# Multi-Pair IRS-Aided Full-Duplex Relay Networks: Rate Analysis and Beamforming Design

The Vi Nguyen, Thi My Tuyen Nguyen, Junsuk Oh, Wonjong Noh, and Sungrae Cho

***Abstract*—The intelligent reflecting surface (IRS) is a promising sixth-generation (6G) communication technology that can improve the wireless communication capacity in a cost-effective and energy-efficient manner by adjusting its passive reflectors to change the signal propagation. In this study, we investigate the multi-pair IRS-aided full-duplex relay (IFDR) network, wherein the relay is equipped with massive antennas. First, we determine the channel estimates of the cascaded channels and corresponding channel statistics for the IFDR system using the linear minimum mean square error (LMMSE) method. Second, we derive a closed-form lower bound for the achievable rate via maximum-ratio combining/transmission (MRC/MRT) processing and the use-and-then-forget (UatF) technique. This expression helps analyze the power scaling laws to provide better insight into the performance gain achieved by the IFDR system. Third, we propose an optimization algorithm to control the IRS phase-shift matrix and maximize the sum spectral efficiency using statistical channel knowledge. In particular, the problem is formulated as a fractional matrix programming problem. Subsequently, it is iteratively solved by using the majorization-minimization (MM) and square penalty methods. This approach can significantly reduce both the training and IRS configuration overhead. Lastly, through experiments, the numerical results show that the proposed control achieves better performance than the benchmark schemes.**

***Index Terms*—Full-duplex relay, beamforming design, intelligent reflecting surface, multi-hop multi-pair transmission**

## I. Introduction

THE sixth-generation (6G) mobile communication networks are envisioned to enable global coverage, massive connectivity, and diversified services [1]. These paradigms impose stringent requirements, including extreme reliability, sub-millisecond latency, and high energy and spectral efficiency. Ensuring reliable and seamless connectivity to users at the cell edge or in obstructed areas is challenging and necessitates advanced techniques for coverage enhancement. Relaying systems and intelligent reflecting surfaces (IRSs) have emerged as potential solutions to this problem. Relaying can effectively extend the network coverage via signal forwarding, whereas IRS passively reconfigures the propagation environment to enhance the signal strength and mitigate the blockage effects. Relaying and IRS technologies both exhibit advantages and limitations. Active relays, such as amplify-and-forward (AF) and decode-and-forward (DF), can improve the achievable rate and coverage via signal decoding, regeneration, and forwarding. However, they are hindered by high power consumption, hardware complexity, and performance bottlenecks due to the weaker hop in two-hop transmission. The performance can be enhanced by increasing the number of relay antennas; however, this leads to an increase in the radio frequency (RF) chains and processing costs. In contrast, the IRS improves system performance by passively reconfiguring the propagation environment with low power consumption and simple hardware. Also, an IRS operates passively in full-duplex mode, bypassing the self-interference inherent in active full-duplex (FD) relays. However, it suffers from double path loss and requires large reflecting arrays to match relay performance, incurring high channel estimation overhead, control complexity, and circuit power consumption. These complementary features motivate a hybrid IRS-relay architecture to balance spectral efficiency, energy efficiency, and hardware complexity. In this system, the relay can compensate for the double path loss and channel state information (CSI) acquisition challenges owing to its powerful signal processing capabilities. In addition, the IRS can enhance channel quality without excessive power consumption and hardware complexity. Therefore, their combination provides improved coverage and energy/spectral efficiency when compared with the deployment of either technology alone. Recent research on IRS and FD relaying extends toward advanced architectures, e.g., stacked intelligent metasurfaces for wave-domain processing [2], [3], and physical layer security/covertness utilizing artificial noise [4], dynamic antenna switching [5], and self-jamming energy harvesting [6]. While these paradigms represent promising extensions, this study focuses specifically on integrating passive IRS into FD-DF relay systems. Toward this end, we present recent studies on the hybrid IRS–relay systems in the half-duplex (HD) and FD relaying modes.

### A. Related Works

*1) IRS-Half Duplex Relaying:* IRS-aided HD relaying systems have gained considerable attention in recent years. In HD relaying, two orthogonal time slots are dedicated to a single transmission. In [7], the authors proposed a hybrid IRS-HD relay system with DF operation for communication between a single source-destination pair. Given the same achievable rate, a hybrid system with a single-antenna relay requires fewer IRS elements than an IRS-only system without compromising the power cost. Moreover, the IRS-only system achieves comparable performance for a large number of IRS elements in the high signal-to-noise ratio (SNR) region. Regarding the

T. V. Nguyen, T. M. T. Nguyen, J. Oh, and S. Cho are with the School of Computer Science and Engineering, Chung-Ang University, Seoul 156-756, South Korea (e-mail: tvnguyen@uclab.re.kr, tuyen@uclab.re.kr, jsoh@uclab.re.kr, srcho@cau.ac.kr). W. Noh is with the School of Software, Hallym University, Chuncheon 24252, South Korea (e-mail: wonjong.noh@hallym.ac.kr).

TABLE I
COMPARISON BETWEEN THIS WORK AND RELATED STUDIES

| Ref. | Duplex | AF/DF | Multi-ant. | Comm. | Imp. CSI | Stat. CSI |
|---|---|---|---|---|---|---|
| [7] | HD | DF | ✗ | SP | ✓ | ✓ |
| [8] | HD | DF | ✗ | SP | ✗ | ✓ |
| [9] | HD | DF | ✗ | SP | ✗ | ✗ |
| [10] | HD | DF | ✗ | SP | ✗ | ✗ |
| [11] | HD | DF | ✗ | SU | ✗ | ✗ |
| [12] | HD | DF | ✓ | SP | ✗ | ✗ |
| [13] | HD | DF | ✓ | MU | ✗ | ✗ |
| [14] | HD/FD | DF | ✓ | MU | ✗ | ✗ |
| [15] | FD | DF | ✗ | SP | ✗ | ✗ |
| [16] | FD | AF | ✗ | SU | ✗ | ✗ |
| **Ours** | **FD** | **DF** | ✓ | **MP** | ✓ | ✓ |

SP = single-pair; MP = multi-pair; SU = BS - single-user; MU = BS - multi-user; Multi-ant. = multi-antenna; Comm. = communication system; Imp. CSI = imperfect CSI; Stat. CSI = statistical CSI.

IRS deployment in the IRS-aided relay system, the studies conducted in [8], [9] demonstrated that the system achieves the best performance when the IRS is placed close to the relay. In [10], the authors proposed joint and integrated IRS-aided relay schemes. The results demonstrated that the joint scheme exhibits better bit error rate performance than the integrated scheme. The authors in [11] investigated the HD-DF relay system with multiple IRSs, wherein the numbers of IRSs and reflecting elements were jointly optimized to maximize the net ergodic rate. A greater number of antennas were used at the relay in the studies [12], [13]. They demonstrated that the system performance can be further improved by jointly optimizing beamforming at the IRS and the relay. However, as transmission in the HD mode occurs over two phases, it results in high end-to-end latency. In addition, the IRS phase shifts must be optimized separately for each hop, which increases the design complexity. Consequently, this may become a major bottleneck in hybrid IRS–HD relay systems.

*2) IRS-Full Duplex Relaying:* Hybrid IRS-FD relaying has been adopted in recent years, which enables simultaneous signal transmission and reception over the same frequency band. The study in [14] demonstrated that IRS-aided FD relaying significantly reduces transmit power compared with IRS-only and relay-only systems by jointly optimizing beamforming at the BS and the relay and IRS phase shifts. Furthermore, the study in [15] demonstrated that the IRS can improve the rate performance in an IRS-aided FD relay system, particularly when the channels to users are weak, on average. The impact of IRS deployment in the IRS-aided FD relay was analyzed in [16], which reported that the signal-to-interference-plus-noise ratio (SINR) with near-relay deployment scales with the square of the number of IRS elements. However, FD relaying suffers from self-interference caused by signal leakage from the transmit to the receive antennas. Unless sufficiently suppressed, it can severely degrade the performance. With advanced FD relay designs, integrating the IRS can further enhance system performance while enabling low-latency and cost-effective transmission.

### *B. Motivation, Contribution, and Organization*

A major limitation of existing studies [9]–[16] is the assumption of perfect and instantaneous CSI, which is difficult to realize in practical IRS-aided relay systems. In particular, pilot signaling is required to acquire IRS-related channels. The training overhead increases with the number of reflecting elements. Moreover, the IRS phase-shift configuration is typically computed in a processing unit and then fed back to the IRS controller, which presents a high communication overhead and processing complexity. To address these challenges, [17] proposed a two-time-scale framework, in which IRS phase shifts are optimized based on statistical CSI in the slow time scale and relay beamforming is optimized based on instantaneous CSI in the fast time scale. While this approach significantly reduces channel estimation and signaling overhead, its application in IRS-aided FD relay systems remains unexplored. In addition, most existing studies [7]–[13] focused on the HD relay systems, which suffer from a $1/2$ pre-log spectral efficiency penalty and require separate, complex IRS phase optimizations for each orthogonal time slot. Furthermore, recent studies [15], [16] integrated IRSs with FD relaying systems, which are restricted to single-pair or single-user downlink scenarios, and bypass the critical inter-user and inter-pair interference inherent in multi-user networks. Additionally, simplified system models were widely adopted, wherein either a single antenna [7]–[11], [15], [16] or only a small number of relay antennas [12]–[14] was considered. To overcome these limitations, this study investigates an IRS-aided FD-DF system considering multi-pair transmission, massive relay antennas, imperfect CSI, and statistical CSI. This joint framework presents three key challenges: 1) channel estimation over cascaded double-fading links is intractable due to non-Gaussian product distributions; 2) deriving closed-form achievable rates based on statistical CSI is difficult owing to multi-dimensional coupling among double-Rician paths, IRS phase shifts, self-interference, and multi-pair interference; and 3) characterizing asymptotic power scaling laws with massive relay antennas is non-trivial. Table I summarizes the key features of the proposed scheme and those of the existing studies. The main contributions of this study are summarized as follows:

- First, we develop a linear minimum mean square error (LMMSE) framework to estimate cascaded source-to-relay and relay-to-destination channels. Unlike conventional designs where training overhead scales with IRS size, our aggregated channel estimation keeps pilot overhead independent of IRS size, requiring only first- and second-order statistics without explicit channel distributions.
- Second, we derive the closed-form lower bound for the end-to-end achievable rate using the maximum-ratio combining/transmission (MRC/MRT) processing, along with the use-and-then-forget (UatF) technique.
- Third, we analyze the power scaling laws corresponding to the number of receive/transmit antennas of the relay and IRS elements. Under certain channel conditions, this analysis demonstrates that as the number of relay

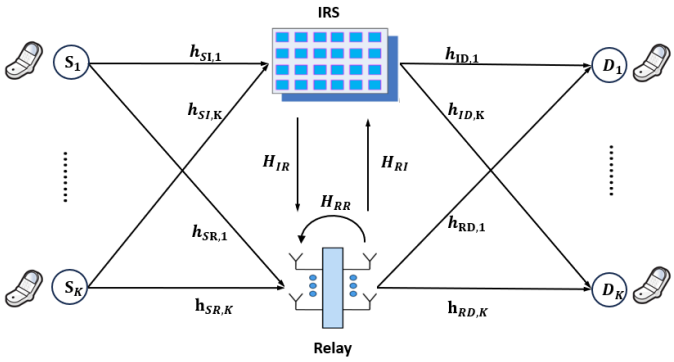


Fig. 1. Multi-pair IRS-aided full-duplex relay (IFDR) system.

receive/transmit antennas $M$ or IRS elements $N$ is large, the transmit power at the sources and relay scales proportionally to $1/M$ or $1/N$, respectively, while maintaining the non-zero total spectral efficiency.

- Fourth, we develop an algorithm to maximize the sum spectral efficiency with respect to the phase shift. The problem can be reformulated as fractional matrix programming. However, the reformulated problem is still non-convex owing to the non-convex constraints. To address this, we employ the majorization-minimization (MM) and square penalty methods. As the design relies on the slowly varying statistical CSI, the optimized phase shifts can be applied once the channel statistics change. This can significantly reduce both the training and IRS configuration overhead.
- Finally, the simulation results demonstrate the superiority of the proposed scheme over the benchmarks. The proposed design outperforms the benchmarks in terms of spectral efficiency, especially for a large number of IRS elements, which also mitigates the loss caused by strong self-interference. Moreover, the IRS demonstrates its advantages in reducing transmit power and the number of relay antennas in hybrid systems.

The remainder of this paper is organized as follows. In Section II, we introduce the system model, channel estimation, and precoding design of the relay. In Section III, we derive a closed-form achievable rate, analyze the power scaling laws, and optimize the IRS phase shifts. Sections IV and V present the simulation results and conclusions, respectively.

## II. System Model

We consider an IRS-aided FD-DF relay system, as shown in Fig. 1. In this system, we jointly deploy an FD-DF relay and an IRS to enable communication between sources $S_k$ and their corresponding destinations $D_k$, $\forall k \in \mathcal{K} \triangleq \{1, ..., K\}$, where $K$ denotes the number of source-destination pairs. All sources and destinations are equipped with a single antenna, the relay is equipped with $M$ receive antennas and $M$ transmit antennas, and the IRS is equipped with $N$ reflecting elements. We assume that there is no direct link between the sources and destinations owing to severe blockage and significant distance. Moreover, the line-of-sight (LoS) components of both the source-to-relay and relay-to-destination channels are assumed to be obstructed by ground obstacles. In this scenario, the IRS is deployed on a high-rise building, thereby enabling it to maintain LoS links with the sources, relay, and destinations. Thus, the IRS can effectively assist both the source-to-relay and relay-to-destination transmissions. We denote the reflection matrix $\mathbf{\Phi} \triangleq \text{diag}\{\phi_1, ..., \phi_N\} \in \mathbb{C}^{N\times N}$, where $\phi_n \triangleq e^{j\theta_n}$ and $\theta_n$ denotes the phase shift of the $n$-th reflecting element. Let $\mathbf{h}_{\text{SR},k} \in \mathbb{C}^{M\times 1}$ and $\mathbf{h}_{\text{SI},k} \in \mathbb{C}^{N\times 1}$ denote the channel vectors from the $k$-th source to the relay's receive antenna array and from the $k$-th source to the IRS, respectively. In addition, let $\mathbf{H}_{\text{IR}} \in \mathbb{C}^{M\times N}$ and $\mathbf{H}_{\text{RI}} \in \mathbb{C}^{N\times M}$ denote the channel matrices from the IRS to the relay's receive antenna array and from the relay's transmit antenna array to the IRS, respectively. Furthermore, let $\mathbf{h}_{\text{RD},k} \in \mathbb{C}^{M\times 1}$ and $\mathbf{h}_{\text{ID},k} \in \mathbb{C}^{N\times 1}$ denote the channel vectors from the relay's transmit antenna array to the $k$-th destination and from the IRS to the $k$-th destination, respectively. The FD relay suffers from self-interference, while the IRS avoids this issue by operating without active RF chains. The self-interference channel includes two components: the direct path from the relay's transmit antenna array to the relay's receive antenna array (represented by $\mathbf{H}_{\text{RR}} \in \mathbb{C}^{M\times M}$) and the reflected path via the IRS (represented by $\mathbf{H}_{\text{IR}}\mathbf{\Phi}\mathbf{H}_{\text{RI}}$).

### A. Channel Model

Owing to the heavy blockage, channels $\mathbf{h}_{\text{SR},k}$ and $\mathbf{h}_{\text{RD},k}$ are modeled via Rayleigh fading, whereas the remaining channels are modeled via Rician fading with virtual LoS components provided by the IRS. In particular, they are modeled as

$$\mathbf{h}_{\text{SR},k} = \sqrt{\eta_{\text{SR},k}}\tilde{\mathbf{h}}_{\text{SR},k}, \quad \mathbf{h}_{\text{RD},k} = \sqrt{\eta_{\text{RD},k}}\tilde{\mathbf{h}}_{\text{RD},k}, \tag{1}$$

$$\mathbf{h}_{\text{SI},k} = \sqrt{\eta_{\text{SI},k}/(1+\varepsilon_k)}\left(\sqrt{\varepsilon_k}\overline{\mathbf{h}}_{\text{SI},k} + \tilde{\mathbf{h}}_{\text{SI},k}\right), \tag{2}$$

$$\mathbf{h}_{\text{ID},k} = \sqrt{\eta_{\text{ID},k}/(1+\tilde{\varepsilon}_k)}\left(\sqrt{\tilde{\varepsilon}_k}\overline{\mathbf{h}}_{\text{ID},k} + \tilde{\mathbf{h}}_{\text{ID},k}\right), \tag{3}$$

$$\mathbf{H}_{\text{IR}} = \sqrt{\eta_{\text{IR}}/(1+\delta)}\left(\sqrt{\delta}\overline{\mathbf{H}}_{\text{IR}} + \tilde{\mathbf{H}}_{\text{IR}}\right), \tag{4}$$

$$\mathbf{H}_{\text{RI}} = \sqrt{\eta_{\text{RI}}/(1+\tilde{\delta})}\left(\sqrt{\tilde{\delta}}\overline{\mathbf{H}}_{\text{RI}} + \tilde{\mathbf{H}}_{\text{RI}}\right), \tag{5}$$

where the parameters $\eta_{\varsigma,k}$ and $\eta_{\varsigma'}$, for $\varsigma \in \{\text{SR}, \text{RD}, \text{SI}, \text{ID}\}$, $\varsigma' \in \{\text{IR}, \text{RI}\}$, represent the large-scale path loss factors, and the parameters $\delta$, $\tilde{\delta}$, $\varepsilon_k$, and $\tilde{\varepsilon}_k$ represent the Rician factors. In addition, $\overline{\mathbf{h}}_{\varsigma,k}$ and $\overline{\mathbf{H}}_{\varsigma'}$ denote the LoS components, while $\tilde{\mathbf{h}}_{\varsigma,k}$ and $\tilde{\mathbf{H}}_{\varsigma'}$ denote the non-line-of-sight (NLoS) components, whose elements are independent and identically distributed (i.i.d.) as $\mathcal{CN}(0,1)$. The IRS comprises a uniform planar array (UPA) with element spacing $d = \lambda/2$, where $\lambda$ denotes the carrier wavelength. In addition, there are $N^{\text{v}}$ and $N^{\text{h}}$ elements along the vertical and horizontal axes, respectively. Thus, $N = N^{\text{v}}N^{\text{h}}$. The LoS components can be modeled as follows. The array response vector for the IRS is given by [18]

$$\mathbf{a}_{\text{I}}(\varphi, \vartheta) = \mathbf{a}_{\text{I}}^{\text{v}}(\varphi, \vartheta) \otimes \mathbf{a}_{\text{I}}^{\text{h}}(\varphi, \vartheta) \in \mathbb{C}^{N\times 1}, \tag{6}$$

where $\varphi$ and $\vartheta$ denote the azimuth and elevation angles, respectively, of the arrival/departure signals at the IRS, $\mathbf{a}_{\text{I}}^{\text{v}}(\varphi, \vartheta) \in \mathbb{C}^{N^{\text{v}}\times 1}$ and $\mathbf{a}_{\text{I}}^{\text{h}}(\varphi, \vartheta) \in \mathbb{C}^{N^{\text{h}}\times 1}$ denote the array

response vectors along the vertical and horizontal directions, respectively. Each element of these vectors is given by

$$\left[\mathbf{a}_{\mathrm{I}}^{\mathrm{v}}\left(\varphi,\vartheta\right)\right]_{n_1}=e^{-j\pi(n_1-1)\sin(\vartheta)},\quad n_1\in\{1,...,N^{\mathrm{v}}\}$$
$$\left[\mathbf{a}_{\mathrm{I}}^{\mathrm{h}}\left(\varphi,\vartheta\right)\right]_{n_2}=e^{-j\pi(n_2-1)\sin(\varphi)\cos(\vartheta)},\quad n_2\in\{1,...,N^{\mathrm{h}}\}.$$

The relay's transmit and receive antenna arrays are uniform linear arrays with antenna spacing $d$. The array response vector for each relay's antenna array is given by [18]

$$\mathbf{a}_{\mathrm{R}}\left(\varphi,\vartheta\right)=\left[1,e^{-j\pi\sin(\vartheta)},...,e^{-j\pi(M-1)\sin(\vartheta)}\right].\quad(7)$$

