Hybrid Relay and Reconfigurable Intelligent Surface Assisted Multiuser MISO Systems

Te-Yi Kan, Ronald Y. Chang, *Senior Member, IEEE*, Feng-Tsun Chien, *Senior Member, IEEE*, Bing-Jia Chen, and H. Vincent Poor, *Life Fellow, IEEE*

***Abstract*—Reconfigurable intelligent surfaces (RISs) are viewed as key enablers for next-generation wireless communications. This paper investigates a multiuser downlink multiple-input single-output (MISO) system in which a multiantenna base station (BS) transmits information to multiple single-antenna users with the aid of both a half-duplex decode-and-forward (DF) relay and a full-duplex RIS. Active beamforming at the BS and the DF relay, as well as passive beamforming at the RIS, are jointly designed for system sum-rate maximization. The design problem is challenging to solve due to coupled beamforming variables. An alternating optimization (AO) based algorithm is proposed to tackle this complex co-design problem. Numerical results demonstrate the superior performance of the proposed hybrid relay–RIS system with a judicious joint beamforming design. Convergence and complexity analysis shows that the convergence rate of the proposed algorithm is dominated by the numbers of users and RIS elements, and the proposed scheme can converge in a few iterations even in the configuration of large numbers of users and RIS elements. Interesting tradeoffs posed in the joint design are discussed. An extension of the proposed design method to a related energy efficiency (EE) optimization problem is also outlined and implemented.**

***Index Terms*—Reconfigurable intelligent surfaces, decode-and-forward, relay, multiuser MISO, beamforming, alternating optimization.**

I. Introduction

RECONFIGURABLE intelligent surfaces (RISs), also known as intelligent reflecting surfaces (IRSs), passive intelligent mirrors (PIMs), and large intelligent surfaces (LISs), have recently been proposed as a key enabler for next-generation green wireless communications [2], [3], [4]. RISs employ a large number of low-cost passive elements, instead of active transmitters, to "reflect" signals to empower smart and reconfigurable radio environments. It is envisioned that deploying RISs is less expensive as compared to installing active transmitters such as base stations (BSs) and relays, and operating RISs is more energy-efficient and eases signal processing and interference management requirements due to its low-cost and passive nature. Thus, RISs are suitable for dense deployment and are promising to help meet the demands of beyond fifth-generation (B5G) and sixth-generation (6G) communications, such as high data rates, energy efficiency, and ubiquitous connectivity. To facilitate a focused discussion, in the following, we introduce selected related topics most relevant to this study, and refer the reader to the comprehensive literature surveys [5], [6], [7], [8], [9], [10] for various other topics concerning RISs.

A. Related Work

There is a wealth of research on RIS focusing on joint transmit beamforming and RIS phase shift optimization for sum-rate maximization (e.g., [11], [12], [13], [14], [15], [16], [17], [18]) and power minimization (e.g., [19], [20], [21], [22]), as introduced in this and the next paragraph, respectively. RIS-enhanced multi-input single-output (MISO) systems were studied for single-user [11] and multiuser [12] scenarios, where transmit beamforming at the BS and the phase shifts at the RIS were optimized by optimizing the received signal power or sum-rate of users. Machine learning-based techniques were also introduced into RIS-aided systems to design transmit beamforming and RIS phase shifts [13], [14]. In [13], the deep neural network (DNN) was utilized to maximize the user's received signal strength in an RIS-aided single-user system. In [14], deep reinforcement learning (DRL) was adopted to maximize the sum-rate of users in multiuser MISO systems. The communication systems considered in the above work are all assumed to be narrow-band systems where the channels are frequency-flat. RIS-enhanced broadband orthogonal frequency division multiplexing (OFDM) systems were also studied [15], [16]. In [15], BS power allocation and RIS reflection coefficients were optimized to improve cell-edge user's data rates in single-user single-input single-output (SISO) OFDM systems. In [16], BS beamforming and RIS reflection were jointly designed for the objective of

Manuscript received 21 July 2022; revised 19 October 2022; accepted 4 February 2023. Date of publication 9 February 2023; date of current version 20 June 2023. This work was supported in part by the National Science and Technology Council (formerly, Ministry of Science and Technology), Taiwan, under Grants MOST 109-2221-E-001-013-MY3, MOST 109-2918-I-001-003, and MOST 111-2221-E-A49-150, and in part by the U.S. National Science Foundation under Grants CCF-1908308 and CNS-2128448. An earlier version of this paper was presented in part at the 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, December 7–11, 2021 [1]. The review of this article was coordinated by Dr. T. Q. Duong. *(Corresponding author: Ronald Y. Chang.)*

Te-Yi Kan and Bing-Jia Chen are with the Research Center for Information Technology Innovation, Academia Sinica, Taipei 11529, Taiwan (e-mail: dexter.ty.kan@gmail.com; b07901088@ntu.edu.tw).

Ronald Y. Chang is with the Research Center for Information Technology Innovation, Academia Sinica, Taipei 11529, Taiwan, and also with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: rchang@citi.sinica.edu.tw).

Feng-Tsun Chien is with the Institute of Electronics, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: ftchien@nycu.edu.tw).

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

Digital Object Identifier 10.1109/TVT.2023.3243489

sum-rate maximization, or, after transformation, mean square error (MSE) minimization, in multiuser MISO systems. RIS-empowered nonorthogonal multiple access (NOMA) systems for signal-to-interference-plus-noise ratio (SINR) improvement were also studied [17], [18]. In [17], a novel combined-channel-strength (CSS) based user ordering scheme was proposed for NOMA, and an algorithm based on block coordinate descent and semidefinite relaxation (SDR) was developed to maximize the minimum SINR by configuring the RIS. In [18], alternating optimization (AO) based sum-rate maximization algorithms were proposed for cases of ideal RIS and non-ideal RIS with fixed reflection amplitude, where the successive convex approximating technique was used to develop the algorithm for the ideal RIS case, and the sequential rank-one constraint relaxation approach was used to develop the algorithm for the non-ideal RIS case.

The potential for RISs to reduce power consumption in communication systems has been explored through joint transmit beamforming and RIS phase shift optimization [19], [20], [21], [22]. In [19], [20], transmit beamforming and RIS beamforming for power minimization subject to user SINR constraints in MISO systems were investigated, where an SDR-based algorithm [19] and a two-stage algorithm [20] were proposed. In [21], [22], transmit power minimization in RIS-empowered NOMA systems was studied. In [21], an access point (AP) transmit power minimization problem with user target rate constraints was examined, and the performance of RIS-aided NOMA and RIS-aided orthogonal multiple access (OMA) was analyzed. It was shown that RIS-aided NOMA is insensitive to the disparity of target rates and outperforms RIS-aided OMA in most cases. In [22], an AO and difference-of-convex algorithm was proposed for the transmit power minimization problem in RIS-empowered NOMA networks. It was shown [21], [22] that incorporating RISs into NOMA brings merits in decoding and differentiating multiple users due to RIS's ability to reconfigure wireless channels.

It is useful to analyze RISs against related, well-established technologies, particularly massive multiple-input multiple-output (MIMO) and relaying, to gain a comprehensive understanding of RISs and maximize their potential. While both RISs and massive MIMO are capable of focusing signal transmissions in specific directions, massive MIMO employs active antennas and RISs employ passive reflecting elements. In [23], a comparison in power scaling laws and near-field behaviors revealed that the signal-to-noise ratio (SNR) increases with the number of antennas for massive MIMO, and with the square of the number of reflecting elements for RIS. However, only in the large number of reflecting elements regime will RIS outperform massive MIMO, due to the passive nature of RIS. In [24], the degrees of freedom (DoFs) of massive MIMO and RIS were studied. It was observed that RIS can achieve a DoF of greater than or equal to one even in the line-of-sight (LoS) condition without multipaths. Thus, RIS may offer additional advantages over massive MIMO in high-frequency or mmWave communication regimes where multipath conditions may not exist.

Comparative studies between RIS and classical relaying technologies such as decode-and-forward (DF) and amplify-and-forward (AF) have been conducted [25], [26], [27], [28]. In [25], comparisons between RIS and DF relaying were conducted analytically and numerically for a SISO system, showing that RIS with sufficient reflecting elements can achieve higher data rates and better energy efficiency in high target data rate regimes than DF relaying. In [26], comparisons between RIS and AF relaying were conducted numerically in an RIS-aided multiuser MISO system. It was shown that replacing the RIS with an AF relay yields higher sum rates but lower energy efficiency due to active power amplification of the AF relay. In [27], RIS and full-duplex/half-duplex AF relaying were compared. It was shown that RIS achieves comparable or even better performance as compared to full-duplex and half-duplex AF relaying when the number of RIS reflecting elements is sufficiently large. In [28], a comprehensive discussion on the subject was presented from various perspectives such as hardware complexity, power consumption, and spectral efficiency. It was similarly concluded that RIS with sufficient reflecting elements has the potential to outperform relay-aided transmission since RIS operates in a full-duplex manner, yet without the loop interference as in full-duplex relays.

The mature relaying technology affords deployment of the new RISs alongside the classical relays in future communication systems. Existing work predominantly focuses on single-user SISO systems assisted by one or more RISs and single-antenna relays [29], [30], [31], [32], [33], [34], [35]. In [29], the ergodic capacity was examined for a single-user SISO system with an RIS and a single-antenna DF relay. In [30], DRL-based joint relay selection and RIS beamforming for throughput maximization was studied for a single-user SISO system with an RIS and multiple single-antenna relays. In [31], joint optimization of power allocation and RIS beamforming to maximize the end-to-end SNR was studied for a single-user SISO system with coexisting RIS and relay. In [32], time allocation and RIS beamforming for maximum achievable rates were considered in a single-user SISO system with an RIS and an RIS controller acting as a single-antenna DF relay. In [33], single-RIS and multi-RIS deployments with the same total number of RIS elements were compared in an RIS and relay assisted single-user SISO system. In [34], RIS beamforming for maximum achievable rates in a single-user SISO system assisted by two DF relays and two RISs was studied. In [35], RIS deployment optimization was performed in a single-user SISO system aided by a DF relay and an RIS for maximum achievable rates. A recent work [36] considered a multiuser MISO system assisted by both an RIS and a multiantenna DF relay, and studied joint beamforming for power minimization with relay and user achievable rate constraints that represent reliable decoding at the relay and users.

B. Motivations

Deploying both RISs and relays is practical and beneficial in future RIS-enabled communications. The resulting system,

termed a *hybrid relay–RIS system*[1] in this paper, offers new design possibilities and performance potential as compared to relay-only and RIS-only systems. Specifically, a hybrid relay–RIS system has the following advantages. First, due to the passive nature of RISs and the high path loss introduced by double reflection of RISs [38], the service range of RIS-only systems may be limited. By employing relays in the vicinity of RISs, the service range could increase in hybrid relay–RIS systems. Second, RISs typically require a sufficiently large number of reflecting elements to achieve comparable performance as relays. A large number of RIS elements makes accurate and low-pilot-overhead channel estimation challenging, since the number of channel coefficients for the cascaded channel (BS–RIS–users) is high and leads to high training overhead. Hybrid relay–RIS systems facilitate robust channel estimation with low training overhead by avoiding massive numbers of RIS elements and exploiting active relays with more powerful signal processing capabilities. Third, due to the passive nature of RISs, hybrid relay–RIS systems could provide an energy-efficient solution without consuming too much transmit power at both the BS and relay, as compared to relay-only systems. Fourth, hybrid relay–RIS systems offer technology interoperability by integrating the new RIS technology with the well-established relay technology.

This work aims to investigate joint beamforming in hybrid relay–RIS systems. Unlike [29], [30], [31], [32], [33], [34], [35], we consider a multiuser system with a multiantenna BS and relay, plus some additional considerations. As compared to [36], we consider a similar hybrid scenario, but have a different problem, focus, and design considerations. Among others, the decoding criterion for the DF relay is expressed in terms of the decoding SINR with a practical decoding scheme in this study, rather than the theoretical achievable rates in [36], resulting in different derivations and solution methods. Moreover, RIS beamforming is performed only at the beginning of the two transmission phases in this study for reduced delay, energy consumption, and information exchange associated with configuring the RIS, rather than in each transmission phase as in [36]. With this consideration, we are motivated to examine the potential tradeoffs involved in RIS design. For example, an RIS designed to help user reception in the first phase may not also benefit relay reception in the first phase or relay transmission towards the users in the second phase, and therefore the combined effect of the two phases on the users is unclear. A joint BS, relay, and RIS design that meets the established requirements is not straightforward and requires further research.

C. Contributions

In this paper, we consider the coexistence of an RIS and a half-duplex DF relay in a downlink multiuser MISO system. The multiantenna BS transmits information to the users in two phases, with the aid of a multiantenna half-duplex DF relay and a (full-duplex) RIS. The main contributions of this paper are summarized as follows:

- The considered hybrid relay–RIS system is well-motivated from both technology evolution and practical perspectives. Hybrid relay–RIS systems introduce new design considerations not found in traditional multi-relay systems, including the fact that relays in multi-relay systems do not interact, whereas the RIS and the relay interact in hybrid systems. The complex co-design problem of BS beamforming, relay beamforming, and RIS beamforming for sum-rate maximization in this specific system is investigated. The problem is challenging to solve as the beamforming variables are highly coupled. An AO-based algorithm is employed to decouple beamforming vectors and solve the resulting subproblems iteratively.
- The proposed hybrid system and co-design scheme demonstrate superior sum-rate (and energy efficiency) performance, as validated by numerical results. Attributes of the proposed scheme such as complexity and convergence are examined. Interesting tradeoffs involved in designing an RIS so that it simultaneously serves multiple, possibly conflicting, purposes across the two transmission phases (i.e., assisting BS transmission towards the users and the relay in the first phase, and assisting relay transmission towards the users in the second phase) in the hybrid relay–RIS system are discussed.

The remainder of the paper is organized as follows. Section II introduces the downlink multiuser MISO system and formulates the sum-rate optimization problem. Section III presents the proposed joint beamforming algorithm for the hybrid system. Section IV presents the complexity and convergence analysis. Section V presents the simulation results and discussion. Finally, Section VI concludes the paper.

