Intelligent Reflecting Surfaces and Classical Relays: Coexistence and Co-Design

Te-Yi Kan[1], Ronald Y. Chang[1], and Feng-Tsun Chien[2]

[1]Research Center for Information Technology Innovation, Academia Sinica, Taiwan

[2]Institute of Electronics, National Yang Ming Chiao Tung University, Taiwan

Email: dexter.ty.kan@gmail.com, rchang@citi.sinica.edu.tw, ftchien@mail.nctu.edu.tw

Abstract—**This paper investigates a multiuser downlink communication system with coexisting intelligent reflecting surface (IRS) and classical half-duplex decode-and-forward (DF) relay. In this system, the IRS and the DF relay interact with each other and assist transmission simultaneously. In particular, active beamforming at the base station (BS) and at the DF relay, and passive beamforming at the IRS, are jointly designed to maximize the sum-rate of all users. The sum-rate maximization problem is non-convex due to the coupled beamforming vectors. We propose an alternating optimization (AO) based algorithm to tackle this complex co-design problem. Numerical validation and discussion on the superiority of the coexistence system and the tradeoffs therein are presented.**

I. INTRODUCTION

Intelligent reflecting surfaces (IRSs), also known as reconfigurable intelligent surfaces (RISs), passive intelligent mirrors (PIMs), and large intelligent surfaces (LISs), have recently been proposed as a key enabler for next-generation wireless communications [1], [2]. IRSs 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 IRSs is less expensive as compared to installing active transmitters such as base stations (BSs) and relays, and operating IRSs is more energy-efficient and eases signal processing and interference management requirements due to its low-cost and passive nature. Thus, IRSs 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.

There is a wealth of research on IRS focusing on joint transmit beamforming and IRS phase shifts optimization for sum-rate maximization (e.g., [3]–[7]) or power minimization (e.g., [8], [9]). IRS-enhanced multi-input single-output (MISO) systems were studied for single-user [3] and multiuser [4] scenarios, where transmit beamforming at the BS and the phase shifts at the IRS were optimized by optimizing the received signal power or sum-rate of users. Machine learning-based techniques were also introduced into IRS-aided systems to adjust transmit beamforming and IRS phase shifts [5]. IRS-enhanced broadband orthogonal frequency division multiplexing (OFDM) systems were also studied [6]. In [6], BS beamforming and IRS reflection were jointly designed for the objective of sum-rate maximization or, after transformation, mean square error (MSE) minimization in multiuser

MISO systems. IRS-empowered non-orthogonal multiple access (NOMA) systems for signal-to-interference-plus-noise ratio (SINR) improvement were also studied [7]. On the other hand, power minimization problems in IRS-aided systems were considered [8], [9]. In [8], transmit beamforming and IRS beamforming for power minimization subject to user SINR constraints in MISO systems were considered. In [9], transmit power minimization in IRS-empowered NOMA systems was studied. It was shown that incorporating IRSs into NOMA brings merits in decoding and differentiating multiple users due to IRS's ability to reconfigure wireless channels.

Similarities and differences between IRS and classical decode-and-forward (DF) and amplify-and-forward (AF) relaying have been explored [10]–[13]. In [10], comparisons between IRS and DF relaying were conducted analytically and numerically for a SISO system, showing that IRS with sufficient reflecting elements can achieve higher data rates and better energy efficiency in high target data rate regimes than DF relaying. In [11], comparisons between IRS and AF relaying were conducted numerically in an IRS-aided multiuser MISO system. It was shown that replacing the IRS with an AF relay yields higher sum rates but lower energy efficiency due to active power amplification of the AF relay. In [12], IRS and full-duplex/half-duplex AF relaying were compared. It was shown that IRS achieves comparable or even better performance as compared to full-duplex and half-duplex AF relaying when the number of IRS reflecting elements is sufficiently large. In [13], 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 IRS with sufficient reflecting elements has the potential to outperform relay-aided transmission since IRS operates in a full-duplex manner, yet without the loop interference as in full-duplex relays.

In this paper, we consider the practical and previously unexamined scenario with *coexisting* IRS and half-duplex DF relay in a multiuser MISO (MU-MISO) system. The consideration is interesting and new, in especially the following two aspects. First, the IRS and relay coexistence system is different from the traditional multi-relay (in this case, two-relay) system, since there was no interaction between the relays in a multi-relay system but there is interaction between IRS and relay in the new system. Second, the IRS reflects signals at both
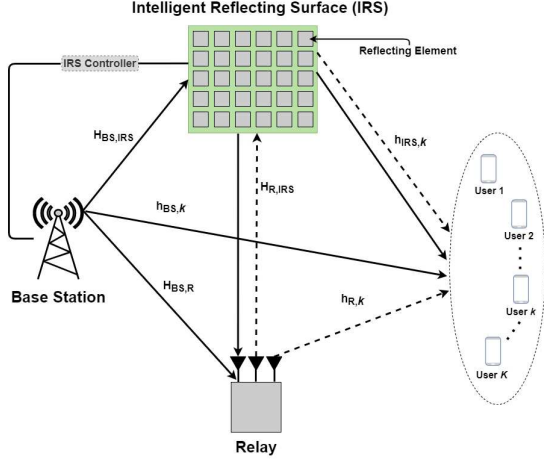
Fig. 1. A coexisting IRS and relay assisted MU-MISO system. Solid/dashed lines indicate transmissions in the first/second phases.

