psi}_2$ is

$$[\boldsymbol{\psi}_2^\star]_m = \exp\Big(j\Big(\angle(r) - \angle([\boldsymbol{v}]_m)\Big)\Big). \quad (21)$$

3) Third-Hop: After receiving the information, R_2 will decode the message and then will retransmit it to D through the direct link and I_2. Assuming successful decoding at R_2, the received signal at the destination can be written as

$$y_D^{(3,3)}(n+2) = \sqrt{p}\Big(h_{R_2D} + \boldsymbol{h}_{I_2D}^T\boldsymbol{\Phi}\boldsymbol{h}_{I_2R_2}\Big)x_s(n) + w_D(n+2), \quad (22)$$

$$y_{R_2}^{(3,3)}(n+1) = \sqrt{p}\Big(h_{R_1R_2} + \boldsymbol{h}_{I_1R_2}^T \boldsymbol{\Psi}_1 \boldsymbol{h}_{I_1R_1} + \boldsymbol{h}_{I_2R_2}^T \boldsymbol{\Psi}_2 \boldsymbol{h}_{I_2R_1} + \boldsymbol{h}_{I_2R_2}^T \boldsymbol{\Psi}_2 \boldsymbol{G} \boldsymbol{\Psi}_1 \boldsymbol{h}_{I_1R_1}\Big) x_s(n) + w_{R_2}(n+1). \tag{17}$$

and the corresponding received SNR at D, assuming perfect phase-optimization at I_2 for $\boldsymbol{\Phi}$, can be expressed as

$$\gamma_D^{(3,3)} = \rho\Big(\big|h_{R_2D}\big| + \sum_{m\in\mathcal{M}}\big|[\boldsymbol{h}_{I_2D}]_m[\boldsymbol{h}_{I_2R_2}]_m\big|\Big)^2, \tag{23}$$

and the corresponding achievable rate can be written as

$$\mathcal{R}^{(3,3)} = \frac{1}{3}\log_2\Big(1 + \min\{\gamma_{R_1}^{(3,3)}, \gamma_{R_2}^{(3,3)}, \gamma_D^{(3,3)}\}\Big), \tag{24}$$

where the spectral efficiency is reduced by a factor of 3 since a new block of data is transmitted every three time-slots.

D. Enhanced transmission with RISs and two-relays

The main setback for the previous scenario is the $\frac{1}{3}$ pre-log factor, which can be costly at high SNRs. Therefore, we further present an enhanced transmission scheme such that while R_2 is transmitting its signal to D, S transmits a new data packet to R_1. Note that for the first two time-slots ($n \in \{1, 2\}$), all equations in the previous subsection hold in terms of SNRs and phase-optimization; as for the subsequent frames (i.e. when $n > 2$), the received signals and transmit powers will change as will be thoroughly explained here.

1) Received signal at R_1: At a given odd-time instant n_o ($n_o \geq 3$), both S and R_2 transmit data to R_1 and D, respectively. Our focus here is on the received signal at R_1.

Clearly, R_1 will receive an interfering signal from R_2 in addition to the desired signal from S, as shown in (25) at the top of the next page, where the superscript in $y_{R_1}^{(3,2)}$ denotes that there are 3 hops, and S transmits a new data packet every two time-slots, p_1 and p_2 are the transmit powers at S and R_2, respectively, with $p_1 + p_2 = p$ to maintain the total transmit power budget, and $\boldsymbol{\Theta}$ and $\boldsymbol{\Phi}$ are the reflection matrices at I_1 and I_2, respectively. The 2nd term in (25) represents the interference from R_2, which can be canceled in different ways. For example, if a global channel state information (CSI) is available, then R_1 can cancel this interference perfectly (assuming perfect channel estimation), since the signal transmitted from R_2, i.e. $x_s(n_o - 2)$, can be viewed as the signal that R_1 transmitted in the previous time-slot. Otherwise, R_1 can estimate the overall effective channel between itself and R_2. This can be performed according to the maximum-likelihood estimation by multiplying the received signal at R_1 with the conjugate of the transmitted signal from R_1 at time $n_o - 1$ (which is $x_s(n_o - 2)^*$), as shown in (26).[5] After performing interference cancellation at R_1, we can rewrite the received signal in (25) as follows:

$$y_{R_1}^{(3,2)}(n_o) = \sqrt{p_1}\Big(h_{SR_1} + \boldsymbol{h}_{I_1R_1}^T \boldsymbol{\Theta} \boldsymbol{h}_{I_1S}\Big)x_s(n_o)$$
$$+ h_e x_s(n_o - 2) + w_{R_1}(n_o), \tag{27}$$

where $h_e = \sqrt{p_2}(h_2 - \hat{h}_2)$ is the residual interference cancellation error at R_1, which is usually assumed to follow

normal distribution such that $h_e \sim \mathcal{CN}(0, \sigma_e^2)$. Assuming perfect phase optimization for $\boldsymbol{\Theta}$ at I_1 to maximize the power of received signal at R_1 from S, we can formulate the received signal-to-interference plus noise ratio (SINR) at R_1 as follows:

$$\gamma_{R_1}^{(3,2)} = \frac{p_1\Big(\big|h_{SR_1}\big| + \sum_{m\in\mathcal{M}}\big|[\boldsymbol{h}_{I_1R_1}]_m[\boldsymbol{h}_{I_1S}]_m\big|\Big)^2}{\sigma_e^2 + \sigma^2}. \tag{28}$$

Next we focus on the received signal at the destination.[6]

2) Received signal at D: While S is transmitting its data to R_1, R_2 transmits the decoded signal from R_1 in the previous time-slot to D. The received signal at D can be expressed as

$$y_D^{(3,2)}(n_o) = \sqrt{p_2}\Big(h_{R_2D} + \boldsymbol{h}_{I_2D}^T \boldsymbol{\Phi} \boldsymbol{h}_{I_2R_2}\Big)x_s(n_o - 2)$$
$$+ \sqrt{p_1}\Big(\boldsymbol{h}_{I_2D}^T \boldsymbol{\Phi} \boldsymbol{G} \boldsymbol{\Theta} \boldsymbol{h}_{I_1S}\Big)x_s(n_o) + w_D(n_o). \tag{29}$$

Note that $x_s(n_o)$ is intended for R_1, and therefore it represents interference to D. As a result, the SINR at D is

$$\gamma_D^{(3,2)} = \frac{p_2\big|h_{R_2D} + \boldsymbol{\phi}^T \boldsymbol{a}\big|^2}{p_1\big|\boldsymbol{\phi}^T \boldsymbol{b}\big|^2 + \sigma^2}, \tag{30}$$

where $\boldsymbol{b} = \text{diag}\{\boldsymbol{h}_{I_2D}\}\boldsymbol{q}$, $\boldsymbol{a} = \text{diag}\{\boldsymbol{h}_{I_2D}\}\boldsymbol{h}_{I_2R_2}$, $\boldsymbol{\phi} = \text{diag}\{\boldsymbol{\Phi}\}$, and $\boldsymbol{q} = \boldsymbol{G}\boldsymbol{\Theta}\boldsymbol{h}_{I_1S}$. Now we can formulate the following optimization problem:

$$\underset{\boldsymbol{\phi}}{\text{minimize}} \quad u(\boldsymbol{\phi}) \triangleq \frac{p_1\big|\boldsymbol{\phi}^T \boldsymbol{b}\big|^2 + \sigma^2}{p_2\big|h_{R_2D} + \boldsymbol{\phi}^T \boldsymbol{a}\big|^2} \tag{31}$$

$$\text{subject to} \quad \big|[\boldsymbol{\phi}]_m\big| = 1, \quad \forall m \in \mathcal{M}. \tag{31a}$$

This problem belongs to fractional programming [20]. As such, we formulate the following parametric program:

$$\underset{\boldsymbol{\phi}}{\text{minimize}} \quad p_1\big|\boldsymbol{\phi}^T \boldsymbol{b}\big|^2 + \sigma^2 - \mu\Big(p_2\big|h_{R_2D} + \boldsymbol{\phi}^T \boldsymbol{a}\big|^2\Big) \tag{32}$$

$$\text{subject to} \quad \big|[\boldsymbol{\phi}]_m\big| = 1, \quad \forall m \in \mathcal{M}. \tag{32a}$$

where $\mu \geq 0$ is an introduced parameter. Although problem (32) is non-convex, it can be solved using the iterative majorization-minimization (MM) method. In particular, we can introduce the following upper-bound of (32) [21], [22]:

$$f(\boldsymbol{\phi}, \mu) \triangleq p_1\big|\boldsymbol{\phi}^T \boldsymbol{b}\big|^2 + \sigma^2 - \mu\Big(p_2\big|h_{R_2D} + \boldsymbol{\phi}^T \boldsymbol{a}\big|^2\Big)$$
$$= \boldsymbol{\phi}^T \boldsymbol{X} \boldsymbol{\phi}^* - 2\mu p_2 \Re\Big\{\boldsymbol{\phi}^T h_{R_2D}^* \boldsymbol{a}\Big\} - \mu p_2 |h_{R_2D}|^2 + \sigma^2$$
$$\leq \lambda_{\max}(\boldsymbol{X})\|\boldsymbol{\phi}\|^2 - 2\Re\Big\{\boldsymbol{\phi}^T \boldsymbol{\alpha}(\tilde{\boldsymbol{\phi}}, \mu)\Big\} + \beta(\tilde{\boldsymbol{\phi}}, \mu), \tag{33}$$

where $\boldsymbol{X} = \big(p_1 \boldsymbol{b}\boldsymbol{b}^H - \mu p_2 \boldsymbol{a}\boldsymbol{a}^H\big)$, $\lambda_{\max}(\boldsymbol{X})$ is the maximum eigenvalue of $\boldsymbol{X}$, $\boldsymbol{L} = \lambda_{\max}(\boldsymbol{X})\boldsymbol{I}_M$, $\boldsymbol{\alpha}(\tilde{\boldsymbol{\phi}}, \mu) = \big((\boldsymbol{L} - \boldsymbol{X})\tilde{\boldsymbol{\phi}}^* + \mu p_2 h_{R_2D}^* \boldsymbol{a}\big)$, $\beta(\tilde{\boldsymbol{\phi}}, \mu) = \big(\tilde{\boldsymbol{\phi}}^T (\boldsymbol{L} - \boldsymbol{X})\tilde{\boldsymbol{\phi}}^* -$

---

[5]Here we assume that this interference signal is suppressed by one of the two methods explained above. Note that another way of suppressing the interference at R_1 is through the passive beamforming at I_1 and I_2. However, we will leave this approach for investigation in our future work.

[6]Note that the received signal at R_2 in the next time-slot (i.e. at time $(n_o + 1)$) will not be affected by this enhanced transmission scheme, since only R_1 will be transmitting data to R_2 with a power budget of p while the source will be silent. Therefore, we have $\gamma_{R_2}^{(3,2)} = \gamma_{R_2}^{(3,3)}$.

$$y_{R_1}^{(3,2)}(n_o) = \sqrt{p_1} \underbrace{\left(h_{SR_1} + \boldsymbol{h}_{I_1 R_1}^T \boldsymbol{\Theta} \boldsymbol{h}_{I_1 S} \right)}_{h_1} x_s(n_o) + \sqrt{p_2} \underbrace{\left(h_{R_1 R_2} + \boldsymbol{h}_{I_1 R_1}^T \boldsymbol{\Theta} \boldsymbol{h}_{I_1 R_2} + \boldsymbol{h}_{I_2 R_1}^T \boldsymbol{\Phi} \boldsymbol{h}_{I_2 R_2} + \boldsymbol{h}_{I_1 R_1}^T \boldsymbol{\Theta} \boldsymbol{G} \boldsymbol{\Phi} \boldsymbol{h}_{I_2 R_2} \right)}_{h_2} x_s(n_o - 2)$$
$$+ w_{R_1}(n_o). \tag{25}$$

$$\mathbb{E}\{y_{R_1}^{(3,2)}(n_o) x_s(n_o - 2)^*\} = \sqrt{p_1} h_1 \underbrace{\mathbb{E}\{x_s(n_o) x_s(n_o - 2)^*\}}_{=0} + \sqrt{p_2} h_2 \underbrace{\mathbb{E}\{x_s(n_o - 2) x_s(n_o - 2)^*\}}_{=1} + \underbrace{\mathbb{E}\{w_{R_1}(n_o) x_s(n_o - 2)^*\}}_{=0} = \sqrt{p_2} h_2. \tag{26}$$

$\mu p_2 |h_{R2D}|^2 + \sigma^2$), and $\tilde{\phi}$ is the solution to ϕ in the previous iteration of the MM scheme. Accordingly, minimizing the upper-bound of (32) can be simplified as:

$$\underset{\phi}{\text{minimize}} \quad \lambda_{\max}(\boldsymbol{X}) \|\phi\|^2 - 2\Re\left\{ \phi^T \boldsymbol{\alpha}(\tilde{\phi}, \mu) \right\} \tag{34}$$
$$\text{subject to} \quad |[\phi]_m| = 1, \quad \forall m \in \mathcal{M}. \tag{34a}$$