Notations: Lowercase letters (e.g., a), boldface lowercase letters (e.g., $\mathbf{a}$), and boldface uppercase letters (e.g., $\mathbf{A}$) denote scalars, vectors, and matrices, respectively. $\mathbf{A}^{\mathrm{H}}$, $\mathbf{A}^{\mathrm{T}}$, $\mathrm{tr}(\mathbf{A})$, and $\mathrm{rank}(\mathbf{A})$ denote the Hermitian, transpose, trace, and rank of matrix $\mathbf{A}$, respectively. $\mathbf{A} \succeq \mathbf{0}$ means that $\mathbf{A}$ is a positive semidefinite (PSD) matrix. $\mathbf{I}_N$ is an $N \times N$ identity matrix. $\mathrm{diag}(\mathbf{a})$ is a diagonal matrix with $\mathbf{a}$ on the main diagonal. $\mathcal{CN}(\mu, \sigma^2)$ denotes the complex Gaussian distribution with mean μ and variance σ^2. $\mathbb{E}[a]$ denotes the mathematical expectation of a. $\mathbb{C}$ and $\mathbb{R}$ denote the sets of complex and real numbers, respectively, and $\mathbb{C}^{m\times n}$ and $\mathbb{R}^{m\times n}$ denote the sets of $m \times n$ complex and real matrices, respectively. $\mathbf{A}[i, j]$ denotes the (i, j)-entry of matrix $\mathbf{A}$, and $\mathbf{a}[i]$ denotes the ith entry of vector $\mathbf{a}$.

II. System Model and Problem Formulation

A. System Model

As shown in Fig. 1, we consider a downlink multiuser MISO communication system that comprises a BS with M antennas, K single-antenna end users, an RIS with N elements, and a classical half-duplex DF relay with L antennas. We consider $K \leq \min\{M, L\}$ to ensure sufficient degrees of freedom. The baseband equivalent channels between two communication nodes among the BS, RIS, relay, and user k are denoted by self-explanatory notations $\mathbf{H}_{\mathrm{BS,R}} \in \mathbb{C}^{L\times M}$, $\mathbf{H}_{\mathrm{BS,RIS}} \in \mathbb{C}^{N\times M}$, $\mathbf{h}_{\mathrm{BS},k} \in \mathbb{C}^{1\times M}$, $\mathbf{H}_{\mathrm{R,RIS}} \in \mathbb{C}^{N\times L}$, $\mathbf{h}_{\mathrm{R},k} \in$

[1]"Hybrid" here refers specifically to the coexistence of relays and RISs, and should not be confused with the hybrid reflecting and sensing RIS (HRIS) in [37].

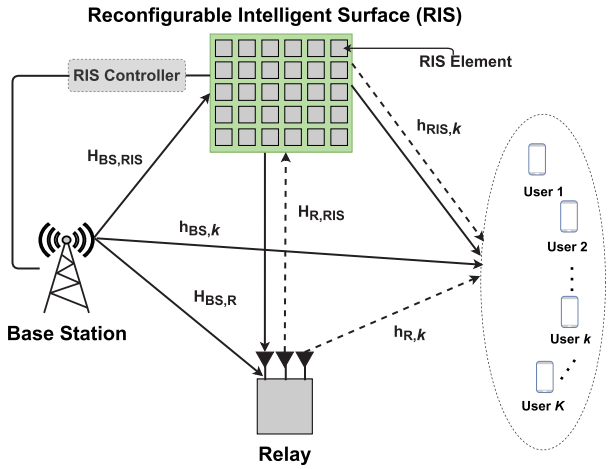


Fig. 1. A hybrid relay–RIS multiuser MISO system. Solid/dashed lines indicate transmissions in the first/second phases.

$\mathbb{C}^{1\times L}$, and $\mathbf{h}_{\mathrm{RIS},k} \in \mathbb{C}^{1\times N}$. Similar to [30], we set the RIS-related channels $\mathbf{H}_{\mathrm{BS,RIS}}, \mathbf{H}_{\mathrm{R,RIS}}, \mathbf{h}_{\mathrm{RIS},k}$ to be quasi-static Rician flat-fading, and the other channels $\mathbf{H}_{\mathrm{BS,R}}, \mathbf{h}_{\mathrm{BS},k}, \mathbf{h}_{\mathrm{R},k}$ quasi-static Rayleigh flat-fading. The $\mathbf{H}_{\mathrm{BS,RIS}}$ is modeled as [29], [30], [39], [40]

$$\mathbf{H}_{\mathrm{BS,RIS}} = \sqrt{\frac{K_1}{K_1+1}}\mathbf{H}_{\mathrm{LoS}} + \sqrt{\frac{1}{K_1+1}}\mathbf{H}_{\mathrm{NLoS}}, \quad (1)$$

where K_1 is the Rician factor, and $\mathbf{H}_{\mathrm{LoS}}$ and $\mathbf{H}_{\mathrm{NLoS}}$ are line-of-sight (LoS) and non-line-of-sight (NLoS) components given by

$$\mathbf{H}_{\mathrm{LoS}} = \kappa_{\mathrm{LoS}}(d_{\mathrm{BS,RIS}}/d_0)^{-\varrho_{\mathrm{LoS}}}\mathbf{a}(\omega)\mathbf{a}^{\mathrm{H}}(\vartheta), \quad (2)$$

$$\mathbf{H}_{\mathrm{NLoS}} = \kappa_{\mathrm{NLoS}}(d_{\mathrm{BS,RIS}}/d_0)^{-\varrho_{\mathrm{NLoS}}}\overline{\mathbf{H}}_{\mathrm{NLoS}}. \quad (3)$$

In (2), κ_{LoS} and ϱ_{LoS} are the large-scale fading constant and path-loss exponent for the LoS channel, respectively; $d_{\mathrm{BS,RIS}}$ is the distance between the BS and RIS; d_0 is the reference distance; $\mathbf{a}(\omega) \triangleq [1, e^{j2\pi\frac{d_{\mathrm{RIS}}}{\lambda}\sin(\omega)}, \ldots, e^{j2\pi\frac{d_{\mathrm{RIS}}}{\lambda}(N-1)\sin(\omega)}]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$ is the array response vector, where d_{RIS} is the RIS inter-element spacing, λ is the wavelength, and ω is the angle of arrival (AoA); and $\mathbf{a}(\vartheta) \in \mathbb{C}^{M\times 1}$ with ϑ being the angle of departure (AoD) is defined in a similar manner. We set $d_{\mathrm{RIS}}/\lambda = 1/2$ [40]. The values of AoA and AoD are randomly generated between 0 and 2π. In (3), κ_{NLoS} and ϱ_{NLoS} are the large-scale fading constant and path-loss exponent for the NLoS channel, respectively, and $\overline{\mathbf{H}}_{\mathrm{NLoS}}$ is the small-scale fading modeled by complex Gaussian with zero mean and unit variance. The $\mathbf{H}_{\mathrm{R,RIS}}$ and $\mathbf{h}_{\mathrm{RIS},k}$ are modeled similarly as in (1). For non-RIS-related channels, the channels are modeled similar to (3) with large-scale fading modeled by $\kappa(d/d_0)^{-\varrho}$ and small-scale fading modeled by complex Gaussian with zero mean and unit variance, where d denotes the distance between the two end nodes. The channel state information (CSI) of all channels is assumed to be perfectly known for the exposition,[2] but the effects of imperfect CSI will be examined numerically. A complete transmission of information consists of two phases within the channel coherence interval. Similar to [30], [33], [34], we consider that the RIS elements are configured by the RIS controller only once in the two transmission phases for simplicity and reduced delay, energy consumption, and information exchange associated with RIS configuration. Next, we elaborate on the two-phase transmission protocol.

1) First Phase: During the first phase, the BS transmits its signal to all users and the relay. Meanwhile, the RIS reflects the incident signal from the BS towards the relay and all users. The received signal at user k in the first phase is given by

$$y_k^{\mathrm{I}} = \mathbf{h}'^{\mathrm{H}}_{\mathrm{BS},k}\mathbf{x} + w_k^{\mathrm{I}}, \quad (4)$$

where $\mathbf{h}'^{\mathrm{H}}_{\mathrm{BS},k} \triangleq \mathbf{h}^{\mathrm{H}}_{\mathrm{RIS},k}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{BS,RIS}} + \mathbf{h}^{\mathrm{H}}_{\mathrm{BS},k}$ is the effective channel from the BS to the end user k, $\boldsymbol{\Theta} \triangleq \mathrm{diag}(\boldsymbol{\theta})$ with $\boldsymbol{\theta} = [\theta_1 \cdots \theta_N]$ accounts for the passive RIS beamforming, and $w_k^{\mathrm{I}} \sim \mathcal{CN}(0, \sigma_k^2)$ denotes the additive white Gaussian noise (AWGN) at user k during the first phase. We consider a practical RIS setting where the nth RIS element θ_n can only take 2^B discrete values, i.e., from the following set [44]

$$\mathcal{F} = \left\{\theta_n \middle| \theta_n = e^{j\psi_n}, \psi_n \in \left\{0, \frac{2\pi}{2^B}, \ldots, \frac{2\pi(2^B-1)}{2^B}\right\}\right\}, \quad (5)$$

where B is the number of bits.

The transmit signal at the BS is $\mathbf{x} = \sum_{k=1}^{K}\mathbf{g}_k s_k = \mathbf{G}\cdot\mathbf{s}$, where $\mathbf{g}_k \in \mathbb{C}^{M\times 1}$ is the beamforming vector for user k, s_k is the data intended for user k with zero mean and unit variance, and $\mathbf{G} \triangleq [\mathbf{g}_1, \ldots, \mathbf{g}_K]$ and $\mathbf{s} \triangleq [s_1, \ldots, s_K]^{\mathrm{T}}$. The data streams from different users are assumed independent. Then, the transmit power at the BS is

$$P_{\mathrm{BS}}^{\mathrm{total}} = \mathbb{E}\left[\|\mathbf{x}\|^2\right] = \mathrm{tr}\left(\mathbf{G}\mathbf{G}^{\mathrm{H}}\right) \le P_{\mathrm{BS}}^{\max}, \quad (6)$$

where $P_{\mathrm{BS}}^{\max}$ is the maximum available power at the BS. It follows that the SINR for user k in the first phase is

$$\gamma_k^{\mathrm{I}} = \frac{\left|\mathbf{h}'^{\mathrm{H}}_{\mathrm{BS},k}\mathbf{g}_k\right|^2}{\sum_{j=1,j\neq k}^{K}\left|\mathbf{h}'^{\mathrm{H}}_{\mathrm{BS},k}\mathbf{g}_j\right|^2 + \sigma_k^2}. \quad (7)$$

On the other hand, the received signal at the relay is

$$\mathbf{y}_{\mathrm{R}} = \mathbf{H}'^{\mathrm{H}}_{\mathrm{BS,R}}\mathbf{x} + \mathbf{w}_{\mathrm{R}}, \quad (8)$$

where $\mathbf{H}'^{\mathrm{H}}_{\mathrm{BS,R}} \triangleq \mathbf{H}^{\mathrm{H}}_{\mathrm{R,RIS}}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{BS,RIS}} + \mathbf{H}^{\mathrm{H}}_{\mathrm{BS,R}}$ is the effective channel from the BS to the relay, and $\mathbf{w}_{\mathrm{R}} \sim \mathcal{CN}(0, \sigma_{\mathrm{R}}^2\mathbf{I}_L)$ is the AWGN at the relay. Matched filtering is employed at the relay to decode all users' signals from the BS. The resulting SINR for user k at the relay is

$$\gamma_{\mathrm{R},k} = \frac{\|\boldsymbol{\alpha}_k\|^4}{\sum_{j=1,j\neq k}^{K}\boldsymbol{\alpha}_k^{\mathrm{H}}\boldsymbol{\alpha}_j\boldsymbol{\alpha}_j^{\mathrm{H}}\boldsymbol{\alpha}_k + \sigma_{\mathrm{R}}^2\|\boldsymbol{\alpha}_k\|^2}, \quad (9)$$

where $\boldsymbol{\alpha}_k \triangleq \mathbf{H}'^{\mathrm{H}}_{\mathrm{BS,R}}\mathbf{g}_k$ is the filter weight for user k. The derivation of (9) is given in the Appendix.

[2]In practice, channel estimation should be performed to obtain the CSI. The reader is referred to [41], [42], [43] and the references therein for state-of-the-art channel estimation schemes in RIS-assisted communications.

2) Second Phase: During the second phase, the relay transmits the signal to all users and the RIS, where the latter also reflects the incident signal towards all users. Here, we assume that the DF relay can perfectly decode the signal for user k if the SINR $\gamma_{\mathrm{R},k}$ exceeds a predefined threshold $\gamma_{\mathrm{R}}^{\mathrm{threshold}}$. The transmit signal at the relay is given by

$$\mathbf{x}_{\mathrm{R}} = \sum_{k=1}^{K} \mathbf{f}_k s_k = \mathbf{F} \cdot \mathbf{s}, \tag{10}$$

where $\mathbf{f}_k \in \mathbb{C}^{L\times 1}$ denotes the beamforming vector at the relay for user k and $\mathbf{F} \triangleq [\mathbf{f}_1, \ldots, \mathbf{f}_K]$. The relay transmit power is

$$P_{\mathrm{R}}^{\mathrm{total}} = \mathbb{E}\left[\|\mathbf{x}_{\mathrm{R}}\|^2\right] = \mathrm{tr}\left(\mathbf{F}\mathbf{F}^{\mathrm{H}}\right) \le P_{\mathrm{R}}^{\max}, \tag{11}$$

where $P_{\mathrm{R}}^{\max}$ denotes the maximum available power at the relay. The received signal at user k in the second phase can be represented by

$$y_k^{\mathrm{II}} = \mathbf{h}'_{\mathrm{R},k}\mathbf{x}_{\mathrm{R}} + w_k^{\mathrm{II}}, \tag{12}$$

where $\mathbf{h}'_{\mathrm{R},k} \triangleq \mathbf{h}_{\mathrm{RIS},k}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{R,RIS}} + \mathbf{h}_{\mathrm{R},k}$ is the effective channel from the relay to user k, and $w_k^{\mathrm{II}} \sim \mathcal{CN}(0, \sigma_k^2)$ is the corresponding AWGN. It follows that the SINR for user k in the second phase is given by

$$\gamma_k^{\mathrm{II}} = \frac{\left|\mathbf{h}'_{\mathrm{R},k}\mathbf{f}_k\right|^2}{\sum_{j=1,j\neq k}^{K}\left|\mathbf{h}'_{\mathrm{R},k}\mathbf{f}_j\right|^2 + \sigma_k^2}. \tag{13}$$

Finally, the resulting total SINR after combining signals from the two phases using the maximal ratio combining (MRC) is

$$\gamma_k = \gamma_k^{\mathrm{I}} + \gamma_k^{\mathrm{II}}. \tag{14}$$

B. Problem Formulation

In this paper, we aim at maximizing the sum-rate of all the users by jointly optimizing the active beamforming at the BS, active beamforming at the relay, and passive beamforming at the RIS, while guaranteeing the BS power constraint, relay power constraint, and RIS configuration constraint. The problem is formulated as

$$\max_{\{\mathbf{g}_k\},\{\mathbf{f}_k\},\boldsymbol{\Theta}} \quad \sum_{k=1}^{K}\log_2\left(1+\gamma_k\right) \tag{15a}$$

$$\text{s.t.} \quad P_{\mathrm{BS}}^{\mathrm{total}} \le P_{\mathrm{BS}}^{\max}, \tag{15b}$$

$$P_{\mathrm{R}}^{\mathrm{total}} \le P_{\mathrm{R}}^{\max}, \tag{15c}$$

$$\theta_n \in \mathcal{F}, \forall n = 1,2,\ldots,N, \tag{15d}$$

$$\gamma_{\mathrm{R},k} \ge \gamma_{\mathrm{R}}^{\mathrm{threshold}}, \forall k = 1,2,\ldots,K. \tag{15e}$$

Constraints (15b) and (15c) denote the transmit power constraint at the BS and relay, respectively. Constraint (15d) is the RIS configuration constraint. Constraint (15e) ensures the decoding SINR requirements for all users at the relay are satisfied. Note that the rate expression in (15a) takes into account the contributions of the direct link between the BS and the users (first phase), RIS-assisted links (first and second phases), and the relay-assisted link (second phase). In addition, the pre-log factor of $1/2$ due to two-phase transmission is dropped in (15a) and all achievable rate expressions in this paper, but is considered in all simulations.