phases of half-duplex DF relaying, and there exist tradeoffs in the design of IRS beamforming which caters for both the end users and the DF relay. The complex co-design problem of BS beamforming, relay beamforming, and IRS beamforming for maximum system sum rates is examined. Numerical validation and discussion on the superiority of and tradeoffs in the coexistence system are presented.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

As shown in Fig. 1, we consider a downlink MU-MISO communication system that comprises a BS with M antennas, K single-antenna end users, an IRS with N reflecting elements, and a classical half-duplex DF relay with L antennas. Here, to ensure a sufficient degree of freedom, we assume $K \leq \min\{M, L\}$. The baseband equivalent channels from the BS to the relay, from the BS to the IRS, from the BS to the end user k, from the relay to the IRS, from the relay to the end user k, and from the IRS to the end user k are denoted by $\mathbf{H}_{\mathrm{BS,R}} \in \mathbb{C}^{L \times M}$, $\mathbf{H}_{\mathrm{BS,IRS}} \in \mathbb{C}^{N \times M}$, $\mathbf{h}_{\mathrm{BS},k} \in \mathbb{C}^{1 \times M}$, $\mathbf{H}_{\mathrm{R,IRS}} \in \mathbb{C}^{N \times L}$, $\mathbf{h}_{\mathrm{R},k} \in \mathbb{C}^{1 \times L}$, and $\mathbf{h}_{\mathrm{IRS},k} \in \mathbb{C}^{1 \times N}$, respectively. All channels are assumed to be quasi-static and Rayleigh flat-fading. The small-scale fading is modeled by complex Gaussian with zero mean and unit variance and the large-scale fading is modeled by $\kappa(d/d_0)^{-\varrho}$, where d is the distance between the two end nodes, d_0 is the reference distance, and κ is a large-scale fading constant. The CSI of all channels is assumed perfectly known at the BS, the IRS controller, and the relay. Due to presence of the relay, there are two phases in one complete transmission of information, for which the first phase and the second phase are assumed to be within the channel coherence interval. Moreover, in this paper, reflecting elements can be adjusted by the IRS controller once in each complete transmission of the information. 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 via direct link. Meanwhile,

the IRS reflects the incident signal from the BS towards the DF relay and all users. The received signal at user k in the first phase is given by

$$y_k^{\mathrm{I}} = \mathbf{h}'_{\mathrm{BS},k}\mathbf{x} + w_k^{\mathrm{I}}, \qquad (1)$$

where $\mathbf{h}'_{\mathrm{BS},k} \triangleq \mathbf{h}_{\mathrm{IRS},k}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{BS,IRS}} + \mathbf{h}_{\mathrm{BS},k}$ is the effective channel from the BS to the end user k, $\boldsymbol{\Theta} \triangleq \mathrm{diag}(\boldsymbol{\theta})$ with $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \cdots & \theta_N \end{bmatrix}$ is the diagonal matrix accounting for the passive IRS beamforming, and $w_k^{\mathrm{I}} \sim \mathcal{CN}\left(0, \sigma_k^2\right)$ denotes the additive white Gaussian noise (AWGN) at user k during the first phase. Note that the n-th reflecting element adjustment θ_n is given by $\theta_n = \beta_n e^{j\psi_n}$ with the amplitude attenuation $\beta_n \leq 1$ and the phase shift $\psi_n \in [0, 2\pi)$.

The transmit signal at the BS is $\mathbf{x} = \sum_{k=1}^{K} \mathbf{g}_k s_k = \mathbf{G} \cdot \mathbf{s}$, where $\mathbf{g}_k \in \mathbb{C}^{M \times 1}$ denotes the beamforming vector for user k with $\mathbf{G} \triangleq [\mathbf{g}_1, \ldots, \mathbf{g}_K]$ and $\mathbf{s} \triangleq [s_1, \ldots, s_K]^{\mathrm{T}}$, s_k is the data intended for user k with zero mean and unit variance. The data streams from different users are assumed independent. Then, the transmit power at the BS is $P_{\mathrm{BS}}^{\mathrm{total}} = \mathbb{E}\left[\|\mathbf{x}\|^2\right] = \mathrm{tr}\left(\mathbf{G}\mathbf{G}^{\mathrm{H}}\right) \leq P_{\mathrm{BS}}^{\max}$, where $P_{\mathrm{BS}}^{\max}$ is the maximum available power at the BS. The SINR for user k in the first phase is

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

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

$$\mathbf{y}_{\mathrm{R}} = \mathbf{H}'_{\mathrm{BS,R}}\mathbf{x} + \mathbf{w}_{\mathrm{R}}, \qquad (3)$$

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

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

where $\boldsymbol{\alpha}_k \triangleq \mathbf{H}'_{\mathrm{BS,R}}\mathbf{g}_k$ is the filter weight for user k.

