

Copyright © 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Optimal Source and Relay Design for Multiuser MIMO AF Relay Communication Systems with Direct Links and Imperfect Channel Information

Zhiqiang He, *Member, IEEE*, Xiaonan Zhang, Yunqiang Bi, Weipeng Jiang, and Yue Rong, *Senior Member, IEEE*

Abstract—In this paper, we propose statistically robust design for multiuser multiple-input multiple-output (MIMO) relay systems with direct source-destination links and imperfect channel state information (CSI). The minimum mean-squared error (MMSE) of the signal waveform estimation at the destination node is adopted as the design criterion. We develop two iterative methods to solve the nonconvex joint source, relay, and receiver optimization problem. Simulation results demonstrate the improved robustness of the proposed algorithms against CSI errors.

Index Terms—multiuser, MIMO relay, robust, channel state information, direct link.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) relay communication has attracted much research interest in recent years for its advantages in increasing the coverage and the capacity of wireless networks. In particular, the amplify-and-forward (AF) relay strategy has been extensively investigated in the literature, as it is simpler to implement than the other relay strategies.

In [1]-[4], the source and relay precoding matrices of linear AF MIMO relay systems have been designed under different criteria. The optimal relay precoding matrix maximizing the source-destination capacity has been developed in [1] for AF MIMO relay systems without the direct source-destination link. In [2], the relay precoding matrix minimizing the mean-squared error (MSE) of the signal waveform estimation at the destination node has been developed. A unified framework for joint source and relay matrices optimizing in linear AF multicarrier MIMO relay communication systems has been developed in [3] under the assumption of a negligible direct source-destination link. An overview of the transceiver optimization problems for AF MIMO relay systems can be found in [4].

Manuscript received March 24, 2015; revised August 27, 2015; accepted October 27, 2015. The associate editor coordinating the review of this paper and approving it for publication was Mohamed Alouini.

This work is supported by National High Technology Research and Development Program of China (863 Program 2015AA01A709), National Natural Science Foundation of China (61171099), and the Australian Research Council's Discovery Projects funding scheme (DP140102131).

Z. He, X. Zhang, Y. Bi, and W. Jiang are with the Key Laboratory of Universal Wireless Communication, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mails: hezq@bupt.edu.cn; xiaobei0545@126.com; b_yunqiang@163.com; jw-pqjty@gmail.com).

Y. Rong is with the Department of Electrical and Computer Engineering, Curtin University, Bentley, WA 6102, Australia (e-mail: y.rong@curtin.edu.au).

The direct source-destination link has been ignored in [2] and [3]. However, the direct signal transmission from source to destination provides a spatial copy of the source signals, and thus, should be considered in the MIMO relay system design. In [1], upper and lower bounds on the capacity of AF MIMO systems have been discussed in the presence of the source-destination link, and a suboptimal structure of the relay precoding matrix has been derived. In [5], source and relay precoding matrices design based on a Tomlinson-Harashima precoder has been studied considering the direct source-destination link. A closed-form design of the relay precoding matrix has been proposed in [6]. Relay precoding matrix design based on a modified power constraint has been proposed in [7]. It has been proven in [8] that the optimal relay precoding matrix has a general beamforming structure for most commonly used objective functions.

The transceiver designs in [1]-[3] and [5]-[8] require the exact channel state information (CSI). However, in real communication systems there is always mismatch between the true and the estimated CSI, due to channel noise, quantization errors, and outdated channel estimates. The performance of the algorithms developed assuming the perfect CSI knowledge will degrade in the presence of such CSI mismatch.

CSI mismatch has been taken into account in the MSE-based transceiver design [9] for MIMO relay systems. In [10], statistically robust source and relay matrices design has been developed considering two imperfect CSI scenarios. In [11], a joint optimization of relay and destination matrices has been proposed considering the imperfect CSI at the relay node. Transceiver optimization for a general multi-hop AF MIMO relay system with Gaussian distributed channel uncertainties has been investigated in [12].

Robust transceiver has been developed in [13] for single user MIMO relay systems considering both the direct link and the CSI mismatch, where only a single data stream is transmitted. Recently, joint source and relay design algorithms have been proposed in [14] which support multiple concurrent data streams and consider both the direct link and the CSI mismatch.

In this paper, we investigate the joint source, relay, and receiver matrices design for multiuser MIMO relay systems with direct source-destination links and CSI mismatch. Since multiple source precoding matrices are involved, the system model and the optimization problem are more complicated than the single user case. To our best knowledge, there is no existing work on the robust transceiver design for multiuser

MIMO relay networks considering both the direct links and the CSI mismatch. The works in [14] and [15] can be viewed as special cases of this paper.

The true CSI of each link is modeled as a Gaussian random matrix with the estimated CSI as the mean value and the well-known Kronecker model is adopted for the covariance of the CSI mismatch. The MMSE of the signal waveform estimation at the destination node is adopted as our design criterion. Since the joint source, relay, and receiver matrices optimization problem is nonconvex, a globally optimal solution is computationally intractable. We develop two iterative algorithms to solve the original optimization problem. The key to solve the nonconvex optimization problem is to convert it to an equivalent non-robust MIMO relay design problem with equivalent channel, source, relay, and receiver matrices. The effect of CSI mismatch is shown in the structure of the optimal robust source and relay matrices. Simulation results demonstrate the improved robustness of the proposed algorithms against CSI errors. Interestingly, the computational complexity of the robust MIMO relay design is in the same order as the non-robust approach.

For multiuser MIMO relay systems with multiple relay nodes, the robust source and relay matrices optimization problems are much more challenging than those in a single-relay system [14] and [15], due to the block diagonal structure of the relay precoding matrix and multiple transmission power constraints at the relay nodes. However, we show that the proposed algorithms can be extended to this more general case.

The rest of the paper is organized as follows. In Section II, the model of a two-hop linear AF multiuser MIMO relay communication system considering the CSI mismatch and the direct source-destination links is introduced. The robust source and relay matrices design algorithms are developed in Section III. In Section IV, we extend the proposed algorithms to multiuser MIMO relay systems with multiple relay nodes. In Section V, we show numerical examples to demonstrate the improved robustness of the proposed approaches against the CSI mismatch. Conclusions are drawn in Section VI.

II. SYSTEM MODEL

We consider a two-hop multiuser MIMO communication system as shown in Fig. 1, where K users (node i , $i = 1, \dots, K$) transmit information to the destination node (node d) with the aid of a relay node (node r). The i th user is equipped with N_{si} , $i = 1, \dots, K$ antennas. The relay and destination nodes are equipped with N_r and N_d antennas, respectively. Using a half-duplex relay, the communication process is completed in two time slots. During the first time slot, the $N_{bi} \times 1$ modulated signal vector \mathbf{s}_i is linearly precoded at the i th user by the source precoding matrix $\mathbf{F}_{1i} \in \mathbb{C}^{N_{si} \times N_{bi}}$. The precoded signal vector

$$\mathbf{x}_i = \mathbf{F}_{1i} \mathbf{s}_i \quad (1)$$

is transmitted to the relay node and the destination node. We denote $N_b = \sum_{i=1}^K N_{bi}$ as the total number of independent data streams from all users. The received signal at the relay

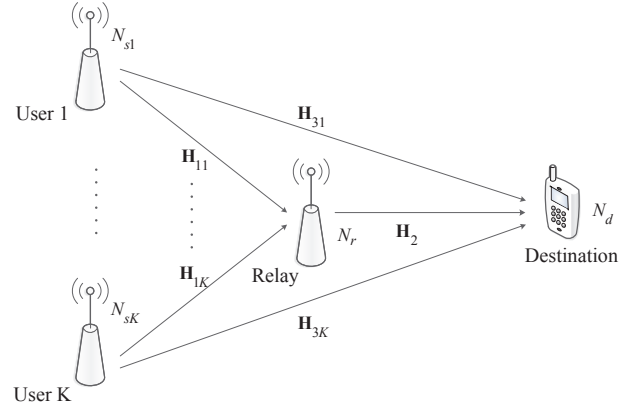


Fig. 1. A Two-Hop Multiuser MIMO Relay Communication System.

node can be written as

$$\mathbf{y}_r = \sum_{i=1}^K \mathbf{H}_{1i} \mathbf{x}_i + \mathbf{n}_r \quad (2)$$

where $\mathbf{H}_{1i} \in \mathbb{C}^{N_r \times N_{si}}$ is the MIMO fading channel matrix of the i th user-relay link, \mathbf{y}_r and \mathbf{n}_r are the received signal vector and the additive Gaussian noise vector at the relay node, respectively.

The received signal at the destination node at the first time slot can be written as

$$\mathbf{y}_{d1} = \sum_{i=1}^K \mathbf{H}_{3i} \mathbf{x}_i + \mathbf{n}_{d1} \quad (3)$$

where $\mathbf{H}_{3i} \in \mathbb{C}^{N_d \times N_{si}}$ is the MIMO fading channel matrix of the i th user-destination link, \mathbf{y}_{d1} and \mathbf{n}_{d1} are the received signal vector and the additive Gaussian noise vector at the destination node, respectively. Substituting (1) into (2) and (3), we have

$$\mathbf{y}_r = \sum_{i=1}^K \mathbf{H}_{1i} \mathbf{F}_{1i} \mathbf{s}_i + \mathbf{n}_r \quad (4)$$

$$\mathbf{y}_{d1} = \sum_{i=1}^K \mathbf{H}_{3i} \mathbf{F}_{1i} \mathbf{s}_i + \mathbf{n}_{d1}. \quad (5)$$

During the second time slot, the users remain silent and the relay node multiplies the received signal vector \mathbf{y}_r by the relay precoding matrix $\mathbf{F}_2 \in \mathbb{C}^{N_r \times N_r}$ and retransmits the signal vector

$$\mathbf{x}_r = \mathbf{F}_2 \mathbf{y}_r \quad (6)$$

to the destination node. The signal vector from the relay node received at the destination node can be written as

$$\mathbf{y}_{d2} = \mathbf{H}_2 \mathbf{x}_r + \mathbf{n}_{d2} \quad (7)$$

where $\mathbf{H}_2 \in \mathbb{C}^{N_d \times N_r}$ is the MIMO fading channel matrix of the relay-destination link, \mathbf{y}_{d2} and \mathbf{n}_{d2} are the received signal vector and the additive Gaussian noise vector at the destination node at the second time slot, respectively.

Substituting (4) and (6) into (7), we obtain

$$\begin{aligned} \mathbf{y}_{d2} &= \mathbf{H}_2 \mathbf{F}_2 \sum_{i=1}^K \mathbf{H}_{1i} \mathbf{F}_{1i} \mathbf{s}_i + \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ &= [\mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_{11} \mathbf{F}_{11}, \dots, \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_{1K} \mathbf{F}_{1K}] \mathbf{s} + \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \end{aligned} \quad (8)$$

where $\mathbf{s} = [\mathbf{s}_1^T, \dots, \mathbf{s}_K^T]^T$, and $(\cdot)^T$ stands for the matrix (vector) transpose.

Combining (5) and (8), the signal received at the destination over two time slots $\mathbf{y} = [\mathbf{y}_{d2}^T, \mathbf{y}_{d1}^T]^T$ can be written as

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_{11} \mathbf{F}_{11}, \dots, \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_{1K} \mathbf{F}_{1K} \\ \mathbf{H}_{31} \mathbf{F}_{11}, \dots, \mathbf{H}_{3K} \mathbf{F}_{1K} \end{bmatrix} \mathbf{s} \\ &\quad + \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ \mathbf{n}_{d1} \end{bmatrix} \\ &= \mathbf{G} \mathbf{s} + \mathbf{v} \end{aligned} \quad (9)$$

where $\mathbf{G} \triangleq \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_{11} \mathbf{F}_{11}, \dots, \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_{1K} \mathbf{F}_{1K} \\ \mathbf{H}_{31} \mathbf{F}_{11}, \dots, \mathbf{H}_{3K} \mathbf{F}_{1K} \end{bmatrix}$ is the equivalent MIMO channel matrix between the source and destination nodes, and $\mathbf{v} \triangleq \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ \mathbf{n}_{d1} \end{bmatrix}$ is the equivalent noise vector. We assume that all noises are independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) with zero mean and unit variance.

In the case of CSI mismatch, the true channel matrices can be modeled as the well-known Gaussian-Kronecker model as in [9] and [14]

$$\mathbf{H}_j \sim \mathcal{CN}(\bar{\mathbf{H}}_j, \boldsymbol{\Theta}_j \otimes \boldsymbol{\Phi}_j), \quad j = 1i, 2, 3i, \quad i = 1, \dots, K \quad (10)$$

where $\bar{\mathbf{H}}_j$ is the estimated channel matrix, $\boldsymbol{\Theta}_j$ and $\boldsymbol{\Phi}_j$ denote the covariance matrix of channel estimation error seen from the transmitter side and the receiver side, respectively, and \otimes stands for the matrix Kronecker product. From (10), we have $\mathbf{H}_j = \bar{\mathbf{H}}_j + \mathbf{A}_{\Phi_j} \mathbf{H}_{w_j} \mathbf{A}_{\Theta_j}^H$, where $\mathbf{A}_{\Phi_j} \mathbf{A}_{\Phi_j}^H = \boldsymbol{\Phi}_j$, $\mathbf{A}_{\Theta_j} \mathbf{A}_{\Theta_j}^H = \boldsymbol{\Theta}_j$, \mathbf{H}_{w_j} is a Gaussian random matrix with i.i.d. zero mean and unit variance entries and is the unknown part in the CSI mismatch. Here, $(\cdot)^H$ denotes the matrix (vector) Hermitian transpose.

In practice, the knowledge of \mathbf{H}_2 and \mathbf{H}_{3i} , $i = 1, \dots, K$, can be obtained at the destination node through channel training, the CSI of \mathbf{H}_{1i} , $i = 1, \dots, K$, can be first obtained at the relay node through channel training and then forwarded to the destination node. In this way, the destination node obtains all CSI required. Then the destination node performs the transceiver optimization and sends the optimized \mathbf{F}_{1i} to the i th source node and \mathbf{F}_2 to the relay node.