It can be envisioned that one or many RISs and relays will coexist in future communication systems since it is highly likely that they will be deployed in the vicinity of each other. It is therefore of critical importance to understand various co-design tradeoffs in hybrid relay–RIS systems, for which there are only few related studies. Note that problem (15) is challenging to solve, as the beamforming variables are highly coupled, rendering a non-convex objective function. To tackle this problem, we leverage the alternating optimization (AO)-based algorithm to decouple the problem into three subproblems and alternately solve them until convergence.

III. The Proposed Joint Beamforming for Hybrid Relay–RIS Systems

In this section, we elaborate on the three subproblems in the AO-based algorithm: BS beamforming optimization, relay beamforming optimization, and RIS beamforming optimization. The algorithm is summarized at the end of the section.

A. BS Beamforming Optimization

Given fixed relay beamforming vectors $\{\mathbf{f}_k\}$ and RIS beamforming matrix $\boldsymbol{\Theta}$, problem (15) becomes

$$\max_{\{\mathbf{g}_k\}} \quad \sum_{k=1}^{K}\log_2\left(C_{1,k}+\gamma_k^{\mathrm{I}}\right) \tag{16a}$$

$$\text{s.t.} \quad P_{\mathrm{BS}}^{\mathrm{total}} \le P_{\mathrm{BS}}^{\max}, \tag{16b}$$

$$\gamma_{\mathrm{R},k} \ge \gamma_{\mathrm{R}}^{\mathrm{threshold}}, \forall k = 1,\ldots,K, \tag{16c}$$

where $C_{1,k} = 1 + \gamma_k^{\mathrm{II}}$, $k = 1,\ldots,K$ are considered as constants here under the AO-based decomposition. Subproblem (16) is still non-convex and difficult. In this subsection, we tackle subproblem (16) by first transforming the problem of finding $\mathbf{g}_k$ into one of finding a rank-one PSD matrix $\mathbf{G}_k = \mathbf{g}_k\mathbf{g}_k^{\mathrm{H}}$ and then relaxing the rank-one constraint, which eventually leads to a convex semidefinite programming (SDP) problem. Toward this end, we develop the following four steps: 1) introducing slack variables, 2) first-order Taylor approximation, 3) alternating optimization, and 4) semidefinite relaxation, as detailed below.

1) Introducing Slack Variables: We introduce slack variables $\{\mathcal{S}_{1,k}\}$, $\{\mathcal{I}_{1,k}\}$, $\{\mathcal{S}_{\mathrm{R},k}\}$, and $\{\mathcal{I}_{\mathrm{R},k}\}$ as follows

$$\frac{1}{\mathcal{S}_{1,k}} \triangleq \left|\mathbf{h}'_{\mathrm{BS},k}\mathbf{g}_k\right|^2 = \mathrm{tr}(\mathbf{G}_k\mathbf{h}''_{\mathrm{BS},k}), \tag{17}$$

$$\mathcal{I}_{1,k} \triangleq \sum_{j=1,j\neq k}^{K}\left|\mathbf{h}'_{\mathrm{BS},k}\mathbf{g}_j\right|^2 + \sigma_k^2 = \sum_{j=1,j\neq k}^{K}\mathrm{tr}(\mathbf{G}_j\mathbf{h}''_{\mathrm{BS},k}) + \sigma_k^2, \tag{18}$$

$$\frac{1}{\mathcal{S}_{\mathrm{R},k}} \triangleq \|\boldsymbol{\alpha}_k\|^4 = [\mathrm{tr}(\mathbf{G}_k\mathbf{H}''_{\mathrm{BS,R}})]^2, \tag{19}$$

$$\mathcal{I}_{\mathrm{R},k} \triangleq \sum_{j=1,j\neq k}^{K} \mathrm{tr}(\mathbf{G}_k\mathbf{H}''_{\mathrm{BS,R}}\mathbf{G}_j\mathbf{H}''_{\mathrm{BS,R}}) + \sigma_{\mathrm{R}}^2\mathrm{tr}(\mathbf{G}_k\mathbf{H}''_{\mathrm{BS,R}}), \tag{20}$$

where $\mathbf{H}''_{\mathrm{BS,R}} \triangleq (\mathbf{H}'_{\mathrm{BS,R}})^{\mathrm{H}}\mathbf{H}'_{\mathrm{BS,R}}$ and $\mathbf{h}''_{\mathrm{BS},k} \triangleq (\mathbf{h}'_{\mathrm{BS},k})^{\mathrm{H}}\mathbf{h}'_{\mathrm{BS},k}$. Substituting these slack variables into (16) leads to the following equivalent problem:

$$\begin{aligned}
\max \quad & \sum_{k=1}^{K} R_{1,k} && \text{(21a)}\\
\text{variables:} \quad & \{\mathbf{G}_k\}, \{R_{1,k}\}, \{\mathcal{S}_{1,k}\}, \{\mathcal{I}_{1,k}\}, \{\mathcal{S}_{\mathrm{R},k}\}, \{\mathcal{I}_{\mathrm{R},k}\} \\
\text{s.t.} \quad & R_{1,k} \leq \log_2\left(C_{1,k} + \frac{1}{\mathcal{S}_{1,k}\mathcal{I}_{1,k}}\right) \triangleq u_1(\mathbf{z}), && \text{(21b)}\\
& \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{G}_k\right) \leq P_{\mathrm{BS}}^{\max}, && \text{(21c)}\\
& v_1(\mathbf{z}_{\mathrm{R}}) \triangleq \frac{1}{\mathcal{S}_{\mathrm{R,k}}\mathcal{I}_{\mathrm{R,k}}} \geq \gamma_{\mathrm{R}}^{\mathrm{threshold}}, && \text{(21d)}\\
& \frac{1}{\mathcal{S}_{1,k}} \leq \mathrm{tr}\left(\mathbf{G}_k\mathbf{h}''_{\mathrm{BS},k}\right), && \text{(21e)}\\
& \mathcal{I}_{1,k} \geq \sum_{j=1,j\neq k}^{K} \mathrm{tr}\left(\mathbf{G}_j\mathbf{h}''_{\mathrm{BS},k}\right) + \sigma_k^2, && \text{(21f)}\\
& \frac{1}{\sqrt{\mathcal{S}_{\mathrm{R},k}}} \leq \mathrm{tr}\left(\mathbf{G}_k\mathbf{H}''_{\mathrm{BS,R}}\right), && \text{(21g)}\\
& \mathcal{I}_{\mathrm{R},k} \geq \sum_{j=1,j\neq k}^{K} \mathrm{tr}\left(\mathbf{G}_k\mathbf{H}''_{\mathrm{BS,R}}\mathbf{G}_j\mathbf{H}''_{\mathrm{BS,R}}\right) \\
& \quad + \sigma_{\mathrm{R}}^2\mathrm{tr}\left(\mathbf{G}_k\mathbf{H}''_{\mathrm{BS,R}}\right), && \text{(21h)}\\
& \mathbf{G}_k \succeq \mathbf{0}, && \text{(21i)}\\
& \mathrm{rank}\left(\mathbf{G}_k\right) = 1, && \text{(21j)}\\
& \forall k = 1, 2, \ldots, K,
\end{aligned}$$

where $u_1(\mathbf{z})$ and $v_1(\mathbf{z}_{\mathrm{R}})$ with $\mathbf{z} = [\mathcal{S}_{1,k}, \mathcal{I}_{1,k}]^{\mathrm{T}}$ and $\mathbf{z}_{\mathrm{R}} = [\mathcal{S}_{\mathrm{R},k}, \mathcal{I}_{\mathrm{R},k}]^{\mathrm{T}}$ respectively defined in (21b) and (21d) are for notational convenience and clarity. Since the optimal solution of (21) meets the constraints (21b) and (21e)−(21h) with equality, the two problems (16) and (21) are equivalent. However, problem (21) is still non-convex and difficult to solve.

2) First-Order Taylor Approximation: To convexify the constraints (21b) and (21d), we utilize the fact that any convex function can be lower bounded by its first-order Taylor approximation. Since $u_1(\mathbf{z})$ and $v_1(\mathbf{z}_{\mathrm{R}})$ are convex, from the first-order Taylor representation, we have

$$u_1(\mathbf{z}) \geq u_1(\mathbf{z}^{\mathrm{loc}}) + \nabla u_1(\mathbf{z}^{\mathrm{loc}})^{\mathrm{T}}(\mathbf{z} - \mathbf{z}^{\mathrm{loc}}) \triangleq R_{1,k}^{\mathrm{low}}, \tag{22}$$

$$v_1(\mathbf{z}_{\mathrm{R}}) \geq v_1(\mathbf{z}_{\mathrm{R}}^{\mathrm{loc}}) + \nabla v_1(\mathbf{z}_{\mathrm{R}}^{\mathrm{loc}})^{\mathrm{T}}(\mathbf{z}_{\mathrm{R}} - \mathbf{z}_{\mathrm{R}}^{\mathrm{loc}}) \triangleq \gamma_{\mathrm{R},k}^{\mathrm{low}}, \tag{23}$$

where $\mathbf{z}^{\mathrm{loc}} = [\mathcal{S}_{1,k}^{\mathrm{loc}}, \mathcal{I}_{1,k}^{\mathrm{loc}}]^{\mathrm{T}}$ and $\mathbf{z}_{\mathrm{R}}^{\mathrm{loc}} = [\mathcal{S}_{\mathrm{R},k}^{\mathrm{loc}}, \mathcal{I}_{\mathrm{R},k}^{\mathrm{loc}}]^{\mathrm{T}}$ are local points at which $u_1(\mathbf{z})$ and $v_1(\mathbf{z}_{\mathrm{R}})$ are differentiable, respectively. The two lower bounds can be carried out as $R_{1,k}^{\mathrm{low}} = \log_2(C_{1,k} + \frac{1}{\mathcal{S}_{1,k}^{\mathrm{loc}}\mathcal{I}_{1,k}^{\mathrm{loc}}}) - \frac{\log_2 e(\mathcal{S}_{1,k} - \mathcal{S}_{1,k}^{\mathrm{loc}})}{C_{1,k}\mathcal{S}_{1,k}^{\mathrm{loc}\,2}\mathcal{I}_{1,k}^{\mathrm{loc}} + \mathcal{S}_{1,k}^{\mathrm{loc}}} - \frac{\log_2 e(\mathcal{I}_{1,k} - \mathcal{I}_{1,k}^{\mathrm{loc}})}{C_{1,k}\mathcal{S}_{1,k}^{\mathrm{loc}}\mathcal{I}_{1,k}^{\mathrm{loc}\,2} + \mathcal{I}_{1,k}^{\mathrm{loc}}}$ and $\gamma_{\mathrm{R},k}^{\mathrm{low}} = \frac{1}{\mathcal{S}_{\mathrm{R},k}^{\mathrm{loc}}\mathcal{I}_{\mathrm{R},k}^{\mathrm{loc}}} - \frac{\mathcal{S}_{\mathrm{R},k} - \mathcal{S}_{\mathrm{R},k}^{\mathrm{loc}}}{\mathcal{S}_{\mathrm{R},k}^{\mathrm{loc}\,2}\mathcal{I}_{\mathrm{R},k}^{\mathrm{loc}}} - \frac{\mathcal{I}_{\mathrm{R},k} - \mathcal{I}_{\mathrm{R},k}^{\mathrm{loc}}}{\mathcal{S}_{\mathrm{R},k}^{\mathrm{loc}}\mathcal{I}_{\mathrm{R},k}^{\mathrm{loc}\,2}}$. Replacing $u(\mathbf{z})$ in (21b) and $v(\mathbf{z})$ in (21d) respectively with their lower bounds in (22) and (23) linearizes and convexifies the constraints (21b) and (21d), converting problem (21) to

$$\begin{aligned}
\max \quad & \sum_{k=1}^{K} R_{1,k} && \text{(24a)}\\
\text{variables:} \quad & \{\mathbf{G}_k\}, \{R_{1,k}\}, \{\mathcal{S}_{1,k}\}, \{\mathcal{I}_{1,k}\}, \{\mathcal{S}_{\mathrm{R},k}\}, \{\mathcal{I}_{\mathrm{R},k}\} \\
\text{s.t.} \quad & R_{1,k} \leq R_{1,k}^{\mathrm{low}}, && \text{(24b)}\\
& \gamma_{\mathrm{R},k}^{\mathrm{low}} \geq \gamma_{\mathrm{R}}^{\mathrm{threshold}}, && \text{(24c)}\\
& \text{(21c) , (21e)-(21j)}, && \text{(24d)}\\
& \forall k = 1, 2, \ldots, K.
\end{aligned}$$

Note that the solution set of (24) is a subset of that of problem (21) and the original problem (16), as lower bounds are used as substitutes in the constraints (21b) and (21d).

