Robust Design for Amplify-and-Forward MIMO Relay Systems With Direct Link and Imperfect Channel Information

Zhiqiang He, *Member, IEEE*, Weipeng Jiang, and Yue Rong, *Senior Member, IEEE*

Abstract—In this paper, we propose statistically robust design for multiple-input multiple-output (MIMO) relay systems with the direct source-destination link 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. In particular, we derive the structure of the optimal relay precoding matrix and show the effect of CSI mismatch on the structure of the optimal robust source and relay matrices. The proposed algorithms generalize the transceiver design of MIMO relay systems with the direct link to the practical scenario of imperfect CSI knowledge. Simulation results demonstrate an improved performance of the proposed algorithms with respect to the conventional methods at various levels of CSI mismatch.

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

I. Introduction

AS a promising technique to improve the reliability and coverage of wireless communication systems, multiple-input multiple-output (MIMO) relay communication has attracted much research interest in recent years [1]. The relay precoding matrix maximizing the source-destination mutual information (MI) of a two-hop MIMO relay system has been developed in [2] and [3]. In [4], the relay precoding matrix that minimizes the mean-squared error (MSE) of the signal waveform estimation was proposed. In [5], a unified framework has been established to optimize the source and relay precoding matrices of linear nonregenerative MIMO relay systems with a broad class of commonly used objective functions.

In the works [4] and [5], the direct source-destination link has been ignored. However, the direct link provides valuable

Z. He 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-mail: hezq@bupt.edu.cn; jwpqjty@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).

spatial diversity, and thus, should be properly considered in the MIMO relay system design. Suboptimal structures of the relay precoding matrix have been derived in [2] and [3] where the direct link is considered. In [6], joint source and relay matrices optimization in the presence of the direct link has been proposed based on the linear minimum MSE (MMSE) receiver. In [7] and [8], source and relay precoding matrices design with the direct link based on Tomlinson-Harashima precoder and diagonalization of the MSE matrix has been studied, respectively. Closed-form designs of the relay precoding matrix have been proposed in [7] and [9].

In the algorithms developed in [6]–[9], the instantaneous channel state information (CSI) of all hops is needed at the node which carries out the optimization procedure. However, in real relay communication systems, the true CSI is unknown, and therefore, has to be estimated. There is always mismatch between the true and the estimated CSI due to channel noise, quantization errors, and outdated channel estimates. Obviously, the performance of the algorithms developed in [6]–[9] will degrade in the presence of such CSI mismatch. MMSE-based optimal relay precoding matrix and destination receiving matrix for a two-hop MIMO relay system have been developed in [10] taking into account the CSI mismatch. The source and relay matrices optimization under CSI mismatch has been investigated in [11] and [12] based on the MMSE criterion. For multi-hop MIMO relay systems, the robust transceiver design has been studied in [13] and [14]. In [15], robust source and relay matrices have been developed considering a broad class of frequently used objective functions in MIMO system design. While the robust MIMO relay design in [10]–[15] ignored the direct link, in [16], the robust source and relay matrices have been developed considering the direct link and using a Tomlinson-Harashima precoder. Due to their advantages in extending the system coverage and enhancing the transmission reliability, relay techniques are considered in modern communication standards such as the LTE-Advanced and WiMax [17]–[19].

In this paper, we investigate statistically robust transceiver design for two-hop MIMO relay problems with the direct source-destination link. The MMSE of the signal waveform estimation at the destination node is adopted as our design criterion. The true CSI of each hop 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 [10]–[16]. Since the robust source and relay

matrices design problem is nonconvex with matrix variables, a globally optimal solution is computationally intractable. To overcome this difficulty, we develop two iterative methods to obtain locally optimal solutions to the original optimization problem by converting it to an equivalent nonrobust MIMO relay design problem with equivalent channel, source, relay, and receiver matrices.

We first develop a Tri-Step method where the source, relay, and receiver matrices are optimized iteratively through solving convex sub-problems. In particular, we derive the structure of the optimal relay precoding matrix and show the effect of CSI mismatch on the structure of the optimal robust source and relay matrices. Moreover, the power allocation at the source and relay nodes is optimally distributed against the CSI mismatch. Then we propose a Bi-Step algorithm where the optimal receiver matrix is substituted into the objective function to obtain an optimization problem only with the source and relay matrices.

Interestingly, the computational complexity of the robust MIMO relay design is in the same order as the nonrobust approach in [6]. Moreover, when the exact CSI is available, the solution of the robust relay design is the same as that in [6]. Therefore, this paper generalizes the MIMO relay design with the direct link to the practical scenario of imperfect CSI knowledge.

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

II. SYSTEM MODEL

We consider a three-node two-hop MIMO communication system as shown in Fig. 1. The source node (node 1) transmits information to the destination node (node 3) with the aid of a relay node (node 2). The ith node is equipped with N_i, $i = 1, 2, 3$, antennas. Using a half-duplex relay, the communication process is completed in two time slots. During the first time slot, the source node transmits a linearly precoded signal vector

$$\mathbf{x}_s = \mathbf{F}_1 \mathbf{s} \tag{1}$$

to both the relay node and the destination node, where $\mathbf{s}$ is the $N_b \times 1$ ($N_b \leq \min(N_1, N_2, N_3)$) source signal vector and $\mathbf{F}_1$ is the $N_1 \times N_b$ source precoding matrix. The received signal vectors at the relay node and the destination node can be written as

$$\mathbf{y}_r = \mathbf{H}_1 \mathbf{x}_s + \mathbf{n}_r \tag{2}$$

$$\mathbf{y}_{d1} = \mathbf{H}_3 \mathbf{x}_s + \mathbf{n}_{d1} \tag{3}$$

where $\mathbf{H}_1$ is the $N_2 \times N_1$ MIMO fading channel matrix between the source and relay nodes, $\mathbf{H}_3$ is the $N_3 \times N_1$ MIMO fading channel matrix between the source and destination nodes, $\mathbf{n}_r$ is the $N_2 \times 1$ noise vector at the relay node, and $\mathbf{n}_{d1}$ is the $N_3 \times 1$ noise vector at the destination node at time slot one.
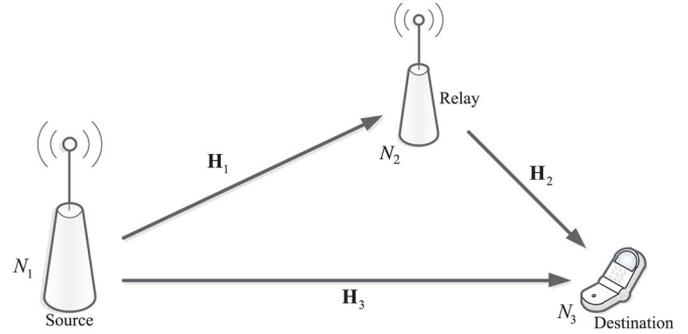
Fig. 1.   A two-hop relay communication system.