Using a linear receiver, the estimated source signal vector at the destination node is given by

$$\hat{\mathbf{s}} = \mathbf{W}^H \mathbf{y} \quad (11)$$

where \mathbf{W} is the $2N_d \times N_b$ receive weight matrix. From (9) and (11), the MSE matrix of the signal waveform estimation at the destination node is a function of \mathbf{W} , $\{\mathbf{F}_{1i}\} \triangleq \{\mathbf{F}_{1i}, i =$

$1, \dots, K\}$, and \mathbf{F}_2 as

$$\begin{aligned} &\mathbf{E}(\mathbf{W}, \{\mathbf{F}_{1i}\}, \mathbf{F}_2) \\ &= E[(\hat{\mathbf{s}} - \mathbf{s})(\hat{\mathbf{s}} - \mathbf{s})^H] \\ &= (\mathbf{W}^H \mathbf{G} - \mathbf{I}_{N_b})(\mathbf{W}^H \mathbf{G} - \mathbf{I}_{N_b})^H + \mathbf{W}^H \mathbf{C}_v \mathbf{W} \\ &= \mathbf{W}^H \mathbf{A} \mathbf{W} - \mathbf{W}^H \mathbf{G} - \mathbf{G}^H \mathbf{W} + \mathbf{I}_{N_b} \end{aligned} \quad (12)$$

where $E[\cdot]$ stands for the statistical expectation with respect to signal and noise, \mathbf{I}_m denotes an $m \times m$ identity matrix, $\mathbf{A} = \mathbf{G} \mathbf{G}^H + \mathbf{C}_v$, and

$$\mathbf{C}_v = E[\mathbf{v} \mathbf{v}^H] = \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{F}_2^H \mathbf{H}_2^H + \mathbf{I}_{N_d} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_d} \end{bmatrix} \quad (13)$$

is the noise covariance matrix. To obtain (12), we assume that $E[\mathbf{s} \mathbf{s}^H] = \mathbf{I}_{N_b}$.

Since the exact CSI is not available at all nodes, there can be a great performance degradation if the estimated channel matrices are simply used to optimize (12), due to the mismatch between \mathbf{H}_j and $\bar{\mathbf{H}}_j$, $j = 1i, 2, 3i, i = 1, \dots, K$. Taking the CSI mismatch into account, we consider the statistical expectation of \mathbf{E} , which is given by

$$E_H[\mathbf{E}(\mathbf{W}, \{\mathbf{F}_{1i}\}, \mathbf{F}_2)] = \mathbf{W}^H \bar{\mathbf{A}} \mathbf{W} - \mathbf{W}^H \bar{\mathbf{G}} - \bar{\mathbf{G}}^H \mathbf{W} + \mathbf{I}_{N_b} \quad (14)$$

where $E_H[\cdot]$ stands for the statistical expectation with respect to the channel matrices, $\bar{\mathbf{A}} \triangleq E_H[\mathbf{A}]$, and $\bar{\mathbf{G}} \triangleq E_H[\mathbf{G}]$.

Since $\{\mathbf{H}_{1i}\}$ and \mathbf{H}_2 are statistically independent, from (9) and (10), we have

$$\begin{aligned} \bar{\mathbf{G}} &= \begin{bmatrix} \bar{\mathbf{H}}_2 \mathbf{F}_2 \bar{\mathbf{H}}_{11} \mathbf{F}_{11}, \dots, \bar{\mathbf{H}}_2 \mathbf{F}_2 \bar{\mathbf{H}}_{1K} \mathbf{F}_{1K} \\ \bar{\mathbf{H}}_{31} \mathbf{F}_{11}, \dots, \bar{\mathbf{H}}_{3K} \mathbf{F}_{1K} \end{bmatrix} \\ \bar{\mathbf{A}} &= E_H[\mathbf{G} \mathbf{G}^H] + E_{H_2}[\mathbf{C}_v]. \end{aligned} \quad (15)$$

Using (13) and the following identity from [16]

$$E_H[\mathbf{H}_i \mathbf{X} \mathbf{H}_i^H] = \bar{\mathbf{H}}_i \mathbf{X} \bar{\mathbf{H}}_i^H + \text{tr}(\mathbf{X} \boldsymbol{\Theta}_i^T) \boldsymbol{\Phi}_i \quad (16)$$

where $\text{tr}(\cdot)$ denotes the matrix trace, we obtain

$$\begin{aligned} E_{H_2}[\mathbf{C}_v] &= \begin{bmatrix} E_{H_2}[\mathbf{H}_2 \mathbf{F}_2 \mathbf{F}_2^H \mathbf{H}_2^H] + \mathbf{I}_{N_d} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_d} \end{bmatrix} \\ &= \bar{\mathbf{C}}_v + \begin{bmatrix} \text{tr}(\mathbf{F}_2 \mathbf{F}_2^H \boldsymbol{\Theta}_2^T) \boldsymbol{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \end{aligned} \quad (17)$$

Here

$$\bar{\mathbf{C}}_v = \begin{bmatrix} \bar{\mathbf{H}}_2 \mathbf{F}_2 \mathbf{F}_2^H \bar{\mathbf{H}}_2^H + \mathbf{I}_{N_d} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_d} \end{bmatrix}.$$

From (10) and (16), we obtain that

$$\begin{aligned} &E_{H_{1i}, H_2} \left(\mathbf{H}_2 \mathbf{F}_2 \sum_{i=1}^K (\mathbf{H}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \mathbf{H}_{1i}^H) \mathbf{F}_2^H \mathbf{H}_2^H \right) \\ &= E_{H_2} \left(\mathbf{H}_2 \mathbf{F}_2 \sum_{i=1}^K E_{H_{1i}} (\mathbf{H}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \mathbf{H}_{1i}^H) \mathbf{F}_2^H \mathbf{H}_2^H \right) \\ &= E_{H_2} \left(\mathbf{H}_2 \mathbf{F}_2 \sum_{i=1}^K (\bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{1i}^H + \alpha_{1i} \boldsymbol{\Phi}_{1i}) \mathbf{F}_2^H \mathbf{H}_2^H \right) \\ &= \bar{\mathbf{H}}_2 \mathbf{F}_2 \sum_{i=1}^K (\bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{1i}^H + \alpha_{1i} \boldsymbol{\Phi}_{1i}) \mathbf{F}_2^H \bar{\mathbf{H}}_2^H + \\ &\quad \text{tr} \left(\mathbf{F}_2 \sum_{i=1}^K (\bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{1i}^H + \alpha_{1i} \boldsymbol{\Phi}_{1i}) \mathbf{F}_2^H \boldsymbol{\Theta}_2^T \right) \boldsymbol{\Phi}_2 \end{aligned} \quad (18)$$

where

$$\alpha_{1i} \triangleq \text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H \boldsymbol{\Theta}_{1i}^T), \quad i = 1, \dots, K. \quad (19)$$

From (9) and (18), we obtain that

$$E_H[\mathbf{G}\mathbf{G}^H] = \bar{\mathbf{G}}\bar{\mathbf{G}}^H + \begin{bmatrix} \sum_{i=1}^K \alpha_{1i} \bar{\mathbf{H}}_2 \mathbf{F}_2 \boldsymbol{\Phi}_{1i} \mathbf{F}_2^H \bar{\mathbf{H}}_2^H + \beta \boldsymbol{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \sum_{i=1}^K \alpha_{3i} \boldsymbol{\Phi}_{3i} \end{bmatrix} \quad (20)$$

where

$$\alpha_{3i} \triangleq \text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H \boldsymbol{\Theta}_{3i}^T), \quad i = 1, \dots, K \quad (21)$$

$$\beta \triangleq \text{tr} \left(\mathbf{F}_2 \sum_{i=1}^K (\bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{1i}^H + \alpha_{1i} \boldsymbol{\Phi}_{1i}) \mathbf{F}_2^H \boldsymbol{\Theta}_2^T \right).$$

Substituting (17) and (20) back into (15), we have

$$\bar{\mathbf{A}} = \bar{\mathbf{G}}\bar{\mathbf{G}}^H + \bar{\mathbf{C}}_v + \mathbf{R} \quad (22)$$

where

$$\mathbf{R} = \begin{bmatrix} \sum_{i=1}^K \alpha_{1i} \bar{\mathbf{H}}_2 \mathbf{F}_2 \boldsymbol{\Phi}_{1i} \mathbf{F}_2^H \bar{\mathbf{H}}_2^H + \alpha_2 \boldsymbol{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \sum_{i=1}^K \alpha_{3i} \boldsymbol{\Phi}_{3i} \end{bmatrix} \quad (23)$$

$$\alpha_2 = \beta + \text{tr}(\mathbf{F}_2 \mathbf{F}_2^H \boldsymbol{\Theta}_2^T). \quad (24)$$

It can be seen from (14) that the CSI mismatch is considered by (23). If the perfect CSI is available, i.e., $\mathbf{H}_j = \bar{\mathbf{H}}_j$ and $\boldsymbol{\Theta}_j = \mathbf{0}$, $j = 1i, 2, 3i, i = 1, \dots, K$, from (23) and (24), there is $\alpha_i = 0$, $i = 1, 2, 3$, and $\mathbf{R} = \mathbf{0}$, then the MSE matrix (14) becomes (12). Therefore, (14) generalizes the MSE matrix from the perfect CSI case to the practical scenario with CSI mismatch.

The transmission power consumed by the i th user and the relay node can be written as $\text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H)$ and $\text{tr}(\mathbf{F}_2 (\sum_{i=1}^K \mathbf{H}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \mathbf{H}_{1i}^H + \mathbf{I}_{N_r}) \mathbf{F}_2^H)$, respectively. However, since the true $\{\mathbf{H}_{1i}\}$ is unknown, we consider the averaged transmission power at the relay node, which is given by

$$E_H \left[\text{tr} \left(\mathbf{F}_2 \left(\sum_{i=1}^K \mathbf{H}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \mathbf{H}_{1i}^H + \mathbf{I}_{N_r} \right) \mathbf{F}_2^H \right) \right] \\ = \text{tr} \left(\mathbf{F}_2 \left(\sum_{i=1}^K (\bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{1i}^H + \alpha_{1i} \boldsymbol{\Phi}_{1i}) + \mathbf{I}_{N_r} \right) \mathbf{F}_2^H \right). \quad (25)$$

From (14) and (25), the robust source, relay, and destination matrices optimization problem can be written as

$$\min_{\mathbf{W}, \{\mathbf{F}_{1i}\}, \mathbf{F}_2} \text{tr}(E_H[\mathbf{E}(\mathbf{W}, \{\mathbf{F}_{1i}\}, \mathbf{F}_2)]) \quad (26)$$

$$\text{s.t. } \text{tr} \left(\mathbf{F}_2 \left(\sum_{i=1}^K (\bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{1i}^H + \alpha_{1i} \boldsymbol{\Phi}_{1i}) + \mathbf{I}_{N_r} \right) \mathbf{F}_2^H \right) \leq P_r \quad (27)$$

$$\text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H) \leq P_{si}, \quad i = 1, \dots, K \quad (28)$$

where P_{si} and P_r are the transmission power available at the i th user and the relay node, respectively. The problem (26)-(28) is nonconvex with matrix variables.

III. PROPOSED ROBUST MIMO RELAY DESIGN ALGORITHMS

In this section, we develop two iterative algorithms namely the Tri-Step and the Bi-Step algorithms to optimize the source, relay, and receive matrices. In the Tri-Step algorithm, the source, relay, and receive matrices are optimized iteratively through solving convex sub-problems. For the Bi-Step algorithm, the optimal receive matrix is substituted into the objective function, so we obtain an optimization problem only with the source and relay matrices. Then, the source and relay matrices are optimized alternately and the receive matrix is calculated after the convergence of the algorithm.

By introducing $\mathbf{P}_1 \triangleq \sum_{i=1}^K \alpha_{1i} \boldsymbol{\Phi}_{1i} + \mathbf{I}_{N_r}$, $\mathbf{P}_2 \triangleq \alpha_2 \boldsymbol{\Phi}_2 + \mathbf{I}_{N_d}$, and $\mathbf{P}_3 \triangleq \sum_{i=1}^K \alpha_{3i} \boldsymbol{\Phi}_{3i} + \mathbf{I}_{N_d}$, (14) can be rewritten as

$$E_H[\mathbf{E}(\mathbf{W}, \{\mathbf{F}_{1i}\}, \mathbf{F}_2)] \\ = [\mathbf{W}_1^H \quad \mathbf{W}_2^H] \mathbf{Z} \mathbf{M} \mathbf{Z} \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} - [\mathbf{W}_1^H \quad \mathbf{W}_2^H] \mathbf{Z} \mathbf{Z}^{-1} \bar{\mathbf{G}} \\ - \bar{\mathbf{G}}^H \mathbf{Z}^{-1} \mathbf{Z} \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} + \mathbf{I}_{N_b} \quad (29)$$

where $\mathbf{Z} \triangleq \text{bd}(\mathbf{P}_2^{\frac{1}{2}}, \mathbf{P}_3^{\frac{1}{2}})$, \mathbf{W}_1 and \mathbf{W}_2 contain the first and the last N_d rows of \mathbf{W} , respectively, and \mathbf{M} is given by (30) shown at the bottom of the page. Here $\text{bd}(\cdot)$ stands for a block diagonal matrix and $(\cdot)^{-1}$ denotes the matrix inversion.