3) Alternating Optimization: The highly coupled BS beamforming matrices $\mathbf{G}_k$, $k = 1, \ldots, K$, still render the constraint (21h) non-convex. AO is adopted within the subproblem to solve $\mathbf{G}_k, k = 1, \ldots, K$, iteratively. Specifically, in the k'-th iteration, only the k'-th beamforming matrix $\mathbf{G}_{k'}$ is solved while keeping other $\mathbf{G}_j$, $j \neq k'$ as constants. As such, constraint (21h) in (24) becomes convex.

4) Semidefinite Relaxation: Finally, we relax the rank-one constraint (21j) in problem (24), which, after combining with the first-order Taylor approximation and alternating optimization, transforms the entire optimization to an SDP problem that can be solved by standard convex optimization software such as CVX [45]. In general, the optimal solutions for $\mathbf{G}_k$'s are not necessarily of rank-one. One typical approach to obtaining rank-one solutions is by using the randomization procedure [46] based on the eigen-decomposition of the optimal $\mathbf{G}_k$.

B. Relay Beamforming Optimization

Given fixed BS beamforming vectors $\{\mathbf{g}_k\}$ and RIS beamforming matrix $\boldsymbol{\Theta}$, the problem of solving the relay beamforming $\{\mathbf{f}_k\}$ in (15) becomes

$$\begin{aligned}
\max_{\{\mathbf{f}_k\}} \quad & \sum_{k=1}^{K} \log_2\left(C_{2,k} + \gamma_k^{\mathrm{II}}\right) && \text{(25a)}\\
\text{s.t.} \quad & P_{\mathrm{R}}^{\mathrm{total}} \leq P_{\mathrm{R}}^{\max}. && \text{(25b)}
\end{aligned}$$

where $C_{2,k} = 1 + \gamma_k^{\mathrm{I}}, k = 1, 2, \ldots, K$ are kept as constants. The subproblem for finding the relay bemforming vector $\mathbf{f}_k$ in (25) is non-convex. Following similar steps developed in Section III-A, we elaborate on how to convexify (25) below.

Let $\mathbf{F}_k = \mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}$. Define the slack variables $\{\mathcal{S}_{2,k}\}$ and $\{\mathcal{I}_{2,k}\}$ as

$$\frac{1}{\mathcal{S}_{2,k}} \triangleq \left|\mathbf{h}'_{\mathrm{R},k}\mathbf{f}_k\right|^2 = \mathrm{tr}\left(\mathbf{F}_k\mathbf{h}''_{\mathrm{R},k}\right), \tag{26}$$

$$\mathcal{I}_{2,k} \triangleq \sum_{j=1,j\neq k}^{K}\left|\mathbf{h}'_{\mathrm{R},k}\mathbf{f}_j\right|^2 + \sigma_k^2 = \sum_{j=1,j\neq k}^{K}\mathrm{tr}\left(\mathbf{F}_j\mathbf{h}''_{\mathrm{R},k}\right) + \sigma_k^2, \tag{27}$$

where $\mathbf{F}_k = \mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}$ and $\mathbf{h}''_{\mathrm{R},k} \triangleq (\mathbf{h}'_{\mathrm{R},k})^{\mathrm{H}}\mathbf{h}'_{\mathrm{R},k}$. With the above slack variables, problem (25) can be equivalently rewritten as

$$\max \quad \sum_{k=1}^{K} R_{2,k} \tag{28a}$$

$$\text{variables:} \quad \{\mathbf{F}_k\},\{R_{2,k}\},\{\mathcal{S}_{2,k}\},\{\mathcal{I}_{2,k}\}$$

$$\text{s.t.} \quad R_{2,k} \le \log_2\left(C_{2,k} + \frac{1}{\mathcal{S}_{2,k}\mathcal{I}_{2,k}}\right) \triangleq u_2(\mathbf{z}), \tag{28b}$$

$$\sum_{k=1}^{K}\mathrm{tr}\left(\mathbf{F}_k\right) \le P_{\mathrm{R}}^{\max}, \tag{28c}$$

$$\frac{1}{\mathcal{S}_{2,k}} \le \mathrm{tr}\left(\mathbf{F}_k\mathbf{h}''_{\mathrm{R},k}\right), \tag{28d}$$

$$\mathcal{I}_{2,k} \ge \sum_{j=1,j\neq k}^{K}\mathrm{tr}\left(\mathbf{F}_j\mathbf{h}''_{\mathrm{R},k}\right) + \sigma_k^2, \tag{28e}$$

$$\mathbf{F}_k \succeq \mathbf{0}, \tag{28f}$$

$$\mathrm{rank}\left(\mathbf{F}_k\right) = 1, \tag{28g}$$

$$\forall k = 1,2,\ldots,K,$$

where $u_2(\mathbf{z})$ defined in (28b) is for notational convenience. Problem (28) is still non-convex. Similarly, we apply first-order Taylor expression on $u_2(\mathbf{z})$ and obtain a lower bound around differentiable local points $\mathcal{S}_{2,k}^{\mathrm{loc}}$ and $\mathcal{I}_{2,k}^{\mathrm{loc}}$

$$u_2(\mathbf{z}) \ge u_2(\mathbf{z}^{\mathrm{loc}}) + \nabla u_2(\mathbf{z}^{\mathrm{loc}})^{\mathrm{T}}(\mathbf{z}-\mathbf{z}^{\mathrm{loc}}) \triangleq R_{2,k}^{\mathrm{low}}, \tag{29}$$

where $R_{2,k}^{\mathrm{low}}$ has the closed-form expression

$$R_{2,k}^{\mathrm{low}} \triangleq \log_2\left(C_{2,k} + \frac{1}{\mathcal{S}_{2,k}^{\mathrm{loc}}\mathcal{I}_{2,k}^{\mathrm{loc}}}\right) - \frac{\log_2 e\left(\mathcal{S}_{2,k}-\mathcal{S}_{2,k}^{\mathrm{loc}}\right)}{C_{2,k}\mathcal{S}_{2,k}^{\mathrm{loc}\,2}\mathcal{I}_{2,k}^{\mathrm{loc}} + \mathcal{S}_{2,k}^{\mathrm{loc}}} - \frac{\log_2 e\left(\mathcal{I}_{2,k}-\mathcal{I}_{2,k}^{\mathrm{loc}}\right)}{C_{2,k}\mathcal{S}_{2,k}^{\mathrm{loc}}\mathcal{I}_{2,k}^{\mathrm{loc}\,2} + \mathcal{I}_{2,k}^{\mathrm{loc}}}. \tag{30}$$

With the above lower bound, (28) can be converted to

$$\max \quad \sum_{k=1}^{K} R_{2,k} \tag{31a}$$

$$\text{variables:} \quad \{\mathbf{F}_k\},\{R_{2,k}\},\{\mathcal{S}_{2,k}\},\{\mathcal{I}_{2,k}\}$$

$$\text{s.t.} \quad R_{2,k} \le R_{2,k}^{\mathrm{low}}, \forall k = 1,2,\ldots,K, \tag{31b}$$

$$(28\mathrm{c}) - (28\mathrm{g}). \tag{31c}$$

The solution set of (31) is the subset of the solution set of (28), since the lower bound $R_{2,k}^{\mathrm{low}}$ is used in the constraint (28b). By relaxing the rank-one constraint (28g), we can solve the resulting SDP problem with CVX. Finally, as in the case in Section III-A, the randomization procedure can be employed to obtain rank-one solutions for $\{\mathbf{f}_k\}$, $k = 1,\ldots,K$.

C. RIS Beamforming Optimization

Given fixed BS beamforming vectors $\{\mathbf{g}_k\}$ and relay beamforming vectors $\{\mathbf{f}_k\}$, the problem of solving the RIS beamforming matrix $\boldsymbol{\Theta}$ (or equivalently its vector form $\boldsymbol{\theta}$) in (15) becomes

$$\max_{\boldsymbol{\Theta}} \quad \sum_{k=1}^{K}\log_2\left(1+\gamma_k\right) \tag{32a}$$

$$\text{s.t.} \quad \theta_n \in \mathcal{F}, \forall n = 1,\ldots,N, \tag{32b}$$

$$\gamma_{\mathrm{R},k} \ge \gamma_{\mathrm{R}}^{\mathrm{threshold}}, \forall k = 1,\ldots,K. \tag{32c}$$

To deal with the non-convex constraint (32b), we first relax the constraint to the set $\{\theta_n|\theta_n = e^{j\psi_n}, \psi_n \in [0,2\pi)\}$. Then, we adopt similar steps as in the previous two subsections to solve the relaxed problem. To facilitate finding the optimal RIS beamforming matrix $\boldsymbol{\Theta}$, we rearrange the order of $\boldsymbol{\Theta}$ in (4) and (12) respectively to

$$y_k^{\mathrm{I}} = \boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,RIS},k}\mathbf{x} + w_k^{\mathrm{I}}, \quad \text{and}$$
$$y_k^{\mathrm{II}} = \boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{R,RIS},k}\mathbf{x}_{\mathrm{R}} + w_k^{\mathrm{II}},$$

where $\boldsymbol{\phi} = [\boldsymbol{\theta} \quad 1]^{\mathrm{H}}$, $\mathbf{H}_{\mathrm{BS,RIS},k} = \begin{bmatrix}\mathrm{diag}(\mathbf{h}_{\mathrm{RIS,k}})\mathbf{H}_{\mathrm{BS,RIS}} \\ \mathbf{h}_{\mathrm{BS},k}\end{bmatrix}$, and $\mathbf{H}_{\mathrm{R,RIS},k} = \begin{bmatrix}\mathrm{diag}(\mathbf{h}_{\mathrm{RIS,k}})\mathbf{H}_{\mathrm{R,RIS}} \\ \mathbf{h}_{\mathrm{R},k}\end{bmatrix}$.

With the reordering, the slack variables $\{\mathcal{S}_{1,k}\}$, $\{\mathcal{I}_{1,k}\}$, $\{\mathcal{S}_{2,k}\}$, and $\{\mathcal{I}_{2,k}\}$ respectively in (17), (18), (26), and (27) can be rewritten as

$$\frac{1}{\mathcal{S}_{1,k}} = \left|\boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,RIS},k}\mathbf{g}_k\right|^2 = \mathrm{tr}\left(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{BS,RIS},k}\mathbf{g}_k\mathbf{g}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,RIS},k}^{\mathrm{H}}\right), \tag{33}$$

$$\mathcal{I}_{1,k} = \sum_{j=1,j\neq k}^{K}\left|\boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,RIS},k}\mathbf{g}_j\right|^2 + \sigma_k^2 = \sum_{j=1,j\neq k}^{K}\mathrm{tr}\left(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{BS,RIS},k}\mathbf{g}_j\mathbf{g}_j^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,RIS},k}^{\mathrm{H}}\right) + \sigma_k^2, \tag{34}$$

$$\frac{1}{\mathcal{S}_{2,k}} = \left|\boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{R,RIS},k}\mathbf{f}_k\right|^2 = \mathrm{tr}(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{R,RIS},k}\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{R,RIS},k}^{\mathrm{H}}), \tag{35}$$

$$\mathcal{I}_{2,k} = \sum_{j=1,j\neq k}^{K}\left|\boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{R,RIS},k}\mathbf{f}_j\right|^2 + \sigma_k^2 = \sum_{j=1,j\neq k}^{K}\mathrm{tr}\left(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{R,RIS},k}\mathbf{f}_j\mathbf{f}_j^{\mathrm{H}}\mathbf{H}_{\mathrm{R,RIS},k}^{\mathrm{H}}\right) + \sigma_k^2, \tag{36}$$

where $\mathbf{\Phi} \triangleq \boldsymbol{\phi}\boldsymbol{\phi}^{\mathrm{H}}$. Similarly, we can rearrange the order of $\mathbf{\Theta}$ in (8), by which the slack variables $\{\mathcal{S}_{\mathrm{R},k}\}$ and $\{\mathcal{I}_{\mathrm{R},k}\}$ in (19) and (20) can be rewritten as

$$\frac{1}{\mathcal{S}_{\mathrm{R},k}} = \left[\sum_{l=1}^{L} \mathrm{tr}\left(\mathbf{\Phi}\mathbf{A}_{k,k,l}\right)\right]^2, \tag{37}$$

$$\begin{aligned}\mathcal{I}_{\mathrm{R},k} &= \sum_{j=1,j\neq k}^{k}\left|\sum_{l=1}^{L}\mathrm{tr}\left(\mathbf{\Phi}\mathbf{A}_{j,k,l}\right)\right|^2 + \sigma_{\mathrm{R}}^2\sum_{l=1}^{L}\mathrm{tr}\left(\mathbf{\Phi}\mathbf{A}_{k,k,l}\right)\\ &= \|\boldsymbol{\beta}_{-k}\|^2 + \sigma_{\mathrm{R}}^2\sum_{l=1}^{L}\mathrm{tr}\left(\mathbf{\Phi}\mathbf{A}_{k,k,l}\right),\end{aligned} \tag{38}$$

where $\mathbf{A}_{j,k,l} \triangleq \mathbf{H}_{\mathrm{BS,R},l}\mathbf{g}_j\mathbf{g}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,R},l}^{\mathrm{H}}$ and $\boldsymbol{\beta}_{-k} \triangleq [b_1 \ \cdots \ b_{k-1} \ \ b_{k+1} \ \cdots \ b_K]^{\mathrm{T}}$ with $b_j \triangleq \sum_{l=1}^{L}\mathrm{tr}(\mathbf{\Phi}\mathbf{A}_{j,k,l})$.

