

# Joint Optimization in Bidirectional Multi-User Multi-Relay MIMO Systems: Non-Robust and Robust Cases

Meng Zhang, Haikē Yi, Hui Yu, Hanwen Luo, and Wen Chen, *Senior Member, IEEE*

**Abstract**—This paper studies the precoder design of bidirectional networks in which each node is equipped with multiple antennas. In contrast to the conventional model of bidirectional relay networks, this paper considers a more general scenario, i.e., a multi-user multi-relay network. First, by assuming that perfect channel state information (CSI) is available, we investigate the precoding design for the relays and users to minimize the sum mean squared error (MSE) and the maximum of single user's MSE, respectively. Then, we consider a more practical scenario where CSI estimation error is taken into account. By means of alternating optimization approach, we decompose the main problem into several decoupled subproblems with tractable solutions. It is shown that, in both the perfect and imperfect CSI cases, the proposed precoding algorithms outperform the existing solutions in terms of MSE and bit-error-rate (BER) performance.

**Index Terms**—Bidirectional, multi-user, multi-relay, precoding, robust.

## I. INTRODUCTION

WITH the continuous burgeoning development of wireless communications technologies, the pursuit for a system with higher performance has never stopped. Initially, people focused on the point-to-point system, and a great deal of remarkable results have been obtained for the single transceiver case [1], [2]. Later, a relay node was introduced into the network with the purpose to enhance the coverage and reliability of the system. However, due to the infeasibility of full-duplex transmission at the relay, most of the relay-based strategies claim their gains at the cost of consuming more time or frequency resources under half-duplex relaying protocols, e.g., a typical two-hop unidirectional amplify-and-forward relay network has to complete one end-to-end transmission with two phases [3]. To increase the spectrum efficiency, a

network coding strategy for bidirectional relay networks has been proposed in [4].

This paper investigates the precoding design for the relays and users in a bidirectional multi-user multi-relay scenario. Since the number of wireless communication users will continuously increase, a single-pair bidirectional relay system can hardly accommodate a vast number of users. Meanwhile, various criteria have been proposed to optimize the system performance, such as minimizing the sum mean squared error (MSE), the maximum user's MSE, the minimum power that can support a certain level of quality-of-service (QoS) requirements, etc. In this paper, we concentrate on the sum MSE criterion as well as maximum user's MSE.

Previous work on MSE optimization can be dated back to the work in [2], where separate designs of receivers and transmitters and their joint design have been investigated. A closed-form MSE-optimal solution for a multiple-input-multiple-output (MIMO) relay network has been obtained in [5], where a diagonal structure has been proven optimal. However, it turns out that the straightforward extension is not optimal in a bidirectional relay scenario [6]. Even so, a suboptimal closed-form solution is still preferred and has been given in [7]. It is worth noting that another relaying protocol named two-path successive relay has been proposed in [8]. This protocol is also designed to increase spectrum efficiency by combining two half-duplex relays into one full-duplex node by means of alternating relaying. However, the drawback of this scheme is obvious since an additional relay has to be deployed, and the associated cost cannot be easily neglected. Additionally, the authors take the direct link from a source to a destination into account in [9] and [10]. In [11], a bidirectional multi-relay network has been studied, and an iterative algorithm is shown to be effective. In addition, the authors investigate sum MSE optimization for the case of multi-relay single-user pair, and gradient-based algorithms have been proposed in [12]. In addition, precoding schemes have been proposed in [13], with an assumption that a base station will exchange messages with multiple single-antenna users through the help of a bidirectional relay. Our previous work in [14] has proposed a joint precoding design for bidirectional multi-user single-relay networks.

Moreover, because perfect channel state information (CSI) is not always available, it is necessary to investigate the relevant precoding designs under imperfect CSI, i.e., a robust precoding scheme. Previous works on the robust precoding for MIMO relay networks fall into two categories: stochastic analysis and

Manuscript received August 28, 2012; revised December 29, 2012 and March 7, 2013; accepted March 10, 2013. Date of publication March 29, 2013; date of current version September 11, 2013. This paper was supported in part by the Shanghai Basic Research Key Project under Grant 11DZ1500206, by the National Key Project of China under Grant 2013ZX03001007-004, by the National 973 Project under Grant 2012CB316106, and by the National Science Foundation of China under Grant 61161130529. This paper was presented in part at the IEEE Global Communications Conference, Anaheim, CA, USA, December 2012. The review of this paper was coordinated by Prof. H. H. Nguyen.

The authors are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: mengzhang@sjtu.edu.cn; yihaike@sjtu.edu.cn; yuhui@sjtu.edu.cn; hwl@sjtu.edu.cn; wenchen@sjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2013.2255898

worst-case analysis. In [15], the precoder structure is analyzed with the existence of stochastic channel estimation error for the MIMO relay networks. In addition, for the MIMO relaying broadcast scenario, the precoding schemes are analyzed under stochastic channel estimation error model in [16]. The capacity performance of a dual-hop MIMO relay network is analyzed in [17]. The robust precoder design for the worst case is proposed in [18], where a multi-antenna user with one multi-antenna relay model is considered. Furthermore, in [19], a robust precoder design is proposed, aiming to minimize the negative effect of stochastic error for bidirectional relay systems.

In this paper, we consider a more general case that involves multiple bidirectional relays and multiple user pairs with individual power constraints applied on each relay and each user. To minimize the MSE, we propose precoding schemes for both perfect and imperfect CSI scenarios. Note that our earlier work in [20] only focuses on sum MSE minimization; in this paper, we further develop the optimization algorithm for minimizing maximum user's MSE under both the perfect and imperfect cases. However, the optimal solutions are hard to obtain since, in [6], the optimal precoder's structure has been proven non-diagonal optimal. Therefore, iterative algorithms are proposed instead. The proposed method has the potential to be applied to large networks [21], [22]. Simulation results show the effectiveness of our proposed algorithms.

The rest of this paper is organized as follows. The system model and problem formulation are described in Section II. The joint design for perfect channel knowledge is proposed in Section III, where minimizing the sum MSE and maximum user's MSE are investigated, respectively. Similarly, we discuss the joint design under imperfect channel knowledge in Section IV. The numerical results are presented in Section V. Finally, the conclusion is made in Section VI.

**Notation:** In this paper, we use bold uppercase and lowercase letters to denote matrices and vectors, respectively.  $(\cdot)^*$ ,  $(\cdot)^T$ ,  $(\cdot)^H$ , and  $(\cdot)^{-1}$  denote the conjugate, transpose, conjugate transpose, and inverse of a matrix or a vector, respectively.  $(\cdot)^{1/2}$  denotes the square root of a square matrix.  $\text{blkdiag}\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$  is the block diagonal matrix formed by matrices  $\mathbf{X}_i$ ,  $i = 1, 2, \dots, N$ .  $\mathbf{I}_N$  is an  $N \times N$  identity matrix.  $\mathbb{E}_{\mathbf{X}}\{\cdot\}$  denotes the statistical expectation over  $\mathbf{X}$ .  $\text{tr}(\cdot)$  and  $\text{Re}[\cdot]$  are the trace of a matrix and the real part of a variable, respectively.  $\text{vec}(\cdot)$  and  $\text{mat}\{\cdot\}$  represent the matrix vectorization and its inverse operation, respectively.  $\otimes$  denotes the Kronecker product.  $\|\cdot\|$  denotes the Frobenius norm.  $\succeq$  represents the property of semidefinite.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a network shown in Fig. 1 that consists of  $R$  relays and  $2K$  users or, equivalently,  $K$  user pairs that are expected to exchange data with the aid of  $R$  relays. Assume that the  $k$ -th user and the  $r$ -th relay are equipped with  $N_k$  and  $M_r$  antennas, respectively. It should be noted that the constraint of  $L_k \leq N_k$  is necessary, where  $L_k$  is the number of data streams transmitted by the  $k$ -th user. Let  $k$  be the index of the user paired with user  $k$ . Additionally,  $\mathbf{H}_{rk}$  denotes the channel between the  $k$ -th user and the  $r$ -th relay in the first phase, and  $\mathbf{G}_{kr}$

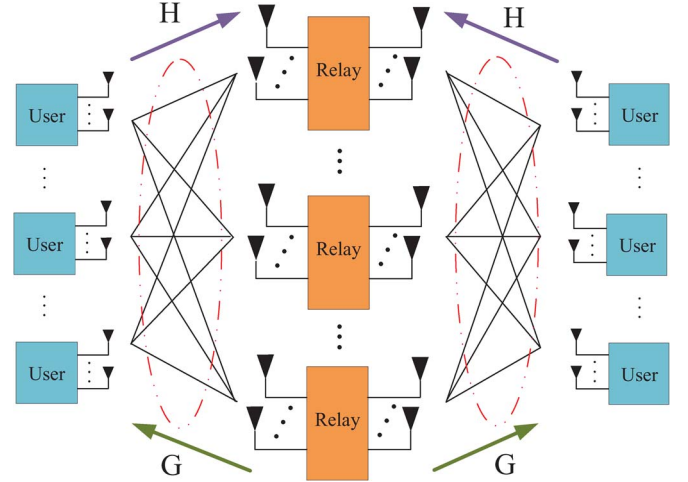


Fig. 1. Multi-user multi-relay bidirectional networks.

denotes the channel between the  $r$ -th relay and the  $k$ -th user in the second phase. The two-hop transmission processing is completed by two phases, i.e., in the first phase, users send their independent messages to the relays, whereas in the second phase, the relays will broadcast the precoded received signal to all users.

In the first phase, the  $k$ -th user transmits precoded signal  $\mathbf{s}_k = \mathbf{A}_k \mathbf{d}_k$ , where  $\mathbf{d}_k \in \mathbb{C}^{L_k \times 1}$  denotes the data information intended for the  $k$ -th user, and  $\mathbf{A}_k$  is an  $N_k \times L_k$  precoding matrix for user  $k$ . Without loss of generality, we assume that  $\mathbb{E}\{\mathbf{d}_k \mathbf{d}_k^H\} = \sigma_d^2 \mathbf{I}_{L_k}$ , where  $\mathbf{I}_{L_k}$  is an  $L_k \times L_k$  identity matrix. We assume that the transmit power at each user side satisfies

$$\sigma_d^2 \text{tr}(\mathbf{A}_k \mathbf{A}_k^H) \leq P_{Sk} \quad \forall k \in \{1, 2, \dots, 2K\} \quad (1)$$

where  $P_{Sk}$  is the power constraint for the  $k$ -th user. The received signal at the  $r$ -th relay is

$$\mathbf{x}_r = \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{A}_k \mathbf{d}_k + \mathbf{n}_r \quad (2)$$

where  $\mathbf{n}_r \in \mathbb{C}^{M_r \times 1}$  is a circular complex additive white Gaussian noise vector at the  $r$ -th relay node with  $\mathbb{E}\{\mathbf{n}_r \mathbf{n}_r^H\} = \sigma_r^2 \mathbf{I}_{M_r}$ .

After multiplying by relay precoding matrix  $\mathbf{F}_r \in \mathbb{C}^{M_r \times M_r}$ , all the relays broadcast their postprocessed signals at the beginning of the second phase, i.e.,

$$\mathbf{y}_r = \sum_{k=1}^{2K} \mathbf{F}_r \mathbf{H}_{rk} \mathbf{A}_k \mathbf{d}_k + \mathbf{F}_r \mathbf{n}_r. \quad (3)$$

The transmitting signal at each relay should not violate the individual relay power constraint, i.e.,

$$\begin{aligned} & \text{tr}(\mathbf{y}_r \mathbf{y}_r^H) \\ &= \text{tr} \left[ \mathbf{F}_r \left( \sigma_d^2 \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{A}_k \mathbf{A}_k^H \mathbf{H}_{rk}^H + \sigma_r^2 \mathbf{I}_{M_r} \right) \mathbf{F}_r^H \right] \leq P_{Rr}. \end{aligned} \quad (4)$$

The received signal at the  $k$ -th user is expressed as

$$\mathbf{r}_k = \sum_{i=1}^{2K} \sum_{r=1}^R \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{ri} \mathbf{A}_i \mathbf{d}_i + \sum_{r=1}^R \mathbf{G}_{kr} \mathbf{F}_r \mathbf{n}_r + \mathbf{n}_k \quad (5)$$

where  $\mathbf{n}_k \in \mathbb{C}^{N_k \times 1}$  is a circular complex additive white Gaussian noise vector with  $\mathbb{E}\{\mathbf{n}_k \mathbf{n}_k^H\} = \sigma_x^2 \mathbf{I}_{N_k}$ .

The Kronecker model is adopted here for the wireless communication channel in which independent and separable correlations between transmit and receive antennas are presumed [23]. For the Rayleigh fading, channel matrices  $\mathbf{H}_{rk}$  and  $\mathbf{G}_{kr}$  can be modeled, respectively, as

$$\mathbf{H}_{rk} = \hat{\mathbf{H}}_{rk} + \Delta \mathbf{H}_{rk} = \Sigma_{H, rk}^{\frac{1}{2}} (\hat{\mathbf{H}}_{W, rk} + \Delta \mathbf{H}_{W, rk}) \Psi_{H, rk}^{\frac{1}{2}} \quad (6)$$

$$\mathbf{G}_{kr} = \hat{\mathbf{G}}_{kr} + \Delta \mathbf{G}_{kr} = \Sigma_{G, kr}^{\frac{1}{2}} (\hat{\mathbf{G}}_{W, kr} + \Delta \mathbf{G}_{W, kr}) \Psi_{G, kr}^{\frac{1}{2}} \quad (7)$$

where  $\hat{\mathbf{H}}_{rk}$  and  $\hat{\mathbf{G}}_{kr}$  are the estimated CSI, and  $\Delta \mathbf{H}_{rk}$  and  $\Delta \mathbf{G}_{kr}$  correspond to the CSI error. It is worth noting that  $\mathbf{H}_{rk} = \hat{\mathbf{H}}_{rk}$  and  $\mathbf{G}_{kr} = \hat{\mathbf{G}}_{kr}$  if perfect channel knowledge is known at each node.  $\Sigma_{H, rk}$  and  $\Sigma_{G, kr}$  are the receive-antenna correlation matrices, and  $\Psi_{H, rk}$  and  $\Psi_{G, kr}$  are the transmit-antenna correlation matrices. Entries of matrices  $\hat{\mathbf{H}}_{W, rk}$ ,  $\Delta \mathbf{H}_{W, rk}$ ,  $\hat{\mathbf{G}}_{W, kr}$ , and  $\Delta \mathbf{G}_{W, kr}$  are independent and identically distributed complex Gaussian random variables with variances  $1 - \sigma_e^2$ ,  $\sigma_e^2$ ,  $1 - \sigma_e^2$ , and  $\sigma_e^2$ .

With the assumption that each user has the knowledge of all estimated CSI and precoder matrices, each user can remove the corresponding self-interference from the received data (5). Therefore, after this self-interference cancelation step, the received signal can be reformulated as

$$\hat{\mathbf{r}}_k = \sum_{i=1}^{2K} \sum_{r=1}^R \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{ri} \mathbf{A}_i \mathbf{d}_i + \sum_{r=1}^R \mathbf{G}_{kr} \mathbf{F}_r \mathbf{n}_r + \mathbf{n}_k - \sum_{r=1}^R \hat{\mathbf{G}}_{kr} \mathbf{F}_r \hat{\mathbf{H}}_{rk} \mathbf{A}_k \mathbf{d}_k. \quad (8)$$

Afterward, the MSE-optimal receiver will be applied to obtain the intended data, and the postprocessed signals can be

expressed as

$$\begin{aligned} \hat{\mathbf{d}}_k &= \mathbf{B}_k \hat{\mathbf{r}}_k \\ &= \sum_{i=1}^{2K} \sum_{r=1}^R \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{ri} \mathbf{A}_i \mathbf{d}_i + \sum_{r=1}^R \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{n}_r \\ &\quad + \mathbf{B}_k \mathbf{n}_k - \sum_{r=1}^R \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{F}_r \hat{\mathbf{H}}_{rk} \mathbf{A}_k \mathbf{d}_k \end{aligned} \quad (9)$$

where  $\mathbf{B}_k$  is the MSE-optimal receiver for the  $k$ -th user. If the transmitting precoding matrices are all fixed,  $\mathbf{B}_k$  is also determined. Consequently, we will calculate the expression for  $\mathbf{B}_k$  first after giving the expression of individual user's MSE. By omitting the terms that involve a high order of channel uncertainties, the MSE for the  $k$ -th user can be expressed as (10) shown below. By setting  $\Delta \mathbf{H}_{rk} = 0$  and  $\Delta \mathbf{G}_{kr} = 0$  in (10), we can obtain the MSE for the perfect CSI case.