Applying the steering vectors in (6) and (7) yields

$$\overline{\boldsymbol{h}}_{\mathrm{SI},k}=\mathbf{a}_{\mathrm{I}}\left(\varphi_{\mathrm{I},k}^{\mathrm{AoA}},\vartheta_{\mathrm{I},k}^{\mathrm{AoA}}\right),\quad \overline{\boldsymbol{h}}_{\mathrm{ID},k}=\mathbf{a}_{\mathrm{I}}\left(\varphi_{\mathrm{I},k}^{\mathrm{AoD}},\vartheta_{\mathrm{I},k}^{\mathrm{AoD}}\right),$$
$$\overline{\mathbf{H}}_{\mathrm{IR}}=\mathbf{a}_{\mathrm{R}}\left(\varphi_{\mathrm{R}}^{\mathrm{AoA}},\vartheta_{\mathrm{R}}^{\mathrm{AoA}}\right)\mathbf{a}_{\mathrm{I}}\left(\varphi_{\mathrm{I}}^{\mathrm{AoD}},\vartheta_{\mathrm{I}}^{\mathrm{AoD}}\right)^{H},$$
$$\overline{\mathbf{H}}_{\mathrm{RI}}=\mathbf{a}_{\mathrm{I}}\left(\varphi_{\mathrm{I}}^{\mathrm{AoA}},\vartheta_{\mathrm{I}}^{\mathrm{AoA}}\right)\mathbf{a}_{\mathrm{R}}\left(\varphi_{\mathrm{R}}^{\mathrm{AoD}},\vartheta_{\mathrm{R}}^{\mathrm{AoD}}\right)^{H},$$

where $(\varphi_{\mathrm{I},k}^{\mathrm{AoA}},\vartheta_{\mathrm{I},k}^{\mathrm{AoA}})$ and $(\varphi_{\mathrm{I},k}^{\mathrm{AoD}},\vartheta_{\mathrm{I},k}^{\mathrm{AoD}})$ denote the azimuth and elevation angles of arrival (AoA) from the $k$-th source to the IRS and angles of departure (AoD) from the IRS to the $k$-th destination, respectively. In addition, $(\varphi_{\mathrm{R}}^{\mathrm{AoA}},\vartheta_{\mathrm{R}}^{\mathrm{AoA}})$ and $(\varphi_{\mathrm{I}}^{\mathrm{AoD}},\vartheta_{\mathrm{I}}^{\mathrm{AoD}})$ denote the azimuth and elevation AoA and AoD from the IRS to the relay's receive antenna array, respectively. Similarly, $(\varphi_{\mathrm{I}}^{\mathrm{AoA}},\vartheta_{\mathrm{I}}^{\mathrm{AoA}})$ and $(\varphi_{\mathrm{R}}^{\mathrm{AoD}},\vartheta_{\mathrm{R}}^{\mathrm{AoD}})$ denote the azimuth and elevation AoA and AoD, respectively, from the relay's transmit antenna array to the IRS. For ease of notation, we denote $\mathbf{a}_{\mathrm{R}}\triangleq\mathbf{a}_{\mathrm{R}}\left(\varphi_{\mathrm{R}}^{\mathrm{AoA}},\vartheta_{\mathrm{R}}^{\mathrm{AoA}}\right)$ and $\mathbf{a}_{\mathrm{I}}\triangleq\mathbf{a}_{\mathrm{I}}(\varphi_{\mathrm{I}}^{\mathrm{AoD}},\vartheta_{\mathrm{I}}^{\mathrm{AoD}})$, $\tilde{\mathbf{a}}_{\mathrm{R}}\triangleq\mathbf{a}_{\mathrm{R}}\left(\varphi_{\mathrm{R}}^{\mathrm{AoD}},\vartheta_{\mathrm{R}}^{\mathrm{AoD}}\right)$, and $\tilde{\mathbf{a}}_{\mathrm{I}}\triangleq\mathbf{a}_{\mathrm{I}}(\varphi_{\mathrm{I}}^{\mathrm{AoA}},\vartheta_{\mathrm{I}}^{\mathrm{AoA}})$. We define the *effective channels*, including the IRS-reflected paths

$$\mathbf{g}_{\mathrm{SR},k}\triangleq\mathbf{h}_{\mathrm{SR},k}+\mathbf{H}_{\mathrm{IR}}\boldsymbol{\Phi}\mathbf{h}_{\mathrm{SI},k}\in\mathbb{C}^{M\times1},\quad(8)$$
$$\mathbf{G}_{\mathrm{RR}}\triangleq\mathbf{H}_{\mathrm{RR}}+\mathbf{H}_{\mathrm{IR}}\boldsymbol{\Phi}\mathbf{H}_{\mathrm{RI}}\in\mathbb{C}^{M\times M},\quad(9)$$
$$\mathbf{g}_{\mathrm{RD},k}^{T}\triangleq\mathbf{h}_{\mathrm{RD},k}^{T}+\mathbf{h}_{\mathrm{ID},k}^{T}\boldsymbol{\Phi}\mathbf{H}_{\mathrm{RI}}\in\mathbb{C}^{1\times M},\quad(10)$$

where the channels $\mathbf{g}_{\mathrm{SR},k}$ and $\mathbf{g}_{\mathrm{RD},k}^{T}$ represent the effective channels from the $k$-th source to the relay's receive antenna array and from the relay's transmit antenna array to the $k$-th destination, respectively. In addition, $\mathbf{G}_{\mathrm{RR}}$ represents the self-interference channel at the relay, including the direct path from the relay's transmit antenna array to the relay's receive antenna array and the reflected path via the IRS. Substituting (1)-(5) into (8) and (10), the effective channels $\mathbf{g}_{\mathrm{SR},k}$ and $\mathbf{g}_{\mathrm{RD},k}$ are expanded as

$$\mathbf{g}_{\mathrm{SR},k}=\underbrace{\sqrt{c_k\delta\varepsilon_k}\overline{\mathbf{H}}_{\mathrm{IR}}\boldsymbol{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k}}_{\mathbf{q}_{\mathrm{SR},k}^{(1)}}+\underbrace{\sqrt{c_k\delta}\overline{\mathbf{H}}_{\mathrm{IR}}\boldsymbol{\Phi}\tilde{\mathbf{h}}_{\mathrm{SI},k}}_{\mathbf{q}_{\mathrm{SR},k}^{(2)}}$$
$$+\underbrace{\sqrt{c_k\varepsilon_k}\tilde{\mathbf{H}}_{\mathrm{IR}}\boldsymbol{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k}}_{\mathbf{q}_{\mathrm{SR},k}^{(3)}}+\underbrace{\sqrt{c_k}\tilde{\mathbf{H}}_{\mathrm{IR}}\boldsymbol{\Phi}\tilde{\mathbf{h}}_{\mathrm{SI},k}}_{\mathbf{q}_{\mathrm{SR},k}^{(4)}}+\underbrace{\sqrt{\eta_{\mathrm{SR},k}}\tilde{\mathbf{h}}_{\mathrm{SR},k}}_{\mathbf{d}_{\mathrm{SR},k}},\quad(11)$$

$$\mathbf{g}_{\mathrm{RD},k}=\underbrace{\sqrt{\tilde{c}_k\tilde{\delta}\tilde{\varepsilon}_k}\overline{\mathbf{H}}_{\mathrm{RI}}^{T}\boldsymbol{\Phi}^{T}\overline{\mathbf{h}}_{\mathrm{ID},k}}_{\mathbf{q}_{\mathrm{RD},k}^{(1)}}+\underbrace{\sqrt{\tilde{c}_k\tilde{\delta}}\overline{\mathbf{H}}_{\mathrm{RI}}^{T}\boldsymbol{\Phi}^{T}\tilde{\mathbf{h}}_{\mathrm{ID},k}}_{\mathbf{q}_{\mathrm{RD},k}^{(2)}}$$
$$+\underbrace{\sqrt{\tilde{c}_k\tilde{\varepsilon}_k}\tilde{\mathbf{H}}_{\mathrm{RI}}^{T}\boldsymbol{\Phi}^{T}\overline{\mathbf{h}}_{\mathrm{ID},k}}_{\mathbf{q}_{\mathrm{RD},k}^{(3)}}+\underbrace{\sqrt{\tilde{c}_k}\tilde{\mathbf{H}}_{\mathrm{RI}}^{T}\boldsymbol{\Phi}^{T}\tilde{\mathbf{h}}_{\mathrm{ID},k}}_{\mathbf{q}_{\mathrm{RD},k}^{(4)}}+\underbrace{\sqrt{\eta_{\mathrm{RD},k}}\tilde{\mathbf{h}}_{\mathrm{RD},k}}_{\mathbf{d}_{\mathrm{RD},k}},\quad(12)$$

where $c_k\triangleq\frac{\eta_{\mathrm{IR}}\eta_{\mathrm{SI},k}}{(1+\delta)(1+\varepsilon_k)}$ and $\tilde{c}_k\triangleq\frac{\eta_{\mathrm{RI}}\eta_{\mathrm{ID},k}}{(1+\tilde{\delta})(1+\tilde{\varepsilon}_k)}$.

### B. Channel Estimation

Following the two-timescale framework [17], the IRS phase shifts are optimized using statistical CSI once per frame and held fixed over multiple coherence blocks. In each block, the relay estimates the channels $\mathbf{g}_{\mathrm{SR},k}$ and $\mathbf{g}_{\mathrm{RD},k}$ using the pilot signals. Let $\tau$ symbols be used for channel estimation, and the remaining part is used for data transmission. During the training interval, all the sources and destinations simultaneously transmit their pilot sequences of $\tau$ symbols to the relay. In particular, let $\sqrt{\tau p_{\mathrm{p}}}\mathbf{l}_{\mathrm{S},k}\in\mathbb{C}^{1\times\tau}$ and $\sqrt{\tau p_{\mathrm{p}}}\mathbf{l}_{\mathrm{D},k}\in\mathbb{C}^{1\times\tau}$ denote the pilot sequences transmitted from the $k$-th source and $k$-th destination, respectively, where $p_{\mathrm{p}}$ represents the transmit power for each pilot symbol. Then, $\mathbf{L}_{\mathrm{S}}\triangleq[\mathbf{l}_{\mathrm{S},1},...,\mathbf{l}_{\mathrm{S},K}]^{T}\in\mathbb{C}^{K\times\tau}$ and $\mathbf{L}_{D}\triangleq[\mathbf{l}_{\mathrm{D},1},...,\mathbf{l}_{\mathrm{D},K}]^{T}\in\mathbb{C}^{K\times\tau}$. We assume that all the pilot sequences are orthogonal, i.e., $\mathbf{L}_{\mathrm{S}}\mathbf{L}_{\mathrm{S}}^{H}=\mathbf{I}_{K},\mathbf{L}_{\mathrm{D}}\mathbf{L}_{\mathrm{D}}^{H}=\mathbf{I}_{K}$, and $\mathbf{L}_{\mathrm{S}}\mathbf{L}_{\mathrm{D}}^{H}=\mathbf{0}_{K}$, which requires $\tau\geq2K$. The received signal matrices at the relay's receive and transmit antenna arrays are given by

$$\mathbf{Y}_{\mathrm{rp}}=\sqrt{\tau p_{\mathrm{p}}}\sum_{k=1}^{K}\mathbf{g}_{\mathrm{SR},k}\mathbf{l}_{\mathrm{S},k}+\sqrt{\tau p_{\mathrm{p}}}\breve{\mathbf{H}}_{\mathrm{RD}}\mathbf{L}_{\mathrm{D}}+\mathbf{N}_{\mathrm{rp}}\in\mathbb{C}^{M\times\tau},$$

$$\mathbf{Y}_{\mathrm{tp}}=\sqrt{\tau p_{\mathrm{p}}}\breve{\mathbf{H}}_{\mathrm{SR}}\mathbf{L}_{\mathrm{S}}+\sqrt{\tau p_{\mathrm{p}}}\sum_{k=1}^{K}\mathbf{g}_{\mathrm{RD},k}\mathbf{l}_{\mathrm{D},k}+\mathbf{N}_{\mathrm{tp}}\in\mathbb{C}^{M\times\tau},$$

where $\breve{\mathbf{H}}_{\mathrm{RD}}\in\mathbb{C}^{M\times K}$ and $\breve{\mathbf{H}}_{\mathrm{SR}}\in\mathbb{C}^{M\times K}$ denote the channel matrices from the destinations to the relay's receive antenna array and from the sources to the relay's transmit antenna array, respectively. In addition, $\mathbf{N}_{\mathrm{rp}}$ and $\mathbf{N}_{\mathrm{tp}}$ denote the additive white Gaussian noise (AWGN) matrices whose elements are i.i.d. and distributed as $\mathcal{CN}(0,\sigma^2)$. During training, the relay's transmit antenna array is temporarily set to receive mode to observe destinations' pilot signal, following standard full-duplex practice [19]. Although the relay's antenna arrays observe superimposed pilot signals from sources and destinations, the desired channels are isolated by exploiting the mutual orthogonality of pilot sequences. In particular, $\frac{1}{\sqrt{\tau p_{\mathrm{p}}}}\mathbf{Y}_{\mathrm{rp}}$ and $\frac{1}{\sqrt{\tau p_{p}}}\mathbf{Y}_{\mathrm{tp}}$ are projected onto $\mathbf{l}_{\mathrm{S},k}$ and $\mathbf{l}_{\mathrm{D},k}$ to obtain the following observations for the $k$-th source's and $k$-th destination's signals, respectively,

$$\tilde{\mathbf{y}}_{\mathrm{rp},k}=\left(1/\sqrt{\tau p_{p}}\right)\mathbf{Y}_{\mathrm{rp}}\mathbf{l}_{\mathrm{S},k}^{H}=\mathbf{g}_{\mathrm{SR},k}+\mathbf{N}_{\mathrm{rp},k},\quad(13)$$
$$\tilde{\mathbf{y}}_{\mathrm{tp},k}=\left(1/\sqrt{\tau p_{p}}\right)\mathbf{Y}_{\mathrm{tp}}\mathbf{l}_{\mathrm{D},k}^{H}=\mathbf{g}_{\mathrm{RD},k}+\mathbf{N}_{\mathrm{tp},k},\quad(14)$$

where $\mathbf{N}_{\mathrm{rp},k}\triangleq\frac{1}{\sqrt{\tau p_{p}}}\mathbf{N}_{\mathrm{rp}}\mathbf{l}_{\mathrm{S},k}^{H}\sim\mathcal{CN}\left(\mathbf{0},\frac{\sigma^2}{\tau p_{p}}\mathbf{I}_{M}\right)$ and $\mathbf{N}_{\mathrm{tp},k}\triangleq\frac{1}{\sqrt{\tau p_{p}}}\mathbf{N}_{\mathrm{tp}}\mathbf{l}_{\mathrm{D},k}^{H}\sim\mathcal{CN}\left(\mathbf{0},\frac{\sigma^2}{\tau p_{p}}\mathbf{I}_{M}\right)$. We employ the LMMSE method for channel estimation [20] (Sec. 12.3, pp. 380).[1] The LMMSE estimator is formed as a linear function of observations, where the weighting coefficients are selected to minimize the mean square error (MSE). Using the LMMSE

[1]LMMSE is practical as it relies on the first- and second-order statistics, which facilitates the closed-form channel estimates. While we adopt standard LMMSE, exploring promising deep learning-based schemes (e.g., [21]) with lower overhead and higher accuracy is left for future work.

estimator with the observations in (13) and (14), we obtain the channel estimates of the effective channels $\mathbf{g}_{\mathrm{SR},k}$ and $\mathbf{g}_{\mathrm{RD},k}$, denoted as $\hat{\mathbf{g}}_{\mathrm{SR},k}$ and $\hat{\mathbf{g}}_{\mathrm{RD},k}$, respectively. Based on [17], we specify the channel estimates in the following proposition.

**Proposition 1.** *(i) The channel estimate of the effective channel $\mathbf{g}_{\mathrm{SR},k}$ is given by*

$$\begin{aligned}\hat{\mathbf{g}}_{\mathrm{SR},k} =& \\ &\underbrace{\sqrt{c_k\delta\varepsilon_k}\,\overline{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k}}_{\hat{\mathbf{q}}^{(1)}_{\mathrm{SR},k}}+\underbrace{\left(Ma_{k,3}+a_{k,4}\right)\sqrt{c_k\delta}\,\overline{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\tilde{\mathbf{h}}_{\mathrm{SI},k}}_{\hat{\mathbf{q}}^{(2)}_{\mathrm{SR},k}}\\ &+\underbrace{\sqrt{c_k\varepsilon_k}\,\mathbf{A}_{\mathrm{SR},k}\tilde{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k}}_{\hat{\mathbf{q}}^{(3)}_{\mathrm{SR},k}}+\underbrace{\sqrt{c_k}\,\mathbf{A}_{\mathrm{SR},k}\tilde{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\tilde{\mathbf{h}}_{\mathrm{SI},k}}_{\hat{\mathbf{q}}^{(4)}_{\mathrm{SR},k}}\\ &+\mathbf{A}_{\mathrm{SR},k}\mathbf{d}_{\mathrm{SR},k}+\mathbf{A}_{\mathrm{SR},k}\mathbf{N}_{\mathrm{rp},k},\end{aligned}\tag{15}$$

*where*

$$\mathbf{A}_{\mathrm{SR},k}=a_{k,3}\mathbf{a}_{\mathrm{R}}\mathbf{a}_{\mathrm{R}}^{H}+a_{k,4}\mathbf{I}_M,\tag{16}$$

$$a_{k,1}=Nc_k\delta,\quad a_{k,2}=Nc_k\left(\varepsilon_k+1\right)+\eta_{\mathrm{SR},k},\tag{17}$$

$$a_{k,3}=\frac{a_{k,1}\frac{\sigma^2}{\tau p_{\mathrm{p}}}}{\left(a_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)\left(a_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}+Ma_{k,1}\right)},\tag{18}$$

$$a_{k,4}=\frac{a_{k,2}}{a_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}}.\tag{19}$$

*(ii) The channel estimate of the effective channel $\mathbf{g}_{\mathrm{RD},k}$ is given by*

$$\begin{aligned}\hat{\mathbf{g}}_{\mathrm{RD},k} =& \\ &\underbrace{\sqrt{\tilde{c}_k\tilde{\delta}\tilde{\varepsilon}_k}\,\overline{\mathbf{H}}^{T}_{\mathrm{RI}}\mathbf{\Phi}^{T}\overline{\mathbf{h}}_{\mathrm{ID},k}}_{\hat{\mathbf{q}}^{(1)}_{\mathrm{RD},k}}+\underbrace{\left(M\tilde{a}_{k,3}+\tilde{a}_{k,4}\right)\sqrt{\tilde{c}_k\tilde{\delta}}\,\overline{\mathbf{H}}^{T}_{\mathrm{RI}}\mathbf{\Phi}^{T}\tilde{\mathbf{h}}_{\mathrm{ID},k}}_{\hat{\mathbf{q}}^{(2)}_{\mathrm{RD},k}}\\ &+\underbrace{\sqrt{\tilde{c}_k\tilde{\varepsilon}_k}\,\mathbf{A}_{\mathrm{RD},k}\tilde{\mathbf{H}}^{T}_{\mathrm{RI}}\mathbf{\Phi}^{T}\overline{\mathbf{h}}_{\mathrm{ID},k}}_{\hat{\mathbf{q}}^{(3)}_{\mathrm{RD},k}}+\underbrace{\sqrt{\tilde{c}_k}\,\mathbf{A}_{\mathrm{RD},k}\tilde{\mathbf{H}}^{T}_{\mathrm{RI}}\mathbf{\Phi}^{T}\tilde{\mathbf{h}}_{\mathrm{ID},k}}_{\hat{\mathbf{q}}^{(4)}_{\mathrm{RD},k}}\\ &+\mathbf{A}_{\mathrm{RD},k}\mathbf{d}_{\mathrm{RD},k}+\mathbf{A}_{\mathrm{RD},k}\mathbf{N}_{\mathrm{tp},k},\end{aligned}\tag{20}$$

*where*

$$\mathbf{A}_{\mathrm{RD},k}=\tilde{a}_{k,3}\tilde{\mathbf{a}}^{*}_{\mathrm{R}}\tilde{\mathbf{a}}^{T}_{\mathrm{R}}+\tilde{a}_{k,4}\mathbf{I}_M,\tag{21}$$

$$\tilde{a}_{k,1}=N\tilde{c}_k\tilde{\delta},\quad \tilde{a}_{k,2}=N\tilde{c}_k\left(\tilde{\varepsilon}_k+1\right)+\eta_{\mathrm{RD},k},\tag{22}$$

$$\tilde{a}_{k,3}=\frac{\tilde{a}_{k,1}\frac{\sigma^2}{\tau p_{\mathrm{p}}}}{\left(\tilde{a}_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)\left(\tilde{a}_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}+M\tilde{a}_{k,1}\right)},\tag{23}$$

$$\tilde{a}_{k,4}=\frac{\tilde{a}_{k,2}}{\tilde{a}_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}}.\tag{24}$$

*Proof.* Please see Appendix A. □

Let $\mathbf{e}_{\mathrm{SR},k}$ and $\mathbf{e}_{\mathrm{RD},k}$ denote the estimation errors of $\mathbf{g}_{\mathrm{SR},k}$ and $\mathbf{g}_{\mathrm{RD},k}$. Thus, $\mathbf{g}_{\mathrm{SR},k}=\hat{\mathbf{g}}_{\mathrm{SR},k}+\mathbf{e}_{\mathrm{SR},k}$, $\mathbf{g}_{\mathrm{RD},k}=\hat{\mathbf{g}}_{\mathrm{RD},k}+\mathbf{e}_{\mathrm{RD},k}$. According to [17], the estimated channel vectors and estimation error vectors are uncorrelated, $\mathbb{E}\{\mathbf{e}_{\mathrm{SR},k}\}=0$, and $\mathbb{E}\{\mathbf{e}_{\mathrm{RD},k}\}=0$. Let $\mathbf{MSE}_{\mathrm{SR},k}\triangleq\mathbb{E}\{\mathbf{e}_{\mathrm{SR},k}\mathbf{e}^{H}_{\mathrm{SR},k}\}$ and $\mathbf{MSE}_{\mathrm{RD},k}\triangleq\mathbb{E}\{\mathbf{e}_{\mathrm{RD},k}\mathbf{e}^{H}_{\mathrm{RD},k}\}$ denote the MSE matrix of the estimation errors $\mathbf{e}_{\mathrm{SR},k}$ and $\mathbf{e}_{\mathrm{RD},k}$, respectively The explicit expression of normalized MSE (NMSE) value is given by the following proposition.