2) Second phase: During the second phase, the relay transmits the signal to all users and the IRS, 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 $\gamma_{\mathrm{R},k}$ for user k at the relay exceeds a predefined threshold $\gamma_{\mathrm{R}}^{\mathrm{th}}$. The transmit signal at the relay is $\mathbf{x}_{\mathrm{R}} = \sum_{k=1}^{K} \mathbf{f}_k s_k = \mathbf{F} \cdot \mathbf{s}$, where $\mathbf{f}_k \in \mathbb{C}^{M \times 1}$ denotes the beamforming vector at the relay for user k and $\mathbf{F} \triangleq [\mathbf{f}_1, \ldots, \mathbf{f}_K]$. The relay's transmit power is $P_{\mathrm{R}}^{\mathrm{total}} = \mathbb{E}\left[\|\mathbf{x}_{\mathrm{R}}\|^2\right] = \mathrm{tr}\left(\mathbf{F}\mathbf{F}^{\mathrm{H}}\right) \leq P_{\mathrm{R}}^{\max}$, where $P_{\mathrm{R}}^{\max}$ denotes the relay's maximum available power. The received signal at user k in the second phase can be represented by

$$y_k^{\mathrm{II}} = \mathbf{h}'_{\mathrm{R},k}\mathbf{x}_{\mathrm{R}} + w_k^{\mathrm{II}}, \qquad (5)$$

where $\mathbf{h}'_{\mathrm{R},k} \triangleq \mathbf{h}_{\mathrm{IRS},k}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{R,IRS}} + \mathbf{h}_{\mathrm{R},k}$ is the effective channel from the relay to user k and $w_k^{\mathrm{II}} \sim \mathcal{CN}\left(0, \sigma_k^2\right)$ is the

corresponding AWGN. It follows that the SINR for user k in the second phase is

$$\gamma_k^{\mathrm{II}} = \frac{|\mathbf{h}'_{\mathrm{R},k}\mathbf{f}_k|^2}{\sum_{j=1,j\neq k}^{K}|\mathbf{h}'_{\mathrm{R},k}\mathbf{f}_j|^2 + \sigma_k^2}. \tag{6}$$

Finally, the resulting SINR after combining the signals from the two phases using the maximal ratio combining is

$$\gamma_k = \gamma_k^{\mathrm{I}} + \gamma_k^{\mathrm{II}}. \tag{7}$$

B. Problem Formulation

In this paper, we aim for maximizing the sum-rate of all the users by jointly optimizing the active beamforming at the BS, the active beamforming at the relay, and the passive beamforming at the IRS. Specifically, the optimization problem is formulated as

$$\max_{\{\mathbf{g}_k\},\{\mathbf{f}_k\},\boldsymbol{\Theta}} \quad \sum_{k=1}^{K}\log_2\left(1+\gamma_k\right) \tag{8a}$$

$$\text{s.t.} \quad P_{\mathrm{BS}}^{\mathrm{total}} \leq P_{\mathrm{BS}}^{\max}, \tag{8b}$$

$$P_{\mathrm{R}}^{\mathrm{total}} \leq P_{\mathrm{R}}^{\max}, \tag{8c}$$

$$|\theta_n| \leq 1, \forall n = 1, 2, \cdots, N, \tag{8d}$$

$$\gamma_{\mathrm{R},k} \geq \gamma_{\mathrm{R}}^{\mathrm{th}}, \forall k = 1, 2, \cdots, K. \tag{8e}$$

where the optimization variables and objective function are given in (8a), $\{\mathbf{g}_k\}$ and $\{\mathbf{f}_k\}$ are the sets of the BS beamforming and the relay beamforming vectors for all users, respectively. Constraints (8b) and (8c) denote the available transmit power at the BS and the DF relay, respectively. Constraint (8d) is the reflection constraint for an ideal IRS. Constraint (8e) ensures the decoding SINR requirements for all users' signals at the relay are satisfied.

Solving the problem in (8) is a challenging task, as the optimization variables are highly coupled, rendering a non-convex problem. To tackle the problem, we propose an alternating optimization (AO)-based algorithm to decouple (8) into three subproblems, as detailed in the next section.

III. AO-BASED BEAMFORMING CO-DESIGN

In this section, we elaborate on the three subproblems in the proposed AO-based algorithm: BS beamforming, relay beamforming, and IRS beamforming. The entire AO algorithm runs through optimizing the BS beamforming (in Sec. III-A), relay beamforming (in Sec. III-B), and IRS beamforming (in Sec. III-C), in an alternate fashion.

A. BS Beamforming Optimization

Given fixed relay beamforming vectors $\{\mathbf{f}_k\}$ and IRS beamforming matrix $\boldsymbol{\Theta}$, the problem in (8) becomes

$$\max_{\{\mathbf{g}_k\}} \quad \sum_{k=1}^{K}\log_2\left(C_{1,k}+\gamma_k^{\mathrm{I}}\right) \tag{9a}$$

$$\text{s.t.} \quad P_{\mathrm{BS}}^{\mathrm{total}} \leq P_{\mathrm{BS}}^{\max}, \tag{9b}$$

$$\gamma_{\mathrm{R},k} \geq \gamma_{\mathrm{R}}^{\mathrm{th}}, k = 1, \ldots, K, \tag{9c}$$