By introducing $\tilde{\mathbf{W}}_1^H \triangleq \mathbf{W}_1^H \mathbf{P}_2^{\frac{1}{2}}$, $\tilde{\mathbf{W}}_2^H \triangleq \mathbf{W}_2^H \mathbf{P}_3^{\frac{1}{2}}$, $\tilde{\mathbf{H}}_2 \triangleq \mathbf{P}_2^{-\frac{1}{2}} \bar{\mathbf{H}}_2$, $\tilde{\mathbf{H}}_{1i} \triangleq \mathbf{P}_1^{-\frac{1}{2}} \bar{\mathbf{H}}_{1i}$, $\tilde{\mathbf{H}}_{3i} \triangleq \mathbf{P}_3^{-\frac{1}{2}} \bar{\mathbf{H}}_{3i}$, $i = 1, \dots, K$, and $\tilde{\mathbf{F}}_2 \triangleq \mathbf{F}_2 \mathbf{P}_1^{\frac{1}{2}}$, (29) can be rewritten as

$$E_H[\mathbf{E}(\tilde{\mathbf{W}}, \{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2)] \\ = [\tilde{\mathbf{W}}_1^H \quad \tilde{\mathbf{W}}_2^H] (\tilde{\mathbf{G}} \tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v) \begin{bmatrix} \tilde{\mathbf{W}}_1 \\ \tilde{\mathbf{W}}_2 \end{bmatrix} \\ - [\tilde{\mathbf{W}}_1^H \quad \tilde{\mathbf{W}}_2^H] \tilde{\mathbf{G}} - \tilde{\mathbf{G}}^H \begin{bmatrix} \tilde{\mathbf{W}}_1 \\ \tilde{\mathbf{W}}_2 \end{bmatrix} + \mathbf{I}_{N_b} \\ = (\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_b}) (\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_b})^H + \tilde{\mathbf{W}}^H \tilde{\mathbf{C}}_v \tilde{\mathbf{W}} \quad (31)$$

where

$$\tilde{\mathbf{G}} = \begin{bmatrix} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_{11} \mathbf{F}_{11}, \dots, \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_{1K} \mathbf{F}_{1K} \\ \tilde{\mathbf{H}}_{31} \mathbf{F}_{11}, \dots, \tilde{\mathbf{H}}_{3K} \mathbf{F}_{1K} \end{bmatrix}, \quad \tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}}_1 \\ \tilde{\mathbf{W}}_2 \end{bmatrix} \\ \tilde{\mathbf{C}}_v = \begin{bmatrix} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \mathbf{I}_{N_d} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_d} \end{bmatrix}.$$

Using (31), the optimization problem (26)-(28) can be equivalently rewritten as

$$\min_{\tilde{\mathbf{W}}, \{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2} \text{tr}(E_H[\mathbf{E}(\tilde{\mathbf{W}}, \{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2)]) \quad (32)$$

$$\text{s.t. } \text{tr} \left(\tilde{\mathbf{F}}_2 \left(\sum_{i=1}^K \tilde{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{1i}^H + \mathbf{I}_{N_r} \right) \tilde{\mathbf{F}}_2^H \right) \leq P_r \quad (33)$$

$$\text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H) \leq P_{si}, \quad i = 1, \dots, K \quad (34)$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{P}_2^{-\frac{1}{2}} \bar{\mathbf{H}}_2 \mathbf{F}_2 \mathbf{P}_1^{\frac{1}{2}} (\mathbf{P}_1^{-\frac{1}{2}} \sum_{i=1}^K \bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{1i}^H \mathbf{P}_1^{-\frac{1}{2}} + \mathbf{I}_{N_r}) \mathbf{P}_1^{\frac{1}{2}} \mathbf{F}_2^H \bar{\mathbf{H}}_2^H \mathbf{P}_2^{-\frac{1}{2}} + \mathbf{I}_{N_d} & \mathbf{P}_2^{-\frac{1}{2}} \bar{\mathbf{H}}_2 \mathbf{F}_2 \sum_{i=1}^K \bar{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{3i}^H \mathbf{P}_3^{-\frac{1}{2}} \\ \mathbf{P}_3^{-\frac{1}{2}} \sum_{i=1}^K \bar{\mathbf{H}}_{3i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{3i}^H \mathbf{P}_3^{-\frac{1}{2}} + \mathbf{I}_{N_d} & \mathbf{P}_3^{-\frac{1}{2}} \sum_{i=1}^K \bar{\mathbf{H}}_{3i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \bar{\mathbf{H}}_{3i}^H \mathbf{P}_3^{-\frac{1}{2}} + \mathbf{I}_{N_d} \end{bmatrix} \quad (30)$$

where the variable substitution of $\tilde{\mathbf{F}}_2$ and $\tilde{\mathbf{H}}_{1i}$, $i = 1, \dots, K$, is used to rewrite the power constraint (27) at the relay node to obtain (33). It is worth noting that the robust precoding matrices design problem (26)-(28) for multiuser MIMO relay systems with imperfect CSI is converted to the problem (32)-(34) for an equivalent multiuser MIMO relay system with direct links and perfect CSI, where the channel matrices are $\tilde{\mathbf{H}}_2$, $\tilde{\mathbf{H}}_{1i}$, $\tilde{\mathbf{H}}_{3i}$, $i = 1, \dots, K$, the source precoding matrices are $\tilde{\mathbf{F}}_{1i}$, $i = 1, \dots, K$, the relay precoding matrix is $\tilde{\mathbf{F}}_2$, and the receive matrix is $\tilde{\mathbf{W}}$.

A. The Tri-Step Algorithm

The problem (32)-(34) is nonconvex with matrix variables and the globally optimal solution is difficult to obtain. In the following, we develop a Tri-Step iterative approach to solve the problem (32)-(34).

Firstly, with given $\{\mathbf{F}_{1i}\}$ and $\tilde{\mathbf{F}}_2$, the weight matrix $\tilde{\mathbf{W}}$ minimizing (32) is the famous Wiener filter [17] (MMSE receiver) given by

$$\tilde{\mathbf{W}} = (\tilde{\mathbf{G}}\tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v)^{-1}\tilde{\mathbf{G}}. \quad (35)$$

Secondly, with given $\tilde{\mathbf{W}}$ and $\{\mathbf{F}_{1i}\}$, $\tilde{\mathbf{F}}_2$ can be updated by solving the following problem

$$\min_{\tilde{\mathbf{F}}_2} tr((\tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_1 - \mathbf{\Pi})(\tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_1 - \mathbf{\Pi})^H + \tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{F}}_2^H\tilde{\mathbf{H}}_2^H + \tilde{\mathbf{W}}_1^H\tilde{\mathbf{W}}_1 + \tilde{\mathbf{W}}_2^H\tilde{\mathbf{W}}_2) \quad (36)$$

$$\text{s.t. } tr(\tilde{\mathbf{F}}_2(\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_r})\tilde{\mathbf{F}}_2^H) \leq P_r \quad (37)$$

where $\tilde{\mathbf{H}}_1 \triangleq [\tilde{\mathbf{H}}_{11}\mathbf{F}_{11}, \dots, \tilde{\mathbf{H}}_{1K}\mathbf{F}_{1K}]$, $\tilde{\mathbf{H}}_2 \triangleq \tilde{\mathbf{W}}_1^H\tilde{\mathbf{H}}_2$, $\mathbf{\Pi} \triangleq \mathbf{I}_{N_b} - \tilde{\mathbf{W}}_2^H\tilde{\mathbf{H}}_3$, $\tilde{\mathbf{H}}_3 \triangleq [\tilde{\mathbf{H}}_{31}\mathbf{F}_{11}, \dots, \tilde{\mathbf{H}}_{3K}\mathbf{F}_{1K}]$.

Using the Lagrange multiplier method, we can solve the problem (36)-(37) and obtain $\tilde{\mathbf{F}}_2$ as

$$\tilde{\mathbf{F}}_2 = \check{\mathbf{H}}_2^H(\check{\mathbf{H}}_2\check{\mathbf{H}}_2^H + \mu\mathbf{I}_{N_b})^{-1}\mathbf{\Pi}\check{\mathbf{H}}_1^H(\check{\mathbf{H}}_1\check{\mathbf{H}}_1^H + \mathbf{I}_{N_r})^{-1} \quad (38)$$

where $\mu \geq 0$ is the Lagrangian multiplier and can be found from the following complementary slackness condition

$$\mu(tr(\tilde{\mathbf{F}}_2(\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_r})\tilde{\mathbf{F}}_2^H) - P_r) = 0. \quad (39)$$

Assuming that $\mu = 0$, from (38) we have

$$\tilde{\mathbf{F}}_2 = \check{\mathbf{H}}_2^H(\check{\mathbf{H}}_2\check{\mathbf{H}}_2^H)^{-1}\mathbf{\Pi}\check{\mathbf{H}}_1^H(\check{\mathbf{H}}_1\check{\mathbf{H}}_1^H + \mathbf{I}_{N_r})^{-1}. \quad (40)$$

If $\tilde{\mathbf{F}}_2$ in (40) satisfies the power constraint (37), then (40) is the solution to the problem (36)-(37). Otherwise, there must be $\mu > 0$ such that

$$tr(\tilde{\mathbf{F}}_2(\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_r})\tilde{\mathbf{F}}_2^H) = P_r. \quad (41)$$

In this case, μ can be obtained by substituting (38) into (41) and solving the following nonlinear equation

$$tr(\check{\mathbf{H}}_2^H(\check{\mathbf{H}}_2\check{\mathbf{H}}_2^H + \mu\mathbf{I}_{N_b})^{-1}\mathbf{\Pi}\check{\mathbf{H}}_1^H(\check{\mathbf{H}}_1\check{\mathbf{H}}_1^H + \mathbf{I}_{N_r})^{-1} \times \check{\mathbf{H}}_1\mathbf{\Pi}^H(\check{\mathbf{H}}_2\check{\mathbf{H}}_2^H + \mu\mathbf{I}_{N_b})^{-1}\check{\mathbf{H}}_2) = P_r. \quad (42)$$

By introducing the singular value decomposition (SVD) of

$$\check{\mathbf{H}}_2 = \mathbf{U}_2\mathbf{\Lambda}_2\mathbf{V}_2^H \quad (43)$$

we obtain from (42) that

$$tr(\mathbf{\Lambda}_2(\mathbf{\Lambda}_2^2 + \mu\mathbf{I}_{N_b})^{-1}\mathbf{U}_2^H\mathbf{\Pi}\check{\mathbf{H}}_1^H(\check{\mathbf{H}}_1\check{\mathbf{H}}_1^H + \mathbf{I}_{N_r})^{-1} \times \check{\mathbf{H}}_1\mathbf{\Pi}^H\mathbf{U}_2(\mathbf{\Lambda}_2^2 + \mu\mathbf{I}_{N_b})^{-1}\mathbf{\Lambda}_2) = P_r. \quad (44)$$

Denoting $\mathbf{\Gamma} \triangleq \mathbf{U}_2^H\mathbf{\Pi}\check{\mathbf{H}}_1^H(\check{\mathbf{H}}_1\check{\mathbf{H}}_1^H + \mathbf{I}_{N_r})^{-1}\check{\mathbf{H}}_1\mathbf{\Pi}^H\mathbf{U}_2$, (44) can be equivalently written as

$$\sum_{i=1}^{N_b} \frac{\lambda_i^2 \gamma_i}{(\lambda_i^2 + \mu)^2} = P_r \quad (45)$$

where λ_i and γ_i are the i th main diagonal elements of $\mathbf{\Lambda}_2$ and $\mathbf{\Gamma}$ respectively. Since the left-hand side of (45) is monotonically decreasing with respect to μ , the bisection method [18] can be applied to solve (45) to obtain μ .

Thirdly, with given $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{F}}_2$, we show that the problem (32)-(34) can be cast as a quadratically constrained quadratic programming (QCQP) problem [18] to optimize $\{\mathbf{F}_{1i}\}$. Let us introduce $\mathbf{D}_{1i} \triangleq \tilde{\mathbf{W}}_1^H\tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_{1i} + \tilde{\mathbf{W}}_2^H\tilde{\mathbf{H}}_{3i}$, $i = 1, \dots, K$, and $\tilde{\mathbf{D}}_{1i}$ as a matrix containing the $(\sum_{j=1}^{i-1} N_{bj} + 1)$ th to the $(\sum_{j=1}^i N_{bj})$ th rows of \mathbf{D}_{1i} . By using the identity of $tr(\mathbf{C}^T\mathbf{D}) = (\text{vec}(\mathbf{C}))^T\text{vec}(\mathbf{D})$ and $\text{vec}(\mathbf{CD}) = (\mathbf{I} \otimes \mathbf{C})\text{vec}(\mathbf{D})$ [19], where $\text{vec}(\mathbf{X})$ stands for a column vector obtained by stacking all columns of \mathbf{X} on top of each other, (32) can be rewritten as

$$\begin{aligned} & tr((\tilde{\mathbf{W}}_1^H\tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_1 + \tilde{\mathbf{W}}_2^H\tilde{\mathbf{H}}_3 - \mathbf{I}_{N_b})(\tilde{\mathbf{W}}_1^H\tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_1 \\ & \quad + \tilde{\mathbf{W}}_2^H\tilde{\mathbf{H}}_3 - \mathbf{I}_{N_b})^H) + t_1 \\ & = tr(([\mathbf{D}_{11}\mathbf{F}_{11}, \dots, \mathbf{D}_{1K}\mathbf{F}_{1K}] - \mathbf{I}_{N_b}) \\ & \quad \times ([\mathbf{D}_{11}\mathbf{F}_{11}, \dots, \mathbf{D}_{1K}\mathbf{F}_{1K}] - \mathbf{I}_{N_b})^H) + t_1 \\ & = \sum_{i=1}^K (tr(\mathbf{D}_{1i}\mathbf{F}_{1i}\mathbf{F}_{1i}^H\mathbf{D}_{1i}^H) - tr(\tilde{\mathbf{D}}_{1i}\mathbf{F}_{1i}) - tr(\mathbf{F}_{1i}^H\tilde{\mathbf{D}}_{1i}^H)) + t_2 \\ & = \sum_{i=1}^K (\mathbf{f}_{1i}^H(\mathbf{I}_{N_{s_i}} \otimes (\mathbf{D}_{1i}^H\mathbf{D}_{1i}))\mathbf{f}_{1i} \\ & \quad - \text{vec}(\tilde{\mathbf{D}}_{1i}^H)\mathbf{f}_{1i} - \mathbf{f}_{1i}^H\text{vec}(\tilde{\mathbf{D}}_{1i}^H)) + t_2 \\ & = \mathbf{f}_1^H\mathbf{\Upsilon}_1\mathbf{f}_1 - \mathbf{d}_1^H\mathbf{f}_1 - \mathbf{f}_1^H\mathbf{d}_1 + t_2 \\ & = (\mathbf{f}_1^H\mathbf{\Upsilon}_1^{\frac{1}{2}} - \mathbf{d}_1^H\mathbf{\Upsilon}_1^{-\frac{1}{2}})(\mathbf{\Upsilon}_1^{\frac{1}{2}}\mathbf{f}_1 - \mathbf{\Upsilon}_1^{-\frac{1}{2}}\mathbf{d}_1) + t_3 \end{aligned} \quad (46)$$

where

$$t_1 \triangleq tr(\tilde{\mathbf{W}}^H\tilde{\mathbf{C}}_v\tilde{\mathbf{W}}), \quad t_2 \triangleq t_1 + N_b,$$

$$\mathbf{f}_1 \triangleq [\mathbf{f}_{11}^H, \dots, \mathbf{f}_{1K}^H]^H, \quad t_3 \triangleq t_2 - \mathbf{d}_1^H\mathbf{\Upsilon}_1^{-1}\mathbf{d}_1,$$

$$\mathbf{f}_{1i} \triangleq \text{vec}(\mathbf{F}_{1i}), \quad \mathbf{d}_1 \triangleq [\text{vec}(\tilde{\mathbf{D}}_{11}^H)^H, \dots, \text{vec}(\tilde{\mathbf{D}}_{1K}^H)^H]^H,$$

$$\mathbf{\Upsilon}_1 \triangleq \text{bd}(\mathbf{I}_{N_{s_1}} \otimes (\mathbf{D}_{1i}^H\mathbf{D}_{1i}), \dots, \mathbf{I}_{N_{s_K}} \otimes (\mathbf{D}_{1K}^H\mathbf{D}_{1K})).$$