**Proposition 2.** *(i) The NMSE value of the estimate $\mathbf{g}_{\mathrm{SR},k}$ is given by:*

$$\begin{aligned}\mathrm{NMSE}_{\mathrm{SR},k}&=\frac{\mathrm{Tr}(\mathbf{MSE}_{\mathrm{SR},k})}{\mathrm{Tr}(\mathrm{Cov}\{\mathbf{g}_{\mathrm{SR},k},\mathbf{g}_{\mathrm{SR},k}\})}\\ &=\frac{\frac{\sigma^2}{\tau p_{\mathrm{p}}}\left(Ma_{k,1}a_{k,2}+a^2_{k,2}+\left(a_{k,1}+a_{k,2}\right)\frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)}{\left(a_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)\left(a_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}+Ma_{k,1}\right)\left(a_{k,1}+a_{k,2}\right)}.\end{aligned}\tag{25}$$

*(ii) The NMSE value of the estimate of $\mathbf{g}_{\mathrm{RD},k}$ is given by:*

$$\begin{aligned}\mathrm{NMSE}_{\mathrm{RD},k}&=\frac{\mathrm{Tr}(\mathbf{MSE}_{\mathrm{RD},k})}{\mathrm{Tr}(\mathrm{Cov}\{\mathbf{g}_{\mathrm{RD},k},\mathbf{g}_{\mathrm{RD},k}\})}\\ &=\frac{\frac{\sigma^2}{\tau p_{\mathrm{p}}}\left(M\tilde{a}_{k,1}\tilde{a}_{k,2}+\tilde{a}^2_{k,2}+\left(\tilde{a}_{k,1}+\tilde{a}_{k,2}\right)\frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)}{\left(\tilde{a}_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)\left(\tilde{a}_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}+M\tilde{a}_{k,1}\right)\left(\tilde{a}_{k,1}+\tilde{a}_{k,2}\right)}.\end{aligned}\tag{26}$$

*Proof.* Please see Appendix B. □

Based on Proposition 2, we investigate the influence of the large number of IRS elements on the NMSE metric, as shown in the following proposition.

**Proposition 3.** *(i) For the SR link, as $N\to\infty$, we have*

$$\lim_{N\to\infty}\mathrm{NMSE}_{\mathrm{SR},k}=0,\quad \lim_{\frac{\sigma^2}{\tau p_{\mathrm{p}}}\to\infty}\mathrm{NMSE}_{\mathrm{SR},k}=1,\quad \lim_{\frac{\sigma^2}{\tau p_{\mathrm{p}}}\to 0}\mathrm{NMSE}_{\mathrm{SR},k}=0.\tag{27}$$

*For $p_{\mathrm{p}}=\frac{E_{\mathrm{p}}}{N}$, where $E_{\mathrm{p}}$ is a constant*

$$\lim_{N\to\infty}\mathrm{NMSE}_{\mathrm{SR},k}<1.\tag{28}$$

*(ii) For the RD link, as $N\to\infty$, we have*

$$\lim_{N\to\infty}\mathrm{NMSE}_{\mathrm{RD},k}=0,\quad \lim_{\frac{\sigma^2}{\tau p_{\mathrm{p}}}\to\infty}\mathrm{NMSE}_{\mathrm{RD},k}=1,\quad \lim_{\frac{\sigma^2}{\tau p_{\mathrm{p}}}\to 0}\mathrm{NMSE}_{\mathrm{RD},k}=0.\tag{29}$$

*For $p_{\mathrm{p}}=\frac{E_{\mathrm{p}}}{N}$, where $E_{\mathrm{p}}$ is a constant*

$$\lim_{N\to\infty}\mathrm{NMSE}_{\mathrm{RD},k}<1.\tag{30}$$

*Proof.* Please see the Appendix C. □

Proposition 3 yields several key insights. First, as $N\to\infty$, the NMSE approaches zero. This is because a larger number of IRS elements yields a larger effective channel gain, which increases the pilot power received at the relay. For a fixed noise power, this leads to a higher effective pilot SNR and hence more accurate channel estimation. Second, in high and low pilot SNR regimes, the NMSE converges to 0 (perfect estimation) and 1 (failed estimation), respectively. Finally, with $p_{\mathrm{p}}=\frac{E_{\mathrm{p}}}{N}$, the pilot power can be reduced proportionally to $1/N$ while the NMSE converges to a constant strictly less than 1.

*Remark* 1. In this study, we assume spatially uncorrelated channels, while the spatial correlation among IRS elements can improve channel estimation accuracy [17]. Due to the passive,

unitary phase shift matrix, the estimation NMSE in (25) and (26) do not depend on IRS configuration. Therefore, IRS phase shifts cannot be optimized to enhance channel estimation. Considering spatial correlation is deferred to future work.

### C. Data Transmission

All the sources $S_k$ transmit their signals $\sqrt{p_{\rm S}}x_k(t)$ to the relay and IRS at the time instance $t$, where $p_{\rm S}$ and $x_k$ denote the transmit power and the transmitted signal of $S_k$, respectively. Meanwhile, the relay transmits $\sqrt{p_{\rm R}}\mathbf{x}_{\rm R}(t)$ to all the destinations at time instance $t$, where $p_{\rm R}$ and $\mathbf{x}_{\rm R} \in \mathbb{C}^{M\times 1}$ denote the transmit power and transmitted signal at the relay, respectively. We assume that $\mathbb{E}\{|x_k(t)|^2\} = 1$ and $\mathbb{E}\{\|\mathbf{x}_{\rm R}(t)\|^2\} = 1$ [19]. Thus, the received signals at the relay and at each destination $k$ are given by

$$\mathbf{y}_{\rm R}(t) = \sqrt{p_{\rm S}}\mathbf{G}_{\rm SR}\mathbf{x}(t) + \sqrt{p_{\rm R}}\mathbf{G}_{\rm RR}\mathbf{x}_{\rm R}(t) + \mathbf{z}_{\rm R}(t), \tag{31}$$

$$y_k(t) = \sqrt{p_{\rm R}}\mathbf{g}_{{\rm RD},k}^T\mathbf{x}_{\rm R}(t) + \sqrt{p_{\rm S}}\mathbf{h}_{{\rm ID},k}^T\mathbf{\Phi}\mathbf{H}_{\rm SI}\mathbf{x}(t) + n_k(t), \quad \forall k \in \mathcal{K} \tag{32}$$

where $\mathbf{x}(t) \triangleq [x_1(t), ..., x_K(t)]^T \in \mathbb{C}^{K\times 1}$ is the vector of the transmitted signal from all the sources; $\mathbf{G}_{\rm SR} \triangleq [\mathbf{g}_{{\rm SR},1}, ..., \mathbf{g}_{{\rm SR},K}] \in \mathbb{C}^{M\times K}$ and $\mathbf{H}_{\rm SI} \triangleq [\mathbf{h}_{{\rm SI},1}, ..., \mathbf{h}_{{\rm SI},K}] \in \mathbb{C}^{N\times K}$. In addition, $\mathbf{z}_{\rm R}(t) \sim \mathcal{CN}(\mathbf{0}, \sigma_{\rm R}^2\mathbf{I}_M)$ and $n_k \sim \mathcal{CN}(0, \sigma_{\rm D}^2)$ denote the AWGN at the relay and at the $k$-th destination, respectively. The second term in (32) represents the signal reflected from all the sources to the $k$-th destination via the IRS. We denote $\tilde{\mathbf{G}}_{\rm RR} = \mathbf{G}_{\rm RR} - \hat{\mathbf{G}}_{\rm RR}$ as the residual self-interference channel matrix after active self-interference cancellation, where $\hat{\mathbf{G}}_{\rm RR}$ is the estimate of the channel $\mathbf{G}_{\rm RR}$. Thus, the residual self-interference vector at time $t$ is given by $\mathbf{z}_{\rm RR}(t) = \tilde{\mathbf{G}}_{\rm RR}\mathbf{x}_{\rm R}(t)$, where $\mathbb{E}\{\mathbf{z}_{\rm RR}(t)\} = \mathbf{0}, \mathbb{E}\{\mathbf{z}_{\rm RR}(t)\mathbf{z}_{\rm RR}^H(t)\} = \sigma_{\rm RR}^2\mathbf{I}_M$. We characterize $\mathbf{z}_{\rm RR}(t)$ only through its first- and second-order statistics. The parameter $\sigma_{\rm RR}^2$ can be understood as the post-cancellation residual level, which is a given hardware-determined system parameter [15], [16]. Thus, after the active self-interference cancellation, the received signal (31) is rewritten as

$$\mathbf{y}_{\rm R}(t) = \sqrt{p_{\rm S}}\mathbf{G}_{\rm SR}\mathbf{x}(t) + \sqrt{p_{\rm R}}\mathbf{z}_{\rm RR}(t) + \mathbf{z}_{\rm R}(t). \tag{33}$$

*Remark* 2. Standard self-interference cancellation techniques remain effective in our system for the following reasons. Although IRS reflections can potentially exacerbate self-interference, existing FD relays effectively suppress self-interference down to near the noise floor using multi-stage cancellation [22]. Furthermore, since $\mathbf{\Phi}$ is optimized on a slow time scale based on statistical CSI, adaptive SIC readily tracks the composite channel $\mathbf{G}_{\rm RR}$. In addition, as a passive device, the IRS introduces no non-linear distortion, ensuring that the residual SI power is determined by hardware imperfections. Finally, our framework explicitly optimizes $\mathbf{\Phi}$ to enhance the desired signal power rather than constructively amplifying SI.

Based on the channel estimates, a *linear receiver* at the relay is used to decode the received signal. Simultaneously, a *linear precoder* is used to forward the signals to their destinations. In particular, the received signal $\mathbf{y}_{\rm R}(t)$ in (33) is separated into $K$ streams by multiplying it by a matrix $\mathbf{W}^T \in \mathbb{C}^{K\times M}$ (depending on the channel estimates), as follows:

$$\begin{aligned}\mathbf{r}(t) &= \mathbf{W}^T\mathbf{y}_{\rm R}(t)\\ &= \sqrt{p_{\rm S}}\mathbf{W}^T\mathbf{G}_{\rm SR}\mathbf{x}(t) + \sqrt{p_{\rm R}}\mathbf{W}^T\mathbf{z}_{\rm RR}(t) + \mathbf{W}^T\mathbf{z}_{\rm R}(t).\end{aligned}$$

The $k$-th stream used to decode the transmitted signal of $S_k$, i.e., the $k$-th element of the vector $\mathbf{r}(t)$, is given as follows:

$$\begin{aligned}r_k(t) = &\underbrace{\sqrt{p_{\rm S}}\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}x_k(t)}_{\text{Desired signal}} + \underbrace{\sqrt{p_{\rm S}}\sum_{i\neq k}\mathbf{w}_k^T\mathbf{g}_{{\rm SR},i}x_i(t)}_{\text{Multi-pair interference}}\\ &+ \underbrace{\sqrt{p_{\rm R}}\mathbf{w}_k^T\mathbf{z}_{\rm RR}(t)}_{\text{Residual self-interference}} + \underbrace{\mathbf{w}_k^T\mathbf{z}_{\rm R}(t)}_{\text{Noise}},\end{aligned} \tag{34}$$

where $\mathbf{w}_k$ denotes the $k$-th column of $\mathbf{W}$. Once the transmitted signals from all the sources are detected, they are precoded by using a linear precoding scheme. The relay broadcasts the signal $\mathbf{x}(t-d)$ to the destinations owing to the processing delay, where $d$ denotes the processing delay [19]. Assuming correct decoding at the relay (see Remark 3), the transmit vector at the relay is expressed as follows:

$$\mathbf{x}_{\rm R}(t) = \mathbf{V}\mathbf{x}(t-d), \tag{35}$$

where $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_K] \in \mathbb{C}^{M\times K}$ denotes a linear precoding matrix that depends on the channel estimates. Subsequently, substituting (35) into (32) yields

$$\begin{aligned}y_k(t) = &\sqrt{p_{\rm R}}\mathbf{g}_{{\rm RD},k}^T\mathbf{v}_k x_k(t-d) + \sqrt{p_{\rm R}}\sum_{i\neq k}\mathbf{g}_{{\rm RD},k}^T\mathbf{v}_i x_i(t-d)\\ &+ \sqrt{p_{\rm S}}\mathbf{h}_{{\rm ID},k}^T\mathbf{\Phi}\mathbf{H}_{\rm SI}\mathbf{x}(t) + n_k(t).\end{aligned} \tag{36}$$

In (36), the received signal comprises the interference intended for other destinations (second term) and the interference reflected from the source via the IRS (third term). Here, $x_k(t-d)$ denotes the signal of interest at the destination $D_k$, whereas $x_k(t)$ is considered as interference [15].

*Remark* 3. The relay signal model in (35) follows the standard DF widely adopted in the literature [19]. We characterize performance via the information-theoretic achievable rate, which assumes capacity-achieving codes that guarantee reliable communication as long as the rate does not exceed capacity. Explicitly modeling the outage or block error probability under imperfect decoding is left for future work.

### D. Relay Precoding Design

We employ MRC/MRT processing at the relay, where MRC is used for reception and MRT for forwarding [19].[2] Accordingly, based on the estimated CSI, the MRC matrix $\mathbf{W}$ and MRT matrix $\mathbf{V}$ are given by

$$\mathbf{W}^T \triangleq \hat{\mathbf{G}}_{\rm SR}^H, \quad \mathbf{V} \triangleq \rho\hat{\mathbf{G}}_{\rm RD}^*, \tag{37}$$

[2]Although ZF and MMSE processing suppress inter-pair interference more effectively, we adopt MRC/MRT because it avoids complex matrix inversions involving the coupled IRS channel, which enables closed-form derivations of the achievable rate. Moreover, MRC/MRT has low complexity and is widely adopted in massive MIMO relaying [19] and two-timescale IRS designs [17]. Analyzing ZF/MMSE processing in IRS-aided networks is left for future work.

where $\hat{\mathbf{G}}_{\rm SR} = [\hat{\mathbf{g}}_{{\rm SR},1}, ..., \hat{\mathbf{g}}_{{\rm SR},K}] \in \mathbb{C}^{M\times K}$, $\hat{\mathbf{G}}_{\rm RD} = [\hat{\mathbf{g}}_{{\rm RD},1}, ..., \hat{\mathbf{g}}_{{\rm RD},K}] \in \mathbb{C}^{M\times K}$, and $\rho$ denotes the normalization coefficient that satisfies the long-term total transmit power constraint at the relay [19],

$$\mathbb{E}\{\|\mathbf{x}_{\rm R}(t)\|^2\} = 1. \tag{38}$$

The value of $\rho$ is obtained as follows:

$$\begin{aligned}\mathbb{E}\{\|\mathbf{x}_{\rm R}(t)\|^2\} &= \mathbb{E}\{\|\mathbf{V}\mathbf{x}(t-d)\|^2\} \overset{(a)}{=} \mathbb{E}\left\{{\rm Tr}\left(\mathbf{V}^H\mathbf{V}\right)\right\} \\ &\overset{(b)}{=} \rho^2\mathbb{E}\left\{{\rm Tr}\left(\hat{\mathbf{G}}_{\rm RD}^T\hat{\mathbf{G}}_{\rm RD}^*\right)\right\},\end{aligned} \tag{39}$$

where (a) holds because the elements of $\mathbf{x}(t-d)$ are i.i.d. with unit variance and $\mathbf{x}(t-d)$ is independent of $\hat{\mathbf{G}}_{\rm RD}$, and (b) is obtained using (37). Therefore, to satisfy (38), $\rho$ is given by $\rho = \sqrt{1\Big/\mathbb{E}\left\{{\rm Tr}\left(\hat{\mathbf{G}}_{\rm RD}^T\hat{\mathbf{G}}_{\rm RD}^*\right)\right\}}$.