where $C_{1,k} = 1 + \gamma_k^{\mathrm{II}}$, $k = 1, \ldots, K$ are considered as constants under the AO-based decomposition. The subproblem (9) is still non-convex and difficult to solve. In this subsection, we tackle the subproblem in (9) by first transforming the problem of finding $\mathbf{g}_k$ into one of finding the rank-one positive semi-definite matrix $\mathbf{G}_k = \mathbf{g}_k\mathbf{g}_k^{\mathrm{H}}$ and then by relaxing the rank-one constraint, which eventually leads to a convex semi-definite 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) semi-definite 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 $\{\mathcal{I}_{\mathrm{R},k}\}$ that satisfy

$$\frac{1}{\mathcal{S}_{1,k}} \triangleq \left|\mathbf{h}'_{\mathrm{BS},k}\mathbf{g}_k\right|^2 \leq \mathrm{tr}(\mathbf{G}_k\mathbf{h}''_{\mathrm{B},k}), \tag{10}$$

$$\mathcal{I}_{1,k} \triangleq \sum_{j\neq k}\left|\mathbf{h}'_{\mathrm{BS},k}\mathbf{g}_j\right|^2 + \sigma_k^2 \geq \sum_{j\neq k}\mathrm{tr}(\mathbf{G}_j\mathbf{h}''_{\mathrm{B},k}) + \sigma_k^2, \tag{11}$$

$$\frac{1}{\mathcal{S}_{\mathrm{R},k}} \triangleq \|\boldsymbol{\alpha}_k\|^4 \leq [\mathrm{tr}(\mathbf{G}_k\mathbf{H}''_{\mathrm{BS},\mathrm{R}})]^2, \tag{12}$$

$$\mathcal{I}_{\mathrm{R},k} \triangleq \sum_{j\neq k}\boldsymbol{\alpha}_k^{\mathrm{H}}\boldsymbol{\alpha}_j\boldsymbol{\alpha}_j^{\mathrm{H}}\boldsymbol{\alpha}_k + \sigma_{\mathrm{R}}^2\|\boldsymbol{\alpha}_k\|^2$$
$$\geq \sum_{j\neq k}\mathrm{tr}(\mathbf{G}_k\mathbf{H}''_{\mathrm{B},\mathrm{R}}\mathbf{G}_j\mathbf{H}''_{\mathrm{B},\mathrm{R}}) + \sigma_{\mathrm{R}}^2\mathrm{tr}(\mathbf{G}_k\mathbf{H}''_{\mathrm{B},\mathrm{R}}) \tag{13}$$

where $\mathbf{H}''_{\mathrm{B},\mathrm{R}} \triangleq \mathbf{H}'^{\mathrm{H}}_{\mathrm{BS},\mathrm{R}}\mathbf{H}'_{\mathrm{BS},\mathrm{R}}$ and $\mathbf{h}''_{\mathrm{B},k} \triangleq \mathbf{h}'^{\mathrm{H}}_{\mathrm{BS},k}\mathbf{h}'_{\mathrm{BS},k}$. Incorporating these slack variables into (9) gives rise to the following equivalent problem

$$\max \quad \sum_{k=1}^{K} R_{1,k} \tag{14a}$$

variables: $\{\mathbf{G}_k\}, \{R_{1,k}\}, \{\mathcal{S}_{1,k}\}, \{\mathcal{I}_{1,k}\}, \{\mathcal{S}_{\mathrm{R},k}\}, \{\mathcal{I}_{\mathrm{R},k}\}$

$$\text{s.t.} \quad R_{1,k} \leq \log_2\left(C_{k,1}+\frac{1}{\mathcal{S}_{1,k}\mathcal{I}_{1,k}}\right) \triangleq u_1(\mathbf{z}), \tag{14b}$$

$$v_1(\mathbf{z}_{\mathrm{R}}) \triangleq \frac{1}{\mathcal{S}_{\mathrm{R},k}\mathcal{I}_{\mathrm{R},k}} \geq \gamma_{\mathrm{R}}^{\mathrm{th}}, \tag{14c}$$

$$(8b), (10), (11), (12), (13), \tag{14d}$$

$$\mathbf{G}_k \succeq 0, \quad \mathrm{rank}\,(\mathbf{G}_k) = 1, \tag{14e}$$

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

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 $\mathbf{z}_{\mathrm{R}} = [\mathcal{S}_{\mathrm{R},k}, \mathcal{I}_{\mathrm{R},k}]^{\mathrm{T}}$ defined in (14b) and (14c), respectively, are for notational convenience and clarity. Note that the optimal solution of (14) meets the constraints (10)–(13) in (14d) with equality. The subproblem in (14) is still non-convex.

2) First-Order Taylor Approximation: To convexify the constraints (14b) and (14c), 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

$$u_1(\mathbf{z}) \geq u_1(\mathbf{z}^{\mathrm{loc}}) + \nabla u_1(\mathbf{z}^{\mathrm{loc}})^{\mathrm{T}}(\mathbf{z} - \mathbf{z}^{\mathrm{loc}}) \triangleq R_{1,k}^{\mathrm{low}} \tag{15}$$

$$v_1(\mathbf{z}_{\mathrm{R}}) \geq v_1(\mathbf{z}_{\mathrm{R}}^{\mathrm{loc}}) + \nabla v_1(\mathbf{z}_{\mathrm{R}}^{\mathrm{loc}})^{\mathrm{T}}(\mathbf{z}_{\mathrm{R}} - \mathbf{z}_{\mathrm{R}}^{\mathrm{loc}}) \triangleq \gamma_{\mathrm{R},k}^{\mathrm{low}}, \tag{16}$$