For the algorithms proposed in Sections III and IV, we adopt an alternating optimization method that iteratively optimizes the receivers filters, relay precoders, and user precoders. We also assume that there exists a central processor that can collect all the uplink and downlink CSI, and it has the second order of statistical knowledge on estimation errors. All the computations of receiver filters, relay precoders, and user precoders are executed at this central processor. After computations, the corresponding results will be sent to each node through control channels. In addition, the optimal solutions are obtained in each subproblem because the other convex-violating parameters are set to be fixed. Furthermore, the sum MSE is definitely to be non-negative or, namely, lower bounded by a fixed value. Therefore, the proposed iterative algorithms for non-robust and robust cases are sure to be convergent.

### III. JOINT DESIGN FOR THE PERFECT CHANNEL KNOWLEDGE

Considering that a diagonal structure is not optimal in bidirectional systems [6] and that, up to now, the optimal closed-form precoding scheme has not been found, we resort to the

$$\begin{aligned} \varepsilon_k &= E \left\{ \left\| \hat{\mathbf{d}}_k - \mathbf{d}_k \right\|^2 \right\} \\ &= \sigma_d^2 \sum_{i=1, i \neq k, \bar{k}}^{2K} \text{tr} \left[ \left( \sum_{r=1}^R \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{ri} \mathbf{A}_i \right)^H \left( \sum_{r=1}^R \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{ri} \mathbf{A}_i \right) \right] \\ &\quad + \sigma_x^2 \text{tr} (\mathbf{B}_k^H \mathbf{B}_k) + \sigma_d^2 \sum_{r=1}^R \text{tr} \left[ (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \Delta \mathbf{H}_{rk} \mathbf{A}_k)^H (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \Delta \mathbf{H}_{rk} \mathbf{A}_k) \right] \\ &\quad + \sigma_d^2 \sum_{r=1}^R \text{tr} \left[ (\mathbf{B}_k \Delta \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{rk} \mathbf{A}_k)^H (\mathbf{B}_k \Delta \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{rk} \mathbf{A}_k) \right] \\ &\quad + \sigma_r^2 \sum_{r=1}^R \text{tr} \left[ (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r)^H (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r) \right] \\ &\quad + \sigma_d^2 \text{tr} \left[ \left( \sum_{r=1}^R \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{r\bar{k}} \mathbf{A}_{\bar{k}} - \mathbf{I}_{L_{\bar{k}}} \right)^H \left( \sum_{r=1}^R \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{r\bar{k}} \mathbf{A}_{\bar{k}} - \mathbf{I}_{L_{\bar{k}}} \right) \right] \end{aligned} \quad (10)$$

precoding design that optimizes the receivers, relay precoders, and user precoders, iteratively. Here, we only consider the scenario that each node has perfect knowledge of all the channels. Then, in Section IV, for the imperfect CSI case, we present an iterative algorithm consisting of three subproblems for the design of receiver filters, relay precoders, and user precoders, respectively.

With the perfect CSI, the sum MSE optimization problem can be written as

$$\begin{aligned} \min_{\mathbf{B}_k, \mathbf{F}_r, \mathbf{A}_k} \quad & \varepsilon = \sum_{k=1}^{2K} \varepsilon_k \\ \text{s.t.} \quad & \sigma_d^2 \text{tr}(\mathbf{A}_k \mathbf{A}_k^H) \leq P_{Sk} \quad \forall k \in \{1, 2, \dots, 2K\} \\ & \text{tr} \left[ \mathbf{F}_r \left( \sigma_d^2 \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{A}_k \mathbf{A}_k^H \mathbf{H}_{rk}^H + \sigma_r^2 \mathbf{I}_{M_r} \right) \mathbf{F}_r^H \right] \leq P_{Rr} \\ & \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (11)$$

where  $\varepsilon_k$  is the individual user's MSE expression shown in (10) by setting  $\Delta \mathbf{H}_{rk} = \Delta \mathbf{G}_{kr} = 0$ .

#### A. Sum MSE Optimization Algorithm

1) *Receiver Design*: In this paper, we adopt an MMSE receiver at each user node, which is also called the Wiener filter. Given the expressions of  $(\partial \varepsilon)/(\partial \mathbf{B}_k^*) = (\partial \varepsilon_k)/(\partial \mathbf{B}_k^*) = 0$ , the optimal MSE receiver can be obtained by setting the derivative of MSE with respect to  $\mathbf{B}_k^*$  to zero. We have

$$\mathbf{B}_k = \sigma_d^2 \sum_{r=1}^R \mathbf{A}_k^H \mathbf{H}_{rk}^H \mathbf{F}_r^H \mathbf{G}_{kr}^H \mathbf{R}_w^{-1} \quad (12)$$

where

$$\begin{aligned} \mathbf{R}_w = \sigma_d^2 \sum_{i=1, i \neq k}^{2K} \sum_{r=1}^R \sum_{s=1}^R \mathbf{G}_{ks} \mathbf{F}_s \mathbf{H}_{si} \mathbf{A}_i \mathbf{A}_i^H \mathbf{H}_{ri}^H \mathbf{F}_r^H \mathbf{G}_{kr}^H \\ + \sigma_r^2 \sum_{r=1}^R \mathbf{G}_{kr} \mathbf{F}_r \mathbf{F}_r^H \mathbf{G}_{kr}^H + \sigma_x^2 \mathbf{I}_{N_k}. \end{aligned} \quad (13)$$

2) *Relay Precoder Design*: Next, given the prefixed MMSE receiver matrices and user precoders, we will design the relay precoders  $\mathbf{F}_r$  ( $\forall r = 1, 2, \dots, R$ ). However, the issue is that jointly optimizing all the relay precoders is nonconvex according to the Hessian matrix derivation. Therefore, by fixing the other relay precoders except for the  $r$ -th relay, this sum MSE optimization problem is convex. By omitting the parts irrelevant to  $\mathbf{F}_r$ , the sum MSE can be reformulated as follows:

$$\begin{aligned} \sum_{k=1}^{2K} \varepsilon_k = \sigma_d^2 \sum_{k=1}^{2K} \sum_{i=1, i \neq k}^{2K} \text{tr} \\ \left( \mathbf{A}_i^H \mathbf{H}_{ri}^H \mathbf{F}_r^H \mathbf{G}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{ri} \mathbf{A}_i \right) \\ + \sum_{k=1}^{2K} \sigma_r^2 \text{tr} \left( \mathbf{F}_r^H \mathbf{G}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \right) \end{aligned}$$

$$\begin{aligned} + \sigma_d^2 \sum_{k=1}^{2K} \sum_{i=1, i \neq k}^{2K} \text{tr} \left( \mathbf{A}_i^H \mathbf{H}_{ri}^H \mathbf{F}_r^H \mathbf{G}_{kr}^H \mathbf{B}_k^H \mathbf{U}_{ir} \right) \\ + \sigma_d^2 \sum_{k=1}^{2K} \sum_{i=1, i \neq k}^{2K} \text{tr} \left( \mathbf{U}_{ir}^H \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{ri} \mathbf{A}_i \right) \\ - \sigma_d^2 \sum_{k=1}^{2K} \text{tr} \left( \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{rk} \mathbf{A}_k \right) \\ - \sigma_d^2 \sum_{k=1}^{2K} \text{tr} \left( \mathbf{A}_k^H \mathbf{H}_{rk}^H \mathbf{F}_r^H \mathbf{G}_{kr}^H \mathbf{B}_k^H \right) \end{aligned} \quad (14)$$

where  $\mathbf{U}_{ir} = \sum_{s=1, s \neq r}^R \mathbf{B}_k \mathbf{G}_{ks} \mathbf{F}_s \mathbf{H}_{si} \mathbf{A}_i$ . Due to the convexity of this subproblem, the Lagrangian multiplier method can be utilized to derive the optimal solution. The Lagrangian function can be expressed as

$$\mathcal{L} = \sum_{k=1}^{2K} \varepsilon_k + \lambda [\text{tr}(\mathbf{F}_r^H \mathbf{R}_x \mathbf{F}_r) - P_{Rr}] \quad (15)$$

where  $\mathbf{R}_x = \sigma_d^2 \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{A}_k \mathbf{A}_k^H \mathbf{H}_{rk}^H + \sigma_r^2 \mathbf{I}_{M_r}$ .

By setting  $(\partial \mathcal{L})/(\partial \mathbf{F}_r^*) = 0$ , additionally with other Karush–Kuhn–Tucker conditions, we obtain

$$\lambda [\text{tr}(\mathbf{F}_r^H \mathbf{R}_x \mathbf{F}_r) - P_{Rr}] = 0 \quad (16)$$

$$\text{tr}(\mathbf{F}_r^H \mathbf{R}_x \mathbf{F}_r) \leq P_{Rr} \quad (17)$$

$$\lambda \geq 0. \quad (18)$$

Then, we have

$$\sum_{k=1}^{2K} \mathbf{R}_{yk} \mathbf{F}_r \mathbf{R}_{xk} + \lambda \mathbf{F}_r \mathbf{R}_x = \mathbf{R}_r. \quad (19)$$

Therefore, the form of optimal solution of the  $r$ -th relay precoder can be written as

$$\mathbf{F}_r^{\text{opt}} = \text{mat} \left\{ \left( \sum_{k=1}^{2K} \mathbf{R}_{xk}^T \otimes \mathbf{R}_{yk} + \lambda \mathbf{R}_x^T \otimes \mathbf{I}_{M_r} \right)^{-1} \text{vec}(\mathbf{R}_r) \right\} \quad (20)$$

where

$$\mathbf{R}_{xk} = \sigma_d^2 \sum_{i=1, i \neq k}^{2K} \mathbf{H}_{ri} \mathbf{A}_i \mathbf{A}_i^H \mathbf{H}_{ri}^H + \sigma_r^2 \mathbf{I}_{M_r} \quad (21)$$

$$\mathbf{R}_{yk} = \mathbf{G}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \mathbf{G}_{kr} \quad (22)$$

$$\begin{aligned} \mathbf{R}_r = \sigma_d^2 \sum_{k=1}^{2K} \left( \mathbf{G}_{kr}^H \mathbf{B}_k^H \mathbf{A}_k^H \mathbf{H}_{rk} \right. \\ \left. - \sum_{i=1, i \neq k}^{2K} \mathbf{G}_{kr}^H \mathbf{B}_k^H \mathbf{U}_{ir} \mathbf{A}_i^H \mathbf{H}_{ri}^H \right). \end{aligned} \quad (23)$$

Omitting the first term of (20),  $\lambda$  should be confined within the interval  $[0, \sqrt{(\text{tr}(\mathbf{R}_r (\mathbf{R}_x^T)^{-1} \mathbf{R}_x (\mathbf{R}_x^T)^{-1} \mathbf{R}_r^T) / P_{Rr})}]$ .



Therefore, a bisection search algorithm is applicable to find the optimal  $\lambda$  and  $\mathbf{F}_r^{\text{opt}}$ .

3) *User Precoder Design*: Here, the user precoders will be optimized with fixed MMSE receivers and relay precoders. Let  $P'_{Rr} = P_{Rr} - \sigma_r^2 \text{tr}(\mathbf{F}_r^H \mathbf{F}_r)$ . The main problem (11) can be modified as

$$\begin{aligned} \min_{\mathbf{A}_k} \quad & \sum_{k=1}^{2K} \text{tr}(\mathbf{R}_{ak} \mathbf{A}_k \mathbf{A}_k^H) \\ & - \sum_{k=1}^{2K} 2\text{Re}[\text{tr}(\mathbf{R}_{bk} \mathbf{A}_k)] + \text{tr}(\mathbf{R}_c) \\ \text{s.t.} \quad & \sigma_d^2 \text{tr}(\mathbf{A}_k^H \mathbf{A}_k) \leq P_{Sk} \quad \forall k \in \{1, 2, \dots, 2K\} \\ & \sum_{k=1}^{2K} \text{tr}(\mathbf{A}_k^H \mathbf{R}_{dkr} \mathbf{A}_k) \leq P'_{Rr} \quad \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (24)$$

where

$$\mathbf{R}_{bk} = \sigma_d^2 \sum_{r=1}^R \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{H}_{rk} \quad (25)$$

$$\begin{aligned} \mathbf{R}_{ak} = \sigma_d^2 \sum_{i=1, i \neq k}^{2K} \left[ \left( \sum_{r=1}^R \mathbf{B}_i \mathbf{G}_{ir} \mathbf{F}_r \mathbf{H}_{rk} \right)^H \right. \\ \left. \times \left( \sum_{r=1}^R \mathbf{B}_i \mathbf{G}_{ir} \mathbf{F}_r \mathbf{H}_{rk} \right) \right] \end{aligned} \quad (26)$$

$$\mathbf{R}_{dkr} = \sigma_d^2 \mathbf{H}_{rk}^H \mathbf{F}_r^H \mathbf{F}_r \mathbf{H}_{rk} \quad (27)$$

$$\begin{aligned} \mathbf{R}_c = \sum_{k=1}^{2K} \left( \sum_{r=1}^R \sigma_r^2 (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{F}_r^H \mathbf{G}_{kr}^H \mathbf{B}_k^H) \right. \\ \left. + \mathbf{I}_{L_k} + \sigma_x^2 (\mathbf{B}_k \mathbf{B}_k^H) \right). \end{aligned} \quad (28)$$

To solve problem (24), we use CVX, which is a software package for specifying and solving convex programs, since the modified problem is a typical quadratically constrained quadratic programming (QCQP) problem. The joint design of sum MSE minimization is summarized in Algorithm 1.

---

**Algorithm 1** Sum MSE optimization algorithm.

---

- 1: Initialization:
  - Set  $\mathbf{A}_k^{(1)} = \sqrt{(P_{Sk}/L_k)} [\mathbf{I}_{L_k} \mathbf{0}_{(N_k-L_k) \times L_k}]^T \quad \forall k \in \{1, 2, \dots, 2K\}$  and  $\mathbf{F}_r^{(1)} = \sqrt{(P_r/\text{tr}(\mathbf{R}_x))} \mathbf{I}_{M_r} \quad \forall r \in \{1, 2, \dots, R\}$ ,  $n = 1$ .
- 2: Iteration:
  - a) Compute  $\mathbf{B}_k^{(n)}$  using (12).
  - b) Compute  $\mathbf{F}_r^{(n)}$  using (20).
  - c) Compute  $\mathbf{A}_k^{(n)}$  by solving the QCQP problem of (24).
  - d) Compute sum MSE  $\varepsilon^{(n)}$ .
- 3: Termination:

The algorithm terminates either when  $\varepsilon^{(n)}$  converges, i.e.,  $|(\varepsilon^{(n)} - \varepsilon^{(n-1)})/\varepsilon^{(n)}| \leq \eta$ , or when  $n \geq N_{\max}$ , where  $\eta$  is a predefined threshold and  $N_{\max}$  is the maximum iteration number;

Output  $\mathbf{B}_k^{\text{opt}} = \mathbf{B}_k^{(n)}$ ,  $\mathbf{F}_r^{\text{opt}} = \mathbf{F}_r^{(n)}$ , and  $\mathbf{A}_k^{\text{opt}} = \mathbf{A}_k^{(n)}$ .  
Else,  $n = n + 1$ , and go to substep 2.

---

### B. Minmax MSE Optimization Algorithm

Here, we will consider minimizing the maximum user's MSE. If we attempt to minimize the sum MSE, the individual user's QoS cannot be guaranteed. Therefore, we will consider minimizing the maximum of users' MSE, which can ensure that the user's MSE is equivalent to each other at the optimal point.