Note that we can ignore t_3 while optimizing \mathbf{f}_1 with given $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{F}}_2$, since it does not depend on the optimization variable \mathbf{f}_1 . By introducing $\mathbf{D}_{2i} \triangleq \tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_{1i}$, $i = 1, \dots, K$, the relay transmit power constraint in (33) can be rewritten as

$$\mathbf{f}_1^H\mathbf{\Upsilon}_2\mathbf{f}_1 + tr(\tilde{\mathbf{F}}_2\tilde{\mathbf{F}}_2^H) \leq P_r \quad (47)$$

where $\mathbf{\Upsilon}_2 \triangleq \text{bd}(\mathbf{I}_{N_{s_1}} \otimes (\mathbf{D}_{21}^H\mathbf{D}_{21}), \dots, \mathbf{I}_{N_{s_K}} \otimes (\mathbf{D}_{2K}^H\mathbf{D}_{2K}))$. Using (46) and (47), the problem (32)-(34) can be equivalently

rewritten as the following QCQP problem

$$\min_{\mathbf{f}_1} (\mathbf{f}_1^H \mathbf{\Upsilon}_1^{\frac{1}{2}} - \mathbf{d}_1^H \mathbf{\Upsilon}_1^{-\frac{1}{2}})(\mathbf{\Upsilon}_1^{\frac{1}{2}} \mathbf{f}_1 - \mathbf{\Upsilon}_1^{-\frac{1}{2}} \mathbf{d}_1) \quad (48)$$

$$\text{s.t. } \mathbf{f}_1^H \mathbf{\Upsilon}_2 \mathbf{f}_1 \leq P_r - \text{tr}(\tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H) \quad (49)$$

$$\mathbf{f}_1^H \tilde{\mathbf{I}}_i \mathbf{f}_1 \leq P_{si}, \quad i = 1, \dots, K \quad (50)$$

where $\tilde{\mathbf{I}}_i \triangleq \text{bd}(\tilde{\mathbf{I}}_{i1}, \dots, \tilde{\mathbf{I}}_{iK})$, with $\tilde{\mathbf{I}}_{ii} = \mathbf{I}_{N_{si}N_{bi}}$ and $\tilde{\mathbf{I}}_{ij} = \mathbf{0}$, $j = 1, \dots, K$, $j \neq i$. The problem (48)-(50) can be efficiently solved by the disciplined convex programming toolbox CVX [20].

TABLE I
PROCEDURE OF THE TRI-STEP ALGORITHM

- 1) Initialize the algorithm with $\mathbf{F}_{1i}^{(0)} = \sqrt{P_{si}/N_{bi}}[\mathbf{I}_{N_{bi}}, \mathbf{0}]^T$, $i = 1, \dots, K$, and $\tilde{\mathbf{F}}_2^{(0)} = \sqrt{P_r/\text{tr}(\check{\mathbf{H}}_1 \check{\mathbf{H}}_1^H + \mathbf{I}_{N_r})} \mathbf{I}_{N_r}$; Set $n = 0$.
- 2) Update $\tilde{\mathbf{W}}^{(n)}$ using $\tilde{\mathbf{F}}_2^{(n)}$ and $\{\mathbf{F}_{1i}^{(n)}\}$ as (35).
- 3) Update $\tilde{\mathbf{F}}_2^{(n+1)}$ using $\tilde{\mathbf{W}}^{(n)}$ and $\{\mathbf{F}_{1i}^{(n)}\}$ as (38).
- 4) Update $\{\mathbf{F}_{1i}^{(n+1)}\}$ using $\tilde{\mathbf{W}}^{(n)}$ and $\tilde{\mathbf{F}}_2^{(n+1)}$ by solving the problem (48)-(50).
- 5) if $(\text{mse}_1^{(n)} - \text{mse}_1^{(n+1)})/\text{mse}_1^{(n)} < \varepsilon$, iteration ends; otherwise go to step (2).

The procedure of applying the Tri-Step iterative algorithm to solve the problem (32)-(34) is listed in Table I, where the superscript (n) denotes the number of iterations, ε is a small positive number close to zero, and $\text{mse}_1^{(n)}$ stands for the value of (32) at the n th iteration.

B. The Bi-Step Algorithm

By substituting (35) back into (31), we have

$$\mathbf{E}_0(\{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2) = \mathbf{I}_{N_b} - \tilde{\mathbf{G}}^H (\tilde{\mathbf{G}} \tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v)^{-1} \tilde{\mathbf{G}} \quad (51)$$

where \mathbf{E}_0 stands for the MSE matrix when a linear MMSE receiver is used at the destination. The source and relay matrices optimization problem can be written as

$$\min_{\{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2} \text{tr}(\mathbf{E}_0(\{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2)) \quad (52)$$

$$\text{s.t. } \text{tr}(\tilde{\mathbf{F}}_2(\check{\mathbf{H}}_1 \check{\mathbf{H}}_1^H + \mathbf{I}_{N_r}) \tilde{\mathbf{F}}_2^H) \leq P_r \quad (53)$$

$$\text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H) \leq P_{si}, \quad i = 1, \dots, K. \quad (54)$$

The problem (52)-(54) is nonconvex with matrix variables and the globally optimal solution is difficult to obtain. In the following, we develop an iterative approach to solve the problem (52)-(54).

It can be shown similar to [6] that for given source precoding matrices $\{\mathbf{F}_{1i}\}$, the optimal $\tilde{\mathbf{F}}_2$ as the solution to the problem (52)-(54) has the structure of

$$\tilde{\mathbf{F}}_2 = \mathbf{T} \mathbf{L} \quad (55)$$

where \mathbf{T} is an $N_r \times N_b$ matrix that remains to be optimized, and

$$\mathbf{L} = (\check{\mathbf{H}}_1^H \check{\mathbf{H}}_1 + \mathbf{Q})^{-1} \check{\mathbf{H}}_1^H, \quad \mathbf{Q} = \check{\mathbf{H}}_3^H \check{\mathbf{H}}_3 + \mathbf{I}_{N_b}.$$

Let us introduce a positive semi-definite (PSD) matrix $\mathbf{\Omega} = \mathbf{L}(\check{\mathbf{H}}_1 \mathbf{Q}^{-1} \check{\mathbf{H}}_1^H + \mathbf{I}_{N_r}) \mathbf{L}^H$ and its eigenvalue decomposition (EVD)

$$\mathbf{\Omega} = \mathbf{U}_\omega \mathbf{\Lambda}_\omega \mathbf{U}_\omega^H \quad (56)$$

where $\mathbf{\Lambda}_\omega$ is the diagonal eigenvalue matrix with eigenvalues $\lambda_{\omega,k}$, $k = 1, \dots, N_b$, arranged in descending order. Let us also introduce the EVD of

$$\check{\mathbf{H}}_2^H \check{\mathbf{H}}_2 = \mathbf{U}_h \mathbf{\Lambda}_h \mathbf{U}_h^H \quad (57)$$

where $\mathbf{\Lambda}_h$ is the diagonal eigenvalue matrix with eigenvalues $\lambda_{h,k}$, $k = 1, \dots, N_r$, arranged in descending order.

Based on the result in [6], \mathbf{T} has the structure of

$$\mathbf{T} = \mathbf{U}_{h,1} \mathbf{\Delta} \mathbf{U}_\omega^H \quad (58)$$

where $\mathbf{U}_{h,1}$ contains the leftmost N_b columns of \mathbf{U}_h , $\mathbf{\Delta}$ is a diagonal matrix and the solution to the following problem

$$\min_{\mathbf{\Delta}} \text{tr} \left((\mathbf{\Delta}^H \mathbf{\Lambda}_{h,1} \mathbf{\Delta} + \mathbf{\Lambda}_\omega^{-1})^{-1} \right) \quad (59)$$

$$\text{s.t. } \text{tr}(\mathbf{\Delta} \mathbf{R}_\omega \mathbf{\Delta}^H) \leq P_r. \quad (60)$$

Here $\mathbf{R}_\omega \triangleq \mathbf{U}_\omega^H \mathbf{L}(\check{\mathbf{H}}_1 \check{\mathbf{H}}_1^H + \mathbf{I}_{N_r}) \mathbf{L}^H \mathbf{U}_\omega$ and $\mathbf{\Lambda}_{h,1}$ contains the largest N_b diagonal elements of $\mathbf{\Lambda}_h$. The problem (59)-(60) can be efficiently solved by the Lagrange multiplier method as

$$|\delta_k|^2 = \frac{1}{\lambda_{\omega,k} \lambda_{h,k}} \left(\sqrt{\frac{\lambda_{\omega,k}^2 \lambda_{h,k}}{\gamma R_k}} - 1 \right)^+, \quad k = 1, \dots, N_b$$

where δ_k is the k th main diagonal element of $\mathbf{\Delta}$, $(x)^+ \triangleq \max(x, 0)$, $R_k \triangleq [\mathbf{R}_\omega]_{k,k}$, and $\gamma > 0$ is the Lagrangian multiplier and the solution to the following nonlinear equation

$$\sum_{k=1}^{N_b} \frac{R_k}{\lambda_{\omega,k} \lambda_{h,k}} \left(\sqrt{\frac{\lambda_{\omega,k}^2 \lambda_{h,k}}{\gamma R_k}} - 1 \right)^+ = P_r.$$

We would like to mention that to obtain the exact solution of the relay precoding matrix in (55), $\mathbf{\Delta}$ in (58) should be considered as a general matrix. However, by limiting $\mathbf{\Delta}$ to be diagonal, a closed-form solution of the relay precoding matrix with a low computational complexity can be obtained similar to [6]. Obviously, such diagonality constraint may degrade the performance of the Bi-Step algorithm. We would like to note that such performance-complexity tradeoff is very useful for practical multiuser MIMO relay communication systems.

Now we start to optimize the source precoding matrices $\{\mathbf{F}_{1i}\}$. Using the matrix inversion lemma, (51) can be rewritten as

$$\begin{aligned} & \text{tr}(\mathbf{E}_0(\{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2)) \\ &= \text{tr}((\mathbf{I}_{N_b} + \tilde{\mathbf{G}}^H \tilde{\mathbf{C}}_v^{-1} \tilde{\mathbf{G}})^{-1}) \\ &= \text{tr}((\mathbf{I}_{2N_d} + \tilde{\mathbf{G}} \tilde{\mathbf{G}}^H \tilde{\mathbf{C}}_v^{-1})^{-1}) + N_b - 2N_d \\ &= \text{tr}((\mathbf{I}_{2N_d} + \sum_{i=1}^K \Psi_i \mathbf{B}_i \Psi_i^H)^{-1}) + N_b - 2N_d \end{aligned} \quad (61)$$

where $\Psi_i \triangleq \tilde{\mathbf{C}}_v^{-\frac{1}{2}} \begin{bmatrix} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_{1i} \\ \tilde{\mathbf{H}}_{3i} \end{bmatrix}$ and $\mathbf{B}_i \triangleq \mathbf{F}_{1i} \mathbf{F}_{1i}^H$, $i = 1, \dots, K$. From (61), for given relay precoding matrix $\tilde{\mathbf{F}}_2$,

TABLE II
PROCEDURE OF THE BI-STEP ALGORITHM

- 1) Initialize the algorithm with $\mathbf{F}_{1i}^{(0)} = \sqrt{P_{si}/N_{bi}}[\mathbf{I}_{N_{bi}}, \mathbf{0}]^T$, $i = 1, \dots, K$; Set $n = 0$.
- 2) Update $\tilde{\mathbf{F}}_2^{(n+1)}$ using $\{\mathbf{F}_{1i}^{(n)}\}$ based on (55) and (58) by solving the problem (59)-(60).
- 3) Update $\{\mathbf{F}_{1i}^{(n+1)}\}$ using $\tilde{\mathbf{F}}_2^{(n+1)}$ by solving the problem (65)-(68).
- 4) if $(\text{mse}_2^{(n)} - \text{mse}_2^{(n+1)})/\text{mse}_2^{(n)} < \varepsilon$, iteration ends; otherwise go to step (2).

$\{\mathbf{B}_i\} \triangleq \{\mathbf{B}_i, i = 1, \dots, K\}$ are optimized by solving the following problem

$$\min_{\{\mathbf{B}_i\}} \text{tr} \left(\left(\mathbf{I}_{2N_d} + \sum_{i=1}^K \Psi_i \mathbf{B}_i \Psi_i^H \right)^{-1} \right) \quad (62)$$

$$\text{s.t.} \quad \sum_{i=1}^K \text{tr}(\mathbf{B}_i \tilde{\mathbf{H}}_{1i}^H \tilde{\mathbf{F}}_2^H \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_{1i}) \leq P_r - \text{tr}(\tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H) \quad (63)$$

$$\text{tr}(\mathbf{B}_i) \leq P_{si}, \quad \mathbf{B}_i \succeq \mathbf{0}, \quad i = 1, \dots, K. \quad (64)$$

Let us introduce a PSD matrix \mathbf{X} with $\mathbf{X} \succeq (\mathbf{I}_{2N_d} + \sum_{i=1}^K \Psi_i \mathbf{B}_i \Psi_i^H)^{-1}$, where $\mathbf{A} \succeq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is a PSD matrix. The problem (62)-(64) can be equivalently converted to the following convex semi-definite programming (SDP) problem by using the Schur complement

$$\min_{\mathbf{X}, \{\mathbf{B}_i\}} \text{tr}(\mathbf{X}) \quad (65)$$

$$\text{s.t.} \quad \begin{pmatrix} \mathbf{X} & & \\ & \mathbf{I}_{2N_d} & \\ & & \mathbf{I}_{2N_d} + \sum_{i=1}^K \Psi_i \mathbf{B}_i \Psi_i^H \end{pmatrix} \succeq \mathbf{0} \quad (66)$$

$$\sum_{i=1}^K \text{tr}(\mathbf{B}_i \tilde{\mathbf{H}}_{1i}^H \tilde{\mathbf{F}}_2^H \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_{1i}) \leq P_r - \text{tr}(\tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H) \quad (67)$$

$$\text{tr}(\mathbf{B}_i) \leq P_{si}, \quad \mathbf{B}_i \succeq \mathbf{0}, \quad i = 1, \dots, K. \quad (68)$$

The problem (65)-(68) can be efficiently solved by the interior-point method [18]. Using the EVD of $\mathbf{B}_i = \mathbf{U}_{bi} \mathbf{\Lambda}_{bi} \mathbf{U}_{bi}^H$, we have $\mathbf{F}_{1i} = \mathbf{U}_{bi} \mathbf{\Lambda}_{bi}^{\frac{1}{2}}$.