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 $R_{1,k}^{\mathrm{low}}$ and $\gamma_{\mathrm{R},k}^{\mathrm{low}}$ can be carried out in closed-form from the gradients $\nabla u_1(\mathbf{z}^{\mathrm{loc}})$ and $\nabla v_1(\mathbf{z}_{\mathrm{R}}^{\mathrm{loc}})$, respectively. Replacing $u_1(\mathbf{z})$ in (14b) and $v_1(\mathbf{z}_{\mathrm{R}})$ in (14c) with their respective lower bounds in (15) and (16) linearizes and convexifies the constraints (14b) and (14c), converting the problem in (14) to

$$\max \quad \sum_{k=1}^{K} R_{1,k} \tag{17a}$$

variables: $\quad \{\mathbf{G}_k\}, \{R_{1,k}\}, \{\mathcal{S}_{1,k}\}, \{\mathcal{I}_{1,k}\}, \{\mathcal{S}_{\mathrm{R},k}\}, \{\mathcal{I}_{\mathrm{R},k}\}$

$$\text{s.t.} \quad R_{1,k} \le R_{1,k}^{\mathrm{low}}, \ \gamma_{\mathrm{R},k}^{\mathrm{low}} \ge \gamma_{\mathrm{R}}^{\mathrm{th}}, \ (14\mathrm{d}), (14\mathrm{e}), \tag{17b}$$
$$\forall k = 1, 2, \cdots, K.$$

Note that the solution set of (17) is a subset of that of the original problem in (14), as lower bounds are used as substitutes in the constraints (14b) and (14c).

3) Alternating Optimization: The interleaving presence of BS beamforming matrices $\mathbf{G}_k$ and $\mathbf{G}_j$ still render the constraint (13) in (14d) 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 $\mathbf{G}_j$, $j \ne k'$ as constants. As such, constraint (13) in (14d) becomes convex.

4) Semi-definite Relaxation: Finally, we relax the rank-one constraint in (14e) which, after combining with the 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 [14]. 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 [15] based on the eigen-decomposition of the optimal $\mathbf{G}_k$.

B. Relay Beamforming Optimization

Given fixed $\{\mathbf{g}_k\}$ (obtained from the steps in Sec. III-A) and $\mathbf{\Theta}$, the problem of solving the relay beamforming $\{\mathbf{f}_k\}$ in (8) becomes

$$\max_{\{\mathbf{f}_k\}} \quad \sum_{k=1}^{K} \log_2 \left(C_{2,k} + \gamma_k^{\mathrm{II}} \right) \tag{18a}$$

$$\text{s.t.} \quad P_{\mathrm{R}}^{\mathrm{total}} \le P_{\mathrm{R}}^{\mathrm{max}}. \tag{18b}$$

where $C_{2,k} = 1 + \gamma_k^{\mathrm{I}}, k = 1, 2, \cdots, K$, are kept as constants. The subproblem for finding the relay bemforming vector $\mathbf{f}_k$ in (18) is non-convex. Following similar steps developed in Sec. III-A, we elaborate on how to convexify (18) below.

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

$$\frac{1}{\mathcal{S}_{2,k}} \triangleq \left| \mathbf{h}_{\mathrm{R},k}' \mathbf{f}_k \right|^2 \le \mathrm{tr} \left(\mathbf{F}_k \mathbf{h}_{\mathrm{R},k}'' \right) \tag{19}$$

$$\mathcal{I}_{2,k} \triangleq \sum_{j \ne k} \left| \mathbf{h}_{\mathrm{R},k}' \mathbf{f}_j \right|^2 + \sigma_k^2 \ge \sum_{j \ne k} \mathrm{tr} \left(\mathbf{F}_j \mathbf{h}_{\mathrm{R},k}'' \right) + \sigma_k^2, \tag{20}$$

where $\mathbf{h}_{\mathrm{R},k}'' \triangleq \mathbf{h}_{\mathrm{R},k}'^{\,\mathrm{H}} \mathbf{h}_{\mathrm{R},k}'$. With the above slack variables, the subproblem (18) can be equivalently rewritten as

$$\max \quad \sum_{k=1}^{K} R_{2,k} \tag{21a}$$

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

$$\text{s.t.} \quad R_{2,k} \le \log_2 \left(C_{2,k} + \frac{1}{\mathcal{S}_{2,k} \mathcal{I}_{2,k}} \right) \triangleq u_2(\mathbf{z}), \tag{21b}$$

$$(8\mathrm{c}), (19), (20), \ \mathbf{F}_k \succeq 0, \ \mathrm{rank}(\mathbf{F}_k) = 1, \tag{21c}$$
$$\forall k = 1, 2, \cdots, K,$$

where $u_2(\mathbf{z})$ is for notational convenience. Similarly, we apply Taylor approximation on $u_2(\mathbf{z})$ and obtain a lower bound $R_{2,k}^{\mathrm{low}}$, which has closed-form expression from $\nabla u_2(\mathbf{z}^{\mathrm{loc}})$. Then, (21) can be converted to

$$\max \quad \sum_{k=1}^{K} R_{2,k} \tag{22a}$$

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

$$\text{s.t.} \quad R_{2,k} \le R_{2,k}^{\mathrm{low}}, \quad (21\mathrm{c}), \tag{22b}$$
$$\forall k = 1, \cdots, K$$

The solution set of (22) is a subset of the solution set of (21), since the lower bound $R_{2,k}^{\mathrm{low}}$ is used in the constraint (21b). By relaxing the rank-one constraint in (21c), we can solve the resulting SDP problem with CVX. Finally, as in the case in Sec. III-A, the randomization procedure can be employed to obtain rank-one solutions for $\{\mathbf{f}_k\}$, $k = 1, \ldots, K$.