With the above slack variables, we have the following equivalent problem:

$$\max \quad \sum_{k=1}^{K} R_{3,k} \tag{39a}$$

$$\begin{aligned}\text{variables:}\quad & \mathbf{\Phi}, \{R_{3,k}\}, \{\mathcal{S}_{1,k}\}, \{\mathcal{I}_{1,k}\}, \{\mathcal{S}_{2,k}\}, \{\mathcal{I}_{2,k}\},\\ & \{\mathcal{S}_{\mathrm{R},k}\}, \{\mathcal{I}_{\mathrm{R},k}\}\end{aligned}$$

$$\text{s.t.} \quad R_{3,k} \le \log_2\left(1 + \frac{1}{\mathcal{S}_{1,k}\mathcal{I}_{1,k}} + \frac{1}{\mathcal{S}_{2,k}\mathcal{I}_{2,k}}\right) \triangleq u_3(\mathbf{z}), \tag{39b}$$

$$v_3(\mathbf{z}_{\mathrm{R}}) \triangleq \frac{1}{\mathcal{S}_{\mathrm{R},k}\mathcal{I}_{\mathrm{R},k}} \ge \gamma_{\mathrm{R}}^{\mathrm{threshold}}, \tag{39c}$$

$$\frac{1}{\mathcal{S}_{1,k}} \le \mathrm{tr}\left(\mathbf{\Phi}\mathbf{H}_{\mathrm{BS,RIS},k}\mathbf{g}_k\mathbf{g}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,RIS},k}^{\mathrm{H}}\right), \tag{39d}$$

$$\mathcal{I}_{1,k} \ge \sum_{j=1,j\neq k}^{K}\mathrm{tr}\left(\mathbf{\Phi}\mathbf{H}_{\mathrm{BS,RIS},k}\mathbf{g}_j\mathbf{g}_j^{\mathrm{H}}\mathbf{H}_{\mathrm{BS,RIS},k}^{\mathrm{H}}\right) + \sigma_k^2, \tag{39e}$$

$$\frac{1}{\mathcal{S}_{2,k}} \le \mathrm{tr}\left(\mathbf{\Phi}\mathbf{H}_{\mathrm{R,RIS},k}\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{R,RIS},k}^{\mathrm{H}}\right), \tag{39f}$$

$$\mathcal{I}_{2,k} \ge \sum_{j=1,j\neq k}^{K}\mathrm{tr}\left(\mathbf{\Phi}\mathbf{H}_{\mathrm{R,RIS},k}\mathbf{f}_j\mathbf{f}_j^{\mathrm{H}}\mathbf{H}_{\mathrm{R,RIS},k}^{\mathrm{H}}\right) + \sigma_k^2, \tag{39g}$$

$$\frac{1}{\sqrt{\mathcal{S}_{\mathrm{R},k}}} \le \sum_{l=1}^{L}\mathrm{tr}\left(\mathbf{\Phi}\mathbf{A}_{k,k,l}\right), \tag{39h}$$

$$\sqrt{\mathcal{I}_{\mathrm{R},k} - \sigma_{\mathrm{R}}^2\sum_{l=1}^{L}\mathrm{tr}\left(\mathbf{\Phi}\mathbf{A}_{k,k,l}\right)} \ge \|\boldsymbol{\beta}_{-k}\|, \tag{39i}$$

$$\mathbf{\Phi} \succeq \mathbf{0}, \tag{39j}$$

$$\mathrm{rank}\left(\mathbf{\Phi}\right) = 1, \tag{39k}$$

$$\mathbf{\Phi}[n,n] = 1, \forall n = 1,2,\ldots,N+1, \quad \forall k = 1,2,\ldots,K, \tag{39l}$$

where $u_3(\mathbf{z})$ and $v_3(\mathbf{z}_{\mathrm{R}})$ respectively defined in (39b) and (39c) are for notational convenience. Problem (39) is still non-convex. The first-order Taylor approximation is adopted to obtain lower bounds of $u_3(\mathbf{z})$ and $v_3(\mathbf{z}_{\mathrm{R}})$ to convexify the two constraints (39b) and (39c). With differentiable local points $\{\mathcal{S}_{1,k}^{\mathrm{loc}}, \mathcal{I}_{1,k}^{\mathrm{loc}}, \mathcal{S}_{2,k}^{\mathrm{loc}}, \mathcal{I}_{2,k}^{\mathrm{loc}}\}$, the lower bound $R_{3,k}^{\mathrm{low}}$ of $u_3(\mathbf{z})$ is

$$u_3(\mathbf{z}) \ge u_3(\mathbf{z}^{\mathrm{loc}}) + \nabla u_3(\mathbf{z}^{\mathrm{loc}})^{\mathrm{T}}(\mathbf{z} - \mathbf{z}^{\mathrm{loc}}) \triangleq R_{3,k}^{\mathrm{low}}, \tag{40}$$

which has the following closed-form expression

$$\begin{aligned}R_{3,k}^{\mathrm{low}} \triangleq\ & \log_2\left(1 + \frac{1}{\mathcal{S}_{1,k}^{\mathrm{loc}}\mathcal{I}_{1,k}^{\mathrm{loc}}} + \frac{1}{\mathcal{S}_{2,k}^{\mathrm{loc}}\mathcal{I}_{2,k}^{\mathrm{loc}}}\right)\\ & - \frac{\log_2 e\left(\mathcal{S}_{1,k} - \mathcal{S}_{1,k}^{\mathrm{loc}}\right)}{\left(1 + \frac{1}{\mathcal{S}_{2,k}^{\mathrm{loc}}\mathcal{I}_{2,k}^{\mathrm{loc}}}\right)\mathcal{S}_{1,k}^{\mathrm{loc}^2}\mathcal{I}_{1,k}^{\mathrm{loc}} + \mathcal{S}_{1,k}^{\mathrm{loc}}}\\ & - \frac{\log_2 e\left(\mathcal{I}_{1,k} - \mathcal{I}_{1,k}^{\mathrm{loc}}\right)}{\left(1 + \frac{1}{\mathcal{S}_{2,k}^{\mathrm{loc}}\mathcal{I}_{2,k}^{\mathrm{loc}}}\right)\mathcal{S}_{1,k}^{\mathrm{loc}}\mathcal{I}_{1,k}^{\mathrm{loc}^2} + \mathcal{I}_{1,k}^{\mathrm{loc}}}\\ & - \frac{\log_2 e\left(\mathcal{S}_{2,k} - \mathcal{S}_{2,k}^{\mathrm{loc}}\right)}{\left(1 + \frac{1}{\mathcal{S}_{1,k}^{\mathrm{loc}}\mathcal{I}_{1,k}^{\mathrm{loc}}}\right)\mathcal{S}_{2,k}^{\mathrm{loc}^2}\mathcal{I}_{2,k}^{\mathrm{loc}} + \mathcal{S}_{2,k}^{\mathrm{loc}}}\\ & - \frac{\log_2 e\left(\mathcal{I}_{2,k} - \mathcal{I}_{2,k}^{\mathrm{loc}}\right)}{\left(1 + \frac{1}{\mathcal{S}_{1,k}^{\mathrm{loc}}\mathcal{I}_{1,k}^{\mathrm{loc}}}\right)\mathcal{S}_{2,k}^{\mathrm{loc}}\mathcal{I}_{2,k}^{\mathrm{loc}^2} + \mathcal{I}_{2,k}^{\mathrm{loc}}}.\end{aligned} \tag{41}$$

The lower bound of $v_3(\mathbf{z}_{\mathrm{R}})$ is the same as (23). Then, (39) can be rewritten as

$$\max \quad \sum_{k=1}^{K} R_{3,k} \tag{42a}$$

$$\begin{aligned}\text{variables:}\quad & \mathbf{\Phi}, \{R_{3,k}\}, \{\mathcal{S}_{1,k}\}, \{\mathcal{I}_{1,k}\}, \{\mathcal{S}_{2,k}\}, \{\mathcal{I}_{2,k}\},\\ & \{\mathcal{S}_{\mathrm{R},k}\}, \{\mathcal{I}_{\mathrm{R},k}\}\end{aligned}$$

$$\text{s.t.} \quad R_{3,k} \le R_{3,k}^{\mathrm{low}}, \tag{42b}$$

$$\gamma_{\mathrm{R},k}^{\mathrm{low}} \ge \gamma_{\mathrm{R}}^{\mathrm{threshold}}, \tag{42c}$$

$$\text{(39d) – (39l)}, \quad \forall k = 1,2,\ldots,K. \tag{42d}$$

As in the previous subsection, we relax the rank-one constraint (39k) and solve the resulting SDP problem. Finally, Gaussian randomization procedure is employed to obtain a rank-one solution $\mathbf{\Phi}$ and the corresponding (continuous-phase) $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$, from which the discrete phase shifts are determined by [16], [18], [44]

$$\angle\theta_n = \underset{\psi_n' \in \left\{0, \frac{2\pi}{2^B}, \ldots, \frac{2\pi(2^B-1)}{2^B}\right\}}{\arg\min} \left|\psi_n' - \psi_n\right|. \tag{43}$$

To ensure convergence, we update the RIS beamforming only when it results in the objective value not decreasing after each iteration, similar to [18].

The entire proposed algorithm, alternately optimizing the BS beamforming (in Section III-A), relay beamforming (in

Algorithm 1: Joint Beamforming for Hybrid Relay–RIS Systems.

Initialization: Randomly initialize points $\mathbf{g}_k(0), \mathbf{f}_k(0), \forall k$, and $\boldsymbol{\theta}(0)$. Let iteration index $t = 1$ and $t_k = 1, \forall k$.
repeat
  BS Beamforming Optimization:
  Given fixed variables $\mathbf{f}_k^{\text{fixed}} = \mathbf{f}_k(t-1), \forall k$ and $\boldsymbol{\theta}^{\text{fixed}} = \boldsymbol{\theta}(t)$.
  for $k \leftarrow 1$ to K **do**
    Given fixed variables $\mathbf{g}_j^{\text{fixed}} = \mathbf{g}_j(t_j - 1), j \neq k$ and local point $\mathbf{g}_k^{\text{loc}} = \mathbf{g}_k(t_k - 1)$.
    Calculate $\mathcal{S}_{1,k}^{\text{loc}}, \mathcal{I}_{1,k}^{\text{loc}}, \mathcal{S}_{\text{R},k}^{\text{loc}}$, and $\mathcal{S}_{\text{R},k}^{\text{loc}}, \forall k$ according to (17)–(20), respectively, with local point $\mathbf{g}_k^{\text{loc}}$ and fixed variables.
    Solve (24) with fixed $\mathbf{g}_j^{\text{fixed}}, j \neq k$ to obtain $\mathbf{G}_k(t_k)$.
    Utilize randomization procedure to generate $\mathbf{g}_k(t_k)$.
    $t_k \leftarrow t_k + 1$
  end for
  Relay Beamforming Optimization:
  Given fixed variables $\mathbf{g}_k^{\text{fixed}} = \mathbf{g}_k(t_k - 1), \forall k$ and $\boldsymbol{\Theta}^{\text{fixed}} = \boldsymbol{\Theta}(t-1)$, and local points $\mathbf{f}_k^{\text{loc}} = \mathbf{f}_k(t-1), \forall k$.
  Calculate $\mathcal{S}_{2,k}^{\text{loc}}$ and $\mathcal{I}_{2,k}^{\text{loc}}, \forall k$ according to (26) and (27), respectively, with local points $\mathbf{f}_k^{\text{loc}}, \forall k$ and fixed variables.
  Solve (31) to obtain $\mathbf{F}_k(t), \forall k$.
  Utilize randomization procedure to generate $\mathbf{f}_k(t)$.
  RIS Beamforming Optimization:
  Given fixed variables $\mathbf{g}_k^{\text{fixed}} = \mathbf{g}_k(t_k - 1)$ and $\mathbf{f}_k^{\text{fixed}} = \mathbf{f}_k(t), \forall k$, and local point $\boldsymbol{\theta}^{\text{loc}} = \boldsymbol{\theta}(t-1)$.
  Calculate $\mathcal{S}_{1,k}^{\text{loc}}, \mathcal{I}_{1,k}^{\text{loc}}, \mathcal{S}_{2,k}^{\text{loc}}, \mathcal{I}_{2,k}^{\text{loc}}, \mathcal{S}_{\text{R,k}}^{\text{loc}}$, and $\mathcal{I}_{\text{R,k}}^{\text{loc}}, \forall k$ according to (33) - (38), respectively, with $\boldsymbol{\theta}^{\text{loc}}$ and fixed variables.
  Solve (42) to obtain $\boldsymbol{\Phi}(t)$.
  Utilize randomization procedure to generate $\boldsymbol{\theta}(t)$.
  Obtain discrete phase shifts by (43).
  $t \leftarrow t + 1$
until the objective value converges
Output: $\{\mathbf{g}_k\}$, $\{\mathbf{f}_k\}$, and $\boldsymbol{\theta}$.

Section III-B), and RIS beamforming (in Section III-C), is summarized in Algorithm 1.

IV. Complexity and Convergence Analysis

A. Complexity

In the proposed algorithm, the BS beamforming, the relay beamforming, and the RIS beamforming are transformed into SDP and solved iteratively. According to [47], the worst-case computational complexity of solving a generic SDP using CVX is given by $\mathcal{O}(\max\{m,n\}^4 n^{1/2}\log(1/\epsilon))$, where m is the number of constraints, n is the size of the PSD matrix in the objective, and ϵ is a constant reflecting the solution accuracy of the CVX solver.

For solving the BS beamforming optimization in (24), since AO is adopted within the BS beamforming optimization to obtain one $\mathbf{g}_k$ in an iteration, K SDPs need to be solved to obtain all $\mathbf{g}_k$'s. In each iteration, the SDP consists of K, K, 1, 1, $K-1$, 1, and K constraints in (24b), (24c), (21c), (21e)–(21h), respectively, which add up to $m = 4K + 2$ for each SDP. With $n = M$, the worst-case complexity is given by

$$\mathcal{O}\left(K \max\{4K+2, M\}^4 M^{1/2}\log(1/\epsilon)\right). \tag{44}$$

For solving the relay beamforming optimization in (31), we have $m = 3K + 1$ and $n = L$, yielding the worst-case complexity

$$\mathcal{O}\left(\max\{3K+1, L\}^4 L^{1/2}\log(1/\epsilon)\right). \tag{45}$$

Similarly, the worst-case complexity for the RIS beamforming optimization is

$$\mathcal{O}\left(\max\{8K, N\}^4 N^{1/2}\log(1/\epsilon)\right). \tag{46}$$

Summing up (44), (45), and (46), it follows that the worst-case complexity of the entire proposed algorithm is

$$\mathcal{O}\Big(N_{\text{iter}}\Big(K\max\{4K+2,M\}^4M^{1/2}+\max\{3K+1,L\}^4L^{1/2} +\max\{8K,N\}^4N^{1/2}\Big)\log(1/\epsilon)\Big), \tag{47}$$

where N_{iter} is the number of iterations.