The procedure of using the Bi-Step iterative algorithm to solve the problem (52)-(54) is listed in Table II, where $\text{mse}_2^{(n)}$ stands for the value of (52) at the n th iteration.

IV. EXTENSION TO MULTIUSER MIMO RELAY SYSTEMS WITH MULTIPLE RELAY NODES

In this section, we extend the proposed Tri-Step and Bi-Step algorithms to multiuser MIMO relay systems with direct links and multiple relay nodes.

A. System Model

We consider a system with K users, L relay nodes, and one destination. The signal vector received at the l th relay node at the first time slot is

$$\mathbf{y}_{rl} = \sum_{i=1}^K \mathbf{H}_{1li} \mathbf{F}_{1i} s_i + \mathbf{n}_{rl}, \quad l = 1, \dots, L \quad (69)$$

where \mathbf{H}_{1li} is the MIMO channel matrix between the l th relay node and the i th user, \mathbf{n}_{rl} is the noise vector at the l th relay.

The signal vector received at the destination through the direct links at the first time slot is the same as (5).

The signal transmitted by the l th relay node at the second time slot is

$$\mathbf{x}_{rl} = \mathbf{F}_{2l} \mathbf{y}_{rl}, \quad l = 1, \dots, L \quad (70)$$

where \mathbf{F}_{2l} is the precoding matrix at the l th relay node. The signal received at the destination node through L relay nodes at the second time slot is

$$\mathbf{y}_{d2} = \sum_{l=1}^L \mathbf{H}_{2l} \mathbf{x}_{rl} + \mathbf{n}_{d2}. \quad (71)$$

From (69)-(71), we have

$$\mathbf{y}_{d2} = \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \mathbf{s} + \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \quad (72)$$

where $\mathbf{H}_2 = [\mathbf{H}_{21}, \dots, \mathbf{H}_{2L}]$, $\mathbf{F}_2 = \text{bd}(\mathbf{F}_{21}, \dots, \mathbf{F}_{2L})$, $\mathbf{H}_1 = \begin{pmatrix} \mathbf{H}_{111} & \dots & \mathbf{H}_{11K} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{1L1} & \dots & \mathbf{H}_{1LK} \end{pmatrix}$, and $\mathbf{n}_r = [\mathbf{n}_{r1}^T, \dots, \mathbf{n}_{rL}^T]^T$.

From (5) and (72), the received signals over two time slots are

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_{d1} \\ \mathbf{y}_{d2} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \\ \mathbf{H}_3 \end{bmatrix} \mathbf{F}_1 \mathbf{s} + \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ \mathbf{n}_{d1} \end{bmatrix} = \mathbf{G} \mathbf{s} + \mathbf{v} \quad (73)$$

where $\mathbf{G} \triangleq \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \\ \mathbf{H}_3 \end{bmatrix} \mathbf{F}_1$ is the equivalent MIMO channel matrix between the source and destination nodes, and $\mathbf{v} \triangleq \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ \mathbf{n}_{d1} \end{bmatrix}$ is the equivalent noise vector. It can be seen from (73) that in systems with multiple relay nodes, \mathbf{F}_2 is a block diagonal matrix. This is different to the single-relay system, where there is no constraint on the structure of \mathbf{F}_2 .

As systems with one relay node, the MSE of the signal waveform estimation at the destination is given by

$$\text{MSE} = \text{tr}((\mathbf{W}^H \mathbf{G} - \mathbf{I}_{N_b})(\mathbf{W}^H \mathbf{G} - \mathbf{I}_{N_b})^H + \mathbf{W}^H \mathbf{C}_v \mathbf{W}) \quad (74)$$

where $\mathbf{C}_v = E[\mathbf{v} \mathbf{v}^H]$ is the noise covariance matrix.

We assume the true channel matrices are $\mathbf{H}_j \sim \mathcal{CN}(\bar{\mathbf{H}}_j, \mathbf{\Theta}_j \otimes \mathbf{\Phi}_j)$, $j = 1, 2, 3$, $l = 1, \dots, L$, $i = 1, \dots, K$. Now we work out $E_H[\mathbf{W}^H \mathbf{A} \mathbf{W} - \mathbf{W}^H \mathbf{G} - \mathbf{G}^H \mathbf{W} + \mathbf{I}_{N_b}]$, where

$$\begin{aligned} \mathbf{A} &= \mathbf{G} \mathbf{G}^H + \mathbf{C}_v \\ &= \begin{pmatrix} \mathbf{B} & \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_3^H \\ \mathbf{H}_3 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{F}_2^H \mathbf{H}_2^H & \mathbf{H}_3 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_3^H + \mathbf{I}_{N_d} \end{pmatrix} \end{aligned}$$

and $\mathbf{B} = \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{F}_2^H \mathbf{H}_2^H + \mathbf{H}_2 \mathbf{F}_2 \mathbf{F}_2^H \mathbf{H}_2^H + \mathbf{I}_{N_d}$. According to (16), we have

$$\begin{aligned} &E_H[\mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{F}_2^H \mathbf{H}_2^H] \\ &= E_H \left[\sum_{k=1}^K \left(\sum_{l=1}^L \mathbf{H}_{2l} \mathbf{F}_{2l} \mathbf{H}_{1lk} \mathbf{F}_{1k} \right) \left(\sum_{l=1}^L \mathbf{F}_{1k}^H \mathbf{H}_{1lk}^H \mathbf{F}_{2l}^H \mathbf{H}_{2l}^H \right) \right] \\ &= \sum_{i=1}^K \left(\sum_{l=1}^L \tilde{\mathbf{H}}_{2l} \mathbf{F}_{2l} \tilde{\mathbf{H}}_{1li} \mathbf{F}_{1i} \right) \left(\sum_{l=1}^L \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{1li}^H \mathbf{F}_{2l}^H \tilde{\mathbf{H}}_{2l}^H \right) \\ &+ \sum_{i=1}^K \sum_{l=1}^L \alpha_{1li} \tilde{\mathbf{H}}_{2l} \mathbf{F}_{2l} \tilde{\mathbf{\Phi}}_{1li} \mathbf{F}_{2l}^H \tilde{\mathbf{H}}_{2l}^H + \sum_{l=1}^L \beta_l \tilde{\mathbf{\Phi}}_{2l} \quad (75) \end{aligned}$$

where $\beta_l = \text{tr}(\mathbf{F}_{2l} \sum_{i=1}^K (\tilde{\mathbf{H}}_{1li} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{1li}^H + \alpha_{1li} \Phi_{1li}) \mathbf{F}_{2l}^H \Theta_{2l}^T)$, $\alpha_{1li} = \text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H \Theta_{1li}^T)$. And we obtain that

$$\begin{aligned} & E_H[\mathbf{H}_2 \mathbf{F}_2 \mathbf{F}_2^H \mathbf{H}_2^H] \\ &= E_H \left[\sum_{l=1}^L \mathbf{H}_{2l} \mathbf{F}_{2l} \mathbf{F}_{2l}^H \mathbf{H}_{2l}^H \right] \\ &= \sum_{l=1}^L \tilde{\mathbf{H}}_{2l} \mathbf{F}_{2l} \mathbf{F}_{2l}^H \tilde{\mathbf{H}}_{2l}^H + \sum_{l=1}^L \text{tr}(\mathbf{F}_{2l} \mathbf{F}_{2l}^H \Theta_{2l}^T) \Phi_{2l} \end{aligned} \quad (76)$$

$$\begin{aligned} & E_H[\mathbf{H}_3 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_3^H] \\ &= E_H \left[\sum_{i=1}^K \mathbf{H}_{3i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \mathbf{H}_{3i}^H \right] \\ &= \sum_{i=1}^K \tilde{\mathbf{H}}_{3i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{3i}^H + \sum_{i=1}^K \alpha_{3i} \Phi_{3i} \end{aligned} \quad (77)$$

where $\alpha_{3i} = \text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H \Theta_{3i}^T)$. Let us introduce $\mathbf{P}_{1l} = \sum_{i=1}^K \alpha_{1li} \Phi_{1li} + \mathbf{I}_{N_l}$, $\mathbf{P}_2 = \sum_{l=1}^L \alpha_{2l} \Phi_{2l} + \mathbf{I}_{N_d}$, $\mathbf{P}_3 = \sum_{i=1}^K \alpha_{3i} \Phi_{3i} + \mathbf{I}_{N_d}$, $\tilde{\mathbf{H}}_{1li} = \mathbf{P}_{1l}^{-\frac{1}{2}} \tilde{\mathbf{H}}_{1li}$ and $\tilde{\mathbf{F}}_{2l} = \mathbf{F}_{2l} \mathbf{P}_{1l}^{\frac{1}{2}}$, where $\alpha_{2l} = \beta_l + \text{tr}(\mathbf{F}_{2l} \mathbf{F}_{2l}^H \Theta_{2l}^T) = \text{tr}(\mathbf{F}_{2l} (\sum_{k=1}^K \tilde{\mathbf{H}}_{1lk} \mathbf{F}_{1k} \mathbf{F}_{1k}^H \tilde{\mathbf{H}}_{1lk}^H + \mathbf{P}_{1l}) \mathbf{F}_{2l}^H \Theta_{2l}^T)$. We have from (75)-(77) that

$$\begin{aligned} & E_H[\mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{F}_2^H \mathbf{H}_2^H + \mathbf{H}_2 \mathbf{F}_2 \mathbf{F}_2^H \mathbf{H}_2^H + \mathbf{I}_{N_d}] \\ &= \sum_{i=1}^K \left(\sum_{l=1}^L \tilde{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \tilde{\mathbf{H}}_{1li} \mathbf{F}_{1i} \right) \left(\sum_{l=1}^L \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{1li}^H \tilde{\mathbf{F}}_{2l}^H \tilde{\mathbf{H}}_{2l}^H \right) \\ &+ \sum_{l=1}^L \tilde{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \tilde{\mathbf{F}}_{2l}^H \tilde{\mathbf{H}}_{2l}^H + \mathbf{P}_2 \\ &= \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \mathbf{P}_2 \end{aligned} \quad (78)$$

$$E_H[\mathbf{H}_3 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_3^H + \mathbf{I}_{N_d}] = \tilde{\mathbf{H}}_3 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_3^H + \mathbf{P}_3 \quad (79)$$

where $\tilde{\mathbf{H}}_2 = [\tilde{\mathbf{H}}_{21}, \dots, \tilde{\mathbf{H}}_{2L}]$, $\tilde{\mathbf{F}}_2 = \text{bd}(\tilde{\mathbf{F}}_{21}, \dots, \tilde{\mathbf{F}}_{2L})$, $\tilde{\mathbf{H}}_1 = \begin{pmatrix} \tilde{\mathbf{H}}_{111} & \dots & \tilde{\mathbf{H}}_{11K} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{H}}_{1L1} & \dots & \tilde{\mathbf{H}}_{1LK} \end{pmatrix}$, and $\tilde{\mathbf{H}}_3 = [\tilde{\mathbf{H}}_{31}, \dots, \tilde{\mathbf{H}}_{3K}]$.

From (78) and (79) we have

$$\begin{aligned} E_H[\mathbf{A}] &= \begin{pmatrix} \tilde{\mathbf{B}} & \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_3^H \\ \tilde{\mathbf{H}}_3 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H & \tilde{\mathbf{H}}_3 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_3^H + \mathbf{P}_3 \end{pmatrix} \\ &= \mathbf{Z}(\tilde{\mathbf{G}} \tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v) \mathbf{Z}^H \end{aligned} \quad (80)$$

where $\tilde{\mathbf{B}} = \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \mathbf{P}_2$, $\mathbf{Z} = \text{bd}(\mathbf{P}_2^{\frac{1}{2}}, \mathbf{P}_3^{\frac{1}{2}})$, $\tilde{\mathbf{H}}_2 = \mathbf{P}_2^{-\frac{1}{2}} \tilde{\mathbf{H}}_2$, $\tilde{\mathbf{H}}_3 = \mathbf{P}_3^{-\frac{1}{2}} \tilde{\mathbf{H}}_3$, and $\tilde{\mathbf{G}} = \begin{bmatrix} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_3 \end{bmatrix} \mathbf{F}_1$, $\tilde{\mathbf{C}}_v = \begin{pmatrix} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \mathbf{I}_{N_d} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_d} \end{pmatrix}$. From (80), the expectation of the MSE in (74) can be written as

$$\begin{aligned} & E_H[\text{MSE}] \\ &= \text{tr} \left([\mathbf{W}_1^H, \mathbf{W}_2^H] \mathbf{Z} \left(\tilde{\mathbf{G}} \tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v \right) \mathbf{Z}^H [\mathbf{W}_1^H, \mathbf{W}_2^H]^H - \right. \\ & \quad \left. [\mathbf{W}_1^H, \mathbf{W}_2^H] \mathbf{Z} \mathbf{Z}^{-1} \tilde{\mathbf{G}} - \tilde{\mathbf{G}}^H \mathbf{Z}^{-1} \mathbf{Z} [\mathbf{W}_1^H, \mathbf{W}_2^H]^H + \mathbf{I}_{N_d} \right) \\ &= \text{tr} \left(\left(\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_d} \right) \left(\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_d} \right)^H + \tilde{\mathbf{W}}^H \tilde{\mathbf{C}}_v \tilde{\mathbf{W}} \right) \end{aligned} \quad (81)$$

where $\tilde{\mathbf{W}} = [\tilde{\mathbf{W}}_1^H, \tilde{\mathbf{W}}_2^H]^H$, $\tilde{\mathbf{W}}_1 = \mathbf{P}_2^{\frac{1}{2}} \mathbf{W}_1$, and $\tilde{\mathbf{W}}_2 = \mathbf{P}_3^{\frac{1}{2}} \mathbf{W}_2$. The power consumption at the l th relay node is

$$\begin{aligned} & E_H \left[\text{tr} \left(\mathbf{F}_{2l} \left(\sum_{i=1}^K \mathbf{H}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \mathbf{H}_{1i}^H + \mathbf{I}_{N_l} \right) \mathbf{F}_{2l}^H \right) \right] \\ &= \text{tr} \left(\tilde{\mathbf{F}}_{2l} \left(\sum_{i=1}^K \tilde{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{1i}^H + \mathbf{I}_{N_l} \right) \tilde{\mathbf{F}}_{2l}^H \right). \end{aligned} \quad (82)$$

From (81) and (82), the optimization problem can be written as

$$\min_{\tilde{\mathbf{W}}, \mathbf{F}_1, \tilde{\mathbf{F}}_2} E_H[\text{MSE}] \quad (83)$$

$$\text{s.t. } \text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H) \leq P_{si}, \quad i = 1, \dots, K \quad (84)$$

$$\text{tr} \left(\tilde{\mathbf{F}}_{2l} \left(\sum_{i=1}^K \tilde{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{1i}^H + \mathbf{I}_{N_l} \right) \tilde{\mathbf{F}}_{2l}^H \right) \leq P_{rl}, \quad l = 1, \dots, L. \quad (85)$$

Note that constraints in (85) and the block diagonal structure of $\tilde{\mathbf{F}}_2$ make the problem (83)-(85) more challenging than that in single-relay system [14], [15].