During the second time slot, the source node keeps silent, and the relay node linearly precodes $\mathbf{y}_r$ as

$$\mathbf{x}_r = \mathbf{F}_2 \mathbf{y}_r \tag{4}$$

and forwards $\mathbf{x}_r$ to the destination node, where $\mathbf{F}_2$ is the $N_2 \times N_2$ relay precoding matrix. The received signal vector at the destination node is given by

$$\mathbf{y}_{d2} = \mathbf{H}_2 \mathbf{x}_r + \mathbf{n}_{d2} \tag{5}$$

where $\mathbf{H}_2$ is the $N_3 \times N_2$ MIMO fading channel matrix between the relay and destination nodes and $\mathbf{n}_{d2}$ is the $N_3 \times 1$ noise vector at the destination node at time slot two.

Combining (1)–(5), the received signals at the destination node over two consecutive time slots can be written as

$$
\begin{aligned}
\mathbf{y} &= \begin{bmatrix} \mathbf{y}_{d2} \\ \mathbf{y}_{d1} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \\ \mathbf{H}_3 \end{bmatrix} \mathbf{F}_1 \mathbf{s} + \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ \mathbf{n}_{d1} \end{bmatrix} \\
&= \mathbf{G} \mathbf{s} + \mathbf{v} \tag{6}
\end{aligned}
$$

where $\mathbf{G}$ is the $2N_3 \times N_b$ equivalent MIMO channel matrix between the source and destination nodes, $\mathbf{v}$ is the $2N_3 \times 1$ equivalent noise vector at the destination node, and they are given respectively by

$$\mathbf{G} \triangleq \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \\ \mathbf{H}_3 \mathbf{F}_1 \end{bmatrix}, \quad \mathbf{v} \triangleq \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{n}_r + \mathbf{n}_{d2} \\ \mathbf{n}_{d1} \end{bmatrix}. \tag{7}$$

We assume that all noises are independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) with zero mean and unit variance, and $\mathbf{H}_i$, $i = 1, 2, 3$, are independent.

In the case of CSI mismatch,[1] the true channel matrices can be modeled as the well-known Gaussian-Kronecker model

---

[1] Such CSI mismatch may be introduced by channel estimation and/or outdated channel estimates. Channel estimation procedures/algorithms for MIMO relay systems have been studied, for example in [20] and [21]. However, channel estimation is not the focus of this work.

in [10]–[16]

$$\mathbf{H}_i \sim \mathcal{CN}(\bar{\mathbf{H}}_i, \boldsymbol{\Theta}_i \otimes \boldsymbol{\Phi}_i), \quad i = 1, 2, 3 \tag{8}$$

where $\bar{\mathbf{H}}_i$ is the estimated channel matrix, $\boldsymbol{\Theta}_i$ and $\boldsymbol{\Phi}_i$ 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 (8), we have $\mathbf{H}_i = \bar{\mathbf{H}}_i + \mathbf{A}_{\Phi_i} \mathbf{H}_{w_i} \mathbf{A}_{\Theta_i}^H$, $i = 1, 2, 3$, where $\mathbf{A}_{\Phi_i} \mathbf{A}_{\Phi_i}^H = \boldsymbol{\Phi}_i$, $\mathbf{A}_{\Theta_i} \mathbf{A}_{\Theta_i}^H = \boldsymbol{\Theta}_i^T$, $\mathbf{H}_{w_i}$ 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, $(\cdot)^T$ and $(\cdot)^H$ stand for the matrix (vector) transpose and Hermitian transpose, respectively. The dimensions of $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_3$ are $N_1 \times N_1$, $\boldsymbol{\Theta}_2$ and $\boldsymbol{\Phi}_1$ have a dimension of $N_2 \times N_2$, while $\boldsymbol{\Phi}_2$ and $\boldsymbol{\Phi}_3$ are $N_3 \times N_3$ matrices.

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

$$\hat{\mathbf{s}} = \mathbf{W}^H \mathbf{y} \tag{9}$$

where $\mathbf{W}$ is the $2N_3 \times N_b$ receiver weight matrix. From (6) and (9), the MSE matrix of the signal waveform estimation at the destination node is a function of $\mathbf{W}, \mathbf{F}_1, \mathbf{F}_2$ as

$$
\begin{aligned}
&\mathbf{E}(\mathbf{W}, \mathbf{F}_1, \mathbf{F}_2) \\
&= E\left[(\hat{\mathbf{s}} - \mathbf{s})(\hat{\mathbf{s}} - \mathbf{s})^H\right] \\
&= \left(\mathbf{W}^H \mathbf{G} - \mathbf{I}_{N_b}\right) \left(\mathbf{W}^H \mathbf{G} - \mathbf{I}_{N_b}\right)^H + \mathbf{W}^H \mathbf{C}_v \mathbf{W} \\
&= \mathbf{W}^H \mathbf{A} \mathbf{W} - \mathbf{W}^H \mathbf{G} - \mathbf{G}^H \mathbf{W} + \mathbf{I}_{N_b}
\end{aligned}
\tag{10}
$$