B. Convergence

Define $\mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^i\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i)$, $\mathcal{R}_{\mathbf{g}}(\{\mathbf{g}_k^i\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i)$, $\mathcal{R}_{\mathbf{f}}(\{\mathbf{g}_k^i\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i)$, and $\mathcal{R}_{\boldsymbol{\Theta}}(\{\mathbf{g}_k^i\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i)$ as the objective value of (15), (24), (31), and (42), respectively, in the ith iteration. For the BS beamforming optimization in the $(i+1)$th iteration, given fixed relay beamforming vectors $\{\mathbf{f}_k^i\}$ and fixed RIS beamforming matrix $\boldsymbol{\Theta}^i$, we have [18]

$$\begin{aligned}\mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^i\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i) &\overset{(a)}{=} \mathcal{R}_{\mathbf{g}}(\{\mathbf{g}_k^i\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i)\\ &\overset{(b)}{\leq} \mathcal{R}_{\mathbf{g}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i)\\ &\overset{(c)}{\leq} \mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i),\end{aligned} \tag{48}$$

where (a) holds since the first-order Taylor approximation is tight at the local points $\{\mathbf{g}_k^i\}$ in problem (24), (b) holds since $\{\mathbf{g}_k^{i+1}\}$ solves problem (24) optimally in the $(i+1)$th iteration, and (c) holds since the optimal objective value of problem (24) is a lower bound of that of problem (16) and problem (15) given $\{\mathbf{g}_k^{i+1}\}$, $\{\mathbf{f}_k^i\}$, and $\boldsymbol{\Theta}^i$. Similarly, for the relay beamforming optimization in the $(i+1)$th iteration, given fixed BS beamforming vectors $\{\mathbf{g}_k^{i+1}\}$ and fixed RIS beamforming matrix $\boldsymbol{\Theta}^i$, we have

$$\begin{aligned}\mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i) &= \mathcal{R}_{\mathbf{f}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^i\},\boldsymbol{\Theta}^i)\\ &\leq \mathcal{R}_{\mathbf{f}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^{i+1}\},\boldsymbol{\Theta}^i)\\ &\leq \mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^{i+1}\},\boldsymbol{\Theta}^i).\end{aligned} \tag{49}$$

For the RIS beamforming optimization in the $(i+1)$th iteration, since we update the RIS beamforming only when the objective

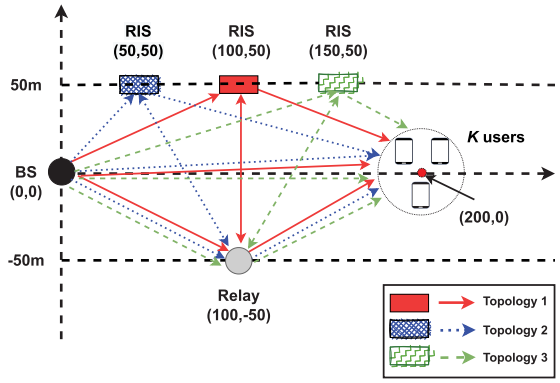


Fig. 2. Simulation scenarios.

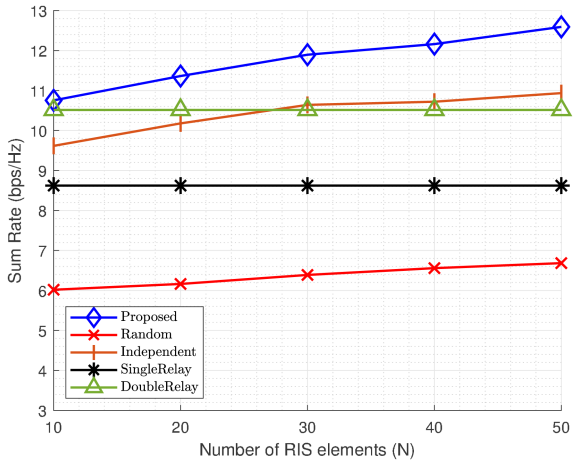


Fig. 3. Sum-rate vs. number of RIS elements for Topology 1, with $(M, N, L, K) = (8, N, 4, 4)$.

value does not decrease after each iteration as stated in Section II-I-C, given fixed BS beamforming vectors $\{\mathbf{g}_k^{i+1}\}$ and fixed relay beamforming vectors $\{\mathbf{f}_k^{i+1}\}$, we have

$$\mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^{i+1}\},\boldsymbol{\Theta}^i) \leq \mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^{i+1}\},\boldsymbol{\Theta}^{i+1}). \tag{50}$$

It follows from (48)–(50) that

$$\mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^{i}\},\{\mathbf{f}_k^{i}\},\boldsymbol{\Theta}^i) \leq \mathcal{R}_{\mathbf{g},\mathbf{f},\boldsymbol{\Theta}}(\{\mathbf{g}_k^{i+1}\},\{\mathbf{f}_k^{i+1}\},\boldsymbol{\Theta}^{i+1}), \tag{51}$$

which suggests that the objective value is non-decreasing after each iteration of the proposed algorithm. Since the optimal sum rate in (15) is finite, (51) ensures convergence of the proposed algorithm.

V. Simulation Results and Discussion

A. Simulation Settings

We simulate Topologies 1–3 shown in Fig. 2. The BS and relay are located at fixed coordinates (0,0) and $(100, -50)$, with varying RIS locations at (100, 50), (50, 50), and (150, 50) in Topologies 1–3, respectively. The K users are randomly distributed within a circle centered at (200, 0) with radius 10. The units are all meters. Topology 1 is the default topology while Topology 2 and Topology 3 represent RIS placement nearer to and farther from the BS, respectively. We set the channel parameters as follows: for RIS-related channels, $K_1 = 10$ dB, $\kappa_{\text{LoS}} = \kappa_{\text{NLoS}} = 10^{-0.5}$, $\varrho_{\text{LoS}} = 2$, $\varrho_{\text{NLoS}} = 2.5$, and $d_0 = 1$ m; for non-RIS-related channels, $\kappa = 10^{-4.5}$ for all, $\varrho = 2.5$ for $\mathbf{H}_{\text{BS,R}}$, $\mathbf{h}_{\text{R},k}$ and $\varrho = 3.5$ for $\mathbf{h}_{\text{BS},k}$, and $d_0 = 1$ m. We set the power constraints $P_{\text{BS}}^{\max} = P_{\text{R}}^{\max} = 10$ mW, SINR threshold $\gamma_{\text{R}}^{\text{threshold}} = 0.01$, noise powers $\sigma_k^2 = \sigma_{\text{R}}^2 = 2 \times 10^{-5}$, and finite-resolution RIS $B = 2$. The proposed scheme is compared with the following benchmarks: 1) *Random:* RIS beamforming takes phase shifts randomly out of $\mathcal{F}$, with BS and relay beamforming optimized in such case; 2) *Independent:* RIS assists transmission in the first phase but not in the second phase (i.e., disabled in the second phase), with BS, relay, and RIS beamforming optimized in such case. It is termed *independent* because there is no interaction between RIS and relay in forwarding information in the second phase; 3) *SingleRelay:* there is no RIS; only the relay assists transmission, with BS and relay beamforming optimized; and 4) *DoubleRelay:* the RIS is replaced by another DF relay, with BS beamforming and relay beamforming for both relays optimized. For fairness, each relay in this scheme has only half of the available powers. Note that 1)–2) are hybrid relay–RIS schemes, and 3)–4) are relay-only schemes.

B. Sum-Rate Performance in Topology 1

Fig. 3 compares the sum-rate vs. number of RIS elements performance, with $(M, N, L, K) = (8, N, 4, 4)$. Several observations can be made. First, the sum-rate increases with N for all hybrid schemes, with smaller gains for Random than for the proposed and Independent schemes. Second, the proposed scheme achieves the highest sum-rate due to the hybrid setting and the judicious joint beamforming design. Comparing the proposed scheme and DoubleRelay confirms the advantages of a well-optimized hybrid relay–RIS system over the traditional two-relay system. Third, comparing the proposed and Independent schemes, where the RIS beamforming is optimized with and without considering the second-phase transmission, respectively, it is observed that Independent achieves higher SINR for end users in the first phase, as shown in Table I, but significantly lower SINR for end users in the second phase, since the RIS is disabled and there is no performance gain from RIS-assisted links in the second phase. This illustrates a tradeoff in the optimization of RIS beamforming serving multiple, possibly conflicting, purposes. Overall, Independent exhibits a performance loss as compared to the proposed scheme, suggesting the effectiveness of the proposed joint design in both phases. Fourth, comparing Independent and Random, it is observed that optimized RIS beamforming in the first phase only still presents advantages over random RIS beamforming in both phases, since the higher SINR in the first phase outweighs the lower SINR in the second phase for Independent, as shown in Table I. The fact that Random has inferior performance suggests the importance of a judicious joint design in the considered scenario. Comparing Independent and DoubleRelay, it is observed that Independent is advantageous over DoubleRelay when N is larger. Fifth,

TABLE I
PHASE I AND PHASE II SINR FOR END USERS, FOR TOPOLOGY 1 WITH NETWORK CONFIGURATIONS $(M, N, L, K) = (8, 30, 4, 4)$

SINR		Phase I SINR				Phase II SINR			
		User 1	User 2	User 3	User 4	User 1	User 2	User 3	User 4
Scheme	Proposed	2.677E+02	3.942E+02	3.724E+02	6.082E+02	1.279E+03	5.372E+02	5.685E+02	7.845E+02
	Random	2.703E+00	7.718E+00	2.281E+00	1.652E+00	3.167E+01	1.656E+03	3.366E+02	1.311E+03
	Independent	7.489E+02	8.362E+02	4.084E+02	1.392E+03	7.168E+01	8.468E+01	7.415E+01	7.949E+01
	SingleRelay	4.173E−05	4.897E−05	4.426E−05	4.926E−05	5.039E+01	6.464E+01	5.104E+01	6.202E+01
	DoubleRelay	4.861E−05	5.073E−05	4.962E−05	5.189E−05	6.773E+01	7.217E+01	6.623E+01	7.222E+01

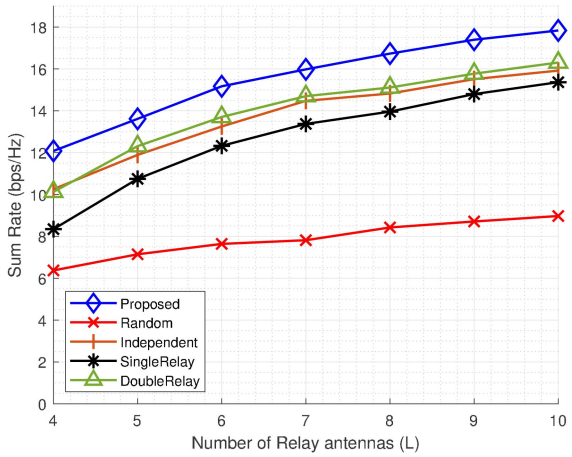


Fig. 4. Sum-rate vs. number of relay antennas for Topology 1, with $(M, N, L, K) = (8, 30, L, 4)$.

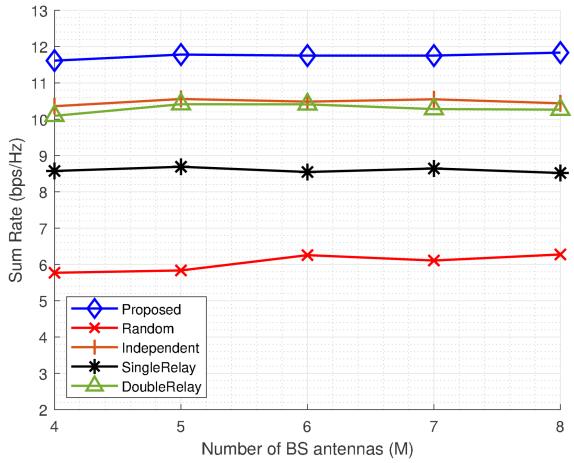


Fig. 5. Sum-rate vs. number of BS antennas for Topology 1, with $(M, N, L, K) = (M, 30, 4, 4)$.

DoubleRelay outperforms SingleRelay for the same total relay transmit power due to the diversity provided by an additional transmission link.

Fig. 4 compares the sum-rate vs. number of relay antennas performance, with $(M, N, L, K) = (8, 30, L, 4)$. As can be seen, the sum-rate of all schemes increases with L, due to the higher degrees of freedom in the second phase. Specifically, the relay with more antennas can better adjust the relay beamforming vectors $\{\mathbf{f}_k\}$ in (13) for the combined RIS-assisted and direct channels from the relay to end users in the second phase. Increasing L generally provides diminishing returns for all schemes. SingleRelay and DoubleRelay benefit from increasing L since their performance is dominated by the second-phase transmission, as shown in Table I.

Fig. 5 plots the sum-rate vs. number of BS antennas performance, with $(M, N, L, K) = (M, 30, 4, 4)$. For all schemes, M has a moderate to negligible effect on sum-rate. This is because the sum-rate is dominated by the second-phase transmission in this topology with a weak direct link, and therefore increasing M, which contributes to higher SINR for end users in the first phase, creates a marginal gain.

Fig. 6 compares the sum-rate vs. number of users performance, with $(M, N, L, K) = (8, 30, 8, K)$. It is seen that the performance of all schemes increases with K until a certain point where the lower transmit power allocated for each user (given fixed BS/relay transmit power constraints) and the heightened multiuser interference negatively affect the performance [48]. The effect is most severe for Random with a non-optimized RIS. An optimized RIS, as in the proposed and Independent schemes, helps mitigate multiuser interference and results in a more steadily increased sum-rate, especially in the regime of larger numbers of users.

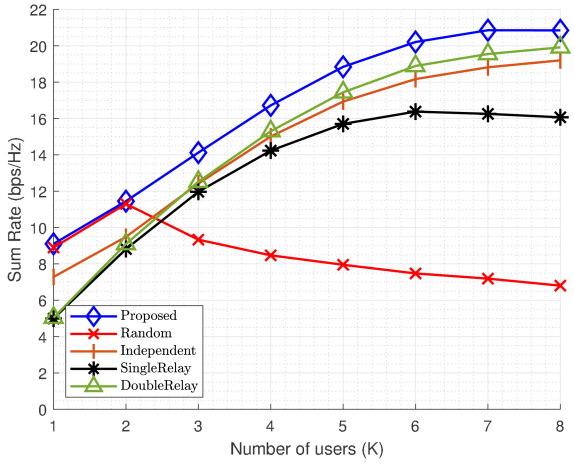


Fig. 6. Sum-rate vs. number of users for Topology 1, with $(M, N, L, K) = (8, 30, 8, K)$.