## III. Performance Analysis

### A. Achievable Rate Analysis

The end-to-end achievable rate of the transmission link $S_k$ to $D_k$ through the IRS and relay is limited by the bottleneck hop. Therefore, the achievable rate is

$$\mathrm{R}_k = \min\{\mathrm{R}_{{\rm SR},k}, \mathrm{R}_{{\rm RD},k}\}, \tag{43}$$

where $\mathrm{R}_{{\rm SR},k}$ and $\mathrm{R}_{{\rm RD},k}$ represent the achievable rates in the first and second hops, respectively. To determine these rates, we adopt the *use-and-then-forget* (UatF) approach for the system with a large number of antennas to derive a lower bound for the achievable rate [23]. The underlying idea involves *using* the channel estimates only for the MRC/MRT schemes and then *forgetting* these estimates during the signal detection. The main advantages of this approach are as follows: (i) it provides an insightful closed-form expression that serves as a lower bound on the rate achievable in practice, and (ii) it does not require the acquisition of instantaneous CSI at the destination. First, we compute $\mathrm{R}_{{\rm SR},k}$ by rewriting the received signal at the relay in (34) by adding and subtracting $\sqrt{p_{\rm S}}\mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}\right\}x_k(t)$

$$\begin{aligned}r_k(t) = &\underbrace{\sqrt{p_{\rm S}}\mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}\right\}x_k(t)}_{\text{Desired signal over average channel}} \\ &+ \underbrace{\sqrt{p_{\rm S}}\left(\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k} - \mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}\right\}\right)x_k(t)}_{\text{Signal leakage}} \\ &+ \underbrace{\sqrt{p_{\rm S}}\sum_{i\neq k}\mathbf{w}_k^T\mathbf{g}_{{\rm SR},i}x_i(t)}_{\text{Multi-pair interference}} + \underbrace{\sqrt{p_{\rm R}}\mathbf{w}_k^T\mathbf{z}_{\rm RR}(t)}_{\text{Self-interference}} + \underbrace{\mathbf{w}_k^T\mathbf{z}_{\rm R}(t)}_{\text{Noise}}.\end{aligned} \tag{44}$$

Here, $\sqrt{p_{\rm S}}\mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}\right\}x_k(t)$ is considered as the true designed signal, which is used for signal detection. In addition, $\sqrt{p_{\rm S}}\left(\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k} - \mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}\right\}\right)x_k(t)$ can be considered as uncorrelated noise. This makes sense when there is substantial channel hardening, i.e., when the channel state fluctuates slightly around its mean and, therefore, the deviation is small. In (44), we observe that the first term (i.e., desired signal)

$$\mathrm{DS}_{{\rm SR},k} \triangleq p_{\rm S}\left|\mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}\right\}\right|^2 = p_{\rm S}\left(\alpha_k^{(1)}\left|f_k(\mathbf{\Phi})\right|^2 + \alpha_k^{(2)}\right)^2, \alpha_k^{(1)} \triangleq Mc_k\delta\varepsilon_k, \alpha_k^{(2)} \triangleq M\left[Nc_k\delta e_{k,2} + \left(Nc_k(\varepsilon_k+1) + \eta_{{\rm SR},k}\right)e_{k,1}\right] \tag{40}$$

$$\begin{aligned}&\mathrm{LS}_{{\rm SR},k} \triangleq p_{\rm S}\mathbb{V}{\rm ar}\left\{\mathbf{w}_k^T\mathbf{g}_{{\rm SR},k}\right\} = p_{\rm S}\left(\beta_k^{(1)}\left|f_k(\mathbf{\Phi})\right|^2 + \beta_k^{(2)}\right), \\ &\beta_k^{(1)} \triangleq Mc_k^2\delta\varepsilon_k\left[N\left(M\delta+\varepsilon_k+1\right)\left((e_{k,2})^2+1\right) + 2\left(Me_{k,1}+e_{k,2}\right)\left(e_{k,2}+1\right)\right] + Mc_k\delta\varepsilon_k\left[\eta_{{\rm SR},k} + \left(\eta_{{\rm SR},k} + \frac{\sigma^2}{\tau p_{\rm p}}\right)(e_{k,2})^2\right], \\ &\beta_k^{(2)} \triangleq M^2N^2c_k^2\delta^2\left(e_{k,2}\right)^2 + MN^2c_k^2\left[2\delta(\varepsilon_k+1)\left(e_{k,2}\right)^2 + (\varepsilon_k+1)^2e_{k,3}\right] + M^2Nc_k^2\left[(2\varepsilon_k+1)\left(e_{k,1}\right)^2 + 2\delta e_{k,1}e_{k,2}\right] \\ &\qquad + MNc_k\left[c_k\left(2\delta\left(e_{k,2}\right)^2 + (2\varepsilon_k+1)e_{k,3}\right) + \left(2\eta_{{\rm SR},k} + \frac{\sigma^2}{\tau p_{\rm p}}\right)\left(\delta\left(e_{k,2}\right)^2 + (\varepsilon_k+1)e_{k,3}\right)\right] + M\eta_{{\rm SR},k}\left(\eta_{{\rm SR},k} + \frac{\sigma^2}{\tau p_{\rm p}}\right)e_{k,3}\end{aligned} \tag{41}$$

$$\begin{aligned}\mathrm{MP}_{{\rm SR},k} &\triangleq p_{\rm S}\sum_{i\neq k}\mathbb{E}\left\{|\mathbf{w}_k^T\mathbf{g}_{{\rm SR},i}|^2\right\} \\ &= p_{\rm S}\sum_{i\neq k}\left(\gamma_{k,i}^{(1)}\left|f_k(\mathbf{\Phi})\right|^2 + \gamma_{k,i}^{(2)}\left|f_i(\mathbf{\Phi})\right|^2 + \gamma_{k,i}^{(3)}\left|f_k(\mathbf{\Phi})\right|^2\left|f_i(\mathbf{\Phi})\right|^2 + \gamma_{k,i}^{(4)}\Re\left\{f_k^H(\mathbf{\Phi})f_i(\mathbf{\Phi})\overline{\mathbf{h}}_{{\rm SI},i}^H\overline{\mathbf{h}}_{{\rm SI},k}\right\} + \gamma_{k,i}^{(5)}\right),\end{aligned} \tag{42}$$

$$\begin{aligned}&\gamma_{k,i}^{(1)} \triangleq Mc_k\delta\varepsilon_k\left[c_i\left(MN\delta + N\varepsilon_i + N + 2Me_{k,1}\right) + \eta_{{\rm SR},i}\right], \\ &\gamma_{k,i}^{(2)} \triangleq Mc_i\delta\varepsilon_i\left[c_ke_{k,2}\left(MN\delta e_{k,2} + N\varepsilon_ke_{k,2} + Ne_{k,2} + 2Me_{k,1}\right) + \left(\eta_{{\rm SR},k} + \frac{\sigma^2}{\tau p_{\rm p}}\right)\left(e_{k,2}\right)^2\right], \\ &\gamma_{k,i}^{(3)} \triangleq M^2c_kc_i\delta^2\varepsilon_k\varepsilon_i, \quad \gamma_{k,i}^{(4)} \triangleq 2M^2c_kc_i\varepsilon_k\varepsilon_ie_{k,1}\delta, \\ &\gamma_{k,i}^{(5)} \triangleq M^2N^2c_kc_i\delta^2\left(e_{k,2}\right)^2 + MN^2c_kc_i\left[\delta\left(\varepsilon_k+\varepsilon_i+2\right)\left(e_{k,2}\right)^2 + (\varepsilon_k+1)(\varepsilon_i+1)e_{k,3}\right] + M^2Nc_kc_ie_{k,1}\left[(\varepsilon_k+\varepsilon_i+1)e_{k,1} + 2\delta e_{k,2}\right] \\ &\qquad + M^2c_kc_i\varepsilon_k\varepsilon_i(e_{k,1})^2|\overline{\mathbf{h}}_{{\rm SI},k}^H\overline{\mathbf{h}}_{{\rm SI},i}|^2 + MN\left\{\left(\eta_{{\rm SR},k} + \frac{\sigma^2}{\tau p_{\rm p}}\right)c_i\left[\delta\left(e_{k,2}\right)^2 + (\varepsilon_i+1)e_{k,3}\right] + \eta_{{\rm SR},i}c_k\left[\delta\left(e_{k,2}\right)^2 + (\varepsilon_k+1)e_{k,3}\right]\right\} \\ &\qquad + M\eta_{{\rm SR},i}\left(\eta_{{\rm SR},k} + \frac{\sigma^2}{\tau p_{\rm p}}\right)e_{k,3}\end{aligned}$$

and the sum of the remaining terms (i.e., effective noise) are uncorrelated. Therefore, by applying the classical result that independent Gaussian noise is the worst-case among all uncorrelated noise for a given variance [19], the achievable rate for the first hop is given in Theorem 1.[3] The achievable rate for the second hop (i.e., $\mathrm{R}_{\mathrm{RD},k}$) is computed similarly.

**Theorem 1.** *(i) The achievable rate for the SR link is given by*

$$\mathrm{R}_{\mathrm{SR},k} = \log_2(1+\mathrm{SINR}_{\mathrm{SR},k}),$$
$$\mathrm{SINR}_{\mathrm{SR},k} = \frac{\mathrm{DS}_{\mathrm{SR},k}}{\mathrm{LS}_{\mathrm{SR},k}+\mathrm{MP}_{\mathrm{SR},k}+\mathrm{SI}_{\mathrm{SR},k}+\mathrm{AN}_{\mathrm{SR},k}}, \quad (49)$$

*where* $\mathrm{DS}_{\mathrm{SR},k}, \mathrm{LS}_{\mathrm{SR},k}, \mathrm{MP}_{\mathrm{SR},k}, \mathrm{SI}_{\mathrm{SR},k}$, *and* $\mathrm{AN}_{\mathrm{SR},k}$ *represent the desired signal, leakage signal, multi-pair interference, self-interference, and additive noise, respectively. Particularly, the first three terms are given in* (40), (41), *and* (42). *In addition, the last two terms are given as* $\mathrm{SI}_{\mathrm{SR},k} \triangleq p_{\mathrm{R}}\sigma_{\mathrm{RR}}^2\mathbb{E}\left\{\|\mathbf{w}_k^T\|^2\right\}, \mathrm{AN}_{\mathrm{SR},k} \triangleq \sigma_{\mathrm{R}}^2\mathbb{E}\left\{\|\mathbf{w}_k^T\|^2\right\}$, *where* $\mathbb{E}\left\{\|\mathbf{w}_k^T\|^2\right\} = \left|\mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\right\}\right|$. *Moreover, we denote* $f_k(\mathbf{\Phi}) \triangleq \mathbf{a}_{\mathrm{I}}^H\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k}$, $e_{k,1} \triangleq a_{k,3}+a_{k,4}$, $e_{k,2} \triangleq Ma_{k,3}+a_{k,4}$, *and* $e_{k,3} \triangleq M\left(a_{k,3}\right)^2+2a_{k,3}a_{k,4}+\left(a_{k,4}\right)^2$.

*(ii) The achievable rate for RD link is given by*

$$\mathrm{R}_{\mathrm{RD},k} = \log_2(1+\mathrm{SINR}_{\mathrm{RD},k}),$$
$$\mathrm{SINR}_{\mathrm{RD},k} = \frac{\mathrm{DS}_{\mathrm{RD},k}}{\mathrm{LS}_{\mathrm{RD},k}+\mathrm{MP}_{\mathrm{RD},k}+\mathrm{RI}_{\mathrm{RD},k}+\sigma_{\mathrm{D}}^2}, \quad (50)$$

*where* $\mathrm{DS}_{\mathrm{RD},k}, \mathrm{LS}_{\mathrm{RD},k}, \mathrm{MP}_{\mathrm{RD},k}$, *and* $\mathrm{RI}_{\mathrm{RD},k}$ *represent the desired signal, leakage signal, multi-pair interference, and reflection interference, respectively. Particularly, these terms are given in* (45), (46), (47), *and* (48), *with* $\tilde{f}_k(\mathbf{\Phi}) \triangleq \overline{\mathbf{h}}_{\mathrm{ID},k}^T\mathbf{\Phi}\tilde{\mathbf{a}}_{\mathrm{I}}$, $\tilde{e}_{k,1} \triangleq \tilde{a}_{k,3}+\tilde{a}_{k,4}$, $\tilde{e}_{k,2} \triangleq M\tilde{a}_{k,3}+\tilde{a}_{k,4}$, *and* $\tilde{e}_{k,3} \triangleq M\left(\tilde{a}_{k,3}\right)^2+2\tilde{a}_{k,3}\tilde{a}_{k,4}+\left(\tilde{a}_{k,4}\right)^2$.

[3]The achievable rate remains a valid achievable-rate lower bound regardless of whether the actual effective noise is Gaussian. Any non-Gaussian disturbance from imperfect relay decoding only loosens the bound relative to the exact mutual information, without compromising its mathematical validity.

$$\mathrm{DS}_{\mathrm{RD},k} \triangleq p_{\mathrm{R}}\left|\mathbb{E}\left\{\mathbf{g}_{\mathrm{RD},k}^T\mathbf{v}_k\right\}\right|^2 = p_{\mathrm{R}}\rho^2\left(\tilde{\alpha}_k^{(1)}\left|\tilde{f}_k(\mathbf{\Phi})\right|^2+\tilde{\alpha}_k^{(2)}\right)^2,$$
$$\tilde{\alpha}_k^{(1)} \triangleq M\tilde{c}_k\tilde{\delta}\tilde{\varepsilon}_k, \quad \tilde{\alpha}_k^{(2)} \triangleq M\left[N\tilde{c}_k\tilde{\delta}\tilde{e}_{k,2}+\left(N\tilde{c}_k(\tilde{\varepsilon}_k+1)+\eta_{\mathrm{RD},k}\right)\tilde{e}_{k,1}\right], \quad \rho^2 = 1\Big/\sum_{l=1}^{K}\left(\tilde{\alpha}_l^{(1)}|\tilde{f}_l(\mathbf{\Phi})|^2+\tilde{\alpha}_l^{(2)}\right) \quad (45)$$

$$\mathrm{LS}_{\mathrm{RD},k} \triangleq p_{\mathrm{R}}\mathbb{V}\mathrm{ar}\left\{\mathbf{g}_{\mathrm{RD},k}^T\mathbf{v}_k\right\} = p_{\mathrm{R}}\rho^2\left(\tilde{\beta}_k^{(1)}\left|\tilde{f}_k(\mathbf{\Phi})\right|^2+\tilde{\beta}_k^{(2)}\right), \quad (46)$$
$$\tilde{\beta}_k^{(1)} \triangleq M\tilde{c}_k^2\tilde{\delta}\tilde{\varepsilon}_k\left[N\left(M\tilde{\delta}+\tilde{\varepsilon}_k+1\right)\left((\tilde{e}_{k,2})^2+1\right)+2\left(M\tilde{e}_{k,1}+\tilde{e}_{k,2}\right)\left(\tilde{e}_{k,2}+1\right)\right]+M\tilde{c}_k\tilde{\delta}\tilde{\varepsilon}_k\left[\eta_{\mathrm{RD},k}+\left(\eta_{\mathrm{RD},k}+\frac{\sigma^2}{\tau p_{\mathrm{P}}}\right)(\tilde{e}_{k,2})^2\right],$$
$$\tilde{\beta}_k^{(2)} \triangleq M^2N^2\tilde{c}_k^2\tilde{\delta}^2\left(\tilde{e}_{k,2}\right)^2+MN^2\tilde{c}_k^2\left[2\tilde{\delta}(\tilde{\varepsilon}_k+1)\left(\tilde{e}_{k,2}\right)^2+\left(\tilde{\varepsilon}_k+1\right)^2\tilde{e}_{k,3}\right]+M^2N\tilde{c}_k^2\left[(2\tilde{\varepsilon}_k+1)\left(\tilde{e}_{k,1}\right)^2+2\tilde{\delta}\tilde{e}_{k,1}\tilde{e}_{k,2}\right]$$
$$+MN\tilde{c}_k\left[\tilde{c}_k\left(2\tilde{\delta}\left(\tilde{e}_{k,2}\right)^2+(2\tilde{\varepsilon}_k+1)\tilde{e}_{k,3}\right)+\left(2\eta_{\mathrm{RD},k}+\frac{\sigma^2}{\tau p_{\mathrm{P}}}\right)\left(\tilde{\delta}\left(\tilde{e}_{k,2}\right)^2+(\tilde{\varepsilon}_k+1)\tilde{e}_{k,3}\right)\right]+M\eta_{\mathrm{RD},k}\left(\eta_{\mathrm{RD},k}+\frac{\sigma^2}{\tau p_{\mathrm{P}}}\right)\tilde{e}_{k,3}$$

$$\mathrm{MP}_{\mathrm{RD},k} \triangleq p_{\mathrm{R}}\sum_{i\neq k}\mathbb{E}\left\{|\mathbf{g}_{\mathrm{RD},k}^T\mathbf{v}_i|^2\right\}$$
$$= p_{\mathrm{R}}\rho^2\sum_{i\neq k}\left(\tilde{\gamma}_{k,i}^{(1)}\left|\tilde{f}_i(\mathbf{\Phi})\right|^2+\tilde{\gamma}_{k,i}^{(2)}\left|\tilde{f}_k(\mathbf{\Phi})\right|^2+\tilde{\gamma}_{k,i}^{(3)}\left|\tilde{f}_i(\mathbf{\Phi})\right|^2\left|\tilde{f}_k(\mathbf{\Phi})\right|^2+\tilde{\gamma}_{k,i}^{(4)}\Re\left\{\tilde{f}_k(\mathbf{\Phi})\tilde{f}_i^H(\mathbf{\Phi})\overline{\mathbf{h}}_{\mathrm{ID},i}^T\overline{\mathbf{h}}_{\mathrm{ID},k}^*\right\}+\tilde{\gamma}_{k,i}^{(5)}\right), \quad (47)$$
$$\tilde{\gamma}_{k,i}^{(1)} \triangleq M\tilde{c}_i\tilde{\delta}\tilde{\varepsilon}_i\left[\tilde{c}_k\left(MN\tilde{\delta}+N\tilde{\varepsilon}_k+N+2M\tilde{e}_{i,1}\right)+\eta_{\mathrm{RD},k}\right],$$
$$\tilde{\gamma}_{k,i}^{(2)} \triangleq M\tilde{c}_k\tilde{\delta}\tilde{\varepsilon}_k\left[\tilde{c}_i\tilde{e}_{i,2}\left(MN\tilde{\delta}\tilde{e}_{i,2}+N\tilde{\varepsilon}_i\tilde{e}_{i,2}+N\tilde{e}_{i,2}+2M\tilde{e}_{i,1}\right)+\left(\eta_{\mathrm{RD},i}+\frac{\sigma^2}{\tau p_{\mathrm{P}}}\right)(\tilde{e}_{i,2})^2\right],$$
$$\tilde{\gamma}_{k,i}^{(3)} \triangleq M^2\tilde{c}_i\tilde{c}_k\tilde{\delta}^2\tilde{\varepsilon}_i\tilde{\varepsilon}_k, \quad \tilde{\gamma}_{k,i}^{(4)} \triangleq 2M^2\tilde{c}_i\tilde{c}_k\tilde{\varepsilon}_i\tilde{\varepsilon}_k\tilde{e}_{i,1}\tilde{\delta},$$
$$\tilde{\gamma}_{k,i}^{(5)} \triangleq M^2N^2\tilde{c}_i\tilde{c}_k\tilde{\delta}^2\left(\tilde{e}_{i,2}\right)^2+MN^2\tilde{c}_i\tilde{c}_k\left[\tilde{\delta}\left(\tilde{\varepsilon}_i+\tilde{\varepsilon}_k+2\right)\left(\tilde{e}_{i,2}\right)^2+(\tilde{\varepsilon}_i+1)(\tilde{\varepsilon}_k+1)\tilde{e}_{i,3}\right]+M^2N\tilde{c}_i\tilde{c}_k\tilde{e}_{i,1}\left[(\tilde{\varepsilon}_i+\tilde{\varepsilon}_k+1)\tilde{e}_{i,1}+2\tilde{\delta}\tilde{e}_{i,2}\right]$$
$$+M^2\tilde{c}_i\tilde{c}_k\tilde{\varepsilon}_i\tilde{\varepsilon}_k(\tilde{e}_{i,1})^2|\overline{\mathbf{h}}_{\mathrm{ID},i}^T\overline{\mathbf{h}}_{\mathrm{ID},k}^*|^2+MN\left[\left(\eta_{\mathrm{RD},i}+\frac{\sigma^2}{\tau p_{\mathrm{P}}}\right)\tilde{c}_k\left(\tilde{\delta}\left(\tilde{e}_{i,2}\right)^2+(\tilde{\varepsilon}_k+1)\tilde{e}_{i,3}\right)+\eta_{\mathrm{RD},k}\tilde{c}_i\left(\tilde{\delta}\left(\tilde{e}_{i,2}\right)^2+(\tilde{\varepsilon}_i+1)\tilde{e}_{i,3}\right)\right]$$
$$+M\eta_{\mathrm{RD},k}\left(\eta_{\mathrm{RD},i}+\frac{\sigma^2}{\tau p_{\mathrm{P}}}\right)\tilde{e}_{i,3}$$

$$\mathrm{RI}_{\mathrm{RD},k} \triangleq p_{\mathrm{S}}\sum_{l=1}^{K}\mathbb{E}\left\{\left|\mathbf{h}_{\mathrm{ID},k}^T\mathbf{\Phi}\mathbf{h}_{\mathrm{SI},l}\right|^2\right\} = p_{\mathrm{S}}\sum_{l=1}^{K}\left(\zeta_{k,l}^{(1)}|\overline{\mathbf{h}}_{\mathrm{ID},k}^T\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},l}|^2+\zeta_{k,l}^{(2)}\right),$$
$$\zeta_{k,l}^{(1)} \triangleq \frac{\eta_{\mathrm{SI},l}\eta_{\mathrm{ID},k}}{(\varepsilon_l+1)(\tilde{\varepsilon}_k+1)}\varepsilon_l\tilde{\varepsilon}_k, \quad \zeta_{k,l}^{(2)} \triangleq \frac{\eta_{\mathrm{SI},l}\eta_{\mathrm{ID},k}}{(\varepsilon_l+1)(\tilde{\varepsilon}_k+1)}N\left(\varepsilon_l+\tilde{\varepsilon}_k+1\right) \quad (48)$$

*Proof.* Please see Appendix D. □

The achievable rate in (43) is rewritten as follows:

$$\mathrm{R}_k = \log_2\left(1 + \min\left\{\mathrm{SINR}_{\mathrm{SR},k}, \mathrm{SINR}_{\mathrm{RD},k}\right\}\right), \tag{51}$$

where the equality follows because the function $\log_2(1+x)$ is monotonically increasing for $x \geq 0$; thus, the minimum operator can be moved inside the logarithm. The *sum spectral efficiency (SSE)* is defined as the sum rate per channel use. During the coherence interval $T$ (symbols), $\tau$ symbols are used for channel estimation, and the remaining symbols are used for data transmission. Therefore, SSE is given by

$$\mathrm{SSE} = \frac{T-\tau}{T}\sum_{k=1}^{K}\mathrm{R}_k. \tag{52}$$

In (51), the relay and destinations decode the desired signals based only on the statistical information of the effective channel gains[4], i.e., $\mathbb{E}\{\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\}$ and $\mathbb{E}\{\mathbf{g}_{\mathrm{RD},k}^T\mathbf{v}_k\}$. Hence, no instantaneous CSI is required at the destinations. To verify how accurately (51) reflects the actual performance, we compare it with the ergodic achievable rate of a genie-aided receiver [19], wherein the relay and the destination $D_k$ perfectly know $\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},l}$ and $\mathbf{g}_{\mathrm{RD},k}^T\mathbf{v}_l$, $\forall l \in \{1,...,K\}$, respectively. The ergodic achievable rate is given by

$$\tilde{\mathrm{R}}_k = \min\left\{\tilde{R}_{\mathrm{SR},k}, \tilde{R}_{\mathrm{RD},k}\right\}, \tag{53}$$

where $\tilde{R}_{\mathrm{SR},k}, \tilde{R}_{\mathrm{RD},k}$ are shown in (54) and (55) at the bottom of this page. In the simulation part, we will show that the performance gap between the achievable rate achieved by (51) and (53) is small.