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

$$\mathbf{C}_v = E[\mathbf{v}\mathbf{v}^H] = \begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{F}_2^H \mathbf{H}_2^H + \mathbf{I}_{N_3} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_3} \end{bmatrix} \tag{11}$$

is the noise covariance matrix. To obtain (10), 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 (10), due to the mismatch between $\mathbf{H}_j$ and $\bar{\mathbf{H}}_i$, $i = 1, 2, 3$. Taking the CSI mismatch into account, we consider the statistical expectation of $\mathbf{E}$, which is given by

$$E_H[\mathbf{E}(\mathbf{W}, \mathbf{F}_1, \mathbf{F}_2)] = \mathbf{W}^H \bar{\mathbf{A}} \mathbf{W} - \mathbf{W}^H \bar{\mathbf{G}} - \bar{\mathbf{G}}^H \mathbf{W} + \mathbf{I}_{N_b} \tag{12}$$

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 $\mathbf{H}_1$ and $\mathbf{H}_2$ are independent, from (7) and (8), we have

$$
\begin{aligned}
\bar{\mathbf{G}} &= \begin{bmatrix} E_{H_2}[\mathbf{H}_2] \mathbf{F}_2 E_{H_1}[\mathbf{H}_1] \mathbf{F}_1 \\ E_{H_3}[\mathbf{H}_3] \mathbf{F}_1 \end{bmatrix} \\
&= \begin{bmatrix} \bar{\mathbf{H}}_2 \mathbf{F}_2 \bar{\mathbf{H}}_1 \mathbf{F}_1 \\ \bar{\mathbf{H}}_3 \mathbf{F}_1 \end{bmatrix} \\
\bar{\mathbf{A}} &= \bar{\mathbf{G}} \bar{\mathbf{G}}^H + \bar{\mathbf{C}}_v + \mathbf{R}
\end{aligned}
\tag{13}
$$

where

$$
\begin{aligned}
\bar{\mathbf{C}}_v &= \begin{bmatrix} \bar{\mathbf{H}}_2 \mathbf{F}_2 \mathbf{F}_2^H \bar{\mathbf{H}}_2^H + \mathbf{I}_{N_3} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_3} \end{bmatrix} \\
\mathbf{R} &= \begin{bmatrix} \alpha_1 \bar{\mathbf{H}}_2 \mathbf{F}_2 \boldsymbol{\Phi}_1 \mathbf{F}_2^H \bar{\mathbf{H}}_2^H + \alpha_2 \boldsymbol{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \alpha_3 \boldsymbol{\Phi}_3 \end{bmatrix}.
\end{aligned}
\tag{14}
$$

The proof of (13) and the definition of α_1, α_2, and α_3 in (14) are given in Appendix A.

It can be seen from (12) that the CSI mismatch is considered by (14). If the perfect CSI is available, i.e., $\bar{\mathbf{H}}_i = \mathbf{H}_i$ and $\boldsymbol{\Theta}_i = \mathbf{0}$, $i = 1, 2, 3$, from (14) and (62), there is $\alpha_i = 0$, $i = 1, 2, 3$, and $\mathbf{R} = \mathbf{0}$, then the MSE matrix (12) becomes (10). Therefore, (12) generalizes the MSE matrix from the perfect CSI case to the practical scenario with CSI mismatch.

From (1) and (4), the transmission power consumed by the source node and the relay node can be written as $tr(\mathbf{F}_1 \mathbf{F}_1^H)$ and $tr(\mathbf{F}_2(\mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H + \mathbf{I}_{N_2})\mathbf{F}_2^H)$, respectively, where $tr(\cdot)$ denotes the matrix trace. However, since the true $\mathbf{H}_1$ is unknown, we consider the averaged transmission power at the relay node, which is given by

$$
\begin{aligned}
&E_H\left[tr(\mathbf{F}_2(\mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H + \mathbf{I}_{N_2})\mathbf{F}_2^H)\right] \\
&= tr\left(\mathbf{F}_2\left(\bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H + \alpha_1 \boldsymbol{\Phi}_1 + \mathbf{I}_{N_2}\right) \mathbf{F}_2^H\right).
\end{aligned}
\tag{15}
$$

From (12) and (15), the robust source, relay, and destination matrices optimization problem can be written as

$$\min_{\mathbf{W}, \mathbf{F}_1, \mathbf{F}_2} \quad tr\left(E_H[\mathbf{E}(\mathbf{W}, \mathbf{F}_1, \mathbf{F}_2)]\right) \tag{16}$$

$$\text{s.t.} \quad tr\left(\mathbf{F}_2\left(\bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H + \alpha_1 \boldsymbol{\Phi}_1 + \mathbf{I}_{N_2}\right) \mathbf{F}_2^H\right) \leq P_2 \tag{17}$$

$$tr\left(\mathbf{F}_1 \mathbf{F}_1^H\right) \leq P_1 \tag{18}$$

where P_1 and P_2 are the transmission power available at the source node and the relay node, respectively. The problem (16)–(18) is nonconvex with matrix variables. Moreover, due to the direct link, this problem is much more challenging to solve than the problem in [15].

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 receiver matrices. In the Tri-Step algorithm, the source, relay, and receiver matrices are optimized iteratively through solving convex sub-problems. For the Bi-Step algorithm, the optimal receiver 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 alternatingly and the receiver matrix is calculated after the convergence of the algorithm.