First, we define  $\mathbf{F} = \text{blkdiag}\{\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_R\}$  and  $\mathbf{A} = \text{blkdiag}\{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_{2K}\}$ . Thus, it is easy to verify that  $\mathbf{F}_r = \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T$ , where  $\mathbf{P}_r$  is the  $M_r \times \sum_{r=1}^R M_r$  permutation matrix with zeros and ones. Similarly, we have  $\mathbf{A}_k = \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T$ , with  $\mathbf{Q}_k \in \mathbb{C}^{N_k \times \sum_{k=1}^{2K} N_k}$  and  $\mathbf{V}_k^T \in \mathbb{C}^{\sum_{k=1}^{2K} L_k \times L_k}$ . For given MMSE receivers, our optimization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{F}, \mathbf{A}} \quad & \max_{k \in \{1, 2, \dots, 2K\}} \varepsilon_k \\ \text{s.t.} \quad & \text{tr} \left[ \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \left( \sigma_d^2 \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \right. \right. \\ & \quad \left. \left. \times \mathbf{Q}_k^T \mathbf{H}_{rk}^H + \sigma_r^2 \mathbf{I}_{M_r} \right) \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T \right] \\ & \leq P_{Rr} \quad \forall r \in \{1, 2, \dots, R\}, \\ & \sigma_d^2 \text{tr}(\mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \mathbf{Q}_k^T) \leq P_{Sk} \quad \forall k \in \{1, 2, \dots, 2K\}. \end{aligned} \quad (29)$$

Then, we can rewrite the MSE expression of the  $k$ -th user, as shown in the following:

$$\begin{aligned} \varepsilon_k = \sigma_d^2 \left\| \sum_{r=1}^R \text{vec}(\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \mathbf{H}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T) - \text{vec}(\mathbf{I}_{L_k}) \right\|^2 \\ + \sigma_r^2 \sum_{r=1}^R \left\| \text{vec}(\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T) \right\|^2 + \sigma_x^2 \|\mathbf{B}_k\|^2 \\ + \sigma_d^2 \sum_{i=1, i \neq k, \bar{k}}^{2K} \left\| \sum_{r=1}^R \text{vec}(\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \mathbf{H}_{ri} \mathbf{Q}_i \mathbf{A} \mathbf{V}_i^T) \right\|^2. \end{aligned} \quad (30)$$

1) *Relay Precoder Design*: By introducing equation  $\text{vec}(\mathbf{F}) = \mathbf{T}_f \mathbf{f}$ , where  $\mathbf{f}$  is defined as

$$\mathbf{f} = \begin{bmatrix} \text{vec}(\mathbf{F}_1) \\ \vdots \\ \text{vec}(\mathbf{F}_R) \end{bmatrix}. \quad (31)$$

$\mathbf{T}_f$  is the transformation matrix formed by ones and zeros, which can be built by observing the nonzero entries of

$\text{vec}(\mathbf{F})$ . Then, with the help of equations  $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$  and  $\text{tr}(\mathbf{X}^H \mathbf{Y} \mathbf{X} \mathbf{W}) = \text{vec}(\mathbf{X})^H (\mathbf{W}^T \otimes \mathbf{Y}) \text{vec}(\mathbf{X})$  in [24], the optimization problem can be written as

$$\begin{aligned} \min_{\mathbf{f}} \quad & \max_{k \in \{1, 2, \dots, 2K\}} \sigma_d^2 \|\mathbf{Z}_{1k} \mathbf{f} - \text{vec}(\mathbf{I}_{L_{\bar{k}}})\|^2 \\ & + \sum_{r=1}^R \sigma_r^2 \|\mathbf{Z}_{3kr} \mathbf{f}\|^2 + \sum_{i=1, i \neq k, \bar{k}}^{2K} \sigma_d^2 \|\mathbf{Z}_{2ki} \mathbf{f}\|^2 + s_{1k} \\ \text{s.t.} \quad & \|\mathbf{D}_r^{\frac{1}{2}} \mathbf{f}\| \leq \sqrt{P_{Rr}} \quad \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (32)$$

where

$$\mathbf{Z}_{1k} = \left( \sum_{r=1}^R \left( (\mathbf{P}_r^T \mathbf{H}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T)^T \otimes (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r) \right) \right) \mathbf{T}_f \quad (33)$$

$$\mathbf{Z}_{2ki} = \left( \sum_{r=1}^R \left( (\mathbf{P}_r^T \mathbf{H}_{ri} \mathbf{Q}_i \mathbf{A} \mathbf{V}_i^T)^T \otimes (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r) \right) \right) \mathbf{T}_f \quad (34)$$

$$\mathbf{Z}_{3kr} = ((\mathbf{P}_r) \otimes (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r)) \mathbf{T}_f \quad (35)$$

$$\begin{aligned} \mathbf{D}_r^{\frac{1}{2}} = & \left( \left( \mathbf{P}_r^T \left( \sigma_d^2 \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \mathbf{Q}_k^T \mathbf{H}_{rk}^H \right. \right. \right. \\ & \left. \left. \left. + \sigma_r^2 \mathbf{I}_{M_r} \right) \mathbf{P}_r \right)^T \otimes (\mathbf{P}_r^T \mathbf{P}_r) \right)^{\frac{1}{2}} \mathbf{T}_f \end{aligned} \quad (36)$$

$$s_{1k} = \sigma_x^2 \|\mathbf{B}_k\|^2. \quad (37)$$

Then, by introducing a new variable  $\tilde{t}$ , the minmax optimization can be expressed as

$$\begin{aligned} \min_{\mathbf{f}, \tilde{t}} \quad & \tilde{t} \\ \text{s.t.} \quad & \sigma_d^2 \|\mathbf{Z}_{1k} \mathbf{f} - \text{vec}(\mathbf{I}_{L_{\bar{k}}})\|^2 + \sum_{i=1, i \neq k, \bar{k}}^{2K} \sigma_d^2 \|\mathbf{Z}_{2ki} \mathbf{f}\|^2 \\ & + \sum_{r=1}^R \sigma_r^2 \|\mathbf{Z}_{3kr} \mathbf{f}\|^2 + s_{1k} \leq \tilde{t}, \quad \|\mathbf{D}_r^{\frac{1}{2}} \mathbf{f}\| \leq \sqrt{P_{Rr}} \\ & \forall k \in \{1, 2, \dots, 2K\}; \quad \forall i \in \{1, 2, \dots, 2K\}, \quad i \neq k, \bar{k}; \\ & \forall r \in \{1, 2, \dots, R\}. \end{aligned} \quad (38)$$

Furthermore, we introduce the set of variables  $t_{1k}, t_{2ki}, t_{3kr}$ . Then, the optimizing problem turns into

$$\begin{aligned} \min_{\mathbf{f}, \tilde{t}, t_{1k}, t_{2ki}, t_{3kr}} \quad & \tilde{t} \\ \text{s.t.} \quad & t_{1k}^2 + \sum_{i=1, i \neq k, \bar{k}}^{2K} t_{2ki}^2 + \sum_{r=1}^R t_{3kr}^2 + s_{1k} \leq \tilde{t} \\ & \sigma_d \|\mathbf{Z}_{1k} \mathbf{f} - \text{vec}(\mathbf{I}_{L_{\bar{k}}})\| \leq t_{1k}, \quad \sigma_d \|\mathbf{Z}_{2ki} \mathbf{f}\| \leq t_{2ki} \\ & \sigma_r \|\mathbf{Z}_{3kr} \mathbf{f}\| \leq t_{3kr}, \quad \|\mathbf{D}_r^{\frac{1}{2}} \mathbf{f}\| \leq \sqrt{P_{Rr}} \\ & \forall k \in \{1, 2, \dots, 2K\} \quad \forall i \in \{1, 2, \dots, 2K\}, \quad i \neq k, \bar{k} \\ & \forall r \in \{1, 2, \dots, R\}. \end{aligned} \quad (39)$$

Moreover, we introduce a new vector  $\mathbf{t}_k \in \mathbb{C}^{(2K-1+R) \times 1}$  defined as

$$\mathbf{t}_k = [t_{1k}, t_{2k1}, t_{2k2}, \dots, t_{2k2K}, t_{3k1}, t_{3k2}, \dots, t_{3kR}]. \quad (40)$$

Using the Schur complement lemma [25], we have

$$\begin{aligned} \min_{\mathbf{f}, \tilde{t}, \mathbf{t}_k} \quad & \tilde{t}, \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{I}_{2K-1+R} & \mathbf{t}_k \\ \mathbf{t}_k^T & \tilde{t} - s_{1k} \end{bmatrix} \succeq 0 \\ & \sigma_d \|\mathbf{Z}_{1k} \mathbf{f} - \text{vec}(\mathbf{I}_{L_{\bar{k}}})\| \leq \mathbf{c}_1^T \mathbf{t}_k, \quad \sigma_d \|\mathbf{Z}_{2ki} \mathbf{f}\| \leq \mathbf{c}_{1+i}^T \mathbf{t}_k \\ & \sigma_r \|\mathbf{Z}_{3kr} \mathbf{f}\| \leq \mathbf{c}_{2K-1+i}^T \mathbf{t}_k, \quad \|\mathbf{D}_r^{\frac{1}{2}} \mathbf{f}\| \leq \sqrt{P_{Rr}} \\ & \forall k \in \{1, 2, \dots, 2K\} \quad \forall i \in \{1, 2, \dots, 2K\}, \quad i \neq k, \bar{k} \\ & \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (41)$$

among which,  $\mathbf{c}_i \in \mathbb{C}^{(2K-1+R) \times 1}$  is a vector with its  $i$ -th element being one, whereas the others are zeros. The problem is a second-order cone programming (SOCP) problem, which can be solved by CVX efficiently.

2) *User Precoder Design:* Define  $\text{vec}(\mathbf{A}) = \mathbf{T}_a \mathbf{a}$ , where  $\mathbf{a} \in \mathbb{C}^{\sum_{k=1}^{2K} (N_k L_k) \times 1}$  and is defined as

$$\mathbf{a} = \begin{bmatrix} \text{vec}(\mathbf{A}_1) \\ \vdots \\ \text{vec}(\mathbf{A}_{2K}) \end{bmatrix}. \quad (42)$$

$\mathbf{T}_a$  is the transformation matrix formed by ones and zeros, which can be built by observing the nonzero entries of  $\text{vec}(\mathbf{A})$ . For fixed  $\mathbf{F}$ , the problem can be formulated as

$$\begin{aligned} \min_{\mathbf{a}} \quad & \max_{k \in \{1, 2, \dots, 2K\}} s_{2k} + \sigma_d^2 \|\mathbf{M}_{1k} \mathbf{a} - \text{vec}(\mathbf{I}_{L_{\bar{k}}})\|^2 \\ & + \sum_{i=1, i \neq k, \bar{k}}^{2K} \sigma_d^2 \|\mathbf{M}_{2ki} \mathbf{a}\|^2 \\ \text{s.t.} \quad & \|\mathbf{N}_{1r}^{\frac{1}{2}} \mathbf{a}\|^2 \leq \tilde{P}_{Rr} \quad \forall r \in \{1, 2, \dots, R\} \\ & \sigma_d^2 \|\mathbf{N}_{2k}^{\frac{1}{2}} \mathbf{a}\|^2 \leq P_{Sk} \quad \forall k \in \{1, 2, \dots, 2K\} \end{aligned} \quad (43)$$

where

$$\mathbf{M}_{1k} = \left( \sum_{r=1}^R ((\mathbf{V}_{\bar{k}}) \otimes (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \mathbf{H}_{r\bar{k}} \mathbf{Q}_{\bar{k}})) \right) \mathbf{T}_a \quad (44)$$

$$\mathbf{M}_{2ki} = \left( \sum_{r=1}^R ((\mathbf{V}_i) \otimes (\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \mathbf{H}_{ri} \mathbf{Q}_i)) \right) \mathbf{T}_a \quad (45)$$

$$\mathbf{N}_{1r}^{\frac{1}{2}} = \left( \sigma_d^2 \sum_{k=1}^{2K} ((\mathbf{V}_k^T \mathbf{V}_k) \otimes (\mathbf{Q}_k^T \mathbf{H}_{rk}^H \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \mathbf{H}_{rk} \mathbf{Q}_k)) \right)^{\frac{1}{2}} \mathbf{T}_a \quad (46)$$

$$\mathbf{N}_{2k}^{\frac{1}{2}} = ((\mathbf{V}_k^T \mathbf{V}_k) \otimes (\mathbf{Q}_k^T \mathbf{Q}_k))^{\frac{1}{2}} \mathbf{T}_a \quad (47)$$

$$s_{2k} = \sigma_x^2 \|\mathbf{B}_k\|^2 + \sigma_r^2 \sum_{r=1}^R \|\mathbf{B}_k \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T\|^2 \quad (48)$$

$$\tilde{P}_{Rr} = P_{Rr} - \sigma_r^2 \text{tr}(\mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T). \quad (49)$$

Then, similar to the relay precoder design, we introduce variables  $\tilde{z}$ ,  $z_{1k}$ , and  $z_{2ki}$ . The user precoder design can be reformulated as

$$\begin{aligned} \min_{\mathbf{a}, \tilde{z}, z_{1k}, z_{2ki}} \quad & \tilde{z}, \\ \text{s.t.} \quad & z_{1k}^2 + \sum_{i=1, i \neq k, \bar{k}}^{2K} z_{2ki}^2 + s_{2k} \leq \tilde{z} \\ & \sigma_d \|\mathbf{M}_{1k} \mathbf{a} - \text{vec}(\mathbf{I}_{L_k})\| \leq z_{1k}, \sigma_d \|\mathbf{M}_{2ki} \mathbf{a}\| \leq z_{2ki} \\ & \|\mathbf{N}_{1r}^{\frac{1}{2}} \mathbf{a}\| \leq \sqrt{\tilde{P}_{Rr}}, \sigma_d \|\mathbf{N}_{2k}^{\frac{1}{2}} \mathbf{a}\| \leq \sqrt{P_{Sk}} \\ & \forall k \in \{1, 2, \dots, 2K\} \forall i \in \{1, 2, \dots, 2K\}, i \neq k, \bar{k} \\ & \forall r \in \{1, 2, \dots, R\}. \end{aligned} \quad (50)$$

Similarly, by introducing  $\mathbf{z}_k \in \mathbb{C}^{(2K-1) \times 1}$ , we have

$$\begin{aligned} \min_{\mathbf{a}, \tilde{z}, \mathbf{z}_k} \quad & \tilde{z} \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{I}_{2K-1} & \mathbf{z}_k \\ \mathbf{z}_k^T & \tilde{z} - s_{2k} \end{bmatrix} \succeq 0 \\ & \sigma_d \|\mathbf{M}_{1k} \mathbf{a} - \text{vec}(\mathbf{I}_{L_k})\| \leq \mathbf{d}_1^T \mathbf{z}_k \\ & \sigma_d \|\mathbf{M}_{2ki} \mathbf{a}\| \leq \mathbf{d}_{1+i}^T \mathbf{z}_k, \|\mathbf{N}_{1r}^{\frac{1}{2}} \mathbf{a}\| \leq \sqrt{\tilde{P}_{Rr}} \end{aligned}$$

$$\begin{aligned} \sigma_d \|\mathbf{N}_{2k}^{\frac{1}{2}} \mathbf{a}\| &\leq \sqrt{P_{Sk}} \\ \forall k \in \{1, 2, \dots, 2K\} \forall i \in \{1, 2, \dots, 2K\}, i \neq k, \bar{k} \\ \forall r \in \{1, 2, \dots, R\}. \end{aligned} \quad (51)$$

Among which,  $\mathbf{d}_i \in \mathbb{C}^{(2K-1) \times 1}$  is a vector with its  $i$ -th element set to one, whereas the others are all zeros. This problem is an SOCP problem, which can be solved by CVX efficiently. The joint design of minmax MSE is summarized in Algorithm 2.

---

**Algorithm 2** Minmax MSE optimization algorithm.

---

1: Initialization:

Set  $\mathbf{A}_k^{(1)} = \sqrt{(P_{Sk}/L_k)} [\mathbf{I}_{L_k} \mathbf{0}_{(N_k-L_k) \times L_k}]^T \quad \forall k \in \{1, 2, \dots, 2K\}$  and  $\mathbf{F}_r^{(1)} = \sqrt{(P_r/\text{tr}(\mathbf{R}_x))} \mathbf{I}_{M_r} \quad \forall r \in \{1, 2, \dots, R\}, n = 1$ .  
Compute  $\mathbf{a}^{(1)}$  and  $\mathbf{f}^{(1)}$  using (42) and (31), respectively.

2: Iteration:

a) Compute  $\mathbf{B}_k^{(n)}$  using (12).  
b) Compute  $\mathbf{f}^{(n)}$  by solving the SOCP problem of (41).  
c) Compute  $\mathbf{a}^{(n)}$  by solving the SOCP problem of (51).  
d) Compute max MSE  $\varepsilon_k^{(n)}$ .

3: Termination:

The algorithm terminates either when  $\varepsilon_k^{(n)}$  converges, i.e.,  $|(\varepsilon_k^{(n)} - \varepsilon_k^{(n-1)})/\varepsilon_k^{(n)}| \leq \eta$ , or when  $n \geq N_{\max}$ , where  $\eta$  is a predefined threshold, and  $N_{\max}$  is the maximum iteration number;  
Output  $\mathbf{B}_k^{\text{opt}} = \mathbf{B}_k^{(n)}$ ,  $\mathbf{f}^{\text{opt}} = \mathbf{f}^{(n)}$ , and  $\mathbf{a}^{\text{opt}} = \mathbf{a}^{(n)}$ .  
Else,  $n = n + 1$ , and go to substep 2.