### B. Power Scaling Law Analysis

We analyze the power scaling laws under specific channel scenarios to determine the power savings when a very large number of relay antennas and IRS elements are employed. First, we analyze the power scaling laws under a large number of relay antennas $M$. We assume that $p_{\mathrm{p}}$ is fixed, $p_{\mathrm{S}} = \frac{E_{\mathrm{S}}}{M}$, and $p_{\mathrm{R}} = \frac{E_{\mathrm{R}}}{M}$, where $E_{\mathrm{S}}$ and $E_{\mathrm{R}}$ are fixed and independent of $M$.

**Proposition 4.** *When $\delta = 0$ and $\tilde{\delta} = 0$, the channels between the IRS and relay are Rayleigh distributed. As $M \to \infty$, the SSE is given by*

$$\mathrm{SSE} \longrightarrow \frac{T-\tau}{T}\sum_{k=1}^{K}\log_2\left(1 + \min\left\{\mathrm{SINR}_{\mathrm{SR},k}, \mathrm{SINR}_{\mathrm{RD},k}\right\}\right),$$

*where $\mathrm{SINR}_{\mathrm{SR},k}$ and $\mathrm{SINR}_{\mathrm{RD},k}$ are given in* (56) *and* (57)*, respectively, as shown at the bottom of this page.*

*Proof.* We first substitute $\delta = 0$, $p_{\mathrm{S}} = \frac{E_{\mathrm{S}}}{M}$, $\tilde{\delta} = 0$, and $p_{\mathrm{R}} = \frac{E_{\mathrm{R}}}{M}$ into Theorem 1. Then, the limit of the SSE can be obtained by keeping all non-zero terms as $M \to \infty$. Note that for the first hop, at the limit, $e_{k,1} = e_{k,2} = \frac{a_{k,2}}{a_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}}$, where $a_{k,2} = Nc_k(\varepsilon_k+1) + \eta_{\mathrm{SR},k}$. For the second hop, $\tilde{e}_{k,1} = \tilde{e}_{k,2} = \frac{\tilde{a}_{k,2}}{\tilde{a}_{k,2}+\frac{\sigma^2}{\tau p_{\mathrm{p}}}}$, where $\tilde{a}_{k,2} = N\tilde{c}_k(\tilde{\varepsilon}_k+1) + \eta_{\mathrm{RD},k}$. □

This proposition provides notable insights: (i) When the channels between the IRS and relay follow Rayleigh fading, the SSE converges to a nonzero limit as $M \to \infty$. This means that the source and relay transmit powers can be proportionally reduced to $1/M$, while maintaining a non-zero SSE. (ii) The self-interference term disappears when large antenna arrays are used, as demonstrated in [19]. Furthermore, it can be observed that the numerators of the SINR expressions scale with $\mathcal{O}(N^2)$, whereas the denominators scale with $\mathcal{O}(N)$. Therefore, the SSE scales logarithmically with $N$ for $M \to \infty$, $p_{\mathrm{S}} = \frac{E_{\mathrm{S}}}{M}$, and $p_{\mathrm{R}} = \frac{E_{\mathrm{R}}}{M}$. This demonstrates that deploying an IRS can potentially increase SSE performance. Note that this result is independent of the IRS configuration. This is because when $\delta = 0$ and $\tilde{\delta} = 0$, all terms containing $\mathbf{\Phi}$ in Theorem 1 vanish as $M \to \infty$. As a result, the SINR expressions in (56) and (57) are independent of the IRS phase matrix $\mathbf{\Phi}$. Therefore, this property holds for any valid IRS phase configuration. In practical scenarios, where $\delta > 0$ (e.g., a typical LoS IRS-relay link), $\mathbf{\Phi}$ critically impacts performance, necessitating the optimization presented in Section III-C.

Second, we analyze the power scaling law as $N \to \infty$. We assume that $p_{\mathrm{p}}$ is fixed, $p_{\mathrm{S}} = \frac{E_{\mathrm{S}}}{N}$, and $p_{\mathrm{R}} = \frac{E_{\mathrm{R}}}{N}$,

[4]In practical deployments, the statistical information, e.g., users' location and angles of arrival/departure, can be obtained. For instance, the statistical information of the IRS links can be estimated by deploying dedicated sensors/receiving circuits integrated into the IRS together with standard mean and covariance matrix estimation methods in [24].

$$\tilde{R}_{\mathrm{SR},k} = \mathbb{E}\left\{\log_2\left(1 + \frac{p_{\mathrm{S}}\left|\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\right|^2}{p_{\mathrm{S}}\sum_{i\neq k}|\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},i}|^2 + p_{\mathrm{R}}\sigma_{\mathrm{RR}}^2\|\mathbf{w}_k^T\|^2 + \sigma_{\mathrm{R}}^2\|\mathbf{w}_k^T\|^2}\right)\right\} \tag{54}$$

$$\tilde{R}_{\mathrm{RD},k} = \mathbb{E}\left\{\log_2\left(1 + \frac{p_{\mathrm{R}}\left|\mathbf{g}_{\mathrm{RD},k}^T\mathbf{v}_k\right|^2}{p_{\mathrm{R}}\sum_{i\neq k}|\mathbf{g}_{\mathrm{RD},k}^T\mathbf{v}_i|^2 + p_{\mathrm{S}}\sum_{l=1}^{K}\left|\mathbf{h}_{\mathrm{ID},k}^T\mathbf{\Phi}\mathbf{h}_{\mathrm{SI},l}\right|^2 + \sigma_{\mathrm{D}}^2}\right)\right\} \tag{55}$$

$$\mathrm{SINR}_{\mathrm{SR},k} = \frac{E_{\mathrm{S}}\left[Nc_k(\varepsilon_k+1) + \eta_{\mathrm{SR},k}\right]^2 e_{k,1}}{E_{\mathrm{S}}Nc_k^2(2\varepsilon_k+1)e_{k,1} + E_{\mathrm{S}}c_k e_{k,1}\sum_{i\neq k}c_i\left[N(\varepsilon_k+\varepsilon_i+1) + \varepsilon_k\varepsilon_i|\overline{\mathbf{h}}_{\mathrm{SI},k}^H\overline{\mathbf{h}}_{\mathrm{SI},i}|^2\right] + \sigma_{\mathrm{R}}^2\left(Nc_k(\varepsilon_k+1) + \eta_{\mathrm{SR},k}\right)} \tag{56}$$

$$\mathrm{SINR}_{\mathrm{RD},k} = \frac{E_{\mathrm{R}}\left[N\tilde{c}_k(\tilde{\varepsilon}_k+1) + \eta_{\mathrm{RD},k}\right]^2\tilde{e}_{k,1}^2}{E_{\mathrm{R}}N\tilde{c}_k^2\left(2\tilde{\varepsilon}_k+1\right)\tilde{e}_{k,1}^2 + E_{\mathrm{R}}\tilde{c}_k\sum_{i\neq k}\tilde{c}_i\tilde{e}_{i,1}^2\left[N(\tilde{\varepsilon}_i+\tilde{\varepsilon}_k+1) + \tilde{\varepsilon}_i\tilde{\varepsilon}_k|\overline{\mathbf{h}}_{\mathrm{ID},i}^T\overline{\mathbf{h}}_{\mathrm{ID},k}^*|^2\right] + \sigma_{\mathrm{D}}^2\sum_{l=1}^{K}\left(N\tilde{c}_l(\tilde{\varepsilon}_l+1) + \eta_{\mathrm{RD},l}\right)\tilde{e}_{l,1}} \tag{57}$$

where $E_{\rm S}$ and $E_{\rm R}$ are fixed and independent of $N$. We consider a special case in which the IRS-relay channels are Rayleigh distributed, and there is no direct reflection between the sources and destinations, i.e., the term $\sqrt{p_{\rm S}}\mathbf{h}_{{\rm ID},k}^T\mathbf{\Phi}\mathbf{H}_{\rm SI}\mathbf{x}(t)$ is not included in (32). This case is practical when the sources and destinations are far apart and the reflection between pairs is sufficiently weak such that it can be neglected.

**Proposition 5.** *When $\delta=0$ and $\tilde{\delta}=0$, as $N\to\infty$, and there is no direct reflection between sources and destinations, then SSE is given by*

$$\text{SSE}\longrightarrow\frac{T-\tau}{T}\sum_{k=1}^{K}\log_2\left(1+\min\left\{\text{SINR}_{{\rm SR},k},\text{SINR}_{{\rm RD},k}\right\}\right),$$

*where $\text{SINR}_{{\rm SR},k}$ and $\text{SINR}_{{\rm RD},k}$ are given in* (58)*, as shown at the bottom of this page.*

*Proof.* We substitute $\delta=0$, $p_{\rm S}=\frac{E_{\rm S}}{N}$, $\tilde{\delta}=0$, and $p_{\rm R}=\frac{E_{\rm R}}{N}$ into Theorem 1. Then, the limit of the SSE can be obtained by keeping all non-zero terms as $N\to\infty$. Note that $a_{k,3}\to 0, e_{k,1}\to 1, e_{k,2}\to 1, e_{k,3}\to 1$, and $\tilde{a}_{k,3}\to 0, \tilde{e}_{k,1}\to 1, \tilde{e}_{k,2}\to 1, \tilde{e}_{k,3}\to 1$, as $N\to\infty$. □

It can be observed that the SSE converges to a non-zero limit when the source and relay transmit powers are reduced proportionally to $N$. In addition, the self-interference term disappears when a large number of IRS elements are used, similar to using a large number of relay antenna arrays. This demonstrates the potential of deploying an IRS in hybrid IRS-relay systems. Moreover, the proposition demonstrates that the numerators of the SINRs scale with $\mathcal{O}(M)$ and the denominators scale with $\mathcal{O}(1)$. This indicates that the SSE scales logarithmically with the number of antennas $M$. With a large $M$, when the environment between the IRS and relay exhibits rich scattering, more spatial multiplexing gain and lower inter-pair interference can be achieved [17].

### C. IRS Phase Shift Optimization

We aim to optimize the IRS phase shift matrix to maximize the SSE. This problem is formulated in the fractional form as

$$\max_{\mathbf{\Phi}}\quad \text{SSE}=\frac{T-\tau}{T}\sum_{k=1}^{K}\log_2\left(1+\min_{m\in\{1,2\}}\left\{\frac{\mathcal{F}_{k,m}(\mathbf{\Phi})}{\mathcal{G}_{k,m}(\mathbf{\Phi})}\right\}\right), \tag{61a}$$

$$\text{s.t.}\quad |\phi_n|=1,\ \forall n\in\{1,\dots,N\}. \tag{61b}$$

where $\mathcal{F}_{k,m}(\mathbf{\Phi})$ and $\mathcal{G}_{k,m}(\mathbf{\Phi}), m\in\{1,2\}$, denote the numerator and denominator of the SINRs derived in (49) and (50). Specifically, $\text{SINR}_{{\rm SR},k}=\frac{\mathcal{F}_{k,1}(\mathbf{\Phi})}{\mathcal{G}_{k,1}(\mathbf{\Phi})}$ and $\text{SINR}_{{\rm RD},k}=\frac{\mathcal{F}_{k,2}(\mathbf{\Phi})}{\mathcal{G}_{k,2}\mathbf{\Phi})}$. To facilitate optimization, we first simplify the expressions, $\mathcal{F}_{k,m}(\mathbf{\Phi})$ and $\mathcal{G}_{k,m}(\mathbf{\Phi})$, by introducing new variables, followed by simplified terms.

*Simplification of $\mathcal{F}_{k,m}(\mathbf{\Phi})$ and $\mathcal{G}_{k,m}(\mathbf{\Phi})$*: We define $\mathbf{r}_k^H\triangleq\mathbf{a}_{\rm I}^H\text{diag}(\overline{\mathbf{h}}_{{\rm SI},k})$, $\boldsymbol{\phi}\triangleq\text{diag}(\mathbf{\Phi})$, and $\tilde{\mathbf{r}}_k^T\triangleq\overline{\mathbf{h}}_{{\rm ID},k}^T\text{diag}(\tilde{\mathbf{a}}_{\rm I})$. Subsequently, we rewrite $f_k(\mathbf{\Phi})=\mathbf{r}_k^H\boldsymbol{\phi}, \tilde{f}_k(\mathbf{\Phi})=\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}$. We then define

$$\mathbf{D}_{k,i}\triangleq\mathbf{r}_k\overline{\mathbf{h}}_{{\rm SI},i}^H\overline{\mathbf{h}}_{{\rm SI},k}\mathbf{r}_i^H,\quad \mathbf{E}_{k,i}\triangleq(\mathbf{D}_{k,i}+\mathbf{D}_{k,i}^H)/2, \tag{62}$$

$$\tilde{\mathbf{D}}_{k,i}\triangleq\tilde{\mathbf{r}}_i^*\overline{\mathbf{h}}_{{\rm ID},i}^T\overline{\mathbf{h}}_{{\rm ID},k}^*\tilde{\mathbf{r}}_k^T,\quad \tilde{\mathbf{E}}_{k,i}\triangleq(\tilde{\mathbf{D}}_{k,i}+\tilde{\mathbf{D}}_{k,i}^H)/2. \tag{63}$$

The expression $\Re\{f_k^H(\mathbf{\Phi})f_i(\mathbf{\Phi})\overline{\mathbf{h}}_{{\rm SI},i}^H\overline{\mathbf{h}}_{{\rm SI},k}\}$ in (42) is rewritten as

$$\begin{aligned}&\Re\left\{\left(\boldsymbol{\phi}^H\mathbf{r}_k\right)\left(\overline{\mathbf{h}}_{{\rm SI},i}^H\overline{\mathbf{h}}_{{\rm SI},k}\right)\left(\mathbf{r}_i^H\boldsymbol{\phi}\right)\right\}=\Re\left\{\boldsymbol{\phi}^H\mathbf{D}_{k,i}\boldsymbol{\phi}\right\}\\&\overset{(a)}{=}\boldsymbol{\phi}^H[(\mathbf{D}_{k,i}+\mathbf{D}_{k,i}^H)/2]\boldsymbol{\phi}=\boldsymbol{\phi}^H\mathbf{E}_{k,i}\boldsymbol{\phi},\end{aligned} \tag{64}$$

where (a) is obtained because $\Re\{c\}=(c+c^H)/2$ and $c\in\mathbb{C}$. Similarly, the expression $\Re\left\{\tilde{f}_k(\mathbf{\Phi})\tilde{f}_i^H(\mathbf{\Phi})\overline{\mathbf{h}}_{{\rm ID},i}^T\overline{\mathbf{h}}_{{\rm ID},k}^*\right\}$ in (47) is rewritten as follows:

$$\Re\left\{\left(\boldsymbol{\phi}^H\tilde{\mathbf{r}}_i^*\right)\left(\overline{\mathbf{h}}_{{\rm ID},i}^T\overline{\mathbf{h}}_{{\rm ID},k}^*\right)\left(\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}\right)\right\}=\boldsymbol{\phi}^H\tilde{\mathbf{E}}_{k,i}\boldsymbol{\phi}. \tag{65}$$

Finally, we define

$$\mathbf{F}_{k,l}\triangleq(\text{diag}(\overline{\mathbf{h}}_{{\rm SI},l}))^H\overline{\mathbf{h}}_{{\rm ID},k}^*\overline{\mathbf{h}}_{{\rm ID},k}^T\text{diag}(\overline{\mathbf{h}}_{{\rm SI},l}). \tag{66}$$

Then, the expression $|\overline{\mathbf{h}}_{{\rm ID},k}^T\mathbf{\Phi}\overline{\mathbf{h}}_{{\rm SI},l}|^2$ in (48) is rewritten as

$$\left(\overline{\mathbf{h}}_{{\rm ID},k}^T\text{diag}(\overline{\mathbf{h}}_{{\rm SI},l})\boldsymbol{\phi}\right)^H\overline{\mathbf{h}}_{{\rm ID},k}^T\text{diag}(\overline{\mathbf{h}}_{{\rm SI},l})\boldsymbol{\phi}=\boldsymbol{\phi}^H\mathbf{F}_{k,l}\boldsymbol{\phi}. \tag{67}$$

Based on the above simplification, $\mathcal{F}_{k,m}(\mathbf{\Phi})$ and $\mathcal{G}_{k,m}(\mathbf{\Phi})$ are rewritten in (59) and (60) at the bottom of this page.