C. IRS Beamforming Optimization

Given fixed BS beamforming $\{\mathbf{g}_k\}$ and relay beamforming $\{\mathbf{f}_k\}$ (respectively obtained from the steps in Sec. III-A and Sec. III-B), the problem of solving the IRS beamforming matrix $\mathbf{\Theta}$ (or equivalently its vector form $\boldsymbol{\theta}$) in (8) becomes

$$\max_{\mathbf{\Theta}} \quad \sum_{k=1}^{K} \log_2 \left(1 + \gamma_k \right) \tag{23a}$$

$$\text{s.t.} \quad |\theta_n| \le 1, \forall n = 1, \ldots, N, \tag{23b}$$

$$\gamma_{\mathrm{R},k} \ge \gamma_{\mathrm{R}}^{\mathrm{th}}, \forall k = 1, \ldots, K. \tag{23c}$$

We adopt similar steps as in Secs. III-A and III-B to solve (23). To facilitate finding the optimal IRS beamforming matrix $\mathbf{\Theta}$, we re-arrange the order of $\mathbf{\Theta}$ in (1) and in (5) respectively to $y_k^{\mathrm{I}} = \boldsymbol{\phi}^{\mathrm{H}} \mathbf{H}_{\mathrm{B},\mathrm{I},k} \mathbf{x} + w_k^{\mathrm{I}}$ and $y_k^{\mathrm{II}} = \boldsymbol{\phi}^{\mathrm{H}} \mathbf{H}_{\mathrm{R},\mathrm{I},k} \mathbf{x}_{\mathrm{R}} + w_k^{\mathrm{II}}$, where $\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\theta} & 1 \end{bmatrix}^{\mathrm{H}}$, $\mathbf{H}_{\mathrm{B},\mathrm{I},k} = \begin{bmatrix} \mathrm{diag}(\mathbf{h}_{\mathrm{IRS},k}) \mathbf{H}_{\mathrm{BS},\mathrm{IRS}} \\ \mathbf{h}_{\mathrm{BS},k} \end{bmatrix}$, and $\mathbf{H}_{\mathrm{R},\mathrm{I},k} = \begin{bmatrix} \mathrm{diag}(\mathbf{h}_{\mathrm{IRS},k}) \mathbf{H}_{\mathrm{R},\mathrm{IRS}} \\ \mathbf{h}_{\mathrm{R},k} \end{bmatrix}$. 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 (10), (11), (19), and (20) can be rewritten as

$$\frac{1}{\mathcal{S}_{1,k}} = \left|\boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{B,I},k}\mathbf{g}_k\right|^2 \leq \mathrm{tr}\left(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{B,I},k}\mathbf{g}_k\mathbf{g}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{B,I},k}^{\mathrm{H}}\right) \quad (24)$$

$$\mathcal{I}_{1,k} \geq \sum_{j\neq k}\mathrm{tr}\left(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{B,I},k}\mathbf{g}_j\mathbf{g}_j^{\mathrm{H}}\mathbf{H}_{\mathrm{B,I},k}^{\mathrm{H}}\right) + \sigma_k^2, \quad (25)$$

$$\frac{1}{\mathcal{S}_{2,k}} = \left|\boldsymbol{\phi}^{\mathrm{H}}\mathbf{H}_{\mathrm{R,I},k}\mathbf{f}_k\right|^2 \leq \mathrm{tr}(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{R,I},k}\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{R,I},k}^{\mathrm{H}}), \quad (26)$$

$$\mathcal{I}_{2,k} \geq \sum_{j=1,j\neq k}^{K}\mathrm{tr}\left(\boldsymbol{\Phi}\mathbf{H}_{\mathrm{R,I},k}\mathbf{f}_j\mathbf{f}_j^{\mathrm{H}}\mathbf{H}_{\mathrm{R,I},k}^{\mathrm{H}}\right) + \sigma_k^2, \quad (27)$$

where $\boldsymbol{\Phi} \triangleq \boldsymbol{\phi}\boldsymbol{\phi}^{\mathrm{H}}$. Similarly, we can re-arrange the order of $\boldsymbol{\Theta}$ in (3), by which the slack variables $\{\mathcal{S}_{\mathrm{R},k}\}$ and $\{\mathcal{I}_{\mathrm{R},k}\}$ in (12) and (13) can be rewritten in terms of $\boldsymbol{\Phi}$.