B. The Tri-Step Algorithm

We can iteratively update $\tilde{\mathbf{W}}$, \mathbf{F}_1 , and $\tilde{\mathbf{F}}_2$ using the Tri-Step algorithm. Updating $\tilde{\mathbf{W}}$ is similar to that in the single-relay system (35). To update a particular $\tilde{\mathbf{F}}_{2l}$, we keep $\tilde{\mathbf{W}}$, \mathbf{F}_1 , and $\tilde{\mathbf{F}}_{2j}$, $j = 1, \dots, L$, $j \neq l$ fixed. Then $E_H[\text{MSE}]$ can be rewritten as follows

$$\begin{aligned} & E_H[\text{MSE}] \\ &= \text{tr} \left(\left(\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_d} \right) \left(\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_d} \right)^H + \tilde{\mathbf{W}}^H \tilde{\mathbf{C}}_v \tilde{\mathbf{W}} \right) \\ &= \text{tr} \left((\hat{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \hat{\mathbf{H}}_{1l} - \mathbf{\Pi}_l) (\hat{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \hat{\mathbf{H}}_{1l} - \mathbf{\Pi}_l)^H + \hat{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \tilde{\mathbf{F}}_{2l}^H \hat{\mathbf{H}}_{2l}^H \right. \\ & \quad \left. + \sum_{i=1, i \neq l}^L \hat{\mathbf{H}}_{2i} \tilde{\mathbf{F}}_{2i} \tilde{\mathbf{F}}_{2i}^H \hat{\mathbf{H}}_{2i}^H + \tilde{\mathbf{W}}_1^H \tilde{\mathbf{W}}_1 + \tilde{\mathbf{W}}_2^H \tilde{\mathbf{W}}_2 \right) \end{aligned} \quad (86)$$

where $\hat{\mathbf{H}}_{1l} = [\hat{\mathbf{H}}_{1l1} \mathbf{F}_{11}, \dots, \hat{\mathbf{H}}_{1lK} \mathbf{F}_{1K}]$, $\hat{\mathbf{H}}_{2l} = \tilde{\mathbf{W}}_1^H \tilde{\mathbf{H}}_{2l}$, $\hat{\mathbf{H}}_3 = [\tilde{\mathbf{H}}_{31} \mathbf{F}_{11}, \dots, \tilde{\mathbf{H}}_{3K} \mathbf{F}_{1K}]$, and $\mathbf{\Pi}_l = \mathbf{I}_{N_d} - \sum_{i=1, i \neq l}^L \hat{\mathbf{H}}_{2i} \tilde{\mathbf{F}}_{2i} \hat{\mathbf{H}}_{1i} - \tilde{\mathbf{W}}_2^H \hat{\mathbf{H}}_3$.

Thus, the problem of optimizing $\tilde{\mathbf{F}}_{2l}$ becomes

$$\begin{aligned} & \min_{\tilde{\mathbf{F}}_{2l}} \text{tr} \left((\hat{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \hat{\mathbf{H}}_{1l} - \mathbf{\Pi}_l) (\hat{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \hat{\mathbf{H}}_{1l} - \mathbf{\Pi}_l)^H \right. \\ & \quad \left. + \hat{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \tilde{\mathbf{F}}_{2l}^H \hat{\mathbf{H}}_{2l}^H \right) \end{aligned} \quad (87)$$

$$\text{s.t. } \text{tr}(\tilde{\mathbf{F}}_{2l} (\hat{\mathbf{H}}_{1l} \hat{\mathbf{H}}_{1l}^H + \mathbf{I}_{N_l}) \tilde{\mathbf{F}}_{2l}^H) \leq P_{rl}. \quad (88)$$

The problem (87)-(88) is similar to the problem (36)-(37). Thus, $\tilde{\mathbf{F}}_{2l}$, $l = 1, \dots, L$, can be optimized by using the Lagrange multiplier method. After we update all $\tilde{\mathbf{F}}_{2l}$, we obtain $\tilde{\mathbf{F}}_2 = \text{bd}(\tilde{\mathbf{F}}_{21}, \dots, \tilde{\mathbf{F}}_{2L})$.

Updating \mathbf{F}_1 is similar to that in single-relay systems. The difference is that $\mathbf{D}_{1i} \triangleq \sum_{l=1}^L \tilde{\mathbf{W}}_1^H \tilde{\mathbf{H}}_{2l} \tilde{\mathbf{F}}_{2l} \tilde{\mathbf{H}}_{1li} + \tilde{\mathbf{W}}_2^H \tilde{\mathbf{H}}_{3i}$ instead, and the power constraint (49) is extended to L power constraints as

$$\mathbf{f}_1^H \Upsilon_{2l} \mathbf{f}_1 \leq P_{rl} - \text{tr}(\tilde{\mathbf{F}}_{2l} \tilde{\mathbf{F}}_{2l}^H), \quad l = 1, \dots, L$$

where $\Upsilon_{2l} \triangleq \text{bd}(\mathbf{I} \otimes (\mathbf{D}_{2l1} \mathbf{D}_{2l1}^H), \dots, \mathbf{I} \otimes (\mathbf{D}_{2lK} \mathbf{D}_{2lK}^H))$ and $\mathbf{D}_{2li} \triangleq \tilde{\mathbf{F}}_{2l} \tilde{\mathbf{H}}_{1li}$. The problem (83)-(85) becomes

$$\min_{\mathbf{f}_1} (\mathbf{f}_1^H \Upsilon_1^{\frac{1}{2}} - \mathbf{d}_1^H \Upsilon_1^{-\frac{1}{2}}) (\Upsilon_1^{\frac{1}{2}} \mathbf{f}_1 - \Upsilon_1^{-\frac{1}{2}} \mathbf{d}_1) \quad (89)$$

$$\text{s.t. } \mathbf{f}_1^H \Upsilon_{2l} \mathbf{f}_1 \leq P_{rl} - \text{tr}(\tilde{\mathbf{F}}_{2l} \tilde{\mathbf{F}}_{2l}^H), \quad l = 1, \dots, L \quad (90)$$

$$\mathbf{f}_1^H \tilde{\mathbf{I}}_i \mathbf{f}_1 \leq P_{si}, \quad i = 1, \dots, K \quad (91)$$

which is still a QCQP problem and can be efficiently solved by the disciplined convex programming toolbox CVX [20].

C. The Bi-Step Algorithm

We can iteratively update \mathbf{F}_1 and $\tilde{\mathbf{F}}_2$ using the Bi-Step algorithm. The optimal $\tilde{\mathbf{W}}$ is obtained after the convergence of \mathbf{F}_1 and $\tilde{\mathbf{F}}_2$ as $\tilde{\mathbf{W}} = (\tilde{\mathbf{G}} \tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v)^{-1} \tilde{\mathbf{G}}$. We use the following technique to handle the power constraints in (85). Considering the power of the signal vector $\mathbf{y} = \tilde{\mathbf{H}}_2 \mathbf{x}_r$ at the destination node, where $\mathbf{x}_r = [\mathbf{x}_{r1}^T \dots \mathbf{x}_{rL}^T]^T$, we have

$$\begin{aligned} & \text{tr} \left(E \left[\left(\tilde{\mathbf{H}}_2 \mathbf{x}_r \right) \left(\tilde{\mathbf{H}}_2 \mathbf{x}_r \right)^H \right] \right) \\ &= \text{tr} \left(\tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \left(\tilde{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_r} \right) \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H \right) \end{aligned} \quad (92)$$

where $N_r = \sum_{l=1}^L N_{rl}$. Note that using the identity of $\text{tr}(\mathbf{A}\mathbf{B}) \leq \text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$, we have

$$\begin{aligned} & \text{tr} \left(E \left[\left(\tilde{\mathbf{H}}_2 \mathbf{x}_r \right) \left(\tilde{\mathbf{H}}_2 \mathbf{x}_r \right)^H \right] \right) \\ &= \text{tr} \left(\tilde{\mathbf{H}}_2 E \left[\mathbf{x}_r \mathbf{x}_r^H \right] \tilde{\mathbf{H}}_2^H \right) \\ &\leq \text{tr} \left(\tilde{\mathbf{H}}_2 \tilde{\mathbf{H}}_2^H \right) \text{tr} \left(E \left[\mathbf{x}_r \mathbf{x}_r^H \right] \right). \end{aligned} \quad (93)$$

Since $\text{tr} \left(E \left[\mathbf{x}_r \mathbf{x}_r^H \right] \right) = \sum_{l=1}^L \text{tr} \left(E \left[\mathbf{x}_{rl} \mathbf{x}_{rl}^H \right] \right) \leq \sum_{l=1}^L P_{rl}$, together with (93), we have

$$\text{tr} \left(\hat{\mathbf{F}}_2 \left(\tilde{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_r} \right) \hat{\mathbf{F}}_2^H \right) \leq \sum_{l=1}^L P_{rl} \text{tr} \left(\tilde{\mathbf{H}}_2 \tilde{\mathbf{H}}_2^H \right) \quad (94)$$

where $\hat{\mathbf{F}}_2 = \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2$. The received signal in (73) can be written as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_{d1} \\ \mathbf{y}_{d2} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_3 \end{bmatrix} \mathbf{F}_1 \mathbf{s} + \begin{bmatrix} \hat{\mathbf{F}}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ \mathbf{n}_{d1} \end{bmatrix}. \quad (95)$$

Now we can use the Bi-Step algorithm to design $\hat{\mathbf{F}}_2$ and \mathbf{F}_1 in the system of (95), where we can treat the second-hop channel as ‘‘identity matrix’’. In particular, with fixed \mathbf{F}_1 , we only have the power constraint (94), and we can write the structure of $\hat{\mathbf{F}}_2 = \mathbf{T}\mathbf{L}$ according to (55).

Similar to (61), to update the source precoding matrices $\{\mathbf{F}_{1i}\}$, (83) can be written as

$$\begin{aligned} & \text{tr} \left(\mathbf{E}_0 \left(\mathbf{F}_1, \tilde{\mathbf{F}}_2 \right) \right) \\ &= \text{tr} \left(\left(\mathbf{I}_{2N_d} + \sum_{i=1}^K \Psi_i \mathbf{B}_i \Psi_i^H \right)^{-1} \right) + N_b - 2N_d \end{aligned} \quad (96)$$

where $\Psi_i = \mathbf{C}_v^{-\frac{1}{2}} \begin{bmatrix} \sum_{l=1}^L \hat{\mathbf{F}}_{2l} \tilde{\mathbf{H}}_{1li} \\ \tilde{\mathbf{H}}_{3i} \end{bmatrix}$ and $\mathbf{B}_i \triangleq \mathbf{F}_{1i} \mathbf{F}_{1i}^H, i = 1, \dots, K$. Thus, \mathbf{B}_i can be optimized by solving the following problem

$$\min_{\{\mathbf{B}_i\}} \text{tr} \left(\left(\mathbf{I}_{2N_d} + \sum_{i=1}^K \Psi_i \mathbf{B}_i \Psi_i^H \right)^{-1} \right) \quad (97)$$

$$\begin{aligned} & \text{s.t. } \sum_{k=1}^K \text{tr} \left(\mathbf{B}_{1k} \left(\sum_{l=1}^L \hat{\mathbf{F}}_{2l} \tilde{\mathbf{H}}_{lk} \right)^H \left(\sum_{l=1}^L \hat{\mathbf{F}}_{2l} \tilde{\mathbf{H}}_{lk} \right) \right) \\ & \leq \sum_{l=1}^L P_{rl} \text{tr} \left(\tilde{\mathbf{H}}_2 \tilde{\mathbf{H}}_2^H \right) - \text{tr} \left(\hat{\mathbf{F}}_2 \hat{\mathbf{F}}_2^H \right) \end{aligned} \quad (98)$$

$$\text{tr}(\mathbf{B}_i) \leq P_{si}, \mathbf{B}_i \succeq \mathbf{0}, \quad i = 1, \dots, K \quad (99)$$

which is similar to the problem (62)-(64) and can be equivalently converted to an SDP problem by using the Schur complement.

V. NUMERICAL EXAMPLES

In this section, we study the performance of the proposed robust source and relay precoding matrices design algorithms for multiuser MIMO relay systems with direct source-destination links and imperfect CSI through numerical simulations. For simplicity, we consider a system with two users, where all nodes are equipped with multiple antennas. The extension to systems with $K > 2$ users is straight-forward. We simulate a flat Rayleigh fading environment. For notational simplicity, we show the setup of channel matrices in a single-relay relay, which can be easily extended to systems with multiple relay nodes. The estimated channel matrices $\tilde{\mathbf{H}}_{1i}$, $\tilde{\mathbf{H}}_2$, and $\tilde{\mathbf{H}}_{3i}$ have i.i.d. complex Gaussian entries with zero-mean and variances σ_{1i}^2 , σ_2^2 , and $\sigma_{3i}^2, i = 1, \dots, K$, respectively. All noises are i.i.d. AWGN with zero mean and unit variance. We define

$$\begin{aligned} \text{SNR}_{1i} &\triangleq \frac{\sigma_{1i}^2 P_{si}}{N_r}, & \text{SNR}_2 &\triangleq \frac{\sigma_2^2 P_2}{N_d}, \\ \text{SNR}_{3i} &\triangleq \frac{\sigma_{3i}^2 P_{si}}{N_d}, & & i = 1, \dots, K \end{aligned}$$

as the signal-to-noise ratio (SNR) for the source i -relay, relay-destination, and source i -destination links, respectively. For simplicity, we assume $N_{b1} = N_{b2} = N_{s1} = N_{s2} = 2, N_r = 4, N_d = 4$ and the same SNR for the source-relay links and the relay-destination link, i.e., $\text{SNR}_{11} = \text{SNR}_{12} = \text{SNR}_2 = \text{SNR}$, throughout the simulations. Due to a larger pass loss, we set $\text{SNR}_{31} = \text{SNR}_{32} = \text{SNR} - \Delta_{\text{SNR}}$. QPSK constellations are used to modulate the source symbols. All simulation results are averaged over 1000 independent channel realizations.