$$\text{SINR}_{{\rm SR},k}=\frac{E_{\rm S}Mc_k(\varepsilon_k+1)}{E_{\rm S}c_k(\varepsilon_k+1)+E_{\rm S}\sum_{i=1,i\neq k}^{K}c_i(\varepsilon_i+1)+\sigma_{\rm R}^2},\quad \text{SINR}_{{\rm RD},k}=\frac{E_{\rm R}M\tilde{c}_k\left(\tilde{\varepsilon}_k+1\right)}{E_{\rm R}\tilde{c}_k(\tilde{\varepsilon}_k+1)+E_{\rm R}\sum_{i=1,i\neq k}^{K}\tilde{c}_i(\tilde{\varepsilon}_i+1)+\sigma_{\rm D}^2} \tag{58}$$

$$\begin{aligned}\mathcal{F}_{k,1}(\boldsymbol{\phi})=p_{\rm S}\left(\alpha_k^{(1)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2+\alpha_k^{(2)}\right)^2,\quad \mathcal{G}_{k,1}(\boldsymbol{\phi})=p_{\rm S}\left(\beta_k^{(1)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2+\beta_k^{(2)}\right)+p_{\rm S}\sum_{i\neq k}\Bigg(\gamma_{k,i}^{(1)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2+\gamma_{k,i}^{(2)}|\mathbf{r}_i^H\boldsymbol{\phi}|^2+\gamma_{k,i}^{(3)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2|\mathbf{r}_i^H\boldsymbol{\phi}|^2\\+\gamma_{k,i}^{(4)}\boldsymbol{\phi}^H\mathbf{E}_{k,i}\boldsymbol{\phi}+\gamma_{k,i}^{(5)}\Bigg)+(\sigma_{\rm R}^2+p_{\rm R}\sigma_{\rm RR}^2)\left(\alpha_k^{(1)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2+\alpha_k^{(2)}\right)\end{aligned} \tag{59}$$

$$\begin{aligned}\mathcal{F}_{k,2}(\boldsymbol{\phi})=p_{\rm R}\left(\tilde{\alpha}_k^{(1)}|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2+\tilde{\alpha}_k^{(2)}\right)^2,\quad \mathcal{G}_{k,2}(\boldsymbol{\phi})=p_{\rm R}\left(\tilde{\beta}_k^{(1)}|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2+\tilde{\beta}_k^{(2)}\right)+p_{\rm R}\sum_{i\neq k}\Bigg(\tilde{\gamma}_{k,i}^{(1)}|\tilde{\mathbf{r}}_i^T\boldsymbol{\phi}|^2+\tilde{\gamma}_{k,i}^{(2)}|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2+\tilde{\gamma}_{k,i}^{(3)}|\tilde{\mathbf{r}}_i^T\boldsymbol{\phi}|^2|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2\\+\tilde{\gamma}_{k,i}^{(4)}\boldsymbol{\phi}^H\tilde{\mathbf{E}}_{k,i}\boldsymbol{\phi}+\tilde{\gamma}_{k,i}^{(5)}\Bigg)+p_{\rm S}\sum_{l=1}^{K}\left(\tilde{\alpha}_l^{(1)}|\tilde{\mathbf{r}}_l^T\boldsymbol{\phi}|^2+\tilde{\alpha}_l^{(2)}\right)\sum_{l=1}^{K}\left(\zeta_{k,l}^{(1)}\boldsymbol{\phi}^H\mathbf{F}_{k,l}\boldsymbol{\phi}+\zeta_{k,l}^{(2)}\right)+\sigma_{\rm D}^2\sum_{l=1}^{K}\left(\tilde{\alpha}_l^{(1)}|\tilde{\mathbf{r}}_l^T\boldsymbol{\phi}|^2+\tilde{\alpha}_l^{(2)}\right)\end{aligned} \tag{60}$$

Subsequently, Problem (61) can be equivalently reformulated as

$$\max_{\boldsymbol{\phi},\mathbf{s}} \quad \frac{T-\tau}{T}\sum_{k=1}^{K}\log_2\left(1+s_k\right) \tag{68a}$$

$$\text{s.t.} \quad \frac{\mathcal{F}_{k,m}(\boldsymbol{\phi})}{\mathcal{G}_{k,m}(\boldsymbol{\phi})} \geq s_k, s_k \geq 0, \text{ and (61b)} \tag{68b}$$

where $\mathbf{s} \triangleq \{s_k\}$ is the set of auxiliary variables. Problem (68) is non-convex and difficult to solve optimally. Based on the general framework for *fractional matrix programming* [25], we solve it by iteratively solving the following problem

$$\max_{\boldsymbol{\phi},\mathbf{s},\mathbf{t}} \quad \frac{T-\tau}{T}\sum_{k=1}^{K}\log_2\left(1+s_k\right) \tag{69a}$$

$$\text{s.t.} \quad s_k \leq 2\xi_{k,m}^{[\kappa]}t_{k,m} - \left(\xi_{k,m}^{[\kappa]}\right)^2\mathcal{G}_{k,m}(\boldsymbol{\phi}), \quad \forall k,m \tag{69b}$$

$$t_{k,m} \leq \sqrt{\mathcal{F}_{k,m}(\boldsymbol{\phi})}, \quad \forall k,m \tag{69c}$$

$$s_k \geq 0, t_{k,m} \geq 0, \quad \forall k,m \tag{69d}$$

$$\text{(61b)} \tag{69e}$$

where $\mathbf{s} = \{s_k\}, \mathbf{t} = \{t_{k,m}\}$ denote the sets of auxiliary variables that address $\min(\cdot)$ and the fractional functions, respectively. In addition, $\xi_{k,m}^{[\kappa]} = t_{k,m}^{[\kappa]}/\mathcal{G}_{k,m}(\boldsymbol{\phi}^{[\kappa]})$ with $\kappa$ being the iteration index. It can be observed that Problem (69) is non-convex because the constraints in (69b), (69c), and (61b) are non-convex. By applying the MM method [26], we approximate constraints (69b) and (69c) with the convex ones. This can be done by determining the convex upper bound and lower bound for $\mathcal{G}_{k,m}(\boldsymbol{\phi})$ and $\sqrt{\mathcal{F}_{k,m}(\boldsymbol{\phi})}$, respectively.

*Determining a convex upper bound for $\mathcal{G}_{k,m}(\boldsymbol{\phi})$:* We begin by introducing the auxiliary variables $(v_k, \tilde{v}_k, w, u_k)$ such that

$$e^{v_k} = |\mathbf{r}_k^H\boldsymbol{\phi}|^2, \quad e^{\tilde{v}_k} = |\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2, \quad \forall k \tag{70}$$

$$e^{w} = \sum_{l=1}^{K}\left(\tilde{\alpha}_l^{(1)}|\tilde{\mathbf{r}}_l^T\boldsymbol{\phi}|^2 + \tilde{\alpha}_l^{(2)}\right), \tag{71}$$

$$e^{u_k} = \sum_{l=1}^{K}\left(\zeta_{k,l}^{(1)}\boldsymbol{\phi}^H\mathbf{F}_{k,l}\boldsymbol{\phi} + \zeta_{k,l}^{(2)}\right), \forall k. \tag{72}$$

According to [27], any Hermitian matrix can be written as the difference between two positive semi-definite matrices. We denote $\mathbf{E}_{k,i} = \mathbf{E}_{k,i}^{(+)} + \mathbf{E}_{k,i}^{(-)}$, $\tilde{\mathbf{E}}_{k,i} = \tilde{\mathbf{E}}_{k,i}^{(+)} + \tilde{\mathbf{E}}_{k,i}^{(-)}$, where $\mathbf{E}_{k,i}^{(+)} \succeq 0$, $\mathbf{E}_{k,i}^{(-)} \preceq 0$, $\tilde{\mathbf{E}}_{k,i}^{(+)} \succeq 0$, $\tilde{\mathbf{E}}_{k,i}^{(-)} \preceq 0$. Then, we approximate the quadratic terms in (59) and (60) as follows:

$$\begin{aligned}\boldsymbol{\phi}^H\mathbf{E}_{k,i}\boldsymbol{\phi} \leq\ & \boldsymbol{\phi}^H\mathbf{E}_{k,i}^{(+)}\boldsymbol{\phi} + 2\Re\left\{(\boldsymbol{\phi}^{[\kappa]})^H\mathbf{E}_{k,i}^{(-)}\boldsymbol{\phi}\right\} \\ & - (\boldsymbol{\phi}^{[\kappa]})^H\mathbf{E}_{k,i}^{(-)}\boldsymbol{\phi}^{[\kappa]},\end{aligned} \tag{73}$$

$$\begin{aligned}\boldsymbol{\phi}^H\tilde{\mathbf{E}}_{k,i}\boldsymbol{\phi} \leq\ & \boldsymbol{\phi}^H\tilde{\mathbf{E}}_{k,i}^{(+)}\boldsymbol{\phi} + 2\Re\left\{(\boldsymbol{\phi}^{[\kappa]})^H\tilde{\mathbf{E}}_{k,i}^{(-)}\boldsymbol{\phi}\right\} \\ & - (\boldsymbol{\phi}^{[\kappa]})^H\tilde{\mathbf{E}}_{k,i}^{(-)}\boldsymbol{\phi}^{[\kappa]}.\end{aligned} \tag{74}$$

Thus, by combining (70)-(74), we can derive the convex upper bound for $\mathcal{G}_{k,m}(\boldsymbol{\phi})$, which is denoted as $\hat{\mathcal{G}}_{k,m}(\boldsymbol{\phi})$. The bound is given by (77) and (78) at the bottom of the next page.

*Determining a convex lower bound for $\sqrt{\mathcal{F}_{k,m}(\boldsymbol{\phi})}$:* We apply the following approximations

$$|\mathbf{r}_k^H\boldsymbol{\phi}|^2 \geq 2\Re\left\{\left(\mathbf{r}_k^H\boldsymbol{\phi}^{[\kappa]}\right)^*\mathbf{r}_k^H\boldsymbol{\phi}\right\} - |\mathbf{r}_k^H\boldsymbol{\phi}^{[\kappa]}|^2, \tag{75}$$

$$|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2 \geq 2\Re\left\{\left(\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}^{[\kappa]}\right)^*\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}\right\} - |\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}^{[\kappa]}|^2. \tag{76}$$

The approximation presented above is obtained by adopting the inequality $|z|^2 \geq 2\Re\{\bar{z}^*z\} - |\bar{z}|^2$, for a given complex $\bar{z}$. We then derive the convex lower bound for $\sqrt{\mathcal{F}_{k,m}(\boldsymbol{\phi})}$, denoted by $\hat{\mathcal{F}}_{k,m}(\boldsymbol{\phi})$. The bound is given in (79) and (80) at the bottom of the next page. Thus, by combining (77)-(80), the solution to Problem (69) can be obtained by iteratively solving the following problem

$$\max_{\substack{\boldsymbol{\phi},\mathbf{s},\mathbf{t},\\ \mathbf{u},\mathbf{v},\tilde{\mathbf{v}},w}} \quad \frac{T-\tau}{T}\sum_{k=1}^{K}\log_2\left(1+s_k\right) \tag{81a}$$

$$\text{s.t.} \quad s_k \leq 2\xi_{k,m}^{[\kappa]}t_{k,m} - (\xi_{k,m}^{[\kappa]})^2\hat{\mathcal{G}}_{k,m}(\boldsymbol{\phi}), \quad \forall k,m \tag{81b}$$

$$t_{k,m} \leq \hat{\mathcal{F}}_{k,m}(\boldsymbol{\phi}), \quad \forall k,m \tag{81c}$$

$$|\mathbf{r}_k^H\boldsymbol{\phi}|^2 \leq e^{v_k}, |\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2 \leq e^{\tilde{v}_k}, \quad \forall k \tag{81d}$$

$$\sum_{l=1}^{K}\left(\tilde{\alpha}_l^{(1)}|\tilde{\mathbf{r}}_l^T\boldsymbol{\phi}|^2 + \tilde{\alpha}_l^{(2)}\right) \leq e^{w} \tag{81e}$$

$$\sum_{l=1}^{K}\left(\zeta_{k,l}^{(1)}\boldsymbol{\phi}^H\mathbf{F}_{k,l}\boldsymbol{\phi} + \zeta_{k,l}^{(2)}\right) \leq e^{u_k}, \forall k \tag{81f}$$

$$\text{(69d), (61b)} \tag{81g}$$

where $\mathbf{u} = \{u_k\}$, $\mathbf{v} = \{v_k\}$, and $\tilde{\mathbf{v}} = \{\tilde{v}_k\}$ denote the sets of auxiliary variables. However, the above problem cannot be solved directly owing to the non-convex constraints (81d)-(81f), and (61b). To address the non-convexity of (81d)-(81f), we adopt the inequality $e^{\bar{\nu}}(\nu - \bar{\nu} + 1) \leq e^{\nu}$ for a given $\bar{\nu}$. Thus, these constraints can be relaxed as follows:

$$|\mathbf{r}_k^H\boldsymbol{\phi}|^2 \leq e^{v_k^{[\kappa]}}(v_k - v_k^{[\kappa]} + 1), \quad \forall k \tag{82}$$

$$|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2 \leq e^{\tilde{v}_k^{[\kappa]}}(\tilde{v}_k - \tilde{v}_k^{[\kappa]} + 1), \quad \forall k \tag{83}$$

$$\sum_{l=1}^{K}\left(\tilde{\alpha}_l^{(1)}|\tilde{\mathbf{r}}_l^T\boldsymbol{\phi}|^2 + \tilde{\alpha}_l^{(2)}\right) \leq e^{w^{[\kappa]}}(w - w^{[\kappa]} + 1), \tag{84}$$

$$\sum_{l=1}^{K}\left(\zeta_{k,l}^{(1)}\boldsymbol{\phi}^H\mathbf{F}_{k,l}\boldsymbol{\phi} + \zeta_{k,l}^{(2)}\right) \leq e^{u_k^{[\kappa]}}(u_k - u_k^{[\kappa]} + 1), \forall k \tag{85}$$

Then, we need to solve the following problem

$$\max_{\substack{\boldsymbol{\phi},\mathbf{s},\mathbf{t},\\ \mathbf{u},\mathbf{v},\tilde{\mathbf{v}},w}} \quad \frac{T-\tau}{T}\sum_{k=1}^{K}\log_2\left(1+s_k\right) \tag{86a}$$

$$\text{s.t.} \quad \text{(81b), (81c), (82)} - \text{(85), (69d), (61b)} \tag{86b}$$

It remains to address the non-convex constraint (61b). We apply the *square penalty* method [28]. In particular, we first relax (61b) into an inequality constraint. Subsequently, we add a penalty term $\lambda\|\boldsymbol{\phi}\|^2$ to the objective function to enforce the equality at optimality. Here, $\lambda > 0$ and $\|\cdot\|$ denotes the Euclidean norm. However, the penalty term is non-convex, which is addressed by adopting the inequality $\|\mathbf{z}\|^2 \geq 2\Re\{\bar{\mathbf{z}}^H\mathbf{z}\} - \|\bar{\mathbf{z}}\|^2$, for a given complex vector $\bar{\mathbf{z}}$ [28].

Then, Problem (86) becomes

$$\max_{\mathbf{X}} \quad \Psi(\mathbf{X}) \triangleq \frac{T-\tau}{T}\sum_{k=1}^{K}\log_2(1+s_k) + \lambda(2\Re\{(\boldsymbol{\phi}^{[\kappa]})^H\boldsymbol{\phi}\} - \|\boldsymbol{\phi}^{[\kappa]}\|^2) \tag{87a}$$

$$\text{s.t.} \quad (81\text{b}), (81\text{c}), (82)-(85), (69\text{d}) \tag{87b}$$

$$|\phi_n|^2 \le 1, \quad \forall n \tag{87c}$$

where $\mathbf{X} \triangleq \{\boldsymbol{\phi}, \mathbf{s}, \mathbf{t}, \mathbf{u}, \mathbf{v}, \tilde{\mathbf{v}}, w\}$. This problem is convex and can be efficiently solved using the CVX solver [29]. The algorithm for solving Problem (61) is presented in Algorithm 1. The main complexity of the proposed iterative algorithm lies in the interior-point algorithm deployed in the CVX solver. The worst-case complexity is $\mathcal{O}(N^{3.5}\log(1/\varepsilon)I_1I_0)$, where $\varepsilon$, $I_1$, and $I_0$ denote the solution accuracy of the interior-point method, the number of required iterations in the inner and outer loop of Algorithm 1, respectively. According to [28, Theorem 2], the study shows that any local optimal solution to Problem (87) must be a feasible solution to Problem (86) for a sufficiently large penalty $\lambda$. In Algorithm 1, the adaptive rule [28], [30] is adopted. The main idea of this update rule is to start from a small value of $\lambda$ so that the problem is easier to handle. Then the hardness of the problem increases as $\lambda$ is gradually increased. Specifically, the initial value $\lambda$ is set to a small value (i.e., $10^{-3}$), then increased by a scaling factor $c > 1$ (i.e., $\lambda \leftarrow c\lambda$). This value is gradually increased until $\max_n\{||\phi_n|-1|\}$ is less than a predefined threshold or the maximum value $\lambda_{\max}$ is reached.

The convergence analysis is provided in the following. For a given $\lambda$, it can be seen that the sequence of objective functions is non-decreasing, i.e., $\Psi(\mathbf{X}^{[\kappa]}) \le \Psi(\mathbf{X}^{[\kappa+1]})$. In addition, $\Psi(\mathbf{X}^{[\kappa+1]})$ is upper bounded. This is because $\mathcal{F}_{k,m}(\boldsymbol{\phi})$ and $\mathcal{G}_{k,m}(\boldsymbol{\phi})$ take finite values, and the constraint $|\phi_n|^2 \le 1$ represents a bounded region. Thus, the sequence $\Psi(\mathbf{X}^{[\kappa]})$ is convergent. Moreover, based on the MM method, the limit of a sequence $\{\mathbf{X}^{[\kappa]}\}$ is a Karush–Kuhn–Tucker (KKT) point [26, Sec. IV-F].

*Remark* 4. Although the theoretical worst-case complexity is $\mathcal{O}(N^{3.5}\log(1/\varepsilon)I_1I_0)$, the proposed algorithm remains practical. Because statistical CSI evolves slowly, phase shifts are updated quasi-statically on a long timescale. Furthermore, execution is handled by the computationally capable relay rather than the passive IRS, and the algorithm converges in a few iterations ($I_1, I_0$), ensuring manageable runtime.

**Algorithm 1** IRS Phase Shift Design

**Initialization:** Initialize feasible point $\mathbf{X}^{(0)}$, set the initial penalty parameter $\lambda$, maximum value $\lambda_{max}$, scaling factor $c > 1$, and tolerances $\epsilon_1 > 0, \epsilon_0 > 0$.
**Repeat**
  Set the iteration index $\kappa = 0$.
  **Repeat**
    Solve Problem (87) to obtain solution $\mathbf{X}^\star$.
    Update $\mathbf{X}^{[\kappa+1]} \leftarrow \mathbf{X}^\star$ and $\xi_{k,m}^{[\kappa+1]} \leftarrow t_{k,m}^\star/\mathcal{G}_{k,m}(\boldsymbol{\phi}^\star)$.
    $\kappa \leftarrow \kappa + 1$.
  **Until** $|\Psi(\mathbf{X}^{[\kappa]}) - \Psi(\mathbf{X}^{[\kappa-1]})| \le \epsilon_1$.
  Update $\mathbf{X}^{(0)}$ by the currently obtained solution.
  Update $\lambda \leftarrow c\lambda$.
**Until** $\max_n\{||\phi_n|-1|\} \le \epsilon_0$ or $\lambda > \lambda_{max}$.