By introducing $\mathbf{P}_1 \triangleq \alpha_1 \boldsymbol{\Phi}_1 + \mathbf{I}_{N_2}$, $\mathbf{P}_2 \triangleq \alpha_2 \boldsymbol{\Phi}_2 + \mathbf{I}_{N_3}$, and $\mathbf{P}_3 \triangleq \alpha_3 \boldsymbol{\Phi}_3 + \mathbf{I}_{N_3}$, (12) can be rewritten as

$$
\begin{aligned}
&E_H[\mathbf{E}(\mathbf{W}, \mathbf{F}_1, \mathbf{F}_2)] \\
&= [\mathbf{W}_1^H \quad \mathbf{W}_2^H] \mathbf{Z} \mathbf{M} \mathbf{Z} \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} - [\mathbf{W}_1^H \quad \mathbf{W}_2^H] \mathbf{Z} \mathbf{Z}^{-1} \bar{\mathbf{G}} \\
&\quad - \bar{\mathbf{G}}^H \mathbf{Z}^{-1} \mathbf{Z} \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} + \mathbf{I}_{N_b}
\end{aligned}
\tag{19}
$$

where $\mathbf{Z} \triangleq \mathrm{bd}(\mathbf{P}_2^{(1/2)}, \mathbf{P}_3^{(1/2)})$, $\mathbf{W}_1$ and $\mathbf{W}_2$ contain the first and the last N_3 rows of $\mathbf{W}$, respectively, and $\mathbf{M}$ is given by (20). (See equation at the bottom of the page) Here $\mathrm{bd}(\cdot)$ stands for a block diagonal matrix and $(\cdot)^{-1}$ denotes the matrix inversion.

By introducing $\tilde{\mathbf{W}}_1^H \triangleq \mathbf{W}_1^H \mathbf{P}_2^{(1/2)}$, $\tilde{\mathbf{W}}_2^H \triangleq \mathbf{W}_2^H \mathbf{P}_3^{\frac{1}{2}}$, $\tilde{\mathbf{H}}_i \triangleq \mathbf{P}_i^{-\frac{1}{2}} \bar{\mathbf{H}}_i$, $i = 1, 2, 3$, and $\tilde{\mathbf{F}}_2 \triangleq \mathbf{F}_2 \mathbf{P}_1^{\frac{1}{2}}$, (19) can be rewritten as

$$E_H \left[\mathbf{E}(\tilde{\mathbf{W}}, \mathbf{F}_1, \tilde{\mathbf{F}}_2) \right]$$

$$= [\tilde{\mathbf{W}}_1^H \quad \tilde{\mathbf{W}}_2^H] (\tilde{\mathbf{G}} \tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v) \begin{bmatrix} \tilde{\mathbf{W}}_1 \\ \tilde{\mathbf{W}}_2 \end{bmatrix}$$

$$- [\tilde{\mathbf{W}}_1^H \quad \tilde{\mathbf{W}}_2^H] \tilde{\mathbf{G}} - \tilde{\mathbf{G}}^H \begin{bmatrix} \tilde{\mathbf{W}}_1 \\ \tilde{\mathbf{W}}_2 \end{bmatrix} + \mathbf{I}_{N_b}$$

$$= (\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_b})(\tilde{\mathbf{W}}^H \tilde{\mathbf{G}} - \mathbf{I}_{N_b})^H + \tilde{\mathbf{W}}^H \tilde{\mathbf{C}}_v \tilde{\mathbf{W}} \quad (21)$$

where

$$\tilde{\mathbf{G}} = \begin{bmatrix} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1 \\ \tilde{\mathbf{H}}_3 \mathbf{F}_1 \end{bmatrix}, \quad \tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}}_1 \\ \tilde{\mathbf{W}}_2 \end{bmatrix}$$

$$\tilde{\mathbf{C}}_v = \begin{bmatrix} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \mathbf{I}_{N_3} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_3} \end{bmatrix}.$$

Using (21), the optimization problem (16)–(18) can be equivalently rewritten as

$$\min_{\tilde{\mathbf{W}}, \mathbf{F}_1, \tilde{\mathbf{F}}_2} \quad tr \left(E_H \left[\mathbf{E}(\tilde{\mathbf{W}}, \mathbf{F}_1, \tilde{\mathbf{F}}_2) \right] \right) \quad (22)$$

$$\text{s.t.} \quad tr \left(\tilde{\mathbf{F}}_2 \left(\tilde{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_2} \right) \tilde{\mathbf{F}}_2^H \right) \leq P_2 \quad (23)$$

$$tr \left(\mathbf{F}_1 \mathbf{F}_1^H \right) \leq P_1 \quad (24)$$

where the variable substitution of $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{F}}_2$ is used to rewrite the power constraint (17) at the relay node to obtain (23). It is worth noting that the robust precoding matrices design problem (16)–(18) for MIMO relay systems with imperfect CSI is converted to the problem (22)–(24) for an equivalent MIMO relay system with direct link and perfect CSI, where the channel matrices are $\tilde{\mathbf{H}}_i$, $i = 1, 2, 3$, the source precoding matrix is $\mathbf{F}_1$, the relay precoding matrix is $\tilde{\mathbf{F}}_2$, and the receiver matrix is $\tilde{\mathbf{W}}$.

A. The Tri-Step Algorithm

The problem (22)–(24) 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 (22)–(24). Firstly, with given $\mathbf{F}_1$ and $\tilde{\mathbf{F}}_2$, the weight matrix $\tilde{\mathbf{W}}$ minimizing (22) is the famous Wiener filter [23] (MMSE receiver) given by

$$\tilde{\mathbf{W}} = (\tilde{\mathbf{G}} \tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v)^{-1} \tilde{\mathbf{G}}. \quad (25)$$

Secondly, with given $\tilde{\mathbf{W}}$ and $\mathbf{F}_1$, $\tilde{\mathbf{F}}_2$ can be updated by solving the following problem

$$\min_{\tilde{\mathbf{F}}_2} \quad tr \left((\breve{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \breve{\mathbf{H}}_1 - \mathbf{\Pi})(\breve{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \breve{\mathbf{H}}_1 - \mathbf{\Pi})^H + \breve{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \breve{\mathbf{H}}_2^H \right) \quad (26)$$

$$\text{s.t.} \quad tr \left(\tilde{\mathbf{F}}_2 \left(\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2} \right) \tilde{\mathbf{F}}_2^H \right) \leq P_2 \quad (27)$$

where $\breve{\mathbf{H}}_1 \triangleq \tilde{\mathbf{H}}_1 \mathbf{F}_1$, $\breve{\mathbf{H}}_2 \triangleq \tilde{\mathbf{W}}_1^H \tilde{\mathbf{H}}_2$, $\mathbf{\Pi} \triangleq \mathbf{I}_{N_b} - \tilde{\mathbf{W}}_2^H \tilde{\mathbf{H}}_3 \mathbf{F}_1$.