We simulate a multiuser MIMO relay system where the channel estimation error at the transmitter side is uncorrelated, i.e., $\Theta_{1i} = \Theta_{3i} = \sigma_e^2 \mathbf{I}_{N_{si}}, i = 1, \dots, K$ and $\Theta_2 = \sigma_e^2 \mathbf{I}_{N_r}$, where σ_e^2 measures the variance of the channel estimation error. We obtain from (19), (21), and (24) that for this case, $\alpha_{1i} = \alpha_{3i} = \sigma_e^2 \text{tr}(\mathbf{F}_{1i} \mathbf{F}_{1i}^H), i = 1, \dots, K$, and $\alpha_2 = \sigma_e^2 \text{tr}(\mathbf{F}_2 (\sum_{i=1}^K (\tilde{\mathbf{H}}_{1i} \mathbf{F}_{1i} \mathbf{F}_{1i}^H \tilde{\mathbf{H}}_{1i}^H + \alpha_{1i} \Phi_{1i}) + \mathbf{I}_{N_r}) \mathbf{F}_2^H)$. It can be shown similar to [10] that $\text{tr}(\mathbf{E}_0(\{\mathbf{F}_{1i}\}, \tilde{\mathbf{F}}_2))$ decreases with respect to α_{1i}, α_2 , and $\alpha_{3i}, i = 1, \dots, K$. Therefore,

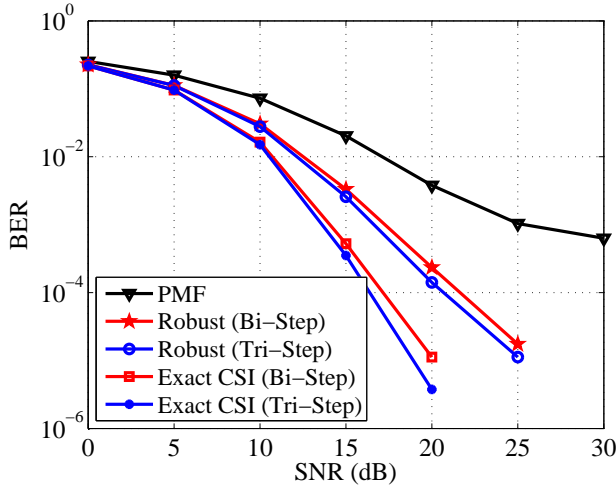


Fig. 2. Example 1: BER versus SNR, $\sigma_e^2 = 0.01$, $\Delta_{\text{SNR}} = 10\text{dB}$.

considering the power constraints (27) and (28), the optimal solution occurs at $\alpha_{1i} = \alpha_{3i} = \sigma_e^2 P_{si}$, $i = 1, \dots, K$, and $\alpha_2 = \sigma_e^2 P_2$. The covariance matrix of channel estimation error at the receiver side is set as

$$\Phi_{1i} = \begin{bmatrix} 1 & \phi_{1i} & \phi_{1i}^2 & \phi_{1i}^3 \\ \phi_{1i} & 1 & \phi_{1i} & \phi_{1i}^2 \\ \phi_{1i}^2 & \phi_{1i} & 1 & \phi_{1i} \\ \phi_{1i}^3 & \phi_{1i}^2 & \phi_{1i} & 1 \end{bmatrix}, \Phi_2 = \Phi_{3i} = \begin{bmatrix} 1 & \phi_2 & \phi_2^2 & \phi_2^3 \\ \phi_2 & 1 & \phi_2 & \phi_2^2 \\ \phi_2^2 & \phi_2 & 1 & \phi_2 \\ \phi_2^3 & \phi_2^2 & \phi_2 & 1 \end{bmatrix}$$

In the simulations, we choose $\phi_{1i} = \phi_2 = 0.45$, $i = 1, 2$.

We compare the performance of the following five systems: (1) The nonrobust design using the pseudo match-and-forward (PMF) algorithm; (2) The proposed robust design using the Bi-Step algorithm; (3) The proposed robust design with the Tri-Step algorithm; (4) The Bi-Step algorithm with the exact CSI knowledge; (5) The Tri-Step algorithm with the exact CSI knowledge. In the PMF algorithm, the MMSE receiver is deployed at the destination node, and we set $\mathbf{F}_{1i} = \sqrt{P_{si}/N_{si}} \mathbf{I}_{N_{si}}$, $i = 1, \dots, K$, and $\mathbf{F}_2 = \sqrt{P_r / \text{tr}((\mathbf{H}_1 \mathbf{H}_2)^H (\mathbf{H}_1 \mathbf{H}_1^H + \mathbf{I}_{N_r}) \mathbf{H}_1 \mathbf{H}_2) (\mathbf{H}_1 \mathbf{H}_2)^H}$, where $\mathbf{H}_1 \triangleq [\mathbf{H}_{11} \mathbf{F}_{11}, \dots, \mathbf{H}_{1K} \mathbf{F}_{1K}]$. For the robust design, the initialization of the Tri-Step and Bi-Step algorithms is listed in Table I and Table II, respectively.

In the first example, we simulate the scenario where $\Delta_{\text{SNR}} = 10\text{dB}$ and $\sigma_e^2 = 0.01$. The BER performance of the above algorithms is shown in Fig. 2. It can be seen that the proposed robust Bi-Step and Tri-Step algorithms outperform the PMF algorithm, indicating the gain of the robust gain. For both the proposed robust design and the system with the exact CSI, the Tri-Step algorithm performs better than the Bi-Step algorithm especially at high SNR. This is because of the approximation of Δ in (58) to be diagonal.

In the second example, we compare the BER and MSE performance of all algorithms with $\Delta_{\text{SNR}} = 20\text{dB}$ and $\sigma_e^2 = 0.001$. The simulation results are illustrated in Figs. 3 and 4. It can be seen from Fig. 3 that by considering the CSI mismatch, the proposed robust algorithms significantly improve the system BER performance compared with the PMF algorithm. In fact, the BER yielded by the robust algorithms

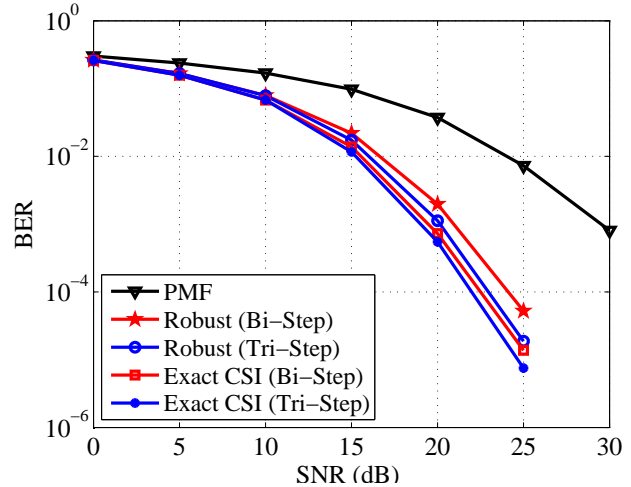


Fig. 3. Example 2: BER versus SNR, $\sigma_e^2 = 0.001$, $\Delta_{\text{SNR}} = 20\text{dB}$.

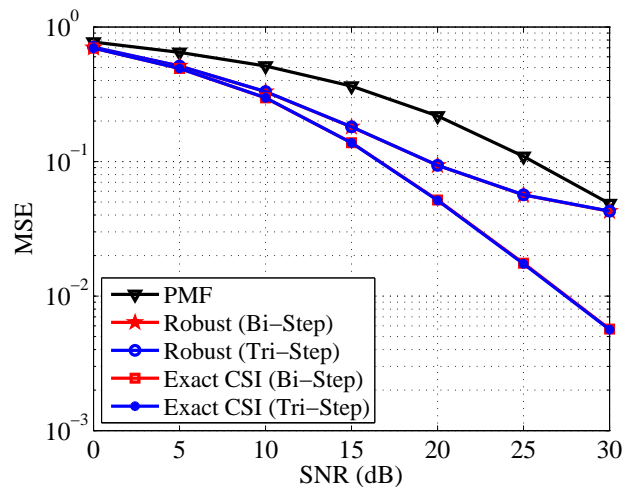


Fig. 4. Example 2: MSE versus SNR, $\sigma_e^2 = 0.001$, $\Delta_{\text{SNR}} = 20\text{dB}$.

is very close to that of the system with the exact CSI. This verifies the importance of considering the direct links and the CSI mismatch in the transceiver design. Interestingly, it can be seen from Fig. 4 that both the robust algorithms and the system with the exact CSI have almost identical MSE.

In the third example, the BER and MSE performance of the five algorithms tested with $\Delta_{\text{SNR}} = 30\text{dB}$ and $\sigma_e^2 = 0.01$ is shown in Figs. 5 and 6. By comparing Fig. 5 with Fig. 2, we can see that the gap between the robust algorithms and the system with the exact CSI knowledge increases with Δ_{SNR} , indicating the impact of pathloss of the direct links on the system performance in the case of CSI mismatch. Moreover, it can be seen from Fig. 6 that the MSE performance of the Bi-Step algorithm is very close to that of the Tri-Step algorithm for both systems with CSI mismatch and the perfect CSI.

In the fourth simulation example, we compare the BER and MSE performance of the above algorithms with $\Delta_{\text{SNR}} = 20\text{dB}$ and different σ_e^2 as 0.1, 0.01, and 0.001, respectively. We can observe from Figs. 7 and 8 that as expected, the proposed robust algorithms have a better performance as σ_e^2 decreases.

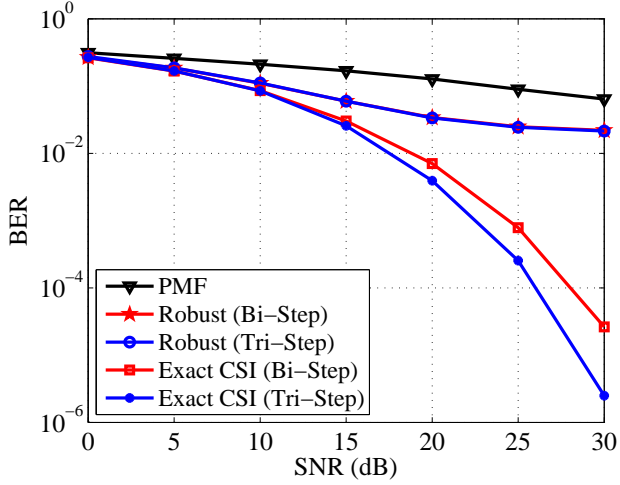


Fig. 5. Example 3: BER versus SNR, $\sigma_e^2 = 0.01$, $\Delta_{\text{SNR}} = 30\text{dB}$.

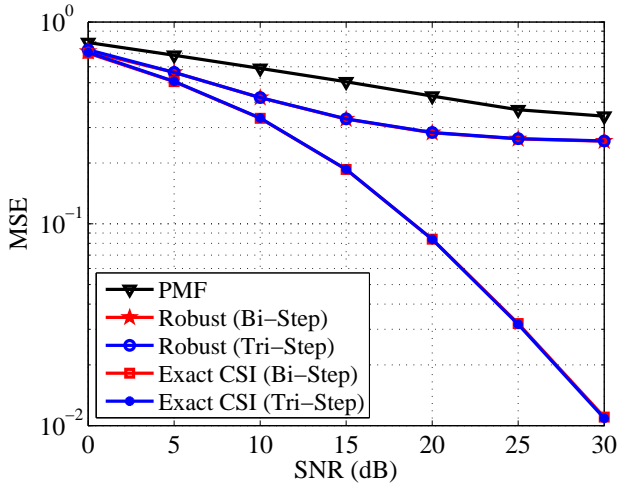


Fig. 6. Example 3: MSE versus SNR, $\sigma_e^2 = 0.01$, $\Delta_{\text{SNR}} = 30\text{dB}$.

At $\sigma_e^2 = 0.001$, the BER and MSE performance of the robust algorithms is very close to that of the system with the exact CSI. Interestingly, it can also be seen that as σ_e^2 decreases, the gap between the BER performance of the Tri-Step and Bi-Step algorithms increases at high SNR.

In the fifth simulation example, we compare the MSE performance of the Bi-Step and Tri-Step algorithms in multi-relay systems, where σ_e^2 is set as 0.1, 0.01, 0.001 respectively, and Δ_{SNR} is fixed to 20dB. We assume there are $L = 2$ relay nodes equipped with $N_{r1} = N_{r2} = 2$ antennas. The other simulation parameters are similar to those in the single-relay system. The MSE performance of the proposed algorithms is shown in Fig. 9. Comparing Fig. 8 with Fig. 9, it can be clearly seen that by incorporating multiple relay nodes, the system MSE can be greatly reduced.

In the next example, we compare the performance of the robust Bi-Step algorithm with the non-robust Bi-Step algorithm, where the source and relay matrices are designed by the Bi-Step algorithm, but using only the estimated channel matrices without considering the robust design. Figs. 10 and

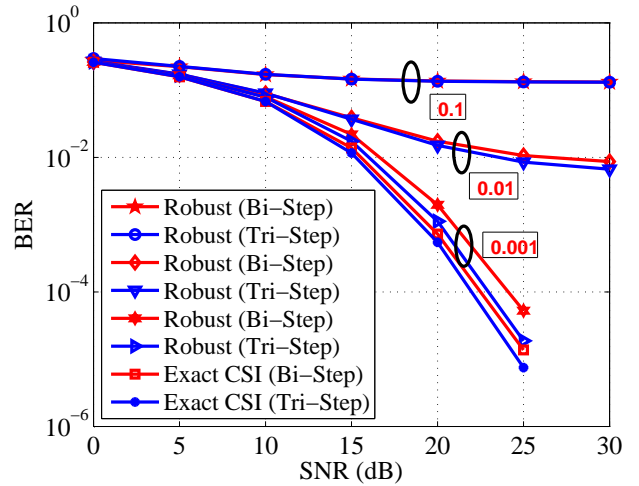


Fig. 7. Example 4: BER versus SNR, $\Delta_{\text{SNR}} = 20\text{dB}$.