For a given value of μ, the term $\lambda_{\max}(\boldsymbol{X}) \|\phi\|^2$ is a constant. As such, the optimal phase for the mth element at any given iteration of the MM scheme can be given as follows:

$$[\phi]_m^\star = \exp\left(j\angle([\boldsymbol{\alpha}(\tilde{\phi}, \mu)]_m) \right)^*, \tag{35}$$

and the value of μ is updated after each iteration as follows:

$$\mu = \frac{p_1 |(\phi^\star)^T \boldsymbol{b}|^2 + \sigma^2}{\left(p_2 |h_{R_2 D} + (\phi^\star)^T \boldsymbol{a}|^2 \right)}. \tag{36}$$

However, before the MM algorithm starts, we initialize μ based on any feasible solution for ϕ in (36), and then utilize both μ and ϕ to find $\boldsymbol{\alpha}(\phi, \mu)$.

Proposition: The value of $u(\phi)$ in (31) is monotonically non-increasing with k, where $k > 0$ is the number of iterations for the MM scheme.

Proof: see Appendix A.

For each iteration of the MM scheme, we find the corresponding values of μ and $\boldsymbol{\alpha}$, then we optimize the phase-shifts based on (35). The same procedure will be repeated until convergence or reaching a maximum number of iterations. It follows that the achievable rate utilizing this enhanced transmission scheme can be expressed as follows:

$$\mathcal{R}^{(3,2)} = \frac{1}{2} \log_2\left(1 + \min\left\{ \gamma_{R_1}^{(3,2)}, \gamma_{R_2}^{(3,2)}, \gamma_D^{(3,2)} \right\} \right). \tag{37}$$

III. RESULTS AND DISCUSSION

We start by introducing the wireless channels adopted in our work. All links from and to the RISs were assumed to experience Rician fading with both line-of-sight (LoS) and non-LoS (NLoS) channels. In particular, $\boldsymbol{G} = \sqrt{\frac{K}{K+1}} \boldsymbol{G}^{\text{LoS}} + \sqrt{\frac{1}{1+K}} \boldsymbol{G}^{\text{NLoS}}$, where K is the Rician factor, $\boldsymbol{G}^{\text{LoS}}$ contains the LoS channels, with each link having a deterministic absolute value of $d_{I_1 I_2}^{-\tilde{\alpha}/2}$, where $d_{I_1 I_2}$ is the distance between two RISs and $\tilde{\alpha}$ is the path-loss exponent for LoS links. In contrast, $\boldsymbol{G}^{\text{NLoS}}$ is the complex Gaussian Rayleigh fading channel, where each link has a zero mean and a variance of $d_{I_1 I_2}^{-\bar{\alpha}}$, where $\bar{\alpha}$ is the path-loss exponent
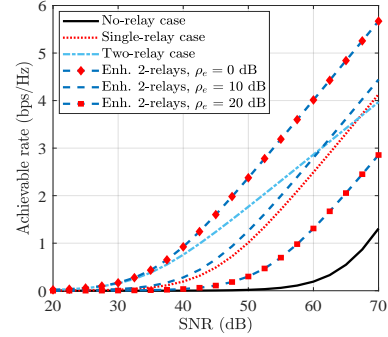

Fig. 2. Achievable rate vs transmit SNR (dB) for different relaying schemes with different levels of INR when $M = 128$.