Using the Lagrange multiplier method, we can solve the problem (26)–(27) and obtain $\tilde{\mathbf{F}}_2$ as

$$\tilde{\mathbf{F}}_2 = \breve{\mathbf{H}}_2^H (\breve{\mathbf{H}}_2 \breve{\mathbf{H}}_2^H + \mu \mathbf{I}_{N_b})^{-1} \mathbf{\Pi} \breve{\mathbf{H}}_1^H (\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2})^{-1} \quad (28)$$

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

$$\mu \left(tr \left(\tilde{\mathbf{F}}_2 \left(\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2} \right) \tilde{\mathbf{F}}_2^H \right) - P_2 \right) = 0. \quad (29)$$

Assuming that $\mu = 0$, from (28) we have

$$\tilde{\mathbf{F}}_2 = \breve{\mathbf{H}}_2^H \left(\breve{\mathbf{H}}_2 \breve{\mathbf{H}}_2^H \right)^{-1} \mathbf{\Pi} \breve{\mathbf{H}}_1^H \left(\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2} \right)^{-1}. \quad (30)$$

If $\tilde{\mathbf{F}}_2$ in (30) satisfies the power constraint (27), then (30) is the solution to the problem (26)–(27). Otherwise, there must be $\mu > 0$ such that

$$tr \left(\tilde{\mathbf{F}}_2 \left(\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2} \right) \tilde{\mathbf{F}}_2^H \right) = P_2. \quad (31)$$

In this case, μ can be obtained by substituting (28) into (31) and solving the following nonlinear equation

$$tr \left(\breve{\mathbf{H}}_2^H \left(\breve{\mathbf{H}}_2 \breve{\mathbf{H}}_2^H + \mu \mathbf{I}_{N_b} \right)^{-1} \mathbf{\Pi} \breve{\mathbf{H}}_1^H \left(\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2} \right)^{-1} \right.$$

$$\left. \times \breve{\mathbf{H}}_1 \mathbf{\Pi}^H \left(\breve{\mathbf{H}}_2 \breve{\mathbf{H}}_2^H + \mu \mathbf{I}_{N_b} \right)^{-1} \breve{\mathbf{H}}_2 \right) = P_2. \quad (32)$$

By introducing the singular value decomposition (SVD) of $\breve{\mathbf{H}}_2 = \mathbf{U}_2 \mathbf{\Lambda}_2 \mathbf{V}_2^H$, we obtain from (32) that

$$tr \left(\mathbf{\Lambda}_2 \left(\mathbf{\Lambda}_2^2 + \mu \mathbf{I}_{N_b} \right)^{-1} \mathbf{U}_2^H \mathbf{\Pi} \breve{\mathbf{H}}_1^H \left(\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2} \right)^{-1} \right.$$

$$\left. \times \breve{\mathbf{H}}_1 \mathbf{\Pi}^H \mathbf{U}_2 \left(\mathbf{\Lambda}_2^2 + \mu \mathbf{I}_{N_b} \right)^{-1} \mathbf{\Lambda}_2 \right) = P_2. \quad (33)$$

Denoting $\mathbf{\Gamma} \triangleq \mathbf{U}_2^H \mathbf{\Pi} \breve{\mathbf{H}}_1^H (\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2})^{-1} \breve{\mathbf{H}}_1 \mathbf{\Pi}^H \mathbf{U}_2$, (33) can be equivalently written as

$$\sum_{i=1}^{N_b} \frac{\lambda_i^2 \gamma_i}{(\lambda_i^2 + \mu)^2} = P_2 \quad (34)$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{P}_2^{-\frac{1}{2}} \bar{\mathbf{H}}_2 \mathbf{F}_2 \mathbf{P}_1^{\frac{1}{2}} (\mathbf{P}_1^{-\frac{1}{2}} \bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H \mathbf{P}_1^{-\frac{1}{2}} + \mathbf{I}_{N_2}) \mathbf{P}_1^{\frac{1}{2}} \mathbf{F}_2^H \bar{\mathbf{H}}_2^H \mathbf{P}_2^{-\frac{1}{2}} + \mathbf{I}_{N_3} & \mathbf{P}_2^{-\frac{1}{2}} \bar{\mathbf{H}}_2 \mathbf{F}_2 \bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_3^H \mathbf{P}_3^{-\frac{1}{2}} \\ \mathbf{P}_3^{-\frac{1}{2}} \bar{\mathbf{H}}_3 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H \mathbf{F}_2^H \bar{\mathbf{H}}_2^H \mathbf{P}_2^{-\frac{1}{2}} & \mathbf{P}_3^{-\frac{1}{2}} \bar{\mathbf{H}}_3 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_3^H \mathbf{P}_3^{-\frac{1}{2}} + \mathbf{I}_{N_3} \end{bmatrix} \quad (20)$$

where λ_i and γ_i are the ith main diagonal elements of $\mathbf{\Lambda}_2$ and $\mathbf{\Gamma}$ respectively. Since the left-hand side of (34) is monotonically decreasing with respect to μ, the bisection method [24] can be applied to solve (34) to obtain μ.

Thirdly, with given $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{F}}_2$, we show that the problem (22)–(24) can be cast as a quadratically constrained quadratic programming (QCQP) problem [24] to optimize $\mathbf{F}_1$. By introducing $\mathbf{D}_1 \triangleq \tilde{\mathbf{W}}_1^H \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 + \tilde{\mathbf{W}}_2^H \tilde{\mathbf{H}}_3$, (22) can be rewritten as

$$tr\left(\left(\tilde{\mathbf{W}}_1^H \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1 + \tilde{\mathbf{W}}_2^H \tilde{\mathbf{H}}_3 \mathbf{F}_1 - \mathbf{I}_{N_b}\right)\left(\tilde{\mathbf{W}}_1^H \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1\right.\right.$$
$$\left.\left. + \tilde{\mathbf{W}}_2^H \tilde{\mathbf{H}}_3 \mathbf{F}_1 - \mathbf{I}_{N_b}\right)^H\right) + t_1$$
$$= tr\left((\mathbf{D}_1\mathbf{F}_1 - \mathbf{I}_{N_b})(\mathbf{D}_1\mathbf{F}_1 - \mathbf{I}_{N_b})^H\right) + t_1$$
$$= tr(\mathbf{D}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{D}_1^H) - tr(\mathbf{D}_1\mathbf{F}_1) - tr(\mathbf{F}_1^H\mathbf{D}_1^H) + t_2 \quad (35)$$

where $t_1 \triangleq tr(\tilde{\mathbf{W}}^H \tilde{\mathbf{C}}_v \tilde{\mathbf{W}})$ and $t_2 \triangleq t_1 + N_b$.