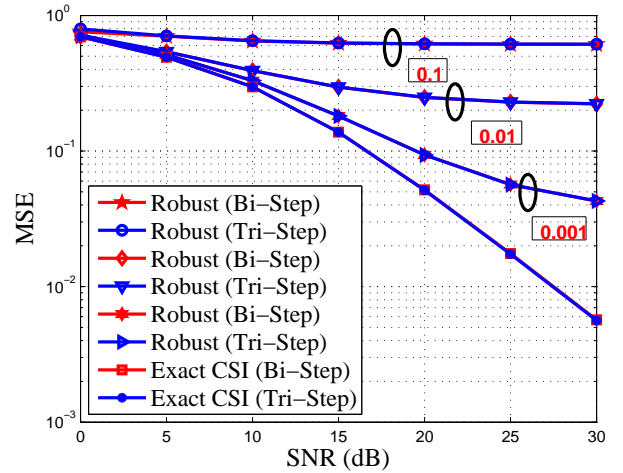


Fig. 8. Example 4: MSE versus SNR, $\Delta_{\text{SNR}} = 20\text{dB}$.

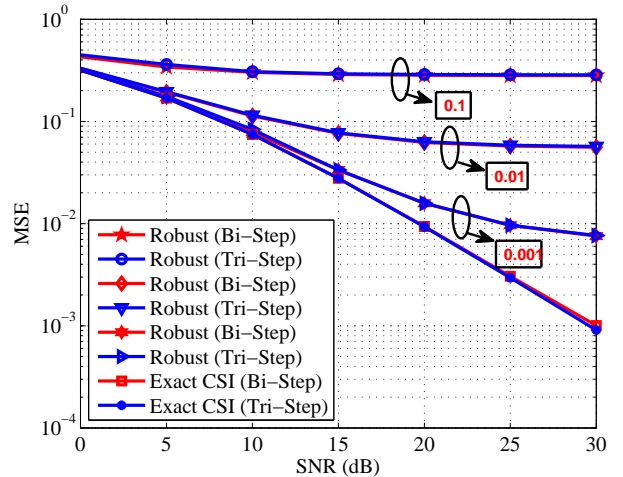


Fig. 9. Example 5: MSE versus SNR, $\Delta_{\text{SNR}} = 20\text{dB}$.

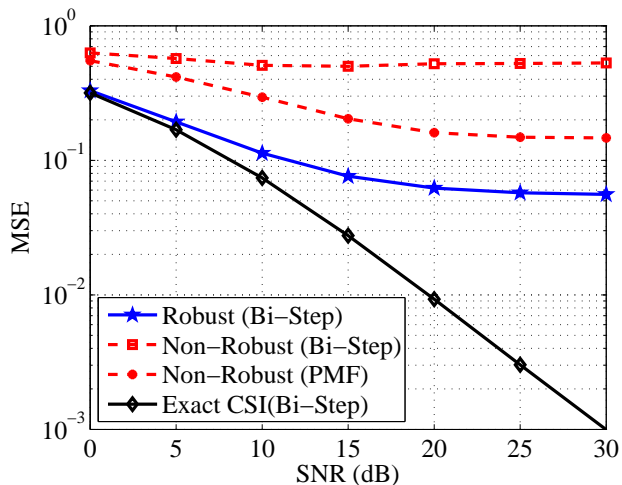


Fig. 10. Example 6: MSE versus SNR, $\Delta_{\text{SNR}} = 20\text{dB}$, $\sigma_e^2 = 0.01$.

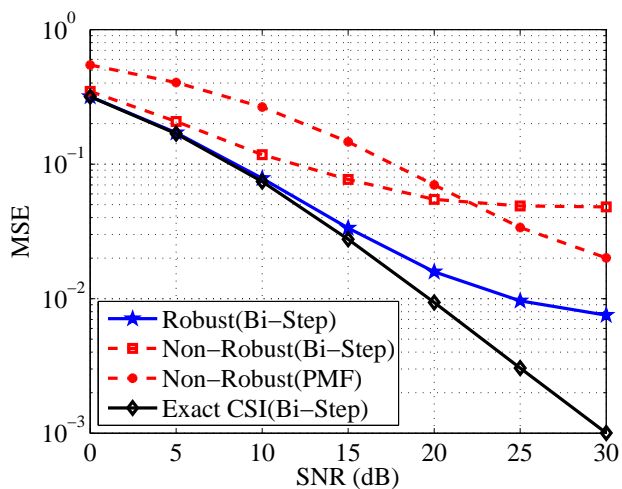


Fig. 11. Example 6: MSE versus SNR, $\Delta_{\text{SNR}} = 20\text{dB}$, $\sigma_e^2 = 0.001$.

11 show the performance comparison with $\sigma_e^2 = 0.01$ and $\sigma_e^2 = 0.001$, respectively. It can be seen that the robust Bi-Step algorithm outperforms the non-robust Bi-Step algorithm for both $\sigma_e^2 = 0.01$ and $\sigma_e^2 = 0.001$. This verifies the benefit of the proposed design which is robust against CSI imperfection. Interestingly, it can be seen from Figs. 10 and 11 that for $\sigma_e^2 = 0.01$, the PMF approach performs better than the non-robust Bi-Step algorithm. When $\sigma_e^2 = 0.001$, the PMF approach yields a smaller MSE than the non-robust Bi-Step algorithm at high SNR. Therefore, the performance of the non-robust Bi-Step algorithm is comparable or even worse than that of the PMF approach. Thus, it is sensible to compare the performance of the proposed algorithms with the PMF approach.

Finally, we compare the computational complexity of the Tri-Step and Bi-Step algorithms. In each iteration of the Tri-Step algorithm, updating $\tilde{\mathbf{W}}$, $\tilde{\mathbf{F}}_2$, and $\{\mathbf{F}_{1i}\}$ involves matrix inversion (35), matrix SVD (43), and solving the QCQP problem (48)-(50). Thus, the per iteration computational complexity order of the Tri-Step algorithm is $\mathcal{O}((\sum_{i=1}^K N_{si}^2)^3)$, which

TABLE III
AVERAGE NUMBER OF ITERATIONS REQUIRED TILL CONVERGENCE BY THE PROPOSED ALGORITHMS

SNR (dB)	5	10	15	20	25	30
Bi-Step (CSI mismatch)	2	2	3	3	4	4
Tri-Step (CSI mismatch)	5	6	7	8	9	10
Bi-Step (exact CSI)	3	2	3	5	6	6
Tri-Step (exact CSI)	6	7	10	19	35	61

is the complexity of solving the QCQP problem (48)-(50) [21]. Similarly, updating $\tilde{\mathbf{F}}_2$ and $\{\mathbf{F}_{1i}\}$ in each iteration of the Bi-Step algorithm involves matrix EVDs (56), (57), and solving the SDP problem (65)-(68). Thus, the Bi-Step algorithm has a per iteration complexity order of $\mathcal{O}((\sum_{i=1}^K N_{si}^2)^{3.5})$, which is the complexity of solving the SDP problem (65)-(68) [21]. Therefore, the Bi-Step algorithm has a higher per-iteration complexity than the Tri-Step algorithm. Interestingly, the robust Tri-Step and Bi-Step algorithms have a similar computational complexity order to their counterparts in non-robust transceiver design for multiuser MIMO relay systems without the direct links [15].

The overall complexity of two algorithms also depends on their convergence speed. Table III shows the average number of iterations required by the proposed Tri-Step and Bi-Step algorithms till convergence with $N_{b1} = N_{b2} = N_{s1} = N_{s2} = 2$, $N_r = N_d = 4$, $\Delta_{\text{SNR}} = 20\text{dB}$, and $\sigma_e^2 = 0.01$. Both algorithms are required to converge up to $\varepsilon = 10^{-3}$. It can be seen from Table III that the number of iterations required by the Tri-Step algorithm increases much faster with SNR than that of the Bi-Step algorithm. This can be explained as follows. In the Tri-Step algorithm, both $\{\mathbf{F}_{1i}\}$ and $\tilde{\mathbf{F}}_2$ need to be initialized, and matrices $\{\mathbf{F}_{1i}\}$, $\tilde{\mathbf{F}}_2$, and $\tilde{\mathbf{W}}$ are optimized in each iteration, while for the Bi-Step algorithm, only $\{\mathbf{F}_{1i}\}$ need to be initialized, and only matrices $\{\mathbf{F}_{1i}\}$ and $\tilde{\mathbf{F}}_2$ need to be updated in each iteration. Therefore, the Bi-Step algorithm converges faster.

VI. CONCLUSION

We have investigated the optimal source, relay, and receiver matrices design for multiuser AF MIMO relay communication systems with direct source-destination links and imperfect CSI. Two iterative algorithms have been developed to design the statistically robust source and relay matrices for the commonly used MMSE criteria. Simulation results show an improved robustness of the proposed algorithms against CSI mismatch. This paper generalizes the multiuser MIMO relay design with direct links to the practical scenario of imperfect CSI knowledge.

ACKNOWLEDGMENT

The authors would like to thank the editor and anonymous reviewers for their valuable comments and suggestions that improved the quality of the paper.

REFERENCES

- [1] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1398-1407, Apr. 2007.

- [2] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, pp. 517-519, July 2008.
- [3] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837-4851, Dec. 2009.
- [4] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A tutorial on the Optimization of amplify-and-forward MIMO relay systems," *IEEE J. Select. Areas Commun.*, vol. 30, pp. 1331-1346, Sep. 2012.
- [5] F.-S. Tseng, M.-Y. Chang, and W.-R. Wu, "Joint Tomlinson-Harashima source and linear relay precoder design in amplify-and-forward MIMO relay systems via MMSE criterion," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 1687-1698, May 2011.
- [6] C. Song, K.-J. Lee, and I. Lee, "MMSE-based MIMO cooperative relaying systems: Closed-form designs and outage behavior," *IEEE J. Select. Areas Commun.*, vol. 30, pp. 1390-1401, Sep. 2012.
- [7] H. Wan, W. Chen, "Joint source and relay design for multiuser MIMO nonregenerative relay networks with direct links," *IEEE Trans. Veh. Technol.*, vol. 61, pp. 2871-2876, July 2012.
- [8] Y. Rong, "Optimal joint source and relay beamforming for MIMO relays with direct link," *IEEE Commun. Lett.*, vol. 14, pp. 390-392, May 2010.
- [9] C. Xing, S. Ma, and Y.-C. 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, pp. 2273-2283, Apr. 2010.
- [10] Y. Rong, "Robust design for linear non-regenerative MIMO relays with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, pp. 2455-2460, May 2011.
- [11] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, pp. 5473-5484, Nov. 2011.
- [12] C. Xing, S. Ma, Z. Fei, Y.-C. Wu, and H. V. Poor, "A general robust linear transceiver design for multi-hop amplify-and-forward MIMO relaying systems," *IEEE Trans. Signal Process.*, vol. 61, pp. 1196-1209, Mar. 2013.
- [13] H. Shen, W. Xu, and C. Zhao, "Robust transceiver for AF MIMO relaying with direct link: A globally optimal solution," *IEEE Signal Process. Lett.*, vol. 21, pp. 947-951, Aug. 2014.
- [14] Z. He, W. Jiang, and Y. Rong, "Robust design for amplify-and-forward MIMO relay systems with direct link and imperfect channel information," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 353-363, Jan. 2015.
- [15] M. R. A. Khandaker and Y. Rong, "Joint transceiver optimization for multiuser MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 60, pp. 5977-5986, Nov. 2012.
- [16] A. Gupta and D. Nagar, *Matrix Variate Distributions*. London, U. K.: Chapman & Hall/CRC, 2000.
- [17] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U. K.: Cambridge University Press, 2004.
- [19] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 772-781, Sep. 1978.
- [20] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming (web page and software)," Apr. 2010. Available: <http://cvxr.com/cvx>.
- [21] Y. Nesterov and A. Nemirovsky, "Interior-point polynomial methods in convex programming," volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.

PLACE
PHOTO
HERE

Zhiqiang He (S'01-M'04) received the B.E. degree and Ph.D. degree (with distinction) from Beijing University of Posts and Telecommunications, China, all in signal and information processing, in 1999 and 2004, respectively. Since July 2004, He has been with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, where he is currently an Associate Professor and the director of the Center of Information Theory and Technology.

His research interests include signal and information processing in wireless communications, networking architecture and protocol design, and underwater acoustic communications.

PLACE
PHOTO
HERE

Xiaonan Zhang received her B.E. degree in Communication Engineering from Beijing University of Chemical Technology, China, in 2012. She studied towards the M.S. degree in the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, China, from 2012 to 2015. She is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at State University of New York, Binghamton.

Her current research interests include wireless communications, mobile data offloading and security and privacy in cyber-physical systems.

PLACE
PHOTO
HERE

YunQiang Bi received his B.E. degree in Electronics and Information Engineering from Ocean University of China, China, in 2013. He is now working towards the M.S. degree in the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, China.

His current research interests include wireless communications, underwater acoustic communication, multimedia technology and computer science and technology.

PLACE
PHOTO
HERE

Weipeng Jiang received the B.Eng. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2010. He is currently working towards the Ph.D. degree in the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. His current research interests include wireless communications, cooperative communication systems and interference alignment.

PLACE
PHOTO
HERE

Yue Rong (S'03-M'06-SM'11) received the Ph.D. degree (summa cum laude) in electrical engineering from the Darmstadt University of Technology, Darmstadt, Germany, in 2005.

He was a Post-Doctoral Researcher with the Department of Electrical Engineering, University of California, Riverside, from February 2006 to November 2007. Since December 2007, he has been with the Department of Electrical and Computer Engineering, Curtin University, Bentley, Australia, where he is currently a Full Professor. His research

interests include signal processing for communications, wireless communications, underwater acoustic communications, applications of linear algebra and optimization methods, and statistical and array signal processing. He has published over 130 journal and conference paper in these areas.

Dr. Rong was a recipient of the Best Paper Award at the 2011 International Conference on Wireless Communications and Signal Processing, the Best Paper Award at the 2010 Asia-Pacific Conference on Communications, and the Young Researcher of the Year Award of the Faculty of Science and Engineering at Curtin University in 2010. He is an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He was an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS from 2012 to 2014, a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS special issue on theories and methods for advanced wireless relays, and was a TPC Member for the IEEE ICC, WCSP, IWCMC, and ChinaCom.