---

#### IV. JOINT DESIGN FOR THE IMPERFECT CHANNEL KNOWLEDGE

Here, we treat the imperfect CSI case. The block-fading channel and channel estimation error are modeled according to Section II. The proposed optimization algorithm in Section III cannot be applied here due to its perfect CSI assumption. Instead, robust precoding design should be considered. All the nodes in the network are assumed to have the imperfect channel

$$\begin{aligned} \sum_{k=1}^{2K} \varepsilon_k &= \sigma_d^2 \sum_{k=1}^{2K} \sum_{i=1, i \neq k}^{2K} \text{tr} \left( \mathbf{A}_i^H \hat{\mathbf{H}}_{ri}^H \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{F}_r \hat{\mathbf{H}}_{ri} \mathbf{A}_i \right) \\ &+ \sum_{k=1}^{2K} \sigma_r^2 \text{tr} \left( \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{F}_r \right) + \sigma_e^2 \sum_{k=1}^{2K} \text{tr} \left( \mathbf{B}_k \mathbf{B}_k^H \boldsymbol{\Psi}_{G,kr} \right) \text{tr} \left( \mathbf{F}_r^H \boldsymbol{\Sigma}_{G,kr} \boldsymbol{\Psi}_{G,kr} \mathbf{F}_r \left( \tilde{\mathbf{R}}_{xk} + \tilde{\mathbf{R}}_{mk} \right) \right) \\ &+ \sigma_e^2 \sigma_d^2 \sum_{k=1}^{2K} \sum_{i=1}^{2K} \text{tr} \left( \mathbf{A}_i \mathbf{A}_i^H \boldsymbol{\Psi}_{H,ri} \right) \text{tr} \left( \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{F}_r \boldsymbol{\Sigma}_{H,ri} \right) + \sigma_d^2 \sum_{k=1}^{2K} \sum_{i=1, i \neq k}^{2K} \text{tr} \left( \mathbf{A}_i^H \hat{\mathbf{H}}_{ri}^H \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \mathbf{U}_{ir} \right) \\ &+ \sigma_d^2 \sum_{k=1}^{2K} \sum_{i=1, i \neq k}^{2K} \text{tr} \left( \mathbf{U}_{ir}^H \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{F}_r \hat{\mathbf{H}}_{ri} \mathbf{A}_i \right) - \sigma_d^2 \sum_{k=1}^{2K} \text{tr} \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{F}_r \hat{\mathbf{H}}_{r\bar{k}} \mathbf{A}_{\bar{k}} \right) - \sigma_d^2 \sum_{k=1}^{2K} \text{tr} \left( \mathbf{A}_{\bar{k}}^H \hat{\mathbf{H}}_{r\bar{k}}^H \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k \right) \end{aligned} \quad (52)$$

knowledge for both  $\mathbf{G}_{kr}$  and  $\mathbf{H}_{rk}$ . Consequently, the self-interference cancelation cannot be perfectly canceled at each user side. The sum MSE can be expressed in (52), shown at the bottom of the previous page, where

$$\tilde{\mathbf{R}}_{xk} = \sigma_d^2 \sum_{i=1, i \neq k}^{2K} \hat{\mathbf{H}}_{ri} \mathbf{A}_i \mathbf{A}_i^H \hat{\mathbf{H}}_{ri}^H + \sigma_r^2 \mathbf{I}_{M_r} \quad (53)$$

$$\tilde{\mathbf{R}}_{mk} = \sigma_d^2 \hat{\mathbf{H}}_{rk} \mathbf{A}_k \mathbf{A}_k^H \hat{\mathbf{H}}_{rk}^H \quad (54)$$

$$\tilde{\mathbf{U}}_{ir} = \sum_{s=1, s \neq r}^R \mathbf{B}_k \hat{\mathbf{G}}_{ks} \mathbf{F}_s \hat{\mathbf{H}}_{si} \mathbf{A}_i. \quad (55)$$

Thus, the robust MMSE problem can be modified from (11) as

$$\begin{aligned} \min_{\mathbf{B}_k, \mathbf{F}_r, \mathbf{A}_k} \quad & \varepsilon = \sum_{k=1}^{2K} \varepsilon_k \\ \text{s.t.} \quad & \sigma_d^2 \text{tr}(\mathbf{A}_k \mathbf{A}_k^H) \leq P_{Sk} \quad \forall k \in \{1, 2, \dots, 2K\} \\ & \text{tr}(\mathbf{F}_r \tilde{\mathbf{R}}_x \mathbf{F}_r^H) \leq P_{Rr} \quad \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (56)$$

where

$$\begin{aligned} \tilde{\mathbf{R}}_x = & \sigma_e^2 \sigma_d^2 \sum_{k=1}^{2K} \text{tr}(\mathbf{A}_k \mathbf{A}_k^H \Psi_{H, rk}) \Sigma_{H, rk} \\ & + \sigma_r^2 \mathbf{I}_{M_r} + \sigma_d^2 \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{A}_k \mathbf{A}_k^H \mathbf{H}_{rk}^H. \end{aligned} \quad (57)$$

Although the modified version of the original problem seems to be more complicated, the resolving methods are supposed to be similar, which decouples the main problem into several convex subproblems.

#### A. Robust Sum MSE Optimization Algorithm

1) *Receiver Design*: Similar to Section III, we first deduct the MMSE receiver with the preknown knowledge of the stochastic properties for the estimation errors. Thus, the expression for the MMSE receiver can be shown to be

$$\mathbf{B}_k = \sigma_d^2 \sum_{r=1}^R \mathbf{A}_k^H \hat{\mathbf{H}}_{rk}^H \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \tilde{\mathbf{R}}_w^{-1} \quad (58)$$

where

$$\begin{aligned} \tilde{\mathbf{R}}_w = & \sigma_d^2 \sum_{i=1, i \neq k}^{2K} \sum_{r=1}^R \sum_{s=1}^R \hat{\mathbf{G}}_{ks} \mathbf{F}_s \hat{\mathbf{H}}_{si} \mathbf{A}_i \mathbf{A}_i^H \hat{\mathbf{H}}_{ri}^H \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \\ & + \sigma_r^2 \sum_{r=1}^R \hat{\mathbf{G}}_{kr} \mathbf{F}_r \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H + \sigma_e^2 \sum_{r=1}^R \text{tr} \\ & \left( \mathbf{F}_r^H \Psi_{G, kr} \mathbf{F}_r \tilde{\mathbf{R}}_{xk} \right) + \sigma_e^2 \sigma_d^2 \sum_{r=1}^R \sum_{i=1}^{2K} \text{tr} \\ & \left( \mathbf{A}_i \mathbf{A}_i^H \Psi_{H, ri} \right) \hat{\mathbf{G}}_{kr} \mathbf{F}_r \Sigma_{H, ri} \mathbf{F}_r^H \hat{\mathbf{G}}_{kr}^H \\ & + \sigma_x^2 \mathbf{I}_{N_k}. \end{aligned} \quad (59)$$

2) *Relay Precoder Design*: By using the Lagrange multiplier method, the form of the optimal relay precoder at the  $r$ -th relay node can be written as

$$\mathbf{F}_r^{\text{opt}} = \text{mat} \left\{ \left[ \sum_{k=1}^{2K} \tilde{\mathbf{R}}_{xk}^T \otimes \tilde{\mathbf{R}}_{yk} + \sum_{k=1}^{2K} \sum_{i=1}^{2K} \left( \Sigma_{H, ri}^T \otimes \tilde{\mathbf{R}}_{zki} \right) + \sum_{k=1}^{2K} \tilde{\mathbf{R}}_{mk}^T \otimes \tilde{\mathbf{R}}_{nk} + \lambda \tilde{\mathbf{R}}_x^T \otimes \mathbf{I}_{M_r} \right]^{-1} \text{vec}(\tilde{\mathbf{R}}_r) \right\} \quad (60)$$

where

$$\tilde{\mathbf{R}}_{yk} = \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \hat{\mathbf{G}}_{kr} + \tilde{\mathbf{R}}_{nk} \quad (61)$$

$$\begin{aligned} \tilde{\mathbf{R}}_r = & \sigma_d^2 \sum_{k=1}^{2K} \left( \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \mathbf{A}_k^H \hat{\mathbf{H}}_{rk}^H \right. \\ & \left. - \sum_{i=1, i \neq k}^{2K} \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \tilde{\mathbf{U}}_{ir} \mathbf{A}_i^H \hat{\mathbf{H}}_{ri}^H \right) \end{aligned} \quad (62)$$

$$\tilde{\mathbf{R}}_{zki} = \sigma_d^2 \sigma_e^2 \text{tr}(\mathbf{A}_i \mathbf{A}_i^H \Psi_{H, ri}) \hat{\mathbf{G}}_{kr}^H \mathbf{B}_k^H \mathbf{B}_k \hat{\mathbf{G}}_{kr} \quad (63)$$

$$\tilde{\mathbf{R}}_{nk} = \sigma_e^2 \text{tr}(\mathbf{B}_k^H \mathbf{B}_k \Sigma_{G, kr}) \Psi_{G, kr}. \quad (64)$$

The bisection search method can be employed here to acquire the optimal Lagrangian multiplier  $\lambda$  within its bound

$$[0, \sqrt{(\text{tr}(\tilde{\mathbf{R}}_r (\tilde{\mathbf{R}}_x^T)^{-1} \tilde{\mathbf{R}}_x (\tilde{\mathbf{R}}_x^T)^{-1} \tilde{\mathbf{R}}_r^T) / P_{Rr})}].$$

3) *User Precoder Design*: Here, the precalculated robust version of MMSE receivers and relay precoders are assumed to be given. Let  $P'_{Rr} = P_{Rr} - \sigma_r^2 \text{tr}(\mathbf{F}_r^H \mathbf{F}_r)$ . We formulate the user precoder computing problem as

$$\begin{aligned} \min_{\mathbf{A}_k} \quad & \sum_{k=1}^{2K} \text{tr}(\tilde{\mathbf{R}}_{ak} \mathbf{A}_k \mathbf{A}_k^H) \\ & - \sum_{k=1}^{2K} 2\text{Re}[\text{tr}(\tilde{\mathbf{R}}_{bk} \mathbf{A}_k)] + \text{tr}(\tilde{\mathbf{R}}_c) \\ \text{s.t.} \quad & \sigma_d^2 \text{tr}(\mathbf{A}_k^H \mathbf{A}_k) \leq P_{Sk} \quad \forall k \in \{1, 2, \dots, 2K\} \\ & \sum_{k=1}^{2K} \text{tr}(\mathbf{A}_k^H \tilde{\mathbf{R}}_{dkr} \mathbf{A}_k) \leq P'_{Rr} \quad \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (65)$$

where

$$\begin{aligned} \tilde{\mathbf{R}}_{ak} = & \sigma_d^2 \sum_{i=1, i \neq k}^{2K} \left[ \left( \sum_{r=1}^R \mathbf{B}_i \hat{\mathbf{G}}_{ir} \mathbf{F}_r \hat{\mathbf{H}}_{rk} \right)^H \right. \\ & \left. \times \left( \sum_{r=1}^R \mathbf{B}_i \hat{\mathbf{G}}_{ir} \mathbf{F}_r \hat{\mathbf{H}}_{rk} \right) \right] \\ & + \sigma_e^2 \sigma_d^2 \sum_{k=1}^{2K} \sum_{r=1}^R \text{tr} \\ & \left( \mathbf{F}_r^H \hat{\mathbf{G}}_{ir}^H \mathbf{B}_i^H \mathbf{B}_i \hat{\mathbf{G}}_{ir} \mathbf{F}_r \Sigma_{H, rk} \right) \Psi_{H, rk} \end{aligned}$$



$$+ \sigma_e^2 \sigma_d^2 \sum_{k=1}^{2K} \sum_{r=1}^R \text{tr} \left( \mathbf{B}_i^H \mathbf{B}_i \Sigma_{G,ir} \right) \hat{\mathbf{H}}_{rk}^H \mathbf{F}_r^H \Psi_{G,ir} \mathbf{F}_r \hat{\mathbf{H}}_{rk} \quad (66)$$

$$\tilde{\mathbf{R}}_{bk} = \sigma_d^2 \sum_{r=1}^R \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{F}_r \hat{\mathbf{H}}_{rk} \quad (67)$$

$$\begin{aligned} \tilde{\mathbf{R}}_c = \sum_{k=1}^{2K} \left( \sum_{r=1}^R \sigma_r^2 \mathbf{B}_k \mathbf{G}_{kr} \mathbf{F}_r \mathbf{F}_r^H \mathbf{G}_{kr}^H \mathbf{B}_k^H \right. \\ \left. + \sigma_x^2 \mathbf{B}_k \mathbf{B}_k^H + \sigma_e^2 \sigma_d^2 \sum_{r=1}^R \text{tr} \left( \mathbf{F}_r \mathbf{F}_r^H \Psi_{G,kr} \right) \mathbf{B}_k \Sigma_{G,kr} \mathbf{B}_k^H + \mathbf{I}_{L_k} \right) \end{aligned} \quad (68)$$

$$\begin{aligned} \tilde{\mathbf{R}}_{dkr} = \sigma_d^2 \left( \hat{\mathbf{H}}_{rk}^H \mathbf{F}_r^H \mathbf{F}_r \hat{\mathbf{H}}_{rk} \right. \\ \left. + \sigma_e^2 \text{tr} \left( \mathbf{F}_r^H \mathbf{F}_r \Sigma_{H,kr} \right) \Psi_{H,kr} \right). \end{aligned} \quad (69)$$

It can be observed that the channel estimation errors do not violate the convex properties of the optimization problem. Therefore, this convex QCQP problem can be solved by the CVX software. The robust design of sum MSE minimization is summarized in Algorithm 3.

---

**Algorithm 3** Robust sum MSE optimization algorithm.

---

1: Initialization:

Set  $\mathbf{A}_k^{(1)} = \sqrt{(P_{Sk}/L_k)} [\mathbf{I}_{L_k} \mathbf{0}_{(N_k-L_k) \times L_k}]^T \quad \forall k \in \{1, 2, \dots, 2K\}$  and  $\mathbf{F}_r^{(1)} = \sqrt{(P_r/\text{tr}(\mathbf{R}_x))} \mathbf{I}_{M_r} \quad \forall r \in \{1, 2, \dots, R\}, n = 1.$

2: Iteration:

- a) Compute  $\mathbf{B}_k^{(n)}$  using (58).
- b) Compute  $\mathbf{F}_r^{(n)}$  using (60).
- c) Compute  $\mathbf{A}_k^{(n)}$  by solving the QCQP problem of (65).
- d) Compute sum MSE  $\varepsilon^{(n)}$ .

3: Termination:

The algorithm terminates either when  $\varepsilon^{(n)}$  converges, i.e.,  $|(\varepsilon^{(n)} - \varepsilon^{(n-1)})/\varepsilon^{(n)}| \leq \eta$ , or when  $n \geq N_{\max}$ , where  $\eta$  is a predefined threshold and  $N_{\max}$  is the maximum iteration number.

Output  $\mathbf{B}_k^{\text{opt}} = \mathbf{B}_k^{(n)}$ ,  $\mathbf{F}_r^{\text{opt}} = \mathbf{F}_r^{(n)}$ , and  $\mathbf{A}_k^{\text{opt}} = \mathbf{A}_k^{(n)}$ .  
Else,  $n = n + 1$ , and go to substep 2.

---

### B. Robust Minmax MSE Optimization Algorithm

Here, we will consider minimizing the maximum user's MSE under the imperfect CSI case. We will adopt the receiver design in (58). Under this imperfect scenario, the MSE of the  $k$ -th user can be expressed in (70), shown at the bottom of the next page.