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

$$\max \quad \sum_{k=1}^{K} R_{3,k} \quad (28a)$$

$$\text{variables:} \quad \boldsymbol{\Phi}, \{R_{3,k}, \mathcal{S}_{1,k}, \mathcal{I}_{1,k}, \mathcal{S}_{2,k}, \mathcal{I}_{2,k}, \mathcal{S}_{\mathrm{R},k}, \mathcal{I}_{\mathrm{R},k}\}$$

$$\text{s.t.} \quad R_{3,k} \leq \log_2\left(1 + \frac{1}{\mathcal{S}_{1,k}\mathcal{I}_{1,k}} + \frac{1}{\mathcal{S}_{2,k}\mathcal{I}_{2,k}}\right)$$
$$\triangleq u_3(\mathbf{z}), \quad (28b)$$

$$v_3(\mathbf{z}_{\mathrm{R}}) \triangleq \frac{1}{\mathcal{S}_{\mathrm{R},k}\mathcal{I}_{\mathrm{R},k}} \geq \gamma_{\mathrm{R}}^{\mathrm{th}}, \quad (28c)$$

$$(8d), (12), (13), (24) - (27), \quad (28d)$$

$$\boldsymbol{\Phi} \succeq 0, \quad \mathrm{rank}\left(\boldsymbol{\Phi}\right) = 1, \quad (28e)$$
$$\forall k = 1, \cdots, K,$$

where $u_3(\mathbf{z})$ and $v_3(\mathbf{z}_{\mathrm{R}})$ respectively defined in (28b) and (28c) are for notational convenience. Next, using the first order Taylor approximation, as in the previous sections, lower bounds of $u_3(\mathbf{z})$ and $v_3(\mathbf{z}_{\mathrm{R}})$ can be obtained as $R_{3,k}^{\mathrm{low}}$ and $\gamma_{\mathrm{R},k}^{\mathrm{low}}$, which can be used to convexify the two constraints (28b) and (28c). Then, (28) can be rewritten as

$$\max \quad \sum_{k=1}^{K} R_{3,k} \quad (29a)$$

$$\text{variables:} \quad \boldsymbol{\Phi}, \{R_{3,k}, \mathcal{S}_{1,k}, \mathcal{I}_{1,k}, \mathcal{S}_{2,k}, \mathcal{I}_{2,k}, \mathcal{S}_{\mathrm{R},k}, \mathcal{I}_{\mathrm{R},k}\}$$

$$\text{s.t.} \quad R_{3,k} \leq R_{3,k}^{\mathrm{low}}, \quad \gamma_{\mathrm{R},k}^{\mathrm{low}} \geq \gamma_{\mathrm{R}}^{\mathrm{th}}, (28d), (28e), \quad (29b)$$
$$\forall k = 1, 2, \cdots, K.$$

Relaxing the rank-one constraint (28e) in (29b) converts the problem in (29) to an SDP, which can then be solved using CVX. Finally, Gaussian randomization procedure can be employed to obtain a rank-one solution.

IV. SIMULATION RESULTS AND DISCUSSION

A. Simulation Settings

We simulate the topology shown in Fig. 1 with the BS, IRS, and relay located at coordinates $(0,0)$, $(100,50)$, and $(100,-50)$, respectively, and the K users randomly distributed

within a circle centered at $(0, 200)$ with a radius of 10. The units of locations and radius are all meters. The small-scale fading of all channels follows $\mathcal{CN}(0,1)$. The parameters for large-scale fading of channels $\kappa(d/d_0)^{-\varrho}$ are set as follows: $d_0 = 1$ m; $\kappa = 10^{-4}$ for $\mathbf{H}_{\mathrm{BS,R}}$, $\mathbf{h}_{\mathrm{BS},k}$, and $\mathbf{h}_{\mathrm{R},k}$, and $\kappa = 10^{-0.5}$ for $\mathbf{H}_{\mathrm{BS,IRS}}$, $\mathbf{H}_{\mathrm{R,IRS}}$, and $\mathbf{h}_{\mathrm{IRS},k}$; and $\varrho = 2$ for IRS-aided links and relay-aided links (free space), and $\varrho = 3.5$ for the direct link (severe path loss). The power constraints at the BS and relay are $P_{\mathrm{BS}}^{\mathrm{max}} = 10$ mW and $P_{\mathrm{R}}^{\mathrm{max}} = 10$ mW, respectively. The proposed coexistence and co-design scheme is compared with the following benchmarks: 1) *Relay only:* there is no IRS; only the relay assists transmission with BS and relay beamforming optimized; 2) *Random:* IRS beamforming adopts random phase shifts and fixed amplitudes of one, with BS and relay beamforming optimized; and 3) *Independent:* IRS assists transmission in the first phase but not in the second phase (turned off), with BS, relay, and IRS beamforming optimized in such case. Note that it is termed 'Independent' since there is no interaction between IRS and relay in assisting BS transmission.

B. Results and Discussion

Fig. 2(a) compares the sum-rate performance vs. the number of IRS reflecting elements N, where we set $M = 8$, $L = 4$, and $K = 4$. As can be seen, the sum-rate of all schemes except the Relay only scheme increases as N increases. The proposed scheme outperforms all benchmarks since it fully exploits the coexistence system with optimized beamforming and transmission mechanism. The Independent scheme exhibits performance loss in comparison to the proposed scheme since IRS is turned off in the second phase and thus relaying in the second phase does not acquire performance gain from IRS-enhanced links and IRS beamforming. This suggests that the interaction between IRS and DF relay benefits the system sum-rate. Comparing the Independent and Random schemes, while the Independent scheme only leverages IRS in the first phase, it still outperforms the Random scheme with IRS operating in both phases. This is because, while the Independent scheme achieves a lower SINR in the second phase as compared to the Random scheme, it provides a much higher SINR in the first phase due to judicious IRS beamforming, resulting in an overall higher sum-rate. The Random scheme outperforms the Relay only scheme since IRS, even with random phase shifts, provides additional paths for transmission.