for NLoS channels. Similarly, we have $\boldsymbol{h}_i = \sqrt{\frac{K}{K+1}} \boldsymbol{h}_i^{\text{LoS}} + \sqrt{\frac{1}{1+K}} \boldsymbol{h}_i^{\text{NLoS}}$, where each element of $\boldsymbol{h}_i^{\text{LoS}}$ has a fixed absolute value of $d_i^{-\tilde{\alpha}/2}$; while $\boldsymbol{h}_i^{\text{NLoS}} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M d_i^{-\bar{\alpha}})$, $i \in \{I_1 S, I_1 R, I_1 R_1, I_1 R_2, I_2 R, I_2 R_1, I_2 R_2, I_2 D\}$. In contrast, we assume pure Rayleigh fading between nodes that do not include any of the two RISs, such that $h_j \sim \mathcal{CN}(0, d_j^{-\bar{\alpha}})$, $j \in \{SR, SR_1, R_1 R_2, R_2 D\}$. Moreover, S was located at the origin of a 2D plane such that $(x_S, y_S) = (0, 0)$, while $(x_{I_1}, y_{I_1}) = (60, 20)$, $(x_{I_2}, y_{I_2}) = (240, 20)$, $(x_{R_1}, y_{R_1}) = (60, 0)$, $(x_R, y_R) = (150, 0)$, $(x_{R_2}, y_{R_2}) = (240, 0)$, and $(x_D, y_D) = (300, 0)$, all in meters (see Fig.1). In addition, we set $\tilde{\alpha} = 2.3$, $\bar{\alpha} = 3.5$, $\sigma^2 = 1$, $K = 10$ dB; while the maximum number of iterations to optimize any phase-shift vector was 50 (which was shown to be enough for convergence), and the optimization convergence threshold was set to 10^{-3}. Furthermore, for the enhanced transmission scheme, we have $p_1 = p_2 = \frac{1}{2}p$. We define the interference-to-noise ratio (INR) as $\rho_e = \sigma_e^2/\sigma^2$, while the transmit SNR was defined as p/σ^2.

As demonstrated in Fig. 2, the double-RIS communication without including relaying suffers from a notably low rate performance. This is due to the large loss in signal power due to the lack of active amplification. In contrast, utilizing one or two relays can provide a significant performance gain. To be more specific, and regardless of the value of SNR, adopting the enhanced two-relays transmission is always better than the single-relay case as long as the inter-relay interference is sufficiently suppressed, i.e. $\rho_e \leq 10$ dB. Otherwise, the choice between a single relay and two relays depends purely on the
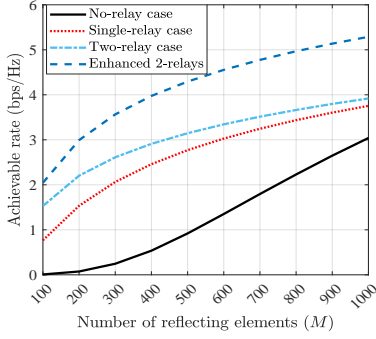
Fig. 3. Achievable rate vs number of reflecting elements per RIS for different relaying schemes when the transmit SNR = 50 dB, and $\rho_e = 0$ dB.

value of SNR. At low SNRs, deploying two relays such that no two nodes can transmit in the network at the same time, can still achieve higher rates than the single relay case despite the $(1/3)$ pre-log penalty; while at high SNRs, the single relay case leads to higher rates.

Finally, Fig. 3 shows the performance of different transmission schemes for a wide range of number of reflecting elements. Once again, our results indicate that adopting two relays to assist the transmission is highly desirable for the double-RIS assisted communication even when the considered RISs are sufficiently large with hundreds of reflecting elements. For example, to achieve 3 bps/Hz, the enhanced two-relay transmission requires 200 reflecting elements per RIS, compared to 600 and 1000 elements for the single-relay and no-relay cases, respectively, given that the inter-relay interference is suppressed to the noise level.

IV. CONCLUSION AND FUTURE WORK

We investigated the DF relay-aided double-RIS reflection channels for coverage extension. Three different relay-aided network architectures were proposed for effective rate maximization under a total power constraint. Our results demonstrated that deploying two relays for the double-RIS channel achieves higher rates at low and medium SNRs; while at high SNRs, deploying a single relay to assist the two RISs is better only if the inter-relay interference was high. The generalization to multihop with arbitrary numbers of RISs and relays is subject to future investigations.

ACKNOWLEDGMENT

APPENDIX A

Let us denote $p_2\left|h_{R_2D} + \boldsymbol{\phi}^T\boldsymbol{a}\right|^2$ by $f_a(\boldsymbol{\phi})$, $p_1\left|\boldsymbol{\phi}^T\boldsymbol{b}\right|^2 + \sigma^2$ by $f_b(\boldsymbol{\phi})$, and the right hand side of (33) by $g(\boldsymbol{\phi}|\tilde{\boldsymbol{\phi}}, \mu)$. Let $\tilde{\boldsymbol{\phi}}$ and $\boldsymbol{\phi}^\star$ be the phase-shift values of I_2 before and after running a single iteration of the MM scheme, and let $\tilde{\mu}$ denote the value of μ that corresponds to $\tilde{\boldsymbol{\phi}}$. Then, from