## IV. Experimental Results

In this section, we validate the effectiveness of the proposed hybrid IRS-relay system with different key system parameters. We consider $K = 3$ source-destination pairs, where the sources and destinations are distributed randomly in the circles centered at (0 m, 0 m) and (140 m, 0 m) with a radius of 10 (m), respectively. The positions of the IRS and the relay are (70 m, -15 m) and (70 m, 15 m), respectively. The path loss coefficients are defined as $\eta_{\varsigma,k} = 10^{-3}d_{\varsigma,k}^{-4}$, for $\varsigma \in \{\text{SR}, \text{RD}\}$, $\eta_{\varsigma',k} = 10^{-3}d_{\varsigma',k}^{-2.5}$, for $\varsigma' \in \{\text{SI}, \text{ID}, \text{IR}, \text{RI}\}$. Unless stated otherwise, we set $\tau = 2K$, $T = 200$, $p_{\text{p}} = p_{\text{S}} = p_{\text{R}} = 10$ (dBm), $\varepsilon_k = \tilde{\varepsilon}_k = 5$ and $\delta = \tilde{\delta} = 10$, $\sigma_{\text{R}}^2 = \sigma_{\text{D}}^2 = \sigma^2 = -94$ (dBm). The residual self-interference level $\sigma_{\text{RR}}^2 = 0$ (dB) with respect to $\sigma^2$. In this evaluation, we compare the proposed method (**Proposed (Stat. CSI - Theoretical)**) with the following methods:

$$\mathcal{G}_{k,1}(\boldsymbol{\phi}) \le \hat{\mathcal{G}}_{k,1}(\boldsymbol{\phi}) \triangleq p_{\text{S}}\left(\beta_k^{(1)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2 + \beta_k^{(2)}\right) + p_{\text{S}}\sum_{i=1,i\neq k}^{K}\Bigg[\gamma_{k,i}^{(1)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2 + \gamma_{k,i}^{(2)}|\mathbf{r}_i^H\boldsymbol{\phi}|^2 + \gamma_{k,i}^{(3)}e^{v_k}e^{v_i} + \gamma_{k,i}^{(4)}\left(\boldsymbol{\phi}^H\mathbf{E}_{k,i}^{(+)}\boldsymbol{\phi} + 2\Re\left\{(\boldsymbol{\phi}^{[\kappa]})^H\mathbf{E}_{k,i}^{(-)}\boldsymbol{\phi}\right\} - (\boldsymbol{\phi}^{[\kappa]})^H\mathbf{E}_{k,i}^{(-)}\boldsymbol{\phi}^{[\kappa]}\right) + \gamma_{k,i}^{(5)}\Bigg] + (\sigma_{\text{R}}^2 + p_{\text{R}}\sigma_{\text{RR}}^2)\left(\alpha_k^{(1)}|\mathbf{r}_k^H\boldsymbol{\phi}|^2 + \alpha_k^{(2)}\right) \tag{77}$$

$$\mathcal{G}_{k,2}(\boldsymbol{\phi}) \le \hat{\mathcal{G}}_{k,2}(\boldsymbol{\phi}) \triangleq p_{\text{R}}\left(\tilde{\beta}_k^{(1)}|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2 + \tilde{\beta}_k^{(2)}\right) + p_{\text{R}}\sum_{i=1,i\neq k}^{K}\Bigg[\tilde{\gamma}_{k,i}^{(1)}|\tilde{\mathbf{r}}_i^T\boldsymbol{\phi}|^2 + \tilde{\gamma}_{k,i}^{(2)}|\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}|^2 + \tilde{\gamma}_{k,i}^{(3)}e^{\tilde{v}_k}e^{\tilde{v}_i} + \tilde{\gamma}_{k,i}^{(4)}\left(\boldsymbol{\phi}^H\tilde{\mathbf{E}}_{k,i}^{(+)}\boldsymbol{\phi} + 2\Re\left\{(\boldsymbol{\phi}^{[\kappa]})^H\tilde{\mathbf{E}}_{k,i}^{(-)}\boldsymbol{\phi}\right\} - (\boldsymbol{\phi}^{[\kappa]})^H\tilde{\mathbf{E}}_{k,i}^{(-)}\boldsymbol{\phi}^{[\kappa]}\right) + \tilde{\gamma}_{k,i}^{(5)}\Bigg] + p_{\text{S}}e^{w}e^{u_k} + \sigma_{\text{D}}^2\sum_{l=1}^{K}\left(\tilde{\alpha}_l^{(1)}|\tilde{\mathbf{r}}_l^T\boldsymbol{\phi}|^2 + \tilde{\alpha}_l^{(2)}\right) \tag{78}$$

$$\sqrt{\mathcal{F}_{k,1}(\boldsymbol{\phi})} \ge \hat{\mathcal{F}}_{k,1}(\boldsymbol{\phi}) \triangleq \sqrt{p_S}\left[\alpha_k^{(1)}\left(2\Re\left\{\left(\mathbf{r}_k^H\boldsymbol{\phi}^{[\kappa]}\right)^*\mathbf{r}_k^H\boldsymbol{\phi}\right\} - |\mathbf{r}_k^H\boldsymbol{\phi}^{[\kappa]}|^2\right) + \alpha_k^{(2)}\right] \tag{79}$$

$$\sqrt{\mathcal{F}_{k,2}(\boldsymbol{\phi})} \ge \hat{\mathcal{F}}_{k,2}(\boldsymbol{\phi}) \triangleq \sqrt{p_R}\left[\tilde{\alpha}_k^{(1)}\left(2\Re\left\{\left(\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}^{[\kappa]}\right)^*\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}\right\} - |\tilde{\mathbf{r}}_k^T\boldsymbol{\phi}^{[\kappa]}|^2\right) + \tilde{\alpha}_k^{(2)}\right] \tag{80}$$

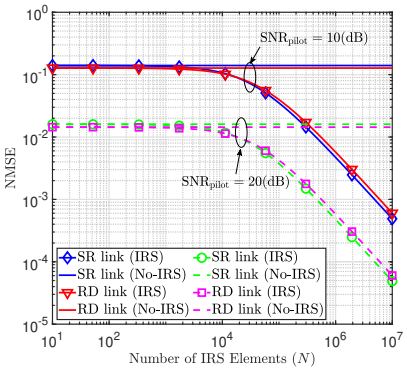


Fig. 2. NMSE performance versus $N$, with different values of $\mathrm{SNR}_{\mathrm{pilot}} = p_{\mathrm{p}}/\sigma^2$.

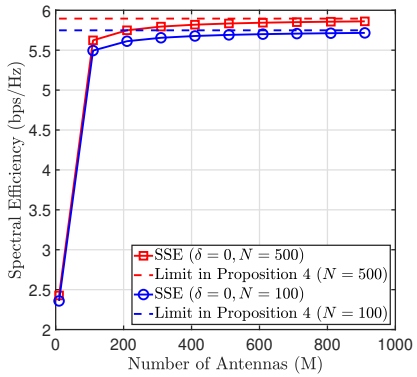


Fig. 3. Power scaling analysis with M for $\delta = 0$.

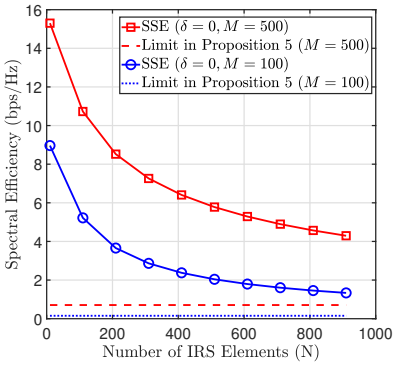


Fig. 4. Power scaling analysis with N for $\delta = 0$.

- **Proposed (Inst. CSI):** To verify the tightness of the UatF lower bound, the closed-form rate (51) is compared with the ergodic rate (53), obtained via Monte Carlo simulations. Here, "Inst. CSI" indicates that the receivers perfectly know the instantaneous effective channel gains (i.e., a genie-aided receiver).
- **Proposed (Stat. CSI - Empirical):** To verify the derivation correctness, the statistical moments in (49) and (50) are validated against Monte Carlo simulations, under the proposed algorithm.
- **FD-relay+IRS (Random Phase):** An IRS-aided FD relay baseline in which the IRS phase-shift coefficients are randomly and independently generated following a uniform distribution over $[0, 2\pi)$, without any optimization.
- **FD-relay+IRS (Fixed Phase):** An IRS-aided FD relay baseline in which all IRS reflecting elements are configured with fixed phase shift.
- **FD-relay (No IRS)**: The considered FD relay system without using IRS.
- **FD-relay (Ngo *et al.*)**: The FD relay system proposed in [19], which adopts the same system settings as those of **FD-relay (No IRS)** scheme. This scheme is used as a benchmark to validate the achievable rate analysis in Theorem 1 with the special case of an FD relay system without IRS.
- **HD-relay+IRS** : The hybrid HD-IRS relay system, where the transmission is performed in two phases: (i) from the sources to the relay and the IRS, and (ii) from the relay and the IRS to the destinations. In the HD mode, no self-interference is present as transmission is performed over two equal adjacent time slots. Consequently, a pre-log factor of $1/2$ is included in the spectral efficiency expression.
- **HD-relay (No IRS)**: The HD relay system without IRS.
- **HD-relay (Ngo *et al.*)**: The HD relay system proposed in [19].
- **IRS-only:** A benchmark in which the source-destination links rely solely on the IRS-reflected path, with the IRS phase-shift matrix optimized to maximize the sum spectral efficiency of this configuration.

Fig. 2 illustrates the NMSE of the LMMSE estimates for the SR and RD links versus $N$. The result is consistent with Proposition 3. Specifically, for small-to-moderate $N$, the IRS-aided curves exhibit no difference compared with the No-IRS benchmark, because the IRS gain is small. As $N$ increases, this gain dominates, driving the NMSE below the benchmark. In addition, increasing $\mathrm{SNR}_{\mathrm{pilot}}$ from 10 to 20 dB scales down both the benchmark and the IRS curves.

Fig. 3 depicts the power scaling with respect to the number of relay antennas $M$. Here, IRS-relay and relay-IRS channels are purely NLoS ($\delta = \tilde{\delta} = 0$), and the transmit power of each source and relay scales proportionally to $1/M$. According to Proposition 4, the figure shows that as the number of relay antennas $M$ increases, the spectral efficiency approaches a non-zero value. This indicates that the transmit power of each source and relay can be proportionally reduced to $1/M$, while the SSE can be maintained to a non-zero value. Furthermore, systems with $N = 500$ achieve higher asymptotic limits than those with $N = 100$. This is because a larger number of IRS elements yields a larger effective channel gain. Fig. 4 depicts the power scaling with the number of IRS elements $N$. Here, IRS-relay and relay-IRS channels are purely NLoS, and the transmit power of each source and relay scales proportionally to $1/N$. This figure shows that SSE tends to approach a non-zero value as $N \to \infty$, which is consistent with Proposition 5. Moreover, systems with more relay antennas ($M = 500$) achieve higher performance and a higher limit than those with fewer ($M = 100$).

Fig. 5 depicts the convergence behavior of the proposed algorithm with $N \in \{50, 200\}$ and $M = 100$. For both cases, the proposed algorithm converges within a small number of iterations, demonstrating fast convergence. Moreover, increasing the number of IRS elements improves the spectral efficiency owing to the enhanced passive beamforming gain. Fig. 6 depicts the SSE as a function of the number of IRS elements. We can see that the small gap between **Proposed (Inst. CSI)** and **Proposed (Stat. CSI - Theoretical)** confirms the bound's tightness. In addition, the overlap with **Proposed (Stat. CSI - Empirical)** validates the derived statistics. As shown, the performance of the proposed schemes increases with the number of IRS elements, whereas the baseline schemes without IRS, random, and fixed phase shift exhibit flat performance. This

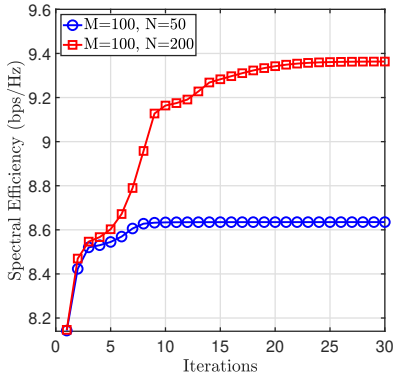


Fig. 5. Convergence of the proposed algorithm.

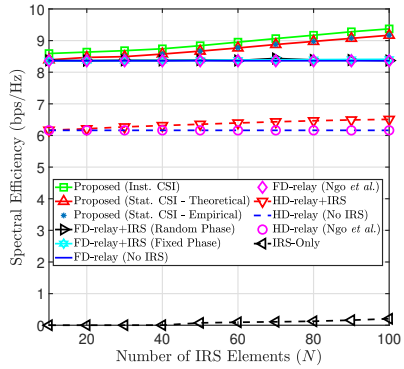


Fig. 6. SSE versus the number of IRS elements $N$ with $M = 100$.

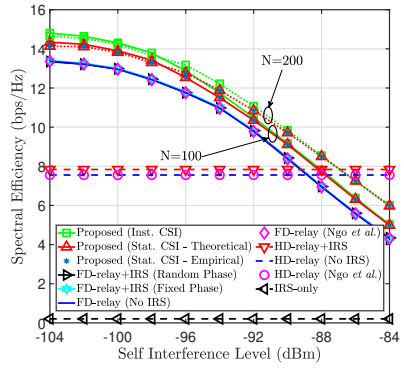


Fig. 7. SSE versus self-interference level, with $M = 200$.

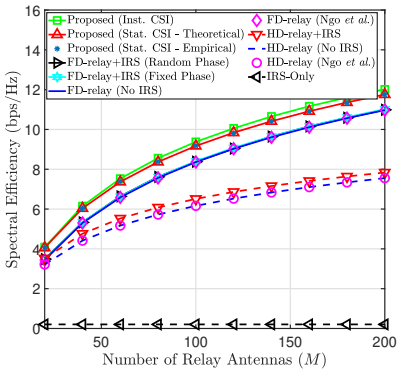


Fig. 8. SSE versus the number of relay transmit/receive antennas $M$, with $N = 100$.

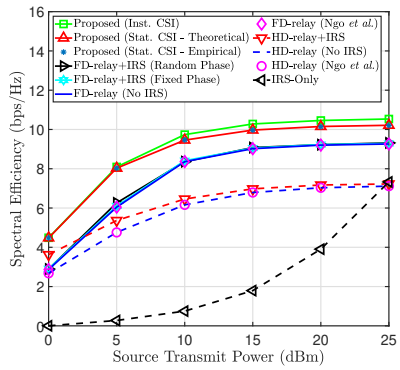


Fig. 9. SSE versus the transmit power at each source $p_{\mathrm{S}}$ (dBm), with $N = 100, M = 200$.

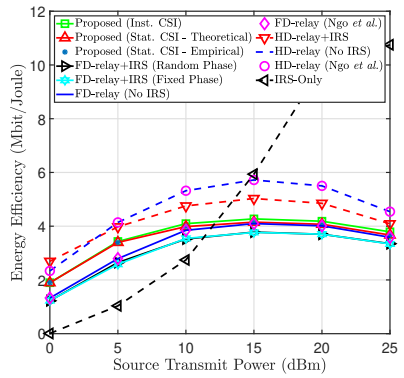


Fig. 10. Energy efficiency of the IRS-aided FD system.

behavior demonstrates the performance gains by deploying a larger IRS array. Furthermore, all schemes outperform the **IRS-Only** baseline, as a passive IRS requires a very large number of elements to match the active gain of the FD relay. In addition, the FD relay schemes consistently outperform their HD counterparts. The performance advantage of the FD schemes over the HD schemes is mainly attributed to the two-phase transmission in the HD mode, which incurs a $1/2$ pre-log penalty. Fig. 7 depicts the impact of residual self-interference on spectral efficiency. It is observed that a higher self-interference level drastically reduces the SSE performance of all FD schemes when compared with the low-interference case. In particular, when residual self-interference is large, self-interference error dominates and causes FD performance to fall below HD. However, increasing $N$ (from 100 to 200) substantially mitigates the performance loss caused by high self-interference. This helps expand the operational regime where FD-IRS retains its superiority over HD-IRS.

Fig. 8 depicts the SSE as a function of the number of relay transmit/receive antennas $M$. It can be observed that the SSE increases with $M$. In addition, the performance gap between the proposed schemes and HD schemes gradually widens as $M$ increases, which indicates that large antenna arrays can be effectively exploited to improve the performance. Furthermore, the proposed FD schemes outperform the FD baselines without IRS. This figure also highlights the energy efficiency of integrating a passive IRS into the FD relay. For a target SE of 10.9 bps/Hz, the **FD-relay (No IRS)** scheme requires $M = 197$ active antennas. The proposed scheme achieves the same SE with only $M = 160$ antennas, saving 37 active antennas. Because passive elements consume much less power than active RF chains, this hardware reduction significantly decreases deployment cost and total energy consumption.

Fig. 9 depicts the SSE versus transmit power. As the transmit power increases, the SSE of all the schemes rapidly increases in the low-power regime (below 15 dBm). When the transmit power exceeds this value, the performance improvement becomes more gradual. The proposed scheme outperforms the benchmark schemes. Fig. 10 illustrates the energy efficiency versus the source transmit power. The energy efficiency is defined as $\mathrm{EE} = \frac{B \cdot \mathrm{SSE}}{P_{\mathrm{total}}}$, where $B$ and $P_{\mathrm{total}}$ represent the bandwidth and total power consumption, respectively [31]. The high performance of the proposed schemes shows that passive beamforming gain effectively offsets the added IRS circuit power. Although HD-relay schemes achieve higher energy efficiency owing to using half the active RF chains, the proposed FD design provides higher spectral efficiency, as shown in Fig. 9. In addition, the IRS-only baseline fails at low power due to double path loss, but achieves the highest EE above 12 dBm by eliminating active circuit consumption.

## V. Conclusion

In this study, we investigated a multi-pair IFDR system with massive relay antennas. We first determined the channel estimate of the cascaded channels and related statistics of the IFDR system, using the LMMSE method. Subsequently, we derived the closed-form lower bound for the achievable rate using MRC/MRT processing and the UatF technique. Second, we analyzed the power scaling laws under large-scale antenna and IRS regimes. Third, we formulated the IRS phase shift design problem for maximizing the sum spectral efficiency and developed an efficient iterative algorithm based on fractional matrix programming, the MM, and square penalty methods. Simulations confirmed that our proposed scheme outperforms the benchmark schemes. While the proposed framework yields tractable rate expressions, several practical limitations remain. Future work will extend this model by: leveraging IRS phase shifts for passive self-interference cancellation, incorporating near-field spherical wave models, developing low-complexity algorithms for large-scale IRS, and considering secrecy constraints and unblocked direct links.

## Appendix A

Using the LMMSE estimator [20] (see Sec. 12.5, p. 382) with the observations provided in (13), we obtain

$$\hat{\mathbf{g}}_{\mathrm{SR},k} = \mathbb{E}\{\mathbf{g}_{\mathrm{SR},k}\} + \mathrm{Cov}\{\mathbf{g}_{\mathrm{SR},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k}\}\mathrm{Cov}^{-1}\{\tilde{\mathbf{y}}_{\mathrm{rp},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k}\} \times (\tilde{\mathbf{y}}_{\mathrm{rp},k} - \mathbb{E}\{\tilde{\mathbf{y}}_{\mathrm{rp},k}\}). \tag{88}$$

By applying [17, Lemma 1], we obtain $\mathrm{Cov}\{\tilde{\mathbf{y}}_{\mathrm{rp},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k}\} = a_{k,1}\mathbf{a}_{\mathrm{R}}\mathbf{a}_{\mathrm{R}}^H + (a_{k,2} + \frac{\sigma^2}{\tau p_{\mathrm{p}}})\mathbf{I}_{\mathrm{M}}$, where $a_{k,1}, a_{k,2}$ are defined in (17). Then, applying the Woodbury's identity (see [20], p. 571) and $\mathbf{a}_{\mathrm{R}}^H\mathbf{a}_{\mathrm{R}} = M$, we obtain

$$\mathrm{Cov}^{-1}\{\tilde{\mathbf{y}}_{\mathrm{rp},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k}\} = \left(a_{k,2} + \frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)^{-1}\mathbf{I}_M - \frac{a_{k,1}\left(a_{k,2} + \frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)^{-2}}{1 + M a_{k,1}\left(a_{k,2} + \frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)^{-1}}\mathbf{a}_{\mathrm{R}}\mathbf{a}_{\mathrm{R}}^H.$$

In addition, we can verify that $\mathrm{Cov}\{\mathbf{g}_{\mathrm{SR},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k}\} = a_{k,1}\mathbf{a}_{\mathrm{R}}\mathbf{a}_{\mathrm{R}}^H + a_{k,2}\mathbf{I}_{\mathrm{M}}$. Then,

$$\mathrm{Cov}\{\mathbf{g}_{\mathrm{SR},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k}\}\mathrm{Cov}^{-1}\{\tilde{\mathbf{y}}_{\mathrm{rp},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k}\} = a_{k,3}\mathbf{a}_{\mathrm{R}}\mathbf{a}_{\mathrm{R}}^H + a_{k,4}\mathbf{I}_M \triangleq \mathbf{A}_{\mathrm{SR},k}, \tag{89}$$

where $a_{k,3}$, $a_{k,4}$, and $\mathbf{A}_{\mathrm{SR},k}$ are defined in (18), (19), and (16), respectively. Next, we can verify that $\mathbb{E}\{\mathbf{g}_{\mathrm{SR},k}\} = \sqrt{c_k\delta\varepsilon_k}\overline{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k}$. Combining this identity and (89), Eq. (88) becomes:

$$\hat{\mathbf{g}}_{\mathrm{SR},k} = \mathbf{A}_{\mathrm{SR},k}\tilde{\mathbf{y}}_{\mathrm{rp},k} + (\mathbf{I}_{\mathrm{M}} - \mathbf{A}_{\mathrm{SR},k})\sqrt{c_k\delta\varepsilon_k}\overline{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k}.$$

This can be simplified to obtain the estimate (15) by applying (11) and (13). The estimate for $\mathbf{g}_{\mathrm{RD},k}$ is obtained similarly.