C. Sum-Rate Performance in Topology 2 and Topology 3

The effect of topology is examined from selected perspectives to provide new insights and avoid duplication. Figs. 7 and 8 (Figs. 9 and 10) present the same comparisons as in Figs. 3 and 4, respectively, but for Topology 2 (Topology 3). Placing the RIS closer to the BS in Topology 2 increases the contribution of the first-phase transmission and reduces the contribution of

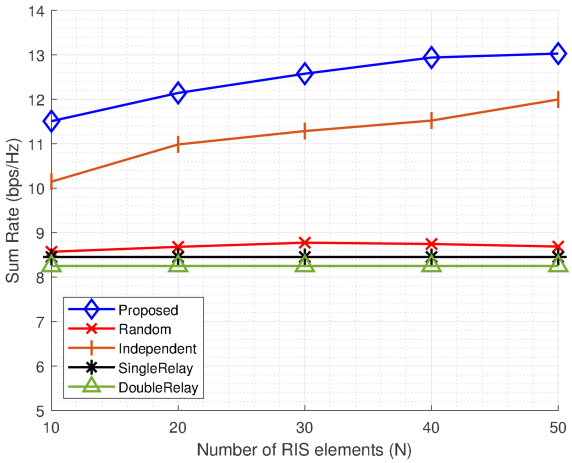


Fig. 7. Sum-rate vs. number of RIS elements for Topology 2, with $(M, N, L, K) = (8, N, 4, 4)$.

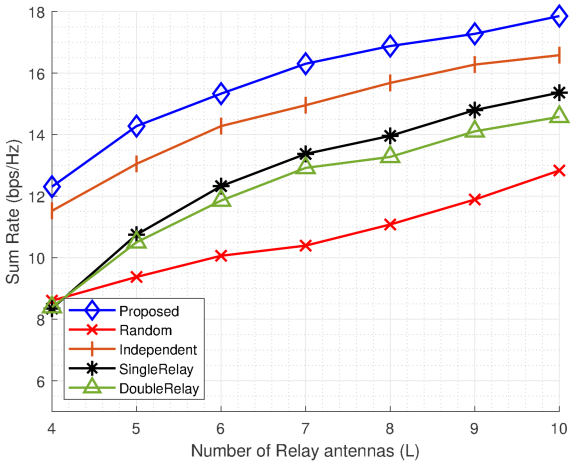


Fig. 8. Sum-rate vs. number of relay antennas for Topology 2, with $(M, N, L, K) = (8, 30, L, 4)$.

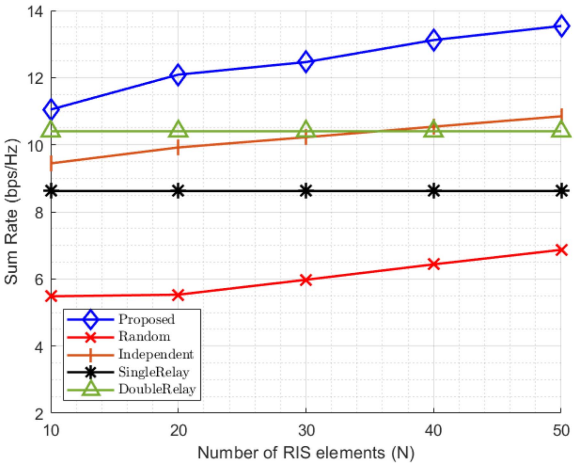


Fig. 9. Sum-rate vs. number of RIS elements for Topology 3, with $(M, N, L, K) = (8, N, 4, 4)$.

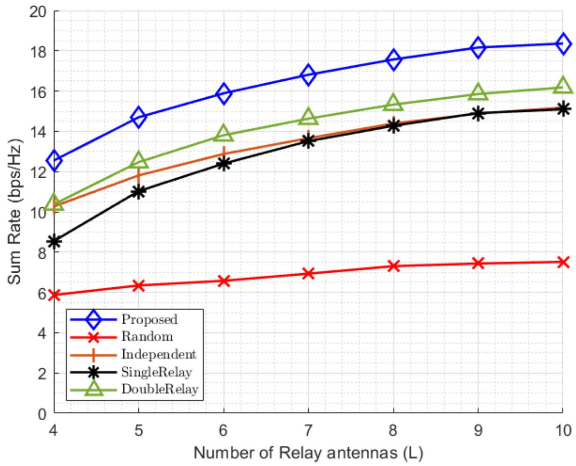


Fig. 10. Sum-rate vs. number of relay antennas for Topology 3, with $(M, N, L, K) = (8, 30, L, 4)$.

the second-phase transmission to the sum-rate. This leads to the increased performance of Independent. The performance of Random also increases, because Random, without optimized RIS beamforming, relies more on BS beamforming in the first phase and this effect is heightened in Topology 2. The performance of DoubleRelay decreases, even to the point where it is outperformed by SingleRelay. This is because, while the additional relay in DoubleRelay provides diversity, it is also farther from the users in Topology 2. Topology 3 shows similar results as compared to Topology 1, but with a slight decrease in the relative performance of Independent and Random due to the reduced contribution of first-phase transmission. The performance of the proposed scheme remains largely stable in Topologies 1–3. This study of different topologies suggests that RIS placement and a judicious co-design have a determining effect on the performance of the hybrid system.

D. Sum-Rate Performance in Topology 1 With Imperfect CSI

The effects of imperfect CSI are examined numerically. Similar to [49], [50], we consider imperfect RIS–user channels since obtaining accurate CSI of these channels is the most challenging due to varying user locations and environmental conditions. The channel from the RIS to user k for beamforming design is modeled as

$$\widehat{\mathbf{h}}_{\mathrm{RIS},k} = \mathbf{h}_{\mathrm{RIS},k} + \Delta\mathbf{h}_{\mathrm{RIS},k}, \tag{52}$$

where $\widehat{\mathbf{h}}_{\mathrm{RIS},k}$ is the estimated channel and $\Delta\mathbf{h}_{\mathrm{RIS},k}$ is the estimation error with $\Delta\mathbf{h}_{\mathrm{RIS},k}[i] \sim \mathcal{CN}(0, \sigma_{\mathrm{e}}^2), \forall i = 1, 2, \ldots, N$. Fig. 11 compares the sum-rate vs. number of RIS elements performance of hybrid relay–RIS schemes under imperfect CSI conditions ($\sigma_{\mathrm{e}}^2 = 0.001$ and 0.01) alongside the perfect CSI condition. As can be seen, the performance of all schemes degrades as σ_{e}^2 increases. The proposed scheme with an optimized RIS is more affected by channel estimation errors than Random, but still achieves higher sum-rates in the same channel conditions. The Independent is relatively unaffected by channel estimation

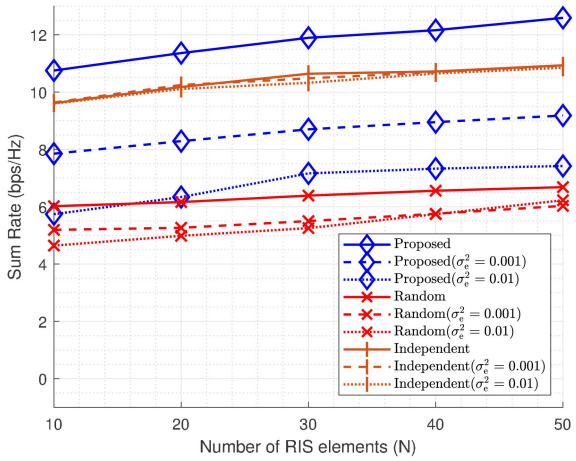


Fig. 11. Sum-rate vs. number of RIS elements for Topology 1 under imperfect RIS–user channel conditions, with $(M, N, L, K) = (8, N, 4, 4)$.

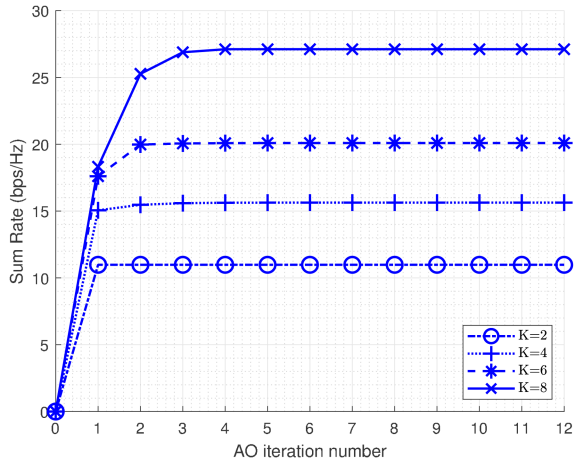


Fig. 12. Convergence behavior of the proposed algorithm for Topology 1, with $(M, N, L, K) = (8, 30, 8, K)$ and $K = 2, 4, 6, 8$.

errors, due to the absence of RIS–user channels in the second phase. In fact, Independent's advantage in the imperfect RIS–user channel condition increases by avoiding RIS–user channels in the second phase.

E. Convergence and Complexity Properties of the Proposed Algorithm for Topology 1

Fig. 12 demonstrates the convergence behavior of the proposed scheme, for $(M, N, L, K) = (8, 30, 8, K)$ with $K = 2, 4, 6, 8$ in Topology 1. The result confirms the convergence analysis in Section IV-B and shows that the proposed scheme converges in few iterations. The number of iterations required to converge increases with K. This is because the number of subproblems that need to be solved to obtain BS beamforming vectors is proportional to K, as described in Section III-A.

The CPU time for the proposed scheme under different network configurations is presented in Table II. As can be seen, CPU time per iteration increases monotonically with K, N, M, or L. This agrees with the complexity analysis provided in

TABLE II
RUNTIME OF THE PROPOSED ALGORITHM FOR TOPOLOGY 1

Number of users (K) $(M, N, L, K) = (8, 30, 8, K)$	2	4	6	8
CPU time per iteration (sec)	5.38	8.38	18.76	25.66
Number of RIS elements (N) $(M, N, L, K) = (8, N, 4, 4)$	20	30	40	50
CPU time per iteration (sec)	11.52	12.14	12.73	14.62
Number of BS antennas (M) $(M, N, L, K) = (M, 30, 4, 4)$	5	6	7	8
CPU time per iteration (sec)	10.10	11.68	11.97	12.23
Number of Relay antennas (L) $(M, N, L, K) = (8, 30, L, 4)$	6	7	8	9
CPU time per iteration (sec)	10.59	10.61	10.81	11.15

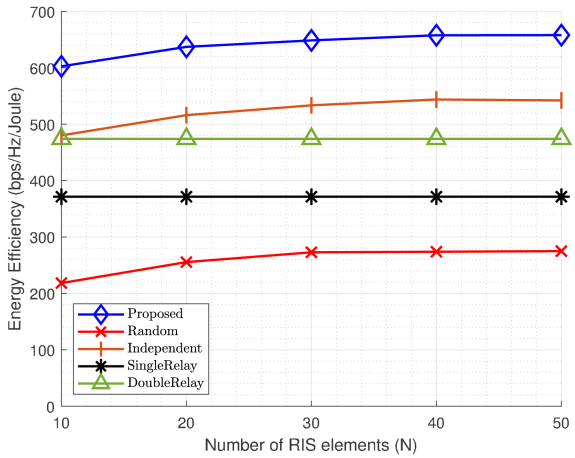


Fig. 13. Energy efficiency vs. number of RIS elements for Topology 1, with $(M, N, L, K) = (8, N, 4, 4)$.

Section IV-A, although the changes are at a more moderate rate than predicted by the worst-case complexity analysis due to the various numerical optimizations in the actual implementation. The fact that the runtime generally increases more rapidly with K and N than with M and L also confirms the analysis in (47).

F. Energy Efficiency Performance in Topology 1

The proposed solution method can be extended to solve an energy efficiency (EE) problem with the objective (15a) being replaced by $\frac{\sum_{k=1}^{K} \log_2(1+\gamma_k)}{P_{\mathrm{BS}}^{\mathrm{total}} + P_{\mathrm{R}}^{\mathrm{total}} + P_{\mathrm{RIS}}^{\mathrm{total}}}$, where $P_{\mathrm{RIS}}^{\mathrm{total}} = N \times P_{\mathrm{RIS}}$ denotes the RIS power consumption, with P_{RIS} being the power consumption of each RIS element [25], [51]. The extension entails first reformulating an equivalent problem with the objective $\sum_{k=1}^{K} \log_2(1 + \gamma_k) - \xi(P_{\mathrm{BS}}^{\mathrm{total}} + P_{\mathrm{R}}^{\mathrm{total}} + P_{\mathrm{RIS}}^{\mathrm{total}})$ via the Dinkelbach's transform [52], [53], where ξ is an auxiliary variable, and then applying a similar AO procedure to obtain ξ, $\{\mathbf{g}_k\}$, $\{\mathbf{f}_k\}$, and $\boldsymbol{\Theta}$ iteratively. Figs. 13 and 14 present the same comparisons as in Figs. 3 and 4, respectively, but in terms of EE, where we set $P_{\mathrm{RIS}} = 0.05$ mW [25], [51]. The comparative relations largely agree with the sum-rate results. Compared to sum-rate, the EE increases more moderately with N for all hybrid schemes in Fig. 13 due to the additional RIS power consumption proportional to N. The additional RIS power consumption also somewhat decreases the advantages

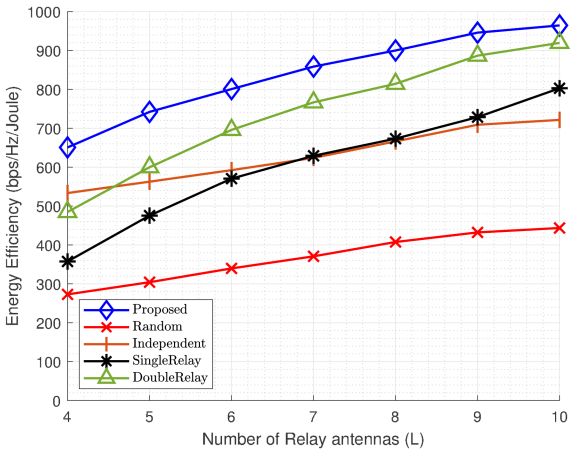


Fig. 14. Energy efficiency vs. number of relay antennas for Topology 1, with $(M, N, L, K) = (8, 30, L, 4)$.

of the proposed and Independent schemes as compared with SingleRelay and DoubleRelay in larger L regimes in Fig. 14.