Compared with the perfect channel case, we can notice that seven additional terms have been added into the MSE expression. With equation  $\mathbb{E}_{\Delta \mathbf{G}_{kr}} \{\Delta \mathbf{G}_{kr} \mathbf{R}_s \Delta \mathbf{G}_{kr}^H\} = \sigma_e^2 \text{tr}(\mathbf{R}_s \Psi_{G,kr}) \Sigma_{G,kr}$ , term (a) can be simplified as (71),

also shown at the bottom of the next page. Similarly, we can transform the other additional terms into the following forms:

$$\begin{aligned} \mathbf{T}_k^{(b)} = \sigma_e^2 \sigma_d^2 \sum_{r=1}^R \underbrace{\left\| \Psi_{H,r\bar{k}}^{\frac{1}{2}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_k^T \right\|^2}_{\alpha_{kr}^{(b)}} \\ \times \underbrace{\left\| \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F}_r \mathbf{F}_r^T \Sigma_{H,r\bar{k}}^{\frac{1}{2}} \right\|^2}_{\beta_{kr}^{(b)}} \end{aligned} \quad (72)$$

$$\begin{aligned} \mathbf{T}_k^{(c)} = \sigma_e^2 \sigma_d^2 \sum_{r=1}^R \underbrace{\left\| \Psi_{G,kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F}_r \mathbf{F}_r^T \hat{\mathbf{H}}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \right\|^2}_{\beta_{kr}^{(c)}} \\ \times \underbrace{\left\| \mathbf{B}_k \Sigma_{G,kr}^{\frac{1}{2}} \right\|^2}_{\alpha_{kr}^{(c)}} \end{aligned} \quad (73)$$

$$\begin{aligned} \mathbf{T}_k^{(d)} = \sigma_e^2 \sigma_d^2 \sum_{r=1}^R \underbrace{\left\| \Psi_{H,rk}^{\frac{1}{2}} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \right\|^2}_{\alpha_{kr}^{(d)}} \\ \times \underbrace{\left\| \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F}_r \mathbf{F}_r^T \Sigma_{H,rk}^{\frac{1}{2}} \right\|^2}_{\beta_{kr}^{(d)}} \end{aligned} \quad (74)$$

$$\begin{aligned} \mathbf{T}_k^{(e)} = \sigma_e^2 \sigma_d^2 \sum_{i=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \underbrace{\left\| \Psi_{H,r\bar{i}}^{\frac{1}{2}} \mathbf{Q}_{\bar{i}} \mathbf{A} \mathbf{V}_i^T \right\|^2}_{\alpha_{kir}^{(e)}} \\ \times \underbrace{\left\| \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F}_r \mathbf{F}_r^T \Sigma_{H,r\bar{i}}^{\frac{1}{2}} \right\|^2}_{\beta_{kir}^{(e)}} \end{aligned} \quad (75)$$

$$\begin{aligned} \mathbf{T}_k^{(f)} = \sigma_e^2 \sigma_d^2 \sum_{i=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \underbrace{\left\| \Psi_{G,kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F}_r \mathbf{F}_r^T \hat{\mathbf{H}}_{ri} \mathbf{Q}_{\bar{i}} \mathbf{A} \mathbf{V}_i^T \right\|^2}_{\beta_{kir}^{(f)}} \\ \times \underbrace{\left\| \mathbf{B}_k \Sigma_{G,kr}^{\frac{1}{2}} \right\|^2}_{\alpha_{kir}^{(f)}} \end{aligned} \quad (76)$$

$$\mathbf{T}_k^{(g)} = \sigma_e^2 \sigma_r^2 \sum_{r=1}^R \underbrace{\left\| \Psi_{G,kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F}_r \mathbf{F}_r^T \right\|^2}_{\beta_{kr}^{(g)}} \underbrace{\left\| \mathbf{B}_k \Sigma_{G,kr}^{\frac{1}{2}} \right\|^2}_{\alpha_{kr}^{(g)}}. \quad (77)$$

The relay power constraints can be reformulated into the form

$$\begin{aligned} \text{tr} \left[ \mathbf{P}_r \mathbf{F}_r \mathbf{F}_r^T \left( \sigma_d^2 \sum_{k=1}^{2K} \mathbf{H}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \mathbf{Q}_k^T \mathbf{H}_{rk}^H + \sigma_r^2 \mathbf{I}_{M_r} \right) \right. \\ \left. \mathbf{P}_r \mathbf{F}_r^H \mathbf{P}_r^T \right] \end{aligned}$$

$$\begin{aligned}
&= \text{tr} \left[ \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \left( \sigma_d^2 \sum_{k=1}^{2K} \hat{\mathbf{H}}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \mathbf{Q}_k^T \hat{\mathbf{H}}_{rk}^H + \sigma_r^2 \mathbf{I}_{M_r} \right) \right. \\
&\quad \left. \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T \right] \\
&+ \sigma_e^2 \sigma_d^2 \sum_{k=1}^{2K} \text{tr} \left( \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \mathbf{Q}_k^T \Psi_{H, rk} \right) \text{tr} \\
&\left( \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \Sigma_{H, rk} \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T \right) \leq P_{Rr}. \tag{78}
\end{aligned}$$

1) *Relay Precoder Design*: First, define the following expressions:

$$\hat{\mathbf{Z}}_{kr}^{(a)} = \left( \left( \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \right)^T \otimes \left( \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \right) \right) \mathbf{T}_f \tag{79}$$

$$\hat{\mathbf{Z}}_{kr}^{(b)} = \left( \left( \mathbf{P}_r^T \Sigma_{H, r\bar{k}}^{\frac{1}{2}} \right)^T \otimes \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \right) \right) \mathbf{T}_f \tag{80}$$

$$\hat{\mathbf{Z}}_{kr}^{(c)} = \left( \left( \mathbf{P}_r^T \hat{\mathbf{H}}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \right)^T \otimes \left( \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \right) \right) \mathbf{T}_f \tag{81}$$

$$\begin{aligned}
\varepsilon_k &= \sigma_d^2 \left\| \sum_{r=1}^R \text{vec} \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \right) - \text{vec}(\mathbf{I}_{L_{\bar{k}}}) \right\|^2 + \sigma_r^2 \sum_{r=1}^R \left\| \text{vec} \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \right) \right\|^2 \\
&+ \sigma_x^2 \|\mathbf{B}_k\|^2 + \sigma_d^2 \sum_{i=1, i \neq k, \bar{k}}^{2K} \left\| \sum_{r=1}^R \text{vec} \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{ri} \mathbf{Q}_i \mathbf{A} \mathbf{V}_i^T \right) \right\|^2 \\
&+ \mathbb{E}_{\Delta \mathbf{G}_{kr}} \left\{ \underbrace{\sigma_d^2 \left\| \sum_{r=1}^R \left( \mathbf{B}_k \Delta \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \right) \right\|^2}_{(a)} \right\} + \mathbb{E}_{\Delta \mathbf{H}_{r\bar{k}}} \left\{ \underbrace{\sigma_d^2 \sum_{r=1}^R \left\| \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \Delta \mathbf{H}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \right\|^2}_{(b)} \right\} \\
&+ \mathbb{E}_{\Delta \mathbf{G}_{kr}} \left\{ \underbrace{\sigma_d^2 \sum_{r=1}^R \left\| \left( \mathbf{B}_k \Delta \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \right) \right\|^2}_{(c)} \right\} + \mathbb{E}_{\Delta \mathbf{H}_{rk}} \left\{ \underbrace{\sigma_d^2 \sum_{r=1}^R \left\| \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \Delta \mathbf{H}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \right\|^2}_{(d)} \right\} \\
&+ \mathbb{E}_{\Delta \mathbf{H}_{ri}} \left\{ \underbrace{\sigma_d^2 \sum_{i=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \left\| \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \Delta \mathbf{H}_{ri} \mathbf{Q}_i \mathbf{A} \mathbf{V}_i^T \right\|^2}_{(e)} \right\} + \mathbb{E}_{\Delta \mathbf{G}_{kr}} \left\{ \underbrace{\sigma_r^2 \sum_{r=1}^R \left\| \left( \mathbf{B}_k \Delta \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \right) \right\|^2}_{(g)} \right\} \\
&+ \mathbb{E}_{\Delta \mathbf{G}_{kr}} \left\{ \underbrace{\sigma_d^2 \sum_{i=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \left\| \mathbf{B}_k \Delta \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{ri} \mathbf{Q}_i \mathbf{A} \mathbf{V}_i^T \right\|^2}_{(f)} \right\} \tag{70}
\end{aligned}$$

$$\begin{aligned}
\mathbf{T}_k^{(a)} &= \mathbb{E}_{\Delta \mathbf{G}_{kr}} \left\{ \sigma_d^2 \sum_{r=1}^R \left\| \mathbf{B}_k \Delta \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \right\|^2 \right\} \\
&= \sigma_d^2 \sum_{r=1}^R \text{tr} \left( \mathbf{B}_k \mathbb{E}_{\Delta \mathbf{G}_{kr}} \left\{ \Delta \mathbf{G}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \mathbf{V}_{\bar{k}} \mathbf{A}^H \mathbf{Q}_{\bar{k}}^H \hat{\mathbf{H}}_{r\bar{k}}^H \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T \Delta \mathbf{G}_{kr}^H \right\} \mathbf{B}_k^H \right) \\
&= \sigma_e^2 \sigma_d^2 \sum_{r=1}^R \underbrace{\left\| \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \right\|^2}_{\beta_{kr}^{(a)}} \underbrace{\left\| \mathbf{B}_k \Sigma_{G, kr}^{\frac{1}{2}} \right\|^2}_{\alpha_{kr}^{(a)}} \tag{71}
\end{aligned}$$

$$\hat{\mathbf{Z}}_{kr}^{(d)} = \left( \left( \mathbf{P}_r^T \Sigma_{H, rk}^{\frac{1}{2}} \right)^T \otimes \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \right) \right) \mathbf{T}_f \quad (82)$$

$$\hat{\mathbf{Z}}_{kir}^{(e)} = \left( \left( \mathbf{P}_r^T \Sigma_{H, ri}^{\frac{1}{2}} \right)^T \otimes \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \right) \right) \mathbf{T}_f \quad (83)$$

$$\hat{\mathbf{Z}}_{kir}^{(f)} = \left( \left( \mathbf{P}_r^T \hat{\mathbf{H}}_{ri} \mathbf{Q}_i \mathbf{A} \mathbf{V}_i^T \right)^T \otimes \left( \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \right) \right) \mathbf{T}_f \quad (84)$$

$$\hat{\mathbf{Z}}_{kr}^{(g)} = \left( \left( \mathbf{P}_r \right) \otimes \left( \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \right) \right) \mathbf{T}_f. \quad (85)$$

Similar to the perfect CSI case, we introduce variables  $\tilde{t}$ ,  $\hat{t}_{1k}$ ,  $\hat{t}_{2ki}$ ,  $\hat{t}_{3kr}$ ,  $\hat{t}_{kr}^{(a)}$ ,  $\hat{t}_{kr}^{(b)}$ ,  $\hat{t}_{kr}^{(c)}$ ,  $\hat{t}_{kr}^{(d)}$ ,  $\hat{t}_{kr}^{(e)}$ ,  $\hat{t}_{kr}^{(f)}$ ,  $\hat{t}_{kr}^{(g)}$   $\forall k \in 1, 2, \dots, 2K$ ;  $\forall i \in 1, 2, \dots, 2K, i \neq k, k$ ; and  $\forall r \in 1, 2, \dots, R$ . Then, the minmax optimization can be expressed as

$$\begin{aligned} & \min_{\mathbf{f}, \tilde{t}, \hat{t}_{1k}, \hat{t}_{2ki}, \hat{t}_{3kr}, \hat{t}_{kr}^{(a)}, \hat{t}_{kr}^{(b)}, \hat{t}_{kr}^{(c)}, \hat{t}_{kr}^{(d)}, \hat{t}_{kr}^{(e)}, \hat{t}_{kr}^{(f)}, \hat{t}_{kr}^{(g)}} \tilde{t} \\ & \text{s.t.} \quad \sum_{I=1, i \neq k, \bar{k}}^{2K} \hat{t}_{2ki}^2 + \sum_{r=1}^R \hat{t}_{3kr}^2 + \sum_{r=1}^R \left( \hat{t}_{kr}^{(a)} \right)^2 + \sum_{r=1}^R \left( \hat{t}_{kr}^{(b)} \right)^2 \\ & \quad + \sum_{r=1}^R \left( \hat{t}_{kr}^{(c)} \right)^2 + \sum_{r=1}^R \left( \hat{t}_{kr}^{(d)} \right)^2 + \sum_{I=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \left( \hat{t}_{kir}^{(e)} \right)^2 \\ & \quad + \sum_{I=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \left( \hat{t}_{kir}^{(f)} \right)^2 + \sum_{r=1}^R \left( \hat{t}_{kr}^{(g)} \right)^2 + \hat{t}_{1k}^2 + s_{1k} \leq \tilde{t}, \\ & \quad \sigma_d \left\| \hat{\mathbf{Z}}_{1k} \mathbf{f} - \text{vec}(\mathbf{I}_{L_{\bar{k}}}) \right\| \leq \hat{t}_{1k}, \sigma_d \left\| \hat{\mathbf{Z}}_{2ki} \mathbf{f} \right\| \leq \hat{t}_{2ki}, \\ & \quad \sigma_r \left\| \hat{\mathbf{Z}}_{3kr} \mathbf{f} \right\| \leq \hat{t}_{3kr}, \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(a)}} \left\| \hat{\mathbf{Z}}_{kr}^{(a)} \mathbf{f} \right\| \leq \hat{t}_{kr}^{(a)}, \\ & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(b)}} \left\| \hat{\mathbf{Z}}_{kr}^{(b)} \mathbf{f} \right\| \leq \hat{t}_{kr}^{(b)}, \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(c)}} \left\| \hat{\mathbf{Z}}_{kr}^{(c)} \mathbf{f} \right\| \leq \hat{t}_{kr}^{(c)}, \\ & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(d)}} \left\| \hat{\mathbf{Z}}_{kr}^{(d)} \mathbf{f} \right\| \leq \hat{t}_{kr}^{(d)}, \sigma_e \sigma_d \sqrt{\alpha_{kir}^{(e)}} \left\| \hat{\mathbf{Z}}_{kir}^{(e)} \mathbf{f} \right\| \leq \hat{t}_{kir}^{(e)}, \\ & \quad \sigma_e \sigma_d \sqrt{\alpha_{kir}^{(f)}} \left\| \hat{\mathbf{Z}}_{kir}^{(f)} \mathbf{f} \right\| \leq \hat{t}_{kir}^{(f)}, \sigma_e \sigma_r \sqrt{\alpha_{kr}^{(g)}} \left\| \hat{\mathbf{Z}}_{kr}^{(g)} \mathbf{f} \right\| \leq \hat{t}_{kr}^{(g)}, \\ & \quad \left\| \hat{\mathbf{D}}_r^{\frac{1}{2}} \mathbf{f} \right\| \leq \sqrt{P_{Rr}} \forall k \in \{1, 2, \dots, 2K\}, \quad \forall i \in \{1, 2, \dots, 2K\}, \\ & \quad i \neq k, \bar{k}, \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (86)$$

where

$$\hat{\mathbf{Z}}_{1k} = \left( \sum_{r=1}^R \left( \left( \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \mathbf{A} \mathbf{V}_{\bar{k}}^T \right)^T \otimes \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \right) \right) \right) \mathbf{T}_f \quad (87)$$

$$\hat{\mathbf{Z}}_{2ki} = \left( \sum_{r=1}^R \left( \left( \mathbf{P}_r^T \hat{\mathbf{H}}_{ri} \mathbf{Q}_i \mathbf{A} \mathbf{V}_i^T \right)^T \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \right) \right) \right) \mathbf{T}_f \quad (88)$$

$$\hat{\mathbf{Z}}_{3kr} = \left( \left( \mathbf{P}_r \right) \otimes \left( \mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \right) \right) \mathbf{T}_f \quad (89)$$

$$\begin{aligned} \hat{\mathbf{D}}_r^{\frac{1}{2}} = & \left[ \left( \mathbf{P}_r^T \left( \sigma_d^2 \sum_{k=1}^{2K} \hat{\mathbf{H}}_{rk} \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \mathbf{Q}_k^T \hat{\mathbf{H}}_{rk}^H \right. \right. \right. \\ & \left. \left. + \sigma_r^2 \mathbf{I}_{M_r} \right) \mathbf{P}_r \right)^T \otimes \left( \mathbf{P}_r^T \mathbf{P}_r \right) \\ & \left. + \sigma_e^2 \sigma_d^2 \sum_{k=1}^{2K} \text{tr} \left( \mathbf{Q}_k \mathbf{A} \mathbf{V}_k^T \mathbf{V}_k \mathbf{A}^H \mathbf{Q}_k^T \Psi_{H, rk} \right) \right. \\ & \left. \times \left( \left( \mathbf{P}_r^T \Sigma_{H, rk} \mathbf{P}_r \right)^T \otimes \left( \mathbf{P}_r^T \mathbf{P}_r \right) \right) \right]^{\frac{1}{2}} \mathbf{T}_f. \quad (90) \end{aligned}$$