Using the identity of $tr(\mathbf{C}^T\mathbf{D}) = (vec(\mathbf{C}))^T vec(\mathbf{D})$ and $vec(\mathbf{CD}) = (\mathbf{I} \otimes \mathbf{C})vec(\mathbf{D})$ [25], where $vec(\mathbf{X})$ stands for a column vector obtained by stacking all columns of $\mathbf{X}$ on top of each other, we have

$$tr\left(\mathbf{D}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{D}_1^H\right) = (vec(\mathbf{F}_1))^H vec\left(\mathbf{D}_1^H\mathbf{D}_1\mathbf{F}_1\right)$$
$$= \mathbf{f}_1^H\left(\mathbf{I}_{N_1} \otimes \left(\mathbf{D}_1^H\mathbf{D}_1\right)\right)\mathbf{f}_1$$

where $\mathbf{f}_1 \triangleq vec(\mathbf{F}_1)$. Then (35) can be rewritten as

$$\mathbf{f}_1^H\left(\mathbf{I}_{N_1} \otimes \left(\mathbf{D}_1^H\mathbf{D}_1\right)\right)\mathbf{f}_1 - \left(vec\left(\mathbf{D}_1^H\right)\right)^H\mathbf{f}_1 - \mathbf{f}_1^H vec\left(\mathbf{D}_1^H\right) + t_2$$
$$= \mathbf{f}_1^H\mathbf{\Upsilon}_1\mathbf{f}_1 - \mathbf{d}_1^H\mathbf{f}_1 - \mathbf{f}_1^H\mathbf{d}_1 + t_2$$
$$= \left(\mathbf{f}_1^H\mathbf{\Upsilon}_1^{\frac{1}{2}} - \mathbf{d}_1^H\mathbf{\Upsilon}_1^{-\frac{1}{2}}\right)\left(\mathbf{\Upsilon}_1^{\frac{1}{2}}\mathbf{f}_1 - \mathbf{\Upsilon}_1^{-\frac{1}{2}}\mathbf{d}_1\right) + t_3 \quad (36)$$

where $t_3 \triangleq t_2 - \mathbf{d}_1^H\mathbf{\Upsilon}_1^{-1}\mathbf{d}_1$, $\mathbf{\Upsilon}_1 \triangleq \mathbf{I}_{N_1} \otimes (\mathbf{D}_1^H\mathbf{D}_1)$, and $\mathbf{d}_1 \triangleq vec(\mathbf{D}_1^H)$. 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 $\mathbf{D}_2 \triangleq \tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_1$, the relay transmit power constraint in (23) can be rewritten as

$$tr\left(\mathbf{D}_2\mathbf{F}_1\mathbf{F}_1^H\mathbf{D}_2^H\right) + tr\left(\tilde{\mathbf{F}}_2\tilde{\mathbf{F}}_2^H\right)$$
$$= \mathbf{f}_1^H(\mathbf{I}_{N_1} \otimes (\mathbf{D}_2^H\mathbf{D}_2)\mathbf{f}_1 + tr\left(\tilde{\mathbf{F}}_2\tilde{\mathbf{F}}_2^H\right)$$
$$= \mathbf{f}_1^H\mathbf{\Upsilon}_2\mathbf{f}_1 + tr\left(\tilde{\mathbf{F}}_2\tilde{\mathbf{F}}_2^H\right) \leq P_2 \quad (37)$$

where $\mathbf{\Upsilon}_2 \triangleq \mathbf{I}_{N_1} \otimes (\mathbf{D}_2^H\mathbf{D}_2)$. Using (35) and (37), the problem (22)–(24) can be equivalently rewritten as the following QCQP problem

$$\min_{\mathbf{f}_1} \quad \left(\mathbf{f}_1^H\mathbf{\Upsilon}_1^{\frac{1}{2}} - \mathbf{d}_1^H\mathbf{\Upsilon}_1^{-\frac{1}{2}}\right)\left(\mathbf{\Upsilon}_1^{\frac{1}{2}}\mathbf{f}_1 - \mathbf{\Upsilon}_1^{-\frac{1}{2}}\mathbf{d}_1\right) \quad (38)$$

$$\text{s.t.} \quad \mathbf{f}_1^H\mathbf{\Upsilon}_2\mathbf{f}_1 \leq \bar{P}_2 \quad (39)$$

$$\mathbf{f}_1^H\mathbf{f}_1 \leq P_1 \quad (40)$$

TABLE I
PROCEDURE OF THE TRI-STEP ALGORITHM

1) Initialize the algorithm with $\mathbf{F}_1^{(0)} = \sqrt{P_1/N_b}[\mathbf{I}_{N_b}, \mathbf{0}]^T$ and $\tilde{\mathbf{F}}_2^{(0)} = \sqrt{P_2/tr\left(\tilde{\mathbf{H}}_1\mathbf{F}_1^{(0)}(\mathbf{F}_1^{(0)})^H\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_2}\right)}\mathbf{I}_{N_2}$; Set $n = 0$.
2) Update $\tilde{\mathbf{W}}^{(n)}$ using $\tilde{\mathbf{F}}_2^{(n)}$ and $\mathbf{F}_1^{(n)}$ as (25).
3) Update $\tilde{\mathbf{F}}_2^{(n+1)}$ using $\tilde{\mathbf{W}}^{(n)}$ and $\mathbf{F}_1^{(n)}$ as (28).
4) Update $\mathbf{F}_1^{(n+1)}$ using $\tilde{\mathbf{W}}^{(n)}$ and $\tilde{\mathbf{F}}_2^{(n+1)}$ by solving the problem (38)-(40).
5) if $\left(\text{mse}_1^{(n)} - \text{mse}_1^{(n+1)}\right)/\text{mse}_1^{(n)} < \varepsilon$, iteration ends; otherwise go to step (2).

where $\bar{P}_2 \triangleq P_2 - tr(\tilde{\mathbf{F}}_2\tilde{\mathbf{F}}_2^H)$. The problem (38)–(40) can be efficiently solved by the disciplined convex programming toolbox CVX [26].