VI. Conclusion

We have considered a multiuser MISO system with a coexisting RIS and DF relay. We investigated the joint BS beamforming, relay beamforming, and RIS beamforming design problem in this hybrid relay–RIS system, and proposed an algorithm to tackle the complex non-convex problem. The proposed joint design scheme demonstrated substantial performance advantages over benchmarks. The main findings include: 1) With optimized beamforming in their respective scenarios, hybrid relay–RIS systems carry the potential to achieve higher sum-rate and EE performance as compared to classical relay-only systems. 2) A judicious joint design in both transmission phases is crucial to fully exploit the hybrid relay–RIS system. 3) There exist tradeoffs in RIS configurations to assist BS transmission in the first phase and relay transmission in the second phase. 4) The proposed algorithm converges in few iterations, and its complexity is dominated by the numbers of users and RIS elements.

Appendix
Derivation of (9)

From (8), the received signal at the relay can be written as

$$\begin{aligned}\mathbf{y}_{\mathrm{R}} &= \mathbf{H}'_{\mathrm{BS,R}}\mathbf{x} + \mathbf{w}_{\mathrm{R}} \\ &= \mathbf{H}'_{\mathrm{BS,R}}\mathbf{g}_k s_k + \sum_{j=1, j\neq k}^{K} \mathbf{H}'_{\mathrm{BS,R}}\mathbf{g}_j s_j + \mathbf{w}_{\mathrm{R}} \\ &= \boldsymbol{\alpha}_k s_k + \sum_{j=1, j\neq k}^{K} \boldsymbol{\alpha}_j s_j + \mathbf{w}_{\mathrm{R}} \\ &= \boldsymbol{\alpha}_k s_k + \mathbf{n}_k, \end{aligned} \tag{53}$$

where $\mathbf{n}_k = \sum_{j=1, j\neq k}^{K} \boldsymbol{\alpha}_j s_j + \mathbf{w}_{\mathrm{R}}$ is the interference plus noise. Let $\mathbf{q}_k$ be the combining filter employed at the relay to decode users' signals from the BS. It follows that the SINR for user k after adopting the matched filter $\mathbf{q}_k = \boldsymbol{\alpha}_k$ is

$$\begin{aligned}\gamma_{\mathrm{R},k} &= \frac{\mathbb{E}\left[\left|\mathbf{q}_k^{\mathrm{H}}\boldsymbol{\alpha}_k s_k\right|^2\right]}{\mathbb{E}\left[\left|\mathbf{q}_k^{\mathrm{H}}\mathbf{n}_k\right|^2\right]} \\ &= \frac{\mathbf{q}_k^{\mathrm{H}}\boldsymbol{\alpha}_k\boldsymbol{\alpha}_k^{\mathrm{H}}\mathbf{q}_k}{\mathbf{q}_k^{\mathrm{H}}\mathbb{E}\left[\mathbf{n}_k\mathbf{n}_k^{\mathrm{H}}\right]\mathbf{q}_k} \\ &\overset{(a)}{=} \frac{\mathbf{q}_k^{\mathrm{H}}\boldsymbol{\alpha}_k\boldsymbol{\alpha}_k^{\mathrm{H}}\mathbf{q}_k}{\mathbf{q}_k^{\mathrm{H}}\left(\sum_{j=1, j\neq k}^{K}\boldsymbol{\alpha}_j\boldsymbol{\alpha}_j^{\mathrm{H}} + \sigma_{\mathrm{R}}^2\mathbf{I}_L\right)\mathbf{q}_k} \\ &\overset{(b)}{=} \frac{\|\boldsymbol{\alpha}_k\|^4}{\sum_{j=1, j\neq k}^{K}\boldsymbol{\alpha}_k^{\mathrm{H}}\boldsymbol{\alpha}_j\boldsymbol{\alpha}_j^{\mathrm{H}}\boldsymbol{\alpha}_k + \sigma_{\mathrm{R}}^2\|\boldsymbol{\alpha}_k\|^2}, \end{aligned} \tag{54}$$

where (a) holds since the interfering signals and noise are uncorrelated, and the users' signals are independent to each other; (b) is achieved by substituting $\mathbf{q}_k = \boldsymbol{\alpha}_k$.

References

[1] T.-Y. Kan, R. Y. Chang, and F.-T. Chien, "Intelligent reflecting surfaces and classical relays: Coexistence and co-design," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM) Workshops*, Madrid, Spain, 2021, pp. 1–6.
[2] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.
[3] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of data transmission with large intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2746–2758, May 2018.
[4] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. F. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.
[5] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
[6] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Net.*, vol. 2019, no. 1, pp. 1–20, May 2019.
[7] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.
[8] S. Gong et al., "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Commun. Surv. Tuts.*, vol. 22, no. 4, pp. 2283–2314, Oct.–Dec. 2020.
[9] M. Jian et al., "Reconfigurable intelligent surfaces for wireless communications: Overview of hardware designs, channel models, and estimation techniques," *Intell. Converged Netw.*, vol. 3, no. 1, pp. 1–32, Mar. 2022.
[10] J. Xu et al., "Reconfiguring wireless environment via intelligent surfaces for 6G: Reflection, modulation, and security," 2022, *arXiv: 2208.10931.*
[11] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, 2018, pp. 1–6.
[12] C. Huang, A. Zappone, M. Debbah, and C. Yuen, "Achievable rate maximization by passive intelligent mirrors," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Calgary, AB, Canada, 2018, pp. 3714–3718.
[13] C. Huang, G. C. Alexandropoulos, C. Yuen, and M. Debbah, "Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, Cannes, France, 2019, pp. 1–5.
[14] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, Aug. 2020.
[15] Y. Yang, S. Zhang, and R. Zhang, "IRS-enhanced OFDM: Power allocation and passive array optimization," in *Proc. IEEE Glob. Commun. Conf.*, Waikoloa, HI, USA, 2019, pp. 1–6.

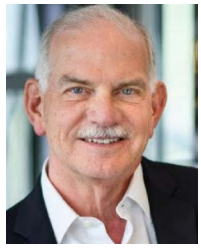
[16] H. Li, R. Liu, M. Liy, Q. Liu, and X. Li, "IRS-enhanced wideband MU-MISO-OFDM communication systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Seoul, South Korea, 2020, pp, 1–6.
[17] G. Yang, X. Xu, and Y.-C. Liang, "Intelligent reflecting surface assisted non-orthogonal multiple access," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Seoul, South Korea, 2020, pp. 1–6.
[18] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, "Exploiting intelligent reflecting surfaces in NOMA networks: Joint beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6884–6898, Oct. 2020.
[19] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.
[20] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5849–5863, Sep. 2020.
[21] B. Zheng, Q. Wu, and R. Zhang, "Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 753–757, Apr. 2020.
[22] M. Fu, Y. Zhou, and Y. Shi, "Intelligent reflecting surface for downlink non-orthogonal multiple access networks," in *Proc. IEEE Glob. Commun. Conf. Workshops*, Waikoloa, HI, USA, 2019, pp. 1–6.
[23] E. Björnson and L. Sanguinetti, "Power scaling laws and near-field behaviors of massive MIMO and intelligent reflecting surfaces," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1306–1324, 2020.
[24] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2526–2537, Nov. 2020.
[25] 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. 2020.
[26] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
[27] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7748–7764, Dec. 2020.
[28] M. Di Renzo et al., "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, 2020.
[29] 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.
[30] C. Huang, G. Chen, Y. Gong, M. Wen, and J. A. Chambers, "Deep reinforcement learning-based relay selection in intelligent reflecting surface assisted cooperative networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1036–1040, May 2021.
[31] 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.
[32] B. Zheng and R. Zhang, "IRS meets relaying: Joint resource allocation and passive beamforming optimization," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 2080–2084, Jun. 2021.
[33] 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. 2022.
[34] Z. Abdullah, S. Kisseleff, K. Ntontin, W. A. Martins, S. Chatzinotas, and B. Ottersten, "Successive decode-and-forward relaying with reconfigurable intelligent surfaces," in *Proc. IEEE Int. Conf. Commun.*, Seoul, South Korea, 2022, pp. 2633–2638.
[35] 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.
[36] 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.*, vol. 21, no. 10, pp. 8216–8229, Oct. 2022.
[37] G. C. Alexandropoulos, N. Shlezinger, I. Alamzadeh, M. F. Imani, H. Zhang, and Y. C. Eldar, "Hybrid reconfigurable intelligent metasurfaces: Enabling simultaneous tunable reflections and sensing for 6G wireless communications," 2021. *arXiv:2104.04690.*
[38] Y. Han, S. Zhang, L. Duan, and R. Zhang, "Cooperative double-IRS aided communication: Beamforming design and power scaling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1206–1210, Aug. 2020.
[39] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.
[40] J. Yuan, Y.-C. Liang, J. Joung, G. Feng, and E. G. Larsson, "Intelligent reflecting surface-assisted cognitive radio system," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 675–687, Jan. 2021.
[41] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, Jun. 2021.
[42] B. Zheng, C. You, W. Mei, and R. Zhang, "A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Commun. Surv. Tuts.*, vol. 24, no. 2, pp. 1035–1071, Apr.–Jun. 2022.
[43] J. An, C. Xu, L. Gan, and L. Hanzo, "Low-complexity channel estimation and passive beamforming for RIS-assisted MIMO systems relying on discrete phase shifts," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1245–1260, Feb. 2022.
[44] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate optimization for intelligent reflecting surface enhanced wireless networks," in *Proc. IEEE Glob. Commun. Conf.*, Waikoloa, HI, USA, 2019, pp. 1–6.
[45] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," Mar. 2014. [Online]. Available: http://cvxr.com/cvx
[46] Y. Huang and D. P. Palomar, "Randomized algorithms for optimal solutions of double-sided QCQP with applications in signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1093–1108, Mar. 2014.
[47] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
[48] X. Ma, S. Guo, H. Zhang, Y. Fang, and D. Yuan, "Joint beamforming and reflecting design in reconfigurable intelligent surface-aided multi-user communication systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3269–3283, May 2021.
[49] G. Zhou et al., "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1658–1662, Oct. 2020.
[50] P. Yang, L. Yang, and S. Wang, "Performance analysis for RIS-aided wireless systems with imperfect CSI," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 588–592, Mar. 2022.
[51] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3236–3251, 2020.
[52] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
[53] K. Shen and W. Yu, "Fractional programming for communication systems–Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.

Te-Yi Kan received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 2019, and the M.S. degree in electrical and computer engineering from the University of California, Los Angeles, CA, USA, in 2020. He is currently working toward the Ph.D. degree in electrical and computer engineering with the University of Southern California, Los Angeles. From 2020 to 2022, he was a Research Assistant with the Research Center for Information Technology Innovation, Academia Sinica, Taipei. His research interests include wireless communication systems and wireless networks.

Ronald Y. Chang (Senior Member, IEEE) received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2000, the M.S. degree in electronics engineering from National Chiao Tung University, Hsinchu, in 2002, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2008. From 2002 to 2003, he was with the Industrial Technology Research Institute, Hsinchu. In 2008, he was a Research Intern with the Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. In 2009, he was involved in the NASA Small Business Innovation Research projects. Since 2010, he has been with the Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan, where he is currently a full Research Fellow (Professor). He was a Visiting Scholar with the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA, in July–August 2018, and with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ, USA, in 2021–2023. His research interests include wireless communications and networking. He was the recipient of the Best Paper Award from the IEEE Wireless Communications and Networking Conference (WCNC), in 2012, and the Outstanding Young Scholar Award from the Ministry of Science and Technology, Taiwan, in 2015 and 2017, respectively. He is currently an Associate Editor for IEEE ACCESS and *APSIPA Transactions on Signal and Information Processing*, and an APSIPA Distinguished Lecturer.

Feng-Tsun Chien (Senior Member, IEEE) received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1995, the M.S. degree in electrical engineering from National Taiwan University, New Taipei, Taiwan, in 1997, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2004. From 2005 to 2021, he was with the Department of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan. Since 2021, he has been with the Institute of Electronics and the Institute of Artificial Intelligence Innovation, National Yang Ming Chiao Tung University, Hsinchu, where he is currently a Professor. He was a Fulbright Scholar with the University of California at Los Angeles, Los Angeles, CA, during the academic year 2016–2017. His research interests include wireless communications, statistical signal processing, machine learning for wireless communications, and game theoretic resource allocation. Dr. Chien is currently an Associate Editor for IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS and *APSIPA Transactions on Signal and Information Processing*.

Bing-Jia Chen received the B.S. degree in 2022 from the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, where he is currently working toward the M.S. degree with the Graduate Institute of Communication Engineering. Since 2021, he has been a Research Assistant with the Research Center for Information Technology Innovation, Academia Sinica, Taipei. His research interests include deep learning for indoor localization, the development of prototypes and communication systems for reconfigurable intelligent surfaces, and machine learning for wireless networks and Internet of Things (IoT) applications.



H. Vincent Poor (Life Fellow, IEEE) received the Ph.D. degree in EECS from Princeton University, Princeton, NJ, USA, in 1977. From 1977 to 1990, he was on the Faculty of the University of Illinois at Urbana-Champaign, Champaign, IL, USA. Since 1990, he has been on the Faculty with Princeton, where he is currently the Michael Henry Strater University Professor. During 2006 to 2016, he was the Dean of Princeton's School of Engineering and Applied Science. He has also held visiting appointments with several other universities, including most recently with the University of California, Berkeley, Berkeley, CA, USA, and University of Cambridge, Cambridge, U.K. His research interests include information theory, machine learning and network science, and their applications in wireless networks, energy systems and related fields. Among his publications in these areas is the recent book *Machine Learning and Wireless Communications* (Cambridge University Press, 2022). Dr. Poor is a Member of the National Academy of Engineering and the National Academy of Sciences and is a foreign Member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. He was the recipient of the IEEE Alexander Graham Bell Medal in 2017.