Then, by introducing a new variable, which can be expressed as  $\hat{\mathbf{t}}_k = [\hat{t}_{1k}, \hat{t}_{2ki}, \hat{t}_{3kr}, \hat{t}_{kr}^{(a)}, \hat{t}_{kr}^{(b)}, \hat{t}_{kr}^{(c)}, \hat{t}_{kr}^{(d)}, \hat{t}_{kr}^{(e)}, \hat{t}_{kir}^{(f)}, \hat{t}_{kir}^{(g)}]^T \in \mathbb{C}^{(2K-1+R(4K+2)) \times 1}$ , where  $\hat{\mathbf{t}}_{2ki}$  is a vector containing all possible values of  $\hat{t}_{2ki}$ , and  $\hat{t}_{kr}^{(b)}, \hat{t}_{kr}^{(c)}, \hat{t}_{kr}^{(d)}, \hat{t}_{kr}^{(e)}, \hat{t}_{kir}^{(f)}, \hat{t}_{kir}^{(g)}$  have the similar definitions, we have

$$\begin{aligned} & \min_{\mathbf{f}, \hat{\mathbf{t}}_k \forall k \in 1, 2, \dots, 2K} \tilde{t}, \\ & \text{s.t.} \quad \begin{bmatrix} \mathbf{I}_{2K+1+R \times (4K+2)} & \hat{\mathbf{t}}_k \\ \hat{\mathbf{t}}_k^T & \tilde{t} - s_{1k} \end{bmatrix} \succeq 0, \\ & \quad \sigma_d \left\| \hat{\mathbf{Z}}_{1k} \mathbf{f} - \text{vec}(\mathbf{I}_{L_{\bar{k}}}) \right\| \leq \mathbf{c}_k^{(1)T} \hat{\mathbf{t}}_k, \sigma_d \left\| \hat{\mathbf{Z}}_{2ki} \mathbf{f} \right\| \leq \mathbf{c}_{ki}^{(2)T} \hat{\mathbf{t}}_k, \\ & \quad \sigma_r \left\| \hat{\mathbf{Z}}_{3kr} \mathbf{f} \right\| \leq \mathbf{c}_{kr}^{(3)T} \hat{\mathbf{t}}_k, \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(a)}} \left\| \hat{\mathbf{Z}}_{kr}^{(a)} \mathbf{f} \right\| \leq \mathbf{c}_{kr}^{(a)T} \hat{\mathbf{t}}_k, \\ & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(b)}} \left\| \hat{\mathbf{Z}}_{kr}^{(b)} \mathbf{f} \right\| \leq \mathbf{c}_{kr}^{(b)T} \hat{\mathbf{t}}_k, \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(c)}} \left\| \hat{\mathbf{Z}}_{kr}^{(c)} \mathbf{f} \right\| \leq \mathbf{c}_{kr}^{(c)T} \hat{\mathbf{t}}_k, \\ & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(d)}} \left\| \hat{\mathbf{Z}}_{kr}^{(d)} \mathbf{f} \right\| \leq \mathbf{c}_{kr}^{(d)T} \hat{\mathbf{t}}_k, \sigma_e \sigma_d \sqrt{\alpha_{kir}^{(e)}} \left\| \hat{\mathbf{Z}}_{kir}^{(e)} \mathbf{f} \right\| \leq \mathbf{c}_{kir}^{(e)T} \hat{\mathbf{t}}_k, \\ & \quad \sigma_e \sigma_d \sqrt{\alpha_{kir}^{(f)}} \left\| \hat{\mathbf{Z}}_{kir}^{(f)} \mathbf{f} \right\| \leq \mathbf{c}_{kir}^{(f)T} \hat{\mathbf{t}}_k, \sigma_e \sigma_r \sqrt{\alpha_{kr}^{(g)}} \left\| \hat{\mathbf{Z}}_{kr}^{(g)} \mathbf{f} \right\| \leq \mathbf{c}_{kr}^{(g)T} \hat{\mathbf{t}}_k, \\ & \quad \left\| \hat{\mathbf{D}}_r^{\frac{1}{2}} \mathbf{f} \right\| \leq \sqrt{P_{Rr}}, \forall k \in \{1, 2, \dots, 2K\} \forall i \in \{1, 2, \dots, 2K\}, \\ & \quad i \neq k, \bar{k} \forall r \in \{1, 2, \dots, R\} \end{aligned} \quad (91)$$

where  $\mathbf{c}_k^{(1)}, \mathbf{c}_{ki}^{(2)}, \mathbf{c}_{kr}^{(3)}, \mathbf{c}_{kr}^{(a)}, \mathbf{c}_{kr}^{(b)}, \mathbf{c}_{kr}^{(c)}, \mathbf{c}_{kr}^{(d)}, \mathbf{c}_{kir}^{(e)}, \mathbf{c}_{kir}^{(f)}, \mathbf{c}_{kr}^{(g)}$  correspond to different columns of  $\mathbf{I}_{2K-1+R \times (4K+2)}$ . This problem is a typical SOCP problem, which can be solved by CVX efficiently.

2) *User Precoder Design:* Here, we aim to optimize the user precoders. Let us define the following expressions first:

$$\hat{\mathbf{M}}_{kr}^{(a)} = \left( \mathbf{V}_{\bar{k}} \otimes \left( \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}} \right) \right) \mathbf{T}_a \quad (92)$$

$$\hat{\mathbf{M}}_{kr}^{(b)} = \left( \mathbf{V}_{\bar{k}} \otimes \left( \Psi_{H, rk}^{\frac{1}{2}} \mathbf{Q}_{\bar{k}} \right) \right) \mathbf{T}_a \quad (93)$$

$$\hat{\mathbf{M}}_{kr}^{(c)} = \left( \mathbf{V}_k \otimes \left( \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{rk} \mathbf{Q}_k \right) \right) \mathbf{T}_a \quad (94)$$

$$\hat{\mathbf{M}}_{kr}^{(d)} = \left( \mathbf{V}_k \otimes \left( \Psi_{H, rk}^{\frac{1}{2}} \mathbf{Q}_k \right) \right) \mathbf{T}_a \quad (95)$$

$$\hat{\mathbf{M}}_{kir}^{(e)} = \left( \mathbf{V}_i \otimes \left( \Psi_{H, ri}^{\frac{1}{2}} \mathbf{Q}_i \right) \right) \mathbf{T}_a \quad (96)$$

$$\hat{\mathbf{M}}_{kir}^{(f)} = \left( \mathbf{V}_i \otimes \left( \Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{ri} \mathbf{Q}_i \right) \right) \mathbf{T}_a. \quad (97)$$

Similar to the perfect CSI case, we introduce variable  $\tilde{z}$ . Then, we have

$$\begin{aligned}
 & \min_{\mathbf{a}, \tilde{z}} \tilde{z}, \\
 & \text{s.t.} \quad \sigma_d^2 \left\| \hat{\mathbf{M}}_{1k} \mathbf{a} - \text{vec}(\mathbf{I}_{L_k}) \right\|^2 + \sum_{i=1, i \neq k, \bar{k}}^{2K} \sigma_d^2 \left\| \hat{\mathbf{M}}_{2ki} \mathbf{a} \right\|^2 \\
 & \quad + \hat{s}_{2k} + \sigma_e^2 \sigma_d^2 \left( \sum_{r=1}^R \alpha_{kr}^{(a)} \left\| \hat{\mathbf{M}}_{kr}^{(a)} \mathbf{a} \right\|^2 + \sum_{r=1}^R \beta_{kr}^{(b)} \left\| \hat{\mathbf{M}}_{kr}^{(b)} \mathbf{a} \right\|^2 \right. \\
 & \quad + \sum_{r=1}^R \alpha_{kr}^{(c)} \left\| \hat{\mathbf{M}}_{kr}^{(c)} \mathbf{a} \right\|^2 + \sum_{r=1}^R \beta_{kr}^{(d)} \left\| \hat{\mathbf{M}}_{kr}^{(d)} \mathbf{a} \right\|^2 \\
 & \quad \left. + \sum_{i=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \left( \beta_{kir}^{(e)} \left\| \hat{\mathbf{M}}_{kir}^{(e)} \mathbf{a} \right\|^2 + \alpha_{kir}^{(f)} \left\| \hat{\mathbf{M}}_{kir}^{(f)} \mathbf{a} \right\|^2 \right) \right) \leq \tilde{z}, \\
 & \quad \left\| \hat{\mathbf{N}}_{1r}^{\frac{1}{2}} \mathbf{a} \right\| \leq \sqrt{\tilde{P}_{Rr}}, \sigma_d \left\| \hat{\mathbf{N}}_{2k}^{\frac{1}{2}} \mathbf{a} \right\| \leq \sqrt{P_{Sk}}, \\
 & \quad \forall k \in \{1, 2, \dots, 2K\}; \forall i \in \{1, 2, \dots, 2K\}, i \neq k, \bar{k}, \\
 & \quad \forall r \in \{1, 2, \dots, R\}, \tag{98}
 \end{aligned}$$

where

$$\hat{\mathbf{M}}_{1k} = \left( \sum_{r=1}^R \left( (\mathbf{V}_{\bar{k}}) \otimes (\mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{r\bar{k}} \mathbf{Q}_{\bar{k}}) \right) \right) \mathbf{T}_a \tag{99}$$

$$\hat{\mathbf{M}}_{2ki} = \left( \sum_{r=1}^R \left( (\mathbf{V}_i) \otimes (\mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{ri} \mathbf{Q}_i) \right) \right) \mathbf{T}_a \tag{100}$$

$$\begin{aligned}
 \hat{\mathbf{N}}_{1r}^{\frac{1}{2}} &= \left[ \sigma_d^2 \sum_{k=1}^{2K} \left( (\mathbf{V}_k^T \mathbf{V}_k) \right. \right. \\
 & \quad \left. \left. \otimes \left( \mathbf{Q}_k^T \hat{\mathbf{H}}_{rk}^H \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \hat{\mathbf{H}}_{rk} \mathbf{Q}_k \right) \right) \right. \\
 & \quad \left. + \sigma_e^2 \sigma_d^2 \sum_{k=1}^{2K} \left( ((\mathbf{V}_k^T \mathbf{V}_k) \otimes (\mathbf{Q}_k^T \mathbf{Q}_k)) \text{tr} \right. \right. \\
 & \quad \left. \left. \left( \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \Sigma_{H, rk} \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T \right) \right) \right]^{\frac{1}{2}} \mathbf{T}_a \tag{101}
 \end{aligned}$$

$$\hat{\mathbf{N}}_{2k}^{\frac{1}{2}} = ((\mathbf{V}_k^T \mathbf{V}_k) \otimes (\mathbf{Q}_k^T \mathbf{Q}_k))^{\frac{1}{2}} \mathbf{T}_a \tag{102}$$

$$\begin{aligned}
 \hat{s}_{2k} &= \sigma_x^2 \|\mathbf{B}_k\|^2 + \sigma_e^2 \sigma_r^2 \sum_{r=1}^R \|\Psi_{G, kr}^{\frac{1}{2}} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T\|^2 \|\mathbf{B}_k \Sigma_{G, kr}^{\frac{1}{2}}\|^2 \\
 & \quad + \sigma_r^2 \sum_{r=1}^R \|\mathbf{B}_k \hat{\mathbf{G}}_{kr} \mathbf{P}_r \mathbf{F} \mathbf{P}_r^T\|^2 \tag{103}
 \end{aligned}$$

$$\tilde{P}_{Rr} = P_{Rr} - \sigma_r^2 \text{tr}(\mathbf{P}_r \mathbf{F} \mathbf{P}_r^T \mathbf{P}_r \mathbf{F}^H \mathbf{P}_r^T). \tag{104}$$

Moreover, for the purpose of transforming this problem into a standard convex optimization form, we introduce a set of

variables  $\tilde{z}, \hat{z}_{1k}, \hat{z}_{2ki}, \hat{z}_{kr}^{(a)}, \hat{z}_{kr}^{(b)}, \hat{z}_{kr}^{(c)}, \hat{z}_{kr}^{(d)}, \hat{z}_{kir}^{(e)}, \hat{z}_{kir}^{(f)}$ , and  $\hat{z}_{kr}^{(g)}$ . Then, (98) can be equivalently rewritten as

$$\begin{aligned}
 & \min_{\mathbf{a}, \tilde{z}, \hat{z}_{1k}, \hat{z}_{2ki}, \hat{z}_{kr}^{(a)}, \hat{z}_{kr}^{(b)}, \hat{z}_{kr}^{(c)}, \hat{z}_{kr}^{(d)}, \hat{z}_{kir}^{(e)}, \hat{z}_{kir}^{(f)}, \hat{z}_{kr}^{(g)}} \tilde{z} \\
 & \text{s.t.} \quad \hat{s}_{2k} + \hat{z}_{1k}^2 + \sum_{i=1, i \neq k, \bar{k}}^{2K} \hat{z}_{2ki}^2 + \sum_{r=1}^R \left( \hat{z}_{kr}^{(a)} \right)^2 \sum_{r=1}^R \left( \hat{z}_{kr}^{(b)} \right)^2 \\
 & \quad + \sum_{r=1}^R \left( \hat{z}_{kr}^{(c)} \right)^2 + \sum_{r=1}^R \left( \hat{z}_{kr}^{(d)} \right)^2 + \sum_{i=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \left( \hat{z}_{kir}^{(e)} \right)^2 \\
 & \quad + \sum_{i=1, i \neq k, \bar{k}}^{2K} \sum_{r=1}^R \left( \hat{z}_{kir}^{(f)} \right)^2 \leq \tilde{z}, \\
 & \quad \sigma_d \left\| \hat{\mathbf{M}}_{1k} \mathbf{a} - \text{vec}(\mathbf{I}_{L_k}) \right\| \leq \hat{z}_{1k}, \sigma_d \left\| \hat{\mathbf{M}}_{2ki} \mathbf{a} \right\| \leq \hat{z}_{2ki}, \\
 & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(a)}} \left\| \hat{\mathbf{M}}_{kr}^{(a)} \mathbf{a} \right\| \leq \hat{z}_{kr}^{(a)}, \sigma_e \sigma_d \sqrt{\beta_{kr}^{(b)}} \left\| \hat{\mathbf{M}}_{kr}^{(b)} \mathbf{a} \right\| \leq \hat{z}_{kr}^{(b)}, \\
 & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(c)}} \left\| \hat{\mathbf{M}}_{kr}^{(c)} \mathbf{a} \right\| \leq \hat{z}_{kr}^{(c)}, \sigma_e \sigma_d \sqrt{\beta_{kr}^{(d)}} \left\| \hat{\mathbf{M}}_{kr}^{(d)} \mathbf{a} \right\| \leq \hat{z}_{kr}^{(d)}, \\
 & \quad \sigma_e \sigma_d \sqrt{\beta_{kir}^{(e)}} \left\| \hat{\mathbf{M}}_{kir}^{(e)} \mathbf{a} \right\| \leq \hat{z}_{kir}^{(e)}, \sigma_e \sigma_d \sqrt{\alpha_{kir}^{(f)}} \left\| \hat{\mathbf{M}}_{kir}^{(f)} \mathbf{a} \right\| \leq \hat{z}_{kir}^{(f)}, \\
 & \quad \left\| \mathbf{N}_{1r}^{\frac{1}{2}} \mathbf{a} \right\| \leq \sqrt{\tilde{P}_{Rr}}, \sigma_d \left\| \mathbf{N}_{2k}^{\frac{1}{2}} \mathbf{a} \right\| \leq \sqrt{P_{Sk}}, \\
 & \quad \forall k \in \{1, 2, \dots, 2K\}, \forall i \in \{1, 2, \dots, 2K\}, i \neq k, \bar{k}, \\
 & \quad \forall r \in \{1, 2, \dots, R\}. \tag{105}
 \end{aligned}$$