Fig. 2(b) compares the sum-rate performance vs. the number of relay antennas L, where we set $M = 8$, $N = 30$, and $K = 4$. The sum-rate of all schemes increases with L due to a higher degree of freedom in the second phase. The proposed scheme achieves the best sum-rate performance in all antenna configurations, with diminishing gains as L increases. This is because the existence and optimization of IRS with $N = 30$ reflecting elements dominates the change in L in terms of the system capacity in the second phase. As a result, increasing the number of relay antennas brings relatively small merits to the proposed scheme. The Independent and Random schemes exhibit a crossing point; specifically, the Random
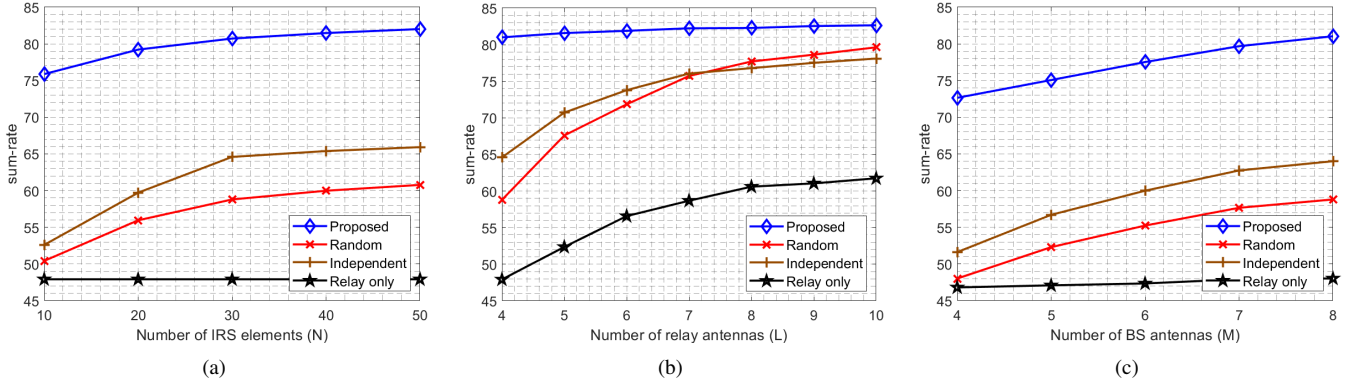
Fig. 2. (a)–(c) Sum-rate vs. the number of IRS reflecting elements, relay antennas, and BS antennas, respectively.

scheme outperforms the Independent scheme in the larger L regime. The reason is that the DF relay with more antennas can better adjust the relay beamforming vectors $\{\mathbf{f}_k\}$ in (6) for the combined IRS-assisted and direct channels from the relay to end users in the Random scheme. In contrast, since the IRS is turned off in the second phase, a relatively small gain is created for the Independent scheme as L increases. More specifically, the higher SINR in the second phase outweighs the lower SINR in the first phase for the Random scheme, as compared to the Independent scheme.

Fig. 2(c) plots the sum-rate performance vs. the number of BS antennas M, where we set $L = 4$, $N = 30$, and $K = 4$. The performance of all schemes improves as M increases. Increasing M however creates an incremental gain for the Relay only scheme due to the weak direct link from the BS to the end users. Other schemes, in contrast, are supported by the combined IRS-assisted and direct links in the first phase instead of only the direct link, and therefore observe larger gains as M increases. The Independent scheme outperforms the Random scheme because the Independent scheme is enhanced by IRS beamforming in the first phase. Without optimized IRS phase shifts, the Random scheme suffers from performance loss. Especially, in the small M region, the Random scheme achieves only a small performance gain as compared to the Relay only scheme. This suggests that the coexistence of an IRS and a relay does not always significantly benefit communications unless a proper co-design is performed.

V. CONCLUSION

We have considered a multiuser MISO system in which an IRS and a DF relay coexist and assist downlink transmission simultaneously. We investigated a co-design problem of BS beamforming, relay beamforming, and IRS beamforming for sum-rate maximization. To tackle the complex non-convex co-design problem, we proposed an algorithm based on alternating optimization and semidefinite relaxation. Simulation results and discussion on our proposed scheme and other benchmarks were presented. It was shown that our proposed algorithm significantly increases the sum-rate of end users.

REFERENCES

[1] 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.

[2] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[3] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *Proc. IEEE GLOBECOM*, 2018, pp. 1–6.

[4] C. Huang, A. Zappone, M. Debbah, and C. Yuen, "Achievable rate maximization by passive intelligent mirrors," in *Proc. IEEE ICASSP*, 2018, pp. 3714–3718.

[5] 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.

[6] H. Li, R. Liu, M. Liy, Q. Liu, and X. Li, "IRS-enhanced wideband MU-MISO-OFDM communication systems," in *Proc. IEEE WCNC*, 2020, pp. 1–6.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov, and S. Shamai, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, 2020.

[14] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[15] 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.