the left hand side of (33), we have $f(\boldsymbol{\phi}^\star, \tilde{\mu}) = f_b(\boldsymbol{\phi}^\star) - \tilde{\mu}f_a(\boldsymbol{\phi}^\star) \overset{(a)}{\leq} g(\boldsymbol{\phi}^\star|\tilde{\boldsymbol{\phi}}, \tilde{\mu}) \overset{(b)}{\leq} g(\tilde{\boldsymbol{\phi}}|\tilde{\boldsymbol{\phi}}, \tilde{\mu}) = f(\tilde{\boldsymbol{\phi}}, \tilde{\mu}) = f_b(\tilde{\boldsymbol{\phi}}) - \tilde{\mu}f_a(\tilde{\boldsymbol{\phi}}) \overset{(c)}{=} 0$, where (a) holds from (33), (b) holds since $\boldsymbol{\phi}^\star$ minimizes $g(\boldsymbol{\phi}|\tilde{\boldsymbol{\phi}}, \tilde{\mu})$, and (c) holds from the definition of μ in (36). Therefore, we have $u(\boldsymbol{\phi}^\star) = f_b(\boldsymbol{\phi}^\star)/f_a(\boldsymbol{\phi}^\star) \leq \tilde{\mu} = f_b(\tilde{\boldsymbol{\phi}})/f_a(\tilde{\boldsymbol{\phi}}) = u(\tilde{\boldsymbol{\phi}})$.

REFERENCES

[1] J. N. Laneman *et al.*, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[2] Minelli *et al.*, "Optimal relay placement in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 998–1009, Feb. 2014.

[3] H. Q. Ngo and E. G. Larsson, "Linear multihop amplify-and-forward relay channels: Error exponent and optimal number of hops," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3834–3842, Nov. 2011.

[4] X. Tan *et al.*, "Increasing indoor spectrum sharing capacity using smart reflect-array," in *IEEE Inter. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[5] N. Rajatheva *et al.*, "White paper on broadband connectivity in 6G," *[Online]. Available: https://arxiv.org/abs/2004.14247*, 2020.

[6] M. Z. Chowdhury *et al.*, "6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 957–975, Aug. 2020.

[7] S. Kisseleff *et al.*, "Reconfigurable intelligent surfaces for smart cities: Research challenges and opportunities," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1781–1797, Dec. 2020.

[8] E. Björnson *et al.*, "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.

[9] M. Di Renzo *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, July 2020.

[10] Z. Abdullah *et al.*, "A hybrid relay and intelligent reflecting surface network and its ergodic performance analysis," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1653–1657, Oct. 2020.

[11] ——, "Optimization of intelligent reflecting surface assisted full-duplex relay networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 363–367, Feb. 2021.

[12] J. Wang *et al.*, "Joint beamforming and reconfigurable intelligent surface design for two-way relay networks," *IEEE Trans. Commun.*, To appear.

[13] X. Ying *et al.*, "Relay aided intelligent reconfigurable surfaces: Achieving the potential without so many antennas," *[Online]. Available: https://arxiv.org/abs/2006.06644*, 2020.

[14] I. Yildirim *et al.*, "Hybrid RIS-empowered reflection and decode-and-forward relaying for coverage extension," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1692–1696, May 2021.

[15] M. Obeed and A. Chaaban, "Joint beamforming design for multiuser MISO downlink aided by a reconfigurable intelligent surface and a relay," *[Online]. Available: https://arxiv.org/abs/2104.08417*, 2021.

[16] C. Huang *et al.*, "Deep reinforcement learning-based relay selection in intelligent reflecting surface assisted cooperative networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1036–1040, May 2021.

[17] Z. Kang *et al.*, "IRS-aided wireless relaying: Optimal deployment and capacity scaling," *[Online]. Available: https://arxiv.org/abs/2105.08495*, 2021.

[18] C. You *et al.*, "Wireless communication via double IRS: Channel estimation and passive beamforming designs," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 431–435, Feb. 2021.

[19] B. Zheng *et al.*, "Efficient channel estimation for double-IRS aided multi-user MIMO system," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3818–3832, June 2021.

[20] W. Dinkelbach, "On nonlinear fractional programming," *Management science*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[21] H. Shen *et al.*, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, Sep. 2019.

[22] Y. Sun *et al.*, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.