The procedure of applying the Tri-Step iterative algorithm to solve the problem (22)–(24) 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 (22) at the nth iteration.

B. The Bi-Step Algorithm

By substituting (25) back into (21), we have

$$\mathbf{E}_0(\mathbf{F}_1, \tilde{\mathbf{F}}_2) = \mathbf{I}_{N_b} - \tilde{\mathbf{G}}^H\left(\tilde{\mathbf{G}}\tilde{\mathbf{G}}^H + \tilde{\mathbf{C}}_v\right)^{-1}\tilde{\mathbf{G}} \quad (41)$$

where $\mathbf{E}_0$ stands for the MSE matrix when a linear MMSE receiver is used at the destination. Interestingly, (41) can be viewed as the minimum MSE matrix for an equivalent two-hop MIMO relay system with channels $\tilde{\mathbf{H}}_i$, $i = 1, 2, 3$, which are exactly known, $\mathbf{F}_1$ is used as the source precoding matrix, and $\tilde{\mathbf{F}}_2$ is chosen as the relay precoding matrix.

The source and relay matrices optimization problem can be written as

$$\min_{\mathbf{F}_1, \tilde{\mathbf{F}}_2} \quad tr\left(\mathbf{E}_0(\mathbf{F}_1, \tilde{\mathbf{F}}_2)\right) \quad (42)$$

$$\text{s.t.} \quad tr\left(\tilde{\mathbf{F}}_2\left(\tilde{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_2}\right)\tilde{\mathbf{F}}_2^H\right) \leq P_2$$
$$(43)$$

$$tr(\mathbf{F}_1\mathbf{F}_1^H) \leq P_1. \quad (44)$$

The problem (42)–(44) 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 (42)–(44).

It can be shown similar to [9] that for given source precoding matrix $\mathbf{F}_1$, the optimal $\tilde{\mathbf{F}}_2$ as the solution to the problem (42)–(44) has the structure of

$$\tilde{\mathbf{F}}_2 = \mathbf{TL} \quad (45)$$

where $\mathbf{T}$ is an $N_2 \times N_b$ matrix that remains to be optimized, and

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

Let us introduce a positive semi-definite (PSD) matrix $\Omega = \mathbf{L}(\breve{\mathbf{H}}_1 \mathbf{Q}^{-1} \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2})\mathbf{L}^H$ and its eigenvalue decomposition (EVD) $\Omega = \mathbf{U}_\omega \mathbf{\Lambda}_\omega \mathbf{U}_\omega^H$, where $\mathbf{\Lambda}_\omega$ is the diagonal eigenvalue matrix with eigenvalues $\lambda_{\omega,k}$, $k = 1, \ldots, N_b$, arranged in descending order. Let us also introduce the EVD of $\tilde{\mathbf{H}}_2^H \tilde{\mathbf{H}}_2 = \mathbf{U}_h \mathbf{\Lambda}_h \mathbf{U}_h^H$, where $\mathbf{\Lambda}_h$ is the diagonal eigenvalue matrix with eigenvalues $\lambda_{h,k}$, $k = 1, \ldots, N_2$, arranged in descending order.

Based on the result in [9], $\mathbf{T}$ has the structure of

$$\mathbf{T} = \mathbf{U}_{h,1} \mathbf{\Delta} \mathbf{U}_\omega^H \tag{46}$$

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

$$\min_{\mathbf{\Delta}} \; tr\left(\left(\mathbf{\Delta}^H \mathbf{\Lambda}_{h,1} \mathbf{\Delta} + \mathbf{\Lambda}_\omega^{-1}\right)^{-1}\right) \tag{47}$$

$$\text{s.t. } tr(\mathbf{\Delta} \mathbf{R}_\omega \mathbf{\Delta}^H) \le P_2. \tag{48}$$

Here $\mathbf{R}_\omega \triangleq \mathbf{U}_\omega^H \mathbf{L}(\breve{\mathbf{H}}_1 \breve{\mathbf{H}}_1^H + \mathbf{I}_{N_2})\mathbf{L}^H \mathbf{U}_\omega$ and $\mathbf{\Lambda}_{h,1}$ contains the largest N_b diagonal elements of $\mathbf{\Lambda}_h$. The problem (47)–(48) can be efficiently solved by the Lagrange multiplier method as

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

where δ_k is the ith 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

$$\sum_{k=1}^{N_b} \frac{R_k}{\lambda_{\omega,k}\lambda_{h,k}} \left(\sqrt{\frac{\lambda_{\omega,k}^2 \lambda_{h,k}}{\gamma R_k}} - 1\right)^+ = P_2.$$

Now we start to optimize the source precoding matrix $\mathbf{F}_1$. Using the matrix inversion lemma, (41) can be rewritten as

$$\mathbf{E}_0(\mathbf{F}_1, \tilde{\mathbf{F}}_2)$$

$$= \left(\mathbf{I}_{N_b} + \tilde{\mathbf{G}}^H \tilde{\mathbf{C}}_v^{-1} \tilde{\mathbf{G}}\right)^{-1}$$

$$= \left(\mathbf{I}_{N_b} + \mathbf{F}_1^H \tilde{\mathbf{H}}_3^H \tilde{\mathbf{H}}_3 \mathbf{F}_1 + \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H\right.$$

$$\left. \times (\tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \mathbf{I}_{N_3})^{-1} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1\right)^{-1}. \tag{49}$$