## Appendix B

We have

$$\begin{aligned}\mathbf{MSE}_{\mathrm{SR},k} &= \mathbb{E}\{\mathbf{e}_{\mathrm{SR},k}\mathbf{e}_{\mathrm{SR},k}^H\} \overset{(a)}{=} \mathrm{Cov}(\mathbf{g}_{\mathrm{SR},k}, \mathbf{g}_{\mathrm{SR},k}) \\ &- \mathrm{Cov}(\mathbf{g}_{\mathrm{SR},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k})\mathrm{Cov}^{-1}(\tilde{\mathbf{y}}_{\mathrm{rp},k}, \tilde{\mathbf{y}}_{\mathrm{rp},k})\mathrm{Cov}(\tilde{\mathbf{y}}_{\mathrm{rp},k}, \mathbf{g}_{\mathrm{SR},k}) \\ &\overset{(b)}{=} \mathrm{Cov}(\mathbf{g}_{\mathrm{SR},k}, \mathbf{g}_{\mathrm{SR},k}) - \mathbf{A}_{\mathrm{SR},k}\mathrm{Cov}(\mathbf{g}_{\mathrm{SR},k}, \mathbf{g}_{\mathrm{SR},k}) \\ &= (\mathbf{I}_{\mathrm{M}} - \mathbf{A}_{\mathrm{SR},k})\mathrm{Cov}(\mathbf{g}_{\mathrm{SR},k}, \mathbf{g}_{\mathrm{SR},k}),\end{aligned} \tag{90}$$

where (a) follows from [20, Eq. (12.21)] and (b) follows from (89) and Lemmma 1 [17]. Then, according to [17], the NMSE of the estimate $\mathbf{g}_{\mathrm{SR},k}$ follows. The NMSE of the estimate $\mathbf{e}_{\mathrm{RD},k}$ is computed similarly. This completes the proof.

## Appendix C

The first limit in (27) is obtained by the fact that the dominant terms in the numerator in (25) scale with $\mathcal{O}(N^2)$, while those in the denominator scale with $\mathcal{O}(N^3)$. The second limit is obtained by the fact that both the numerator and denominator have a highest degree of $\left(\frac{\sigma^2}{\tau p_{\mathrm{p}}}\right)^2$ with identical leading coefficients. The last limit follows by substituting $\frac{\sigma^2}{\tau p_{\mathrm{p}}} = 0$ into (25). For limit in (28), by substituting $p_{\mathrm{p}} = \frac{E_{\mathrm{p}}}{N}$ into (25), the quantities $a_{k,1}$, $a_{k,2}$, and $\frac{\sigma^2}{\tau p_{\mathrm{p}}}$ all scale linearly with $N$ as $N \to \infty$. The $N^3$ factor cancels identically from the numerator and denominator, implying the closed-form limit with the numerator being less than the denominator. A similar argument can be applied to prove (29) and (30).

## Appendix D

In the following, we derive each term of $\mathrm{SINR}_{\mathrm{SR},k}$. Computation for $\mathrm{SINR}_{\mathrm{RD},k}$ can be performed similarly.

*1) Compute* $\mathrm{DS}_{\mathrm{SR},k} = p_{\mathrm{S}}\left|\mathbb{E}\left\{\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\right\}\right|^2$*:* We have

$$\begin{aligned}&\mathbb{E}\{\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\} = \mathbb{E}\{\hat{\mathbf{g}}_{\mathrm{SR},k}^H\mathbf{g}_{\mathrm{SR},k}\} \\ &\overset{(a)}{=} \sum_{\omega=1}^{4}\mathbb{E}\{(\hat{\mathbf{q}}_{\mathrm{SR},k}^{(\omega)})^H\mathbf{q}_{\mathrm{SR},k}^{(\omega)}\} + \eta_{\mathrm{SR},k}\mathbb{E}\{\tilde{\mathbf{h}}_{\mathrm{SR},k}^H\mathbf{A}_{\mathrm{SR},k}^H\tilde{\mathbf{h}}_{\mathrm{SR},k}\} \\ &= c_k\delta\varepsilon_k\overline{\mathbf{h}}_{\mathrm{SI},k}^H\mathbf{\Phi}^H\overline{\mathbf{H}}_{\mathrm{IR}}^H\overline{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k} \\ &+ e_{k,2}c_k\delta\mathbb{E}\{\tilde{\mathbf{h}}_{\mathrm{SI},k}^H\mathbf{\Phi}^H\overline{\mathbf{H}}_{\mathrm{IR}}^H\overline{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\tilde{\mathbf{h}}_{\mathrm{SI},k}\} \\ &+ c_k\varepsilon_k\overline{\mathbf{h}}_{\mathrm{SI},k}^H\mathbf{\Phi}^H\mathbb{E}\{\tilde{\mathbf{H}}_{\mathrm{IR}}^H\mathbf{A}_{\mathrm{SR},k}^H\tilde{\mathbf{H}}_{\mathrm{IR}}\}\mathbf{\Phi}\overline{\mathbf{h}}_{\mathrm{SI},k} \\ &+ c_k\mathbb{E}\{\tilde{\mathbf{h}}_{\mathrm{SI},k}^H\mathbf{\Phi}^H\tilde{\mathbf{H}}_{\mathrm{IR}}^H\mathbf{A}_{\mathrm{SR},k}^H\tilde{\mathbf{H}}_{\mathrm{IR}}\mathbf{\Phi}\tilde{\mathbf{h}}_{\mathrm{SI},k}\} + \eta_{\mathrm{SR},k}\mathrm{Tr}\{\mathbf{A}_{\mathrm{SR},k}^H\} \\ &\overset{(b)}{=} M[|f_k(\mathbf{\Phi})|^2c_k\delta\varepsilon_k + Nc_k\delta e_{k2} + (Nc_k(\varepsilon_k+1) + \eta_{\mathrm{SR},k})e_{k,1}],\end{aligned}$$

where (a) follows by the fact that $\tilde{\mathbf{h}}_{\mathrm{SI},k}$, $\tilde{\mathbf{h}}_{\mathrm{SR},k}$, $\tilde{\mathbf{H}}_{\mathrm{IR}}$ and $\mathbf{N}_{\mathrm{rp},k}$ are independent and have zero mean; (b) follows by applying [17, Lemma 7], the law of iterative expectation, and identities: $\mathbf{\Phi}^H\mathbf{\Phi} = \mathbf{I}_N, \overline{\mathbf{H}}_{\mathrm{IR}}^H\overline{\mathbf{H}}_{\mathrm{IR}} = M\mathbf{a}_{\mathrm{I}}\mathbf{a}_{\mathrm{I}}^H$, $\mathrm{Tr}\left(\mathbf{A}_{\mathrm{SR},k}^H\right) = Me_{k,1}$.

*2) Compute* $\mathrm{SI}_{\mathrm{SR},k}$*:* We have $\mathrm{SI}_{\mathrm{SR},k} = p_{\mathrm{R}}\mathbb{E}\left\{|\mathbf{w}_k^T\mathbf{z}_{\mathrm{RR}}|^2\right\} = p_{\mathrm{R}}\mathbb{E}\left\{\mathbf{w}_k^T\mathbf{z}_{\mathrm{RR}}\mathbf{z}_{\mathrm{RR}}^H\mathbf{w}_k^*\right\} \overset{(c)}{=} p_{\mathrm{R}}\mathbb{E}\left\{\mathbf{w}_k^T\mathbb{E}\left\{\mathbf{z}_{\mathrm{RR}}\mathbf{z}_{\mathrm{RR}}^H\right\}\mathbf{w}_k^*\right\} = p_{\mathrm{R}}\sigma_{\mathrm{RR}}^2\mathbb{E}\left\{\|\mathbf{w}_k^T\|^2\right\}$, where the equality (c) is obtained from the fact that the residual self-interference vector $\mathbf{z}_{\mathrm{RR}}$ is independent of $\mathbf{w}_k$. Next, we compute $\mathbb{E}\left\{\|\mathbf{w}_k^T\|^2\right\}$. As $\mathbf{e}_{\mathrm{SR},k}$ and $\hat{\mathbf{g}}_{\mathrm{SR},k}$ are uncorrelated and $\mathbb{E}\{\mathbf{e}_{\mathrm{SR},k}\} = 0$, thus $\mathbb{E}\{\hat{\mathbf{g}}_{\mathrm{SR},k}^H\mathbf{e}_{\mathrm{SR},k}\} = 0$. Then

$$\mathbb{E}\left\{\hat{\mathbf{g}}_{\mathrm{SR},k}^H\mathbf{g}_{\mathrm{SR},k}\right\} = \mathbb{E}\{\|\hat{\mathbf{g}}_{\mathrm{SR},k}\|^2\}. \tag{91}$$

With the MRC processing, $\mathbb{E}\{\|\mathbf{w}_k^T\|^2\} = \mathbb{E}\{\|\hat{\mathbf{g}}_{\mathrm{SR},k}^H\|^2\}$ and $\mathbb{E}\{\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\} = \mathbb{E}\{\hat{\mathbf{g}}_{\mathrm{SR},k}^H\mathbf{g}_{\mathrm{SR},k}\}$. From (91), we then obtain $\mathbb{E}\{\|\mathbf{w}_k^T\|^2\} = \mathbb{E}\{\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\}$. The computation of $\mathbb{E}\{\mathbf{w}_k^T\mathbf{g}_{\mathrm{SR},k}\}$ has already been obtained in the previous part.

*3) Compute* $\mathrm{AN}_{\mathrm{SR},k} = \sigma_{\mathrm{R}}^2 \mathbb{E}\left\{\|\mathbf{w}_k^T\|^2\right\}$*:* We note that the computation of $\mathbb{E}\left\{\|\mathbf{w}_k^T\|^2\right\}$ has already been obtained in the previous part.

The following calculation is based on [17, Lemmas 8-10].

*4) Compute* $\mathrm{LS}_{\mathrm{SR},k} = p_{\mathrm{S}} \mathbb{V}\mathrm{ar}\left\{\mathbf{w}_k^T \mathbf{g}_{\mathrm{SR},k}\right\}$*:* We have $\mathbb{V}\mathrm{ar}\{\mathbf{w}_k^T \mathbf{g}_{\mathrm{SR},k}\} = \mathbb{E}\{|\hat{\mathbf{g}}_{\mathrm{SR},k}^H \mathbf{g}_{\mathrm{SR},k}|^2\} - |\mathbb{E}\{\hat{\mathbf{g}}_{\mathrm{SR},k}^H \mathbf{g}_{\mathrm{SR},k}\}|^2$. We derive $\mathbb{E}\{|\hat{\mathbf{g}}_{\mathrm{SR},k}^H \mathbf{g}_{\mathrm{SR},k}|^2\}$ as follows:

$$\begin{aligned}
&\mathbb{E}\{\left|\hat{\mathbf{g}}_{\mathrm{SR},k}^H \mathbf{g}_{\mathrm{SR},k}\right|^2\} = \mathbb{E}\{\left|\hat{\mathbf{q}}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},k}\right|^2\} + \mathbb{E}\{\left|\hat{\mathbf{q}}_{\mathrm{SR},k}^H \mathbf{d}_{\mathrm{SR},k}\right|^2\} \\
&+ \mathbb{E}\{\left|\mathbf{d}_{\mathrm{SR},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},k}\right|^2\} + \mathbb{E}\{\left|\mathbf{d}_{\mathrm{SR},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{d}_{\mathrm{SR},k}\right|^2\} \\
&+ \mathbb{E}\{\left|\mathbf{N}_{\mathrm{rp},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},k}\right|^2\} + \mathbb{E}\{\left|\mathbf{N}_{\mathrm{rp},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{d}_{\mathrm{SR},k}\right|^2\} \\
&\quad + 2\Re\{\mathbb{E}\{\hat{\mathbf{q}}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},k} (\mathbf{d}_{\mathrm{SR},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{d}_{\mathrm{SR},k})^H\}\}.
\end{aligned}$$

where $\mathbf{q}_{\mathrm{SR},k} \triangleq \sum_{\omega=1}^4 \mathbf{q}_{\mathrm{SR},k}^{(\omega)}, \hat{\mathbf{q}}_{\mathrm{SR},k} \triangleq \sum_{\omega=1}^4 \hat{\mathbf{q}}_{\mathrm{SR},k}^{(\omega)}$.

*5) Compute* $\mathrm{MP}_{\mathrm{SR},k} = p_{\mathrm{S}} \sum_{i \neq k} \mathbb{E}\left\{|\mathbf{w}_k^T \mathbf{g}_{\mathrm{SR},i}|^2\right\}$*:*

$$\begin{aligned}
&\mathbb{E}\{\left|\hat{\mathbf{g}}_{\mathrm{SR},k}^H \mathbf{g}_{\mathrm{SR},i}\right|^2\} = \mathbb{E}\{\left|\hat{\mathbf{q}}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},i}\right|^2\} + \mathbb{E}\{\left|\hat{\mathbf{q}}_{\mathrm{SR},k}^H \mathbf{d}_{\mathrm{SR},i}\right|^2\} \\
&+ \mathbb{E}\{\left|\mathbf{d}_{\mathrm{SR},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},i}\right|^2\} + \mathbb{E}\{\left|\mathbf{d}_{\mathrm{SR},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{d}_{\mathrm{SR},i}\right|^2\} \\
&+ \mathbb{E}\{\left|\mathbf{N}_{\mathrm{rp},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},i}\right|^2\} + \mathbb{E}\{\left|\mathbf{N}_{\mathrm{rp},k}^H \mathbf{A}_{\mathrm{SR},k}^H \mathbf{d}_{\mathrm{SR},i}\right|^2\}.
\end{aligned}$$

The first term in the above expression is expanded as follows:

$$\begin{aligned}
&\mathbb{E}\{\left|\hat{\mathbf{q}}_{\mathrm{SR},k}^H \mathbf{q}_{\mathrm{SR},i}\right|^2\} = \sum_{\omega=1}^4 \sum_{\psi=1}^4 \mathbb{E}\{|\left(\hat{\mathbf{q}}_{\mathrm{SR},k}^{(\omega)}\right)^H \mathbf{q}_{\mathrm{SR},i}^{(\psi)}|^2\} \\
&+ \sum_{\substack{\omega_1,\psi_1,\,\omega_2,\psi_2 \in \{1,..,4\} \\ (\omega_1,\psi_1) \neq (\omega_2,\psi_2)}} \mathbb{E}\{((\hat{\mathbf{q}}_{\mathrm{SR},k}^{(\omega_1)})^H \mathbf{q}_{\mathrm{SR},i}^{(\psi_1)})((\hat{\mathbf{q}}_{\mathrm{SR},k}^{(\omega_2)})^H \mathbf{q}_{\mathrm{SR},i}^{(\psi_2)})^H\}.
\end{aligned}$$

## References


[1] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas *et al.*, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, Secondquarter 2023.

[2] J. Zhang, Q. Li, K. Fang, Z. Wang, D. Zhang, H. Li, and B. Clerckx, "Spherical stacked intelligent metasurfaces: A paradigm for full-space wave-domain processing," *IEEE Communications Magazine*, 2026.

[3] A. Wei, Z. Wang, K. Fang, J. Zhang, Y. Dong, L. Yang, and Q. Li, "Near-field beam training design for stacked intelligent metasurface assisted communication systems," *IEEE Transactions on Vehicular Technology*, 2026.

[4] Y. Li, R. Zhao, Y. Wang, G. Pan, and C. Li, "Artificial noise aided precoding with imperfect csi in full-duplex relaying secure communications," *IEEE Access*, vol. 6, pp. 44 107–44 119, Aug. 2018.

[5] Y. Li, R. Zhao, L. Fan, and A. Liu, "Antenna mode switching for full-duplex destination-based jamming secure transmission," *IEEE Access*, vol. 6, pp. 9442–9453, Mar. 2018.

[6] Y. Li, R. Zhao, Y. Deng, F. Shu, Z. Nie, and A. H. Aghvami, "Harvest-and-opportunistically-relay: Analyses on transmission outage and covertness," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7779–7795, Dec. 2020.

[7] Z. Abdullah, G. Chen, S. Lambotharan, and J. A. Chambers, "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.

[8] Q. Bie, Y. Liu, Y. Wang, X. Zhao, and X. Y. Zhang, "Deployment optimization of reconfigurable intelligent surface for relay systems," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 221–233, Mar. 2022.

[9] Z. Kang, C. You, and R. Zhang, "IRS-aided wireless relaying: Deployment strategy and capacity scaling," *IEEE Wireless Commun. Lett.*, vol. 11, no. 2, pp. 215–219, Feb. 2021.

[10] I. Yildirim, F. Kilinc, E. Basar, and G. C. Alexandropoulos, "Hybrid RIS-empowered reflection and decode-and-forward relaying for coverage extension," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1692–1696, May 2021.

[11] Q. Sun, P. Qian, W. Duan, J. Zhang, J. Wang, and K.-K. Wong, "Ergodic rate analysis and IRS configuration for multi-IRS dual-hop DF relaying systems," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3224–3228, Oct. 2021.

[12] X. Wang, F. Shu, W. Shi, X. Liang, R. Dong, J. Li, and J. Wang, "Beamforming design for IRS-aided decode-and-forward relay wireless network," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 198–207, Mar. 2022.

[13] T.-Y. Kan, R. Y. Chang, F.-T. Chien, B.-J. Chen, and H. V. Poor, "Hybrid relay and reconfigurable intelligent surface assisted multiuser MISO systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 7653–7668, Jun. 2023.

[14] M. Obeed and A. Chaaban, "Joint beamforming design for multiuser MISO downlink aided by a reconfigurable intelligent surface and a relay," *IEEE Trans. Wireless Commun.*, Oct. 2022.

[15] Z. Abdullah, G. Chen, S. Lambotharan, and J. A. Chambers, "Optimization of intelligent reflecting surface assisted full-duplex relay networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 363–367, Feb. 2020.

[16] Z. Chen, M.-M. Zhao, K. Xu, Y. Cai, and M.-J. Zhao, "Intelligent reflecting surface assisted full-duplex relay systems: Deployment design and beamforming optimization," *IEEE Trans. Commun.*, vol. 72, no. 7, pp. 4493–4508, Jul. 2024.

[17] K. Zhi, C. Pan, H. Ren, K. Wang, M. Elkashlan, M. Di Renzo, R. Schober, H. V. Poor, J. Wang, and L. Hanzo, "Two-timescale design for reconfigurable intelligent surface-aided massive MIMO systems with imperfect CSI," *IEEE Trans. Inf. Theory*, vol. 69, no. 5, pp. 3001–3033, May 2022.

[18] E. Björnson and Ö. T. Demir, *Introduction to multiple antenna communications and reconfigurable surfaces*. Emerald Group Publishing, 2024.

[19] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.

[20] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.

[21] M. Momen-Tayefeh, M. Momen-Tayefeh, and M. Sabbaghian, "Multi-block attention for efficient channel estimation in IRS-assisted mmwave MIMO," *IEEE Trans. Commun.*, Dec. 2025.

[22] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Trans. Microw. Theory Tech.*, vol. 67, no. 7, pp. 3025–3041, Jul. 2019.

[23] E. Björnson, J. Hoydis, L. Sanguinetti *et al.*, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[24] M.-M. Zhao, A. Liu, Y. Wan, and R. Zhang, "Two-timescale beamforming optimization for intelligent reflecting surface aided multiuser communication with QoS constraints," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 6179–6194, Sep. 2021.

[25] M. Soleymani, E. Jorswieck, R. Schober, and L. Hanzo, "A framework for fractional matrix programming problems with applications in FBL MU-MIMO," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 1240–1257, 2026.

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

[27] K. Huang and N. D. Sidiropoulos, "Consensus-ADMM for general quadratically constrained quadratic programming," *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5297–5310, Oct. 2016.

[28] M. Shao, Q. Li, W.-K. Ma, and A. M.-C. So, "A framework for one-bit and constant-envelope precoding over multiuser massive MISO channels," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5309–5324, Oct. 2019.

[29] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0," https://cvxr.com/cvx, Aug. 2012.

[30] X. Mu, Y. Liu, L. Guo, J. Lin, and L. Hanzo, "NOMA-aided joint radar and multicast-unicast communication systems," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1978–1992, Jun. 2022.

[31] E. Björnson, Ö. 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. 2019.