In addition, by introducing a new variable  $\hat{\mathbf{z}}_k = [\hat{z}_{1k}, \hat{z}_{2ki}, \hat{z}_{kr}^{(a)}, \hat{z}_{kr}^{(b)}, \hat{z}_{kr}^{(c)}, \hat{z}_{kr}^{(d)}, \hat{z}_{kir}^{(e)}, \hat{z}_{kir}^{(f)}] \in \mathbb{C}^{(2K-1+4KR) \times 1}$ , it should be noticed that  $\hat{z}_{2ki}$  have all possible values of  $\hat{z}_{2ki}$ , and  $\hat{z}_{kr}^{(a)}, \hat{z}_{kr}^{(b)}, \hat{z}_{kr}^{(c)}, \hat{z}_{kr}^{(d)}, \hat{z}_{kir}^{(e)}, \hat{z}_{kir}^{(f)}$  have similar definitions. Therefore, we have

$$\begin{aligned}
 & \min_{\mathbf{a}, \tilde{z}, \hat{\mathbf{z}}_k} \tilde{z} \\
 & \text{s.t.} \quad \begin{bmatrix} \mathbf{I}_{2K-1+4KR} & \hat{\mathbf{z}}_k \\ \hat{\mathbf{z}}_k^T & \tilde{z} - \hat{s}_{2k} \end{bmatrix} \succeq 0, \\
 & \quad \sigma_d \left\| \hat{\mathbf{M}}_{1k} \mathbf{a} - \text{vec}(\mathbf{I}_{L_k}) \right\| \leq \mathbf{d}_k^{(1)T} \hat{\mathbf{z}}_k, \sigma_d \left\| \hat{\mathbf{M}}_{2ki} \mathbf{a} \right\| \leq \mathbf{d}_{ki}^{(2)T} \hat{\mathbf{z}}_k, \\
 & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(a)}} \left\| \hat{\mathbf{M}}_{kr}^{(a)} \mathbf{a} \right\| \leq \mathbf{d}_{kr}^{(a)T} \hat{\mathbf{z}}_k, \sigma_e \sigma_d \sqrt{\beta_{kr}^{(b)}} \left\| \hat{\mathbf{M}}_{kr}^{(b)} \mathbf{a} \right\| \leq \mathbf{d}_{kr}^{(b)T} \hat{\mathbf{z}}_k, \\
 & \quad \sigma_e \sigma_d \sqrt{\alpha_{kr}^{(c)}} \left\| \hat{\mathbf{M}}_{kr}^{(c)} \mathbf{a} \right\| \leq \mathbf{d}_{kr}^{(c)T} \hat{\mathbf{z}}_k, \sigma_e \sigma_d \sqrt{\beta_{kr}^{(d)}} \left\| \hat{\mathbf{M}}_{kr}^{(d)} \mathbf{a} \right\| \leq \mathbf{d}_{kr}^{(d)T} \hat{\mathbf{z}}_k, \\
 & \quad \sigma_e \sigma_d \sqrt{\beta_{kir}^{(e)}} \left\| \hat{\mathbf{M}}_{kir}^{(e)} \mathbf{a} \right\| \leq \mathbf{d}_{kir}^{(e)T} \hat{\mathbf{z}}_k, \sigma_e \sigma_d \sqrt{\alpha_{kir}^{(f)}} \left\| \hat{\mathbf{M}}_{kir}^{(f)} \mathbf{a} \right\| \leq \mathbf{d}_{kir}^{(f)T} \hat{\mathbf{z}}_k, \\
 & \quad \left\| \mathbf{N}_{1r}^{\frac{1}{2}} \mathbf{a} \right\| \leq \sqrt{\tilde{P}_{Rr}}, \sigma_d \left\| \mathbf{N}_{2k}^{\frac{1}{2}} \mathbf{a} \right\| \leq \sqrt{P_{Sk}}, \\
 & \quad \forall k \in \{1, 2, \dots, 2K\}, \forall i \in \{1, 2, \dots, 2K\}, i \neq k, \bar{k}, \\
 & \quad \forall r \in \{1, 2, \dots, R\} \tag{106}
 \end{aligned}$$

where  $\mathbf{d}_k^{(1)}, \mathbf{d}_{ki}^{(2)}, \mathbf{d}_{kr}^{(3)}, \mathbf{d}_{kr}^{(a)}, \mathbf{d}_{kr}^{(b)}, \mathbf{d}_{kr}^{(c)}, \mathbf{d}_{kr}^{(d)}, \mathbf{d}_{kir}^{(e)}$ , and  $\mathbf{d}_{kir}^{(f)}$  correspond to different columns of  $\mathbf{I}_{2K-1+4KR}$ . This subproblem is also a standard SOCP problem, and a global optimal

solution can be acquired by a software package. The robust design of minmax MSE is summarized in Algorithm 4.

---

**Algorithm 4** Robust minmax MSE optimization algorithm.

---

1: Initialization:

Set  $\mathbf{A}_k^{(1)} = \sqrt{(P_{Sk}/L_k)}[\mathbf{I}_{L_k} \mathbf{0}_{(N_k-L_k) \times L_k}]^T \quad \forall k \in \{1, 2, \dots, 2K\}$  and  $\mathbf{F}_r^{(1)} = \sqrt{(P_r/\text{tr}(\mathbf{R}_x))}\mathbf{I}_{M_r} \quad \forall r \in \{1, 2, \dots, R\}, n = 1$ .  
Compute  $\mathbf{a}^{(1)}$  and  $\mathbf{f}^{(1)}$  using (42) and (31), respectively.

2: Iteration:

a) Compute  $\mathbf{B}_k^{(n)}$  using (58).  
b) Compute  $\mathbf{f}^{(n)}$  by solving the SOCP problem of (91).  
c) Compute  $\mathbf{a}^{(n)}$  by solving the SOCP problem of (106).  
d) Compute max MSE  $\varepsilon_k^{(n)}$ .

3: Termination:

The algorithm terminates either when  $\varepsilon_k^{(n)}$  converges, i.e.,  $|(\varepsilon_k^{(n)} - \varepsilon_k^{(n-1)})/\varepsilon_k^{(n)}| \leq \eta$ , or when  $n \geq N_{\max}$ , where  $\eta$  is a predefined threshold, and  $N_{\max}$  is the maximum iteration number;  
Output  $\mathbf{B}_k^{\text{opt}} = \mathbf{B}_k^{(n)}$ ,  $\mathbf{f}^{\text{opt}} = \mathbf{f}^{(n)}$ , and  $\mathbf{a}^{\text{opt}} = \mathbf{a}^{(n)}$ .  
Else,  $n = n + 1$ , and go to substep 2.

---

## V. NUMERICAL RESULTS AND DISCUSSIONS

Here, numerical results are provided to show the effectiveness of our proposed algorithms. All the channels among relays and users satisfy the Kronecker model. The corresponding channel correlation matrices are generated with the exponential model [15], i.e., the entries of the correlation matrices is given by  $\Sigma_{H, rk}(m, n) = \Sigma_{G, kr}(m, n) = \beta^{|m-n|}$  and  $\Psi_{H, rk}(m, n) = \Psi_{G, kr}(m, n) = \alpha^{|m-n|}$ , where  $\alpha$  and  $\beta$  are the correlation coefficients.

We simulate a bidirectional multi-user multi-relay network in which  $N_k = N \quad \forall k = 1, \dots, 2K$ , and  $M_r = M \quad \forall r = 1, \dots, R$ . A vector  $[K \ N \ R \ M]$  denotes the network with  $K$  user pairs, each equipped with  $N$  antennas, and  $R$  relays, each equipped with  $M$  antennas. We assume that the noise variance coefficients at relay nodes and user nodes are the same, i.e.,  $\sigma_r^2 = \sigma_x^2$ . Transmit power at each user node and relay node are  $P_{Sk} = P_S \quad \forall k = 1, \dots, 2K$ ,  $P_{Rr} = P_R, \forall r = 1, \dots, R$ . The average SNR for the two phases of the bidirectional protocol is defined as  $\text{SNR}_1 = P_S/N/\sigma_r^2$  and  $\text{SNR}_2 = P_R/M/\sigma_x^2$ , and we choose  $\text{SNR}_1 = \text{SNR}_2$  for simplicity. The bit-error-rate (BER) results are averaged over 1000 channel realization and 10 000 symbols transmitted from each user node.

### A. Demonstration of BER and MSE Performance with Perfect Channel Knowledge

First, we consider the perfect CSI case. In this case, the relays and users are assumed to have perfect knowledge of all the CSI. Fig. 2 shows the performance of our proposed two algorithms with various iteration numbers under the case  $[2 \ 2 \ 2 \ 2]$  with a single stream for each user. It can be observed that the minmax MSE algorithm always achieves better performance than the proposed min-sum MSE algorithm under the same iteration

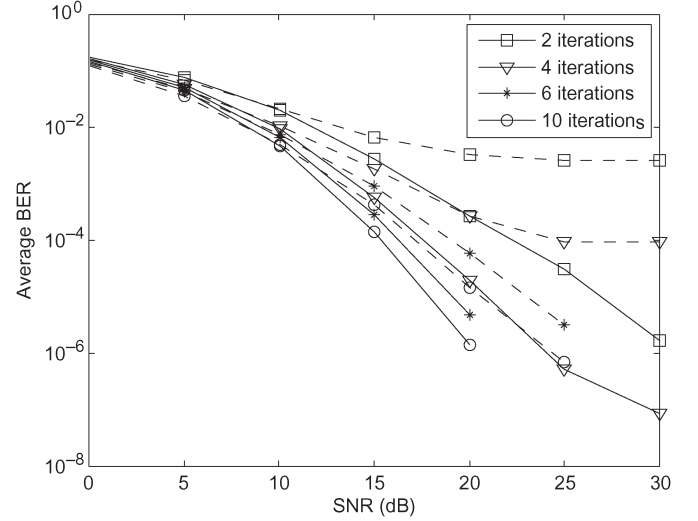


Fig. 2. Perfect channel knowledge. Case  $[2 \ 2 \ 2 \ 2]$  with a single data stream for each user.  $\alpha = 0.5$ , and  $\beta = 0.5$ . The solid line denotes minmax MSE algorithm. The dashed line denotes min-sum MSE algorithm.

number. Generally, the overall BER performance depends on the users' worst MSE. Although min-sum MSE algorithm can achieve a smaller sum MSE value, the imbalanced MSE distribution among different users will largely degrade the system BER performance. On the contrary, the proposed minmax MSE algorithm can guarantee that each user exhibits the same BER performance and, thus, get a lower BER. Furthermore, it can be seen that both algorithms are sensitive to iteration numbers. As the iteration number increases, the BER drops down rapidly.

In addition, in Fig. 3, for the case of  $[2 \ 1 \ 1 \ 4]$ , we compare our proposed min-sum MSE algorithm and minmax MSE algorithm with the algorithm proposed in [26] in terms of BER and sum MSE performance. In [26], the case of multiple user pairs and one bidirectional relay is investigated, and a zero-forcing (ZF)-criterion-based algorithm and an MMSE-criterion-based algorithm are proposed. However, in [26], the structures of transmit–receive beamforming vectors are prefixed, whereas in our algorithms, they are separately designed and can be guaranteed to be optimal in each subproblem. We can observe that our proposed algorithms outperform the algorithm in [26] for both of BER and sum MSE performance. In addition, the min-sum MSE algorithm achieves better sum MSE performance compared with the minmax MSE algorithm since sum MSE is directly optimized by the min-sum MSE algorithm.

Fig. 4 investigates the multi-stream scenario where we consider the case  $[2 \ 2 \ 4 \ 4]$  with dual streams for each user. Moreover, Fig. 5 gives BER performance of  $[4 \ 2 \ 4 \ 4]$ , which involves four pairs of users with a single data stream. Both of the two figures simulate the four- and eight-iteration cases, and better performance can be expected for larger iteration numbers.

We plot the MSE performance in Fig. 6, which is conducted for the case  $[2 \ 2 \ 2 \ 2]$  with a single data stream. It shows that the min-sum MSE algorithm indeed achieve better sum MSE performance compared with the minmax MSE algorithm under the same iteration number. As the iteration number increases, both of the two algorithms tend to reach smaller MSE values.



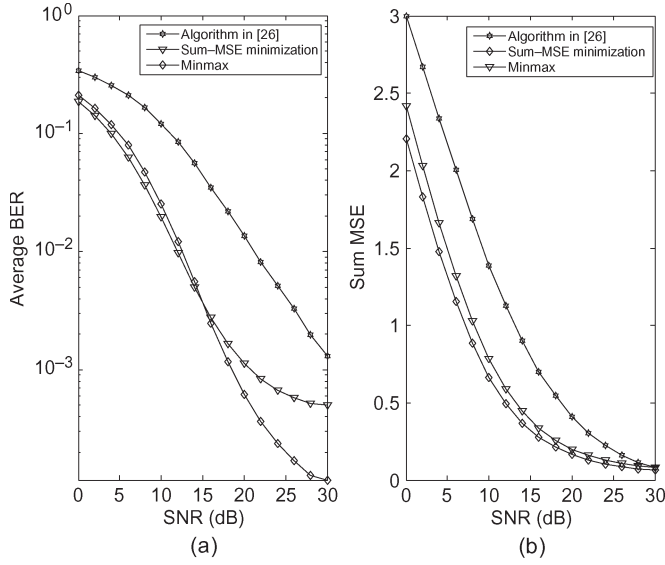


Fig. 3. Perfect channel knowledge. Case [2 1 1 4] with single data stream for each user.  $\alpha = 0.5$ ,  $\beta = 0.5$ . Four iterations.

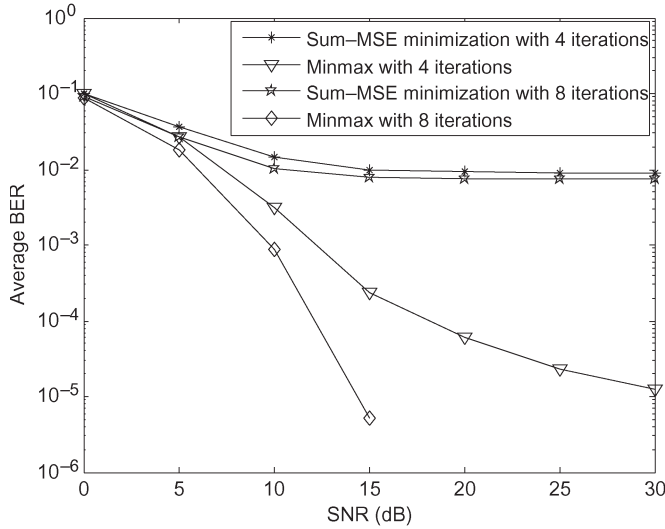


Fig. 4. Perfect channel knowledge. Case [2 2 4 4] with double data streams for each user.  $\alpha = 0.5$ ,  $\beta = 0.5$ .

In addition, it clearly demonstrates that the min-sum MSE algorithm results in a heterogeneous MSE distribution among its users, whereas the minmax MSE algorithm can achieve homogeneous MSE performance.

### B. Demonstration of BER and MSE Performance with Imperfect Channel Knowledge

In Fig. 7, we examine the sum MSE performance for cases of [2 2 2 2] and [2 2 2 4] with the existence of imperfect CSI. It turns out that, with a fixed value of estimation error covariance, the robust min-sum MSE algorithm outperforms the robust minmax MSE algorithm, and as the estimation error covariance decreases, their performance gets better.

Next, we present simulation results for the imperfect channel knowledge case. Similarly, we choose the case [2 2 2 2] with a single data stream for each user in Figs. 8 and 9

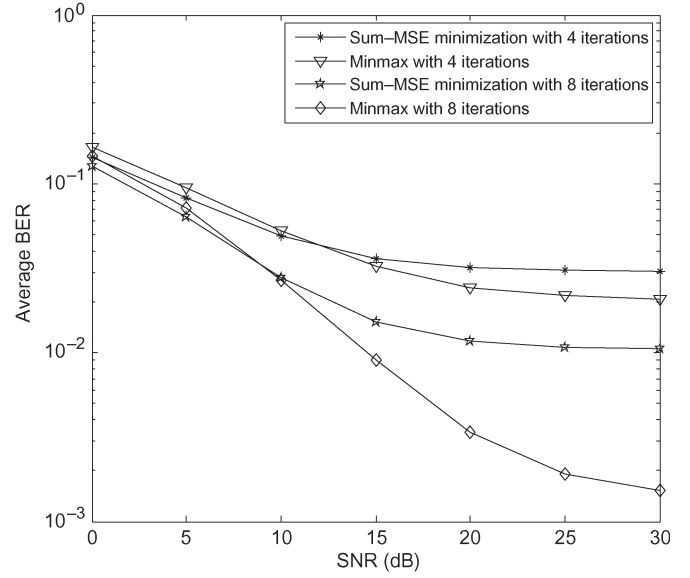


Fig. 5. Perfect channel knowledge. Case [4 2 4 4] with single data stream for each user.  $\alpha = 0.5$ ,  $\beta = 0.5$ .