From (49), for given relay precoding matrix $\tilde{\mathbf{F}}_2$, $\mathbf{F}_1$ is optimized by solving the following problem

$$\min_{\mathbf{F}_1} tr([\mathbf{I}_{N_b} + \mathbf{F}_1^H \mathbf{\Psi} \mathbf{F}_1]^{-1}) \tag{50}$$

$$\text{s.t. } tr\left(\tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \tilde{\mathbf{H}}_1^H \tilde{\mathbf{F}}_2^H\right) \le \bar{P}_2 \tag{51}$$

$$tr(\mathbf{F}_1 \mathbf{F}_1^H) \le P_1 \tag{52}$$

where

$$\mathbf{\Psi} \triangleq \tilde{\mathbf{H}}_3^H \tilde{\mathbf{H}}_3 + \tilde{\mathbf{H}}_1^H \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H \left(\tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H \tilde{\mathbf{H}}_2^H + \mathbf{I}_{N_3}\right)^{-1} \tilde{\mathbf{H}}_2 \tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1.$$

TABLE II
PROCEDURE OF THE BI-STEP ALGORITHM

1) Initialize the algorithm with $\mathbf{F}_1^{(0)} = \sqrt{P_1/N_b}[\mathbf{I}_{N_b}, \mathbf{0}]^T$. Set $n = 0$.
2) Update $\tilde{\mathbf{F}}_2^{(n+1)}$ using $\mathbf{F}_1^{(n)}$ based on (45) and (46) by solving the problem (47)-(48).
3) Update $\mathbf{F}_1^{(n+1)}$ using $\tilde{\mathbf{F}}_2^{(n+1)}$ by solving the problem (53)-(56).
4) if $\left(\text{mse}_2^{(n)} - \text{mse}_2^{(n+1)}\right)/\text{mse}_2^{(n)} < \varepsilon$, iteration ends; otherwise go to step (2).

Let us define $\mathbf{B} \triangleq \mathbf{F}_1 \mathbf{F}_1^H$ and introduce a PSD matrix $\mathbf{X}$ with $\mathbf{X} \succeq (\mathbf{I}_{N_1} + \mathbf{\Psi}^{\frac{1}{2}} \mathbf{B} \mathbf{\Psi}^{\frac{1}{2}})^{-1}$, where $\mathbf{A} \succeq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is a PSD matrix. The problem (50)–(52) can be equivalently converted to the following convex semi-definite programming (SDP) problem by using the Schur complement

$$\min_{\mathbf{X}, \mathbf{B}} tr(\mathbf{X}) \tag{53}$$

$$\text{s.t. } \begin{pmatrix} \mathbf{X} & \mathbf{I}_{N_1} \\ \mathbf{I}_{N_1} & \mathbf{I}_{N_1} + \mathbf{\Psi}^{\frac{1}{2}} \mathbf{B} \mathbf{\Psi}^{\frac{1}{2}} \end{pmatrix} \succeq 0 \tag{54}$$

$$tr\left(\tilde{\mathbf{F}}_2 \tilde{\mathbf{H}}_1 \mathbf{B} \tilde{\mathbf{H}}_1^H \tilde{\mathbf{F}}_2^H\right) \le \bar{P}_2 \tag{55}$$

$$tr(\mathbf{B}) \le P_1, \qquad \mathbf{B} \succeq 0. \tag{56}$$

The problem (53)–(56) can be efficiently solved by the interior-point method [24]. Using the EVD of $\mathbf{B} = \mathbf{U}_b \mathbf{\Lambda}_b \mathbf{U}_b^H$, we have $\mathbf{F}_1 = \mathbf{U}_b \mathbf{\Lambda}_b^{(1/2)}$.

The procedure of using the Bi-Step iterative algorithm to solve the problem (42)–(44) is listed in Table II, where $\text{mse}_2^{(n)}$ stands for the value of (42) at the nth iteration.

IV. NUMERICAL EXAMPLES

In this section, we study the performance of the proposed robust source and relay precoding matrices through numerical simulations. The estimated channel matrices $\bar{\mathbf{H}}_1$, $\bar{\mathbf{H}}_2$, and $\bar{\mathbf{H}}_3$ have i.i.d. complex Gaussian entries with zero-mean and variances of σ_1^2, σ_2^2, and σ_3^2, respectively. All noises are i.i.d. AWGN with zero mean and unit variance. Following [10], we define $\text{SNR}_1 = \sigma_1^2 P_1/N_2$, $\text{SNR}_2 = \sigma_2^2 P_2/N_3$, and $\text{SNR}_3 = \sigma_3^2 P_1/N_3$ as the signal-to-noise ratio (SNR) for the source-relay, relay-destination, and source-destination links, respectively.

We simulate a MIMO relay system where the channel estimation error at the transmitter side is uncorrelated, i.e., $\mathbf{\Theta}_1 = \mathbf{\Theta}_3 = \sigma_e^2 \mathbf{I}_{N_1}$ and $\mathbf{\Theta}_2 = \sigma_e^2 \mathbf{I}_{N_2}$, where σ_e^2 measures the variance of the channel estimation error. We obtain from (62) that for this case, $\alpha_1 = \alpha_3 = \sigma_e^2 tr(\mathbf{F}_1 \mathbf{F}_1^H)$ and $\alpha_2 = \sigma_e^2 tr(\mathbf{F}_2(\bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H + \alpha_1 \mathbf{\Phi}_1 + \mathbf{I}_{N_2})\mathbf{F}_2^H)$. It can be shown similar to [15] that $tr(E_0(\mathbf{F}_1, \tilde{\mathbf{F}}_2))$ decreases with respect to α_1, α_2, and α_3. Therefore, considering the power constraints (17) and (18), the optimal solution occurs at $\alpha_1 = \alpha_3 = \sigma_e^2 P_1$ and $\alpha_2 = \sigma_e^2 P_2$.

For simplicity, in all simulation examples, we set $N_b = N_i = 4$, $i = 1, 2, 3$. The covariance matrix of channel estimation error