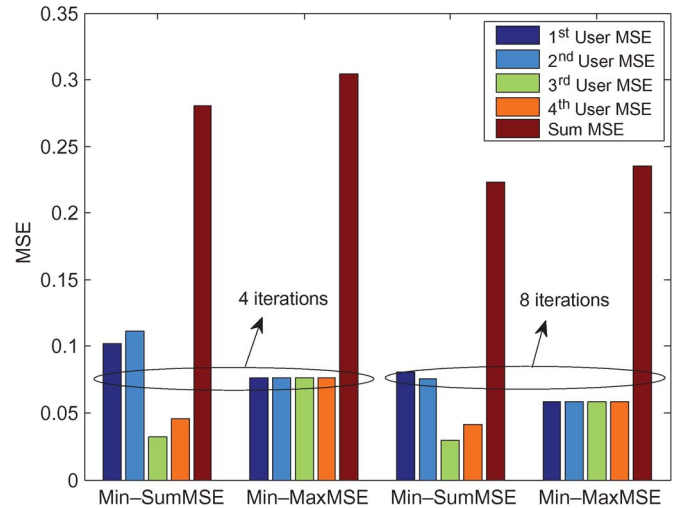


Fig. 6. Perfect channel knowledge. Case [2 2 2 2] with single data stream for each user,  $\alpha = 0.5$ , and  $\beta = 0.5$ . The first ten bars correspond to four iterations, and the last ten bars correspond to eight iterations.

with  $\sigma_e^2 = 0.01$  and  $\sigma_e^2 = 0.005$ , respectively. It can be seen that our proposed robust min-sum MSE algorithm and robust minmax MSE algorithm indeed perform better than their nonrobust versions. In addition, it should be mentioned that, similar to the perfect channel knowledge circumstance, the robust minmax MSE algorithm also exhibit better performance compared with the robust max-sum MSE algorithm. However, it should be mentioned that the BER performance of our proposed nonrobust precoding algorithm will get worse in the high-SNR region. As stated in [27], the ZF-criterion based scheme is optimal among all downlink beamforming strategies at asymptotically optimal at a high SNR. In addition, BER performance of the ZF-criterion-based scheme will be nonincreasing as the SNR increases until it reaches the error floor because it does not take the noise part into account. Since our MSE-optimizing scheme performs better than the ZF scheme

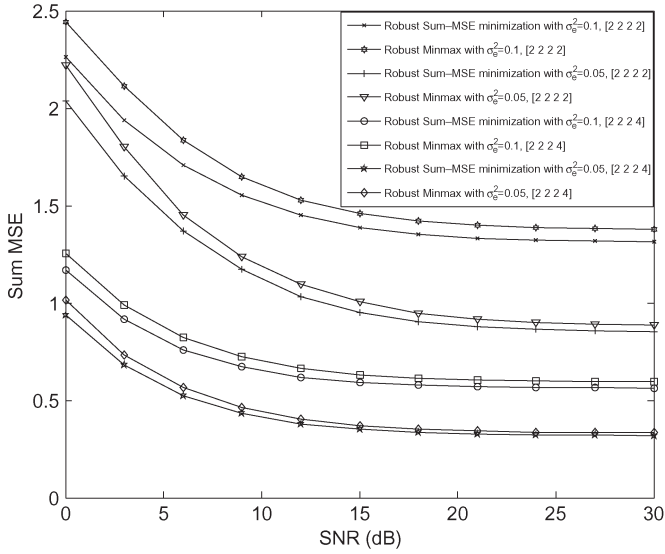


Fig. 7. Imperfect channel knowledge. Cases [2 2 2 2] and [2 2 2 4] with a single data stream for each user.  $\alpha = 0.5$ .  $\beta = 0.5$ . Four iterations.

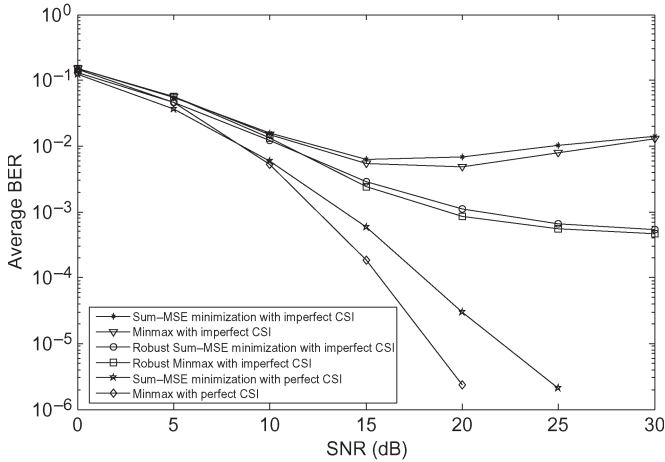


Fig. 8. Imperfect channel knowledge.  $\sigma_e^2 = 0.01$ . Case [2 2 2 2] with a single data stream for each user.  $\alpha = 0.5$ .  $\beta = 0.5$ . Eight iterations.

in the low-SNR region and shall converge to ZF scheme in the high-SNR region, the BER shows getting worse as SNR increases.

We consider the three relays case of [2 2 3 2] in Figs. 10 and 11 with channel estimation errors  $\sigma_e^2 = 0.01$  and  $\sigma_e^2 = 0.02$ , respectively. In addition, in Figs. 12 and 13, we examine the multi-stream case of [2 2 4 4] with double streams for each user with channel estimation errors  $\sigma_e^2 = 0.005$  and  $\sigma_e^2 = 0.01$ , respectively.

## VI. CONCLUSION

We have investigated joint precoding optimization strategies for the bidirectional multi-user multi-relay scenario, where both the perfect and imperfect channel knowledge cases are considered. For the perfect channel knowledge case, iterative algorithms are proposed to optimize iteratively the relay precoders, user precoders, and the MMSE receivers with aims of minimizing sum MSE and maximum user's MSE, respectively. For the imperfect channel knowledge case, the stochastic

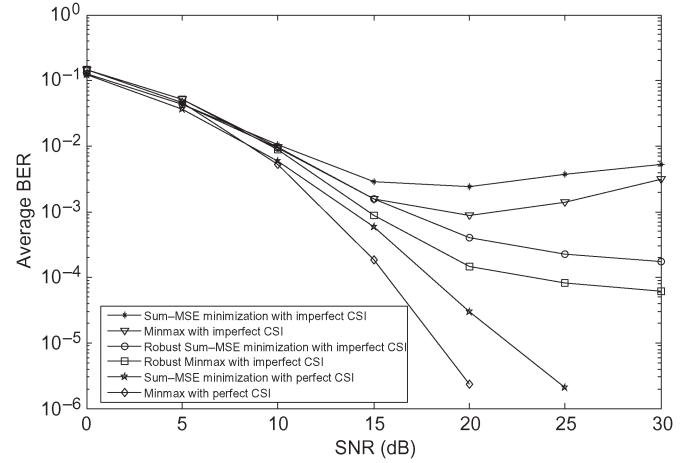


Fig. 9. Imperfect channel knowledge.  $\sigma_e^2 = 0.005$ . Case [2 2 2 2] with a single data stream for each user.  $\alpha = 0.5$ .  $\beta = 0.5$ . Eight iterations.

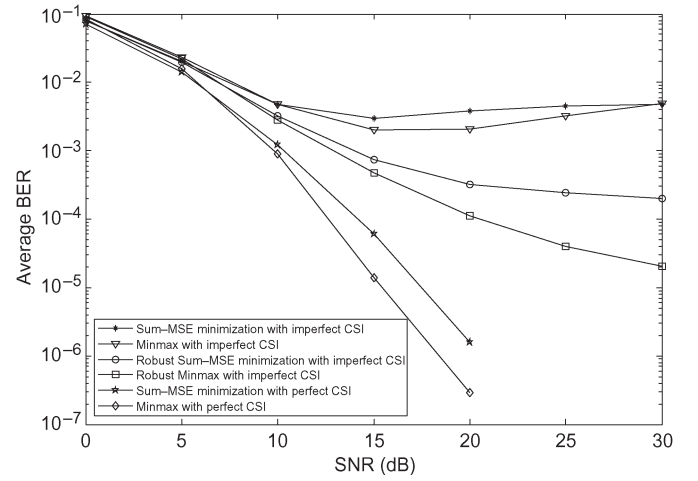


Fig. 10. Imperfect channel knowledge.  $\sigma_e^2 = 0.01$ . Case [2 2 3 2] with a single data stream for each user.  $\alpha = 0.5$ .  $\beta = 0.5$ . Four iterations.

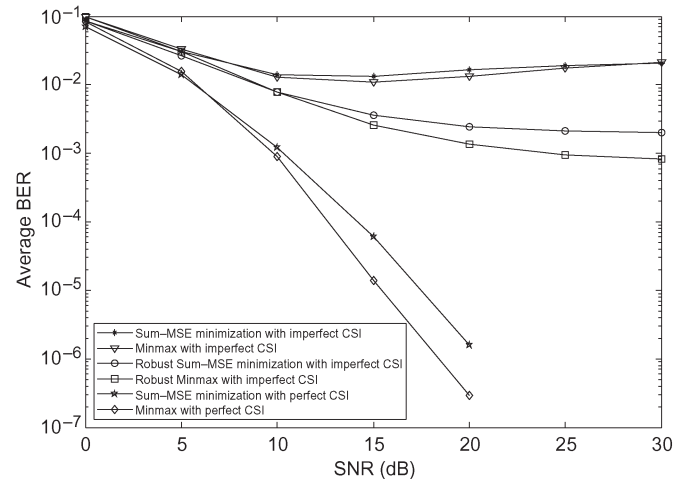


Fig. 11. Imperfect channel knowledge.  $\sigma_e^2 = 0.02$ . Case [2 2 3 2] with a single data stream for each user.  $\alpha = 0.5$ .  $\beta = 0.5$ . Four iterations.

channel estimation error model is considered, and similar iterative algorithms are designed. Simulation results verify the effectiveness of our proposed algorithms.

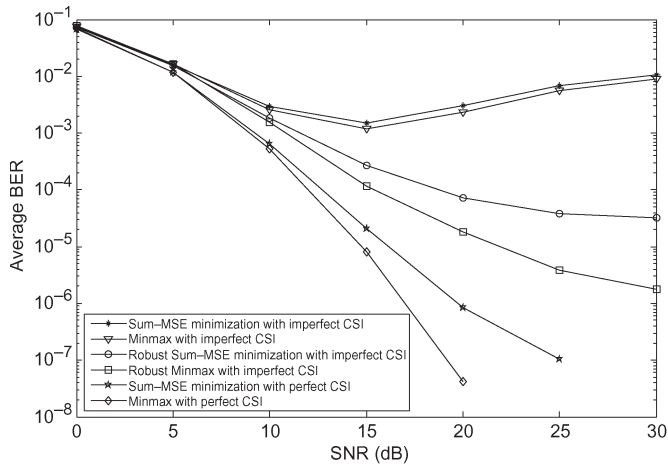


Fig. 12. Imperfect channel knowledge.  $\sigma_e^2 = 0.005$ . Case [2 2 4 4] with double data streams for each user.  $\alpha = 0.5$ .  $\beta = 0.5$ . Eight iterations.

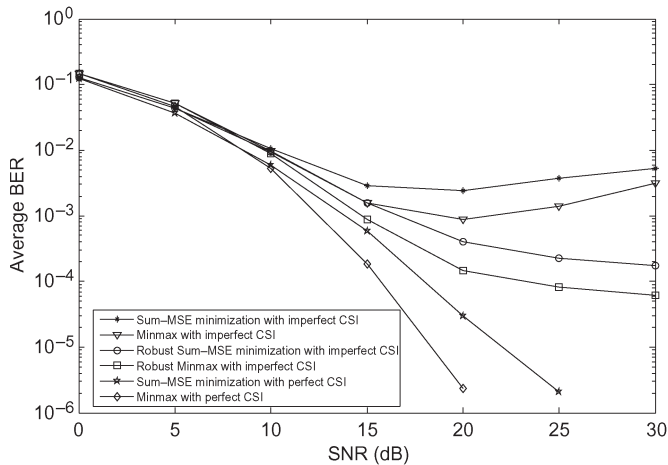


Fig. 13. Imperfect channel knowledge.  $\sigma_e^2 = 0.01$ . Case [2 2 4 4] with double data streams for each user.  $\alpha = 0.5$ .  $\beta = 0.5$ . Eight iterations.

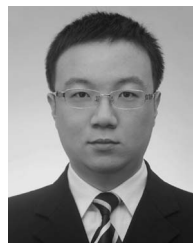
## REFERENCES

- [1] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
- [2] M. Joham, W. Utschick, and J. Nosske, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.
- [3] H. Boelcskei, R. Nabar, O. Oyman, and A. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006.
- [4] C. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, 1961, pp. 611–644.
- [5] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 517–519, Jul. 2008.
- [6] R. Zhang, Y. Liang, C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, Jun. 2009.
- [7] R. Wang and M. Tao, "Joint source and relay precoding designs for MIMO two-way relaying based on MSE criterion," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1352–1365, Mar. 2012.
- [8] Y. Fan, C. Wang, J. Thompson, and H. Poor, "Recovering multiplexing loss through successive relaying using repetition coding," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4484–4493, Dec. 2007.
- [9] H. Wan, W. Chen, and J. Ji, "Efficient linear transmission strategy for MIMO relaying broadcast channels with direct links," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 14–17, Feb. 2012.
- [10] Z. Wang, W. Chen, F. Gao, and J. Li, "Capacity performance of efficient relay beamformings for dual-hop MIMO multi-relay networks with imperfect R-D CSI at relays," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2608–2619, Jul. 2011.
- [11] Y. Rong, "Joint source and relay optimization for two-way MIMO multi-relay networks," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1329–1331, Dec. 2011.
- [12] K. Lee, H. Sung, E. Park, and I. Lee, "Joint optimization for one and two-way MIMO AF multiple-relay systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3671–3681, Dec. 2010.
- [13] C. Sun, C. Yang, Y. Li, and B. Vucetic, "Transceiver optimization for multi-user multi-antenna two-way relay channels," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 2011, pp. 3064–3067.
- [14] H. Yi, J. Zou, H. Luo, H. Yu, and J. Ma, "Joint MMSE precoding design in multi-user two-way MIMO relay systems," in *Proc. IEEE Int. Conf. WCSP*, Nanjing, China, Nov. 2011, pp. 1–5.
- [15] C. Xing, S. Ma, and Y. Wu, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2273–2283, Apr. 2010.
- [16] Z. Wang, W. Chen, and J. Li, "Efficient beamforming for MIMO relaying broadcast channel with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 419–426, Jan. 2012.
- [17] H. Wan and W. Chen, "Joint source and relay design for multi-user MIMO non-regenerative relay networks with direct links," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2871–2876, Jul. 2012.
- [18] P. Ubaidulla and A. Chockalingam, "Robust relay precoder design for MIMO-relay networks," in *Proc. WCNC*, Sydney, Australia, Apr. 2010, pp. 1–6.
- [19] J. Zou, W. Liu, M. Ding, H. Luo, and H. Yu, "Transceiver design for AF MIMO two-way relay systems with imperfect channel estimation," in *Proc. IEEE Global Commun. Conf.*, Houston, TX, USA, Dec. 2011, pp. 1–5.
- [20] M. Zhang, H. Yi, H. Yu, and H. Luo, "Joint optimization in multi-relay multi-user bidirectional systems: Non-robust and robust cases," in *Proc. IEEE Global Commun. Conf.*, Anaheim, CA, USA, Dec. 2012, pp. 5167–5172.
- [21] X. Wang, L. Fu, and C. Hu, "Multicast performance with hierarchical cooperation," *IEEE/ACM Trans. Netw.*, vol. 20, no. 3, pp. 917–930, Jun. 2012.
- [22] X. Wang, W. Huang, S. Wang, J. Zhang, and C. Hu, "Delay and capacity tradeoff analysis for motioncast," *IEEE/ACM Trans. Netw.*, vol. 19, no. 5, pp. 1354–1367, Oct. 2011.
- [23] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [24] R. Horn and C. Johnson, *Topics in Matrix Analysis*. New York, NY, USA: Cambridge Univ. Press, 1991.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [26] J. Joung and A. Sayed, "Multiuser two-way amplify-and-forward relay processing and power control methods for beamforming systems," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1833–1846, Mar. 2010.
- [27] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.



**Meng Zhang** received the B.S. and M.S. degrees in electronic engineering from Southwest Jiaotong University, Chengdu, China, in 2007 and 2010, respectively. He is currently working toward the Ph.D. degree with Shanghai Jiao Tong University, Shanghai, China.

His research interests include relay systems, interference mitigation, and cooperative communications.



**Haike Yi** received the B.S. and M.S. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2010 and 2013, respectively.

He is currently with Marvell Laboratories, Shanghai, China. His research interests include multiple-input-multiple-output optimization, cooperative communications, and nonlinear processing.



**Hui Yu** received the B.S. degree from Tongji University, Shanghai, China, in 1992 and the M.S. degree from Shanghai Jiao Tong University in 1997.

He is currently a Senior Engineer with the Department of Electronic Engineering, Shanghai Jiao Tong University. His research interests include mobile communications, software radio and cognitive radio, and channel coding and modulation for wireless communications.



**Hanwen Luo** received B.S. degree from Shanghai Jiao Tong University, Shanghai, China, in 1977.

He was the Vice Director of the Shanghai Institute of Wireless Communications Technology and the Vice Director of the Institute of Wireless Communication Technology, Shanghai Jiao Tong University, Shanghai, China. He was the leading Specialist with the China 863 High-Tech Program on beyond third-generation wireless communication systems and with the China 973 High-Tech Program on the research of military equipment. He is currently a Full

Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. His research interests include mobile and personal communications.



**Wen Chen** (M'03–SM'11) received the B.S. and M.S. degrees from Wuhan University, Wuhan, China, in 1990 and 1993, respectively, and the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1999.

In 2001, he joined the University of Alberta, Edmonton, AB, Canada, starting as a Postdoctoral Fellow with the Information Research Laboratory and continuing as a Research Associate with the Department of Electrical and Computer Engineering. Since 2006, he has been a Full Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, where he is also the Director of the Institute for Signal Processing and Systems. He is the author of more than 100 papers in IEEE journals and conferences. His research interests include network coding, cooperative communications, cognitive radio, and multiple-input–multiple-output orthogonal frequency-division multiplexing systems.

Dr. Chen was a Researcher of the Japan Society for the Promotion of Sciences from 1999 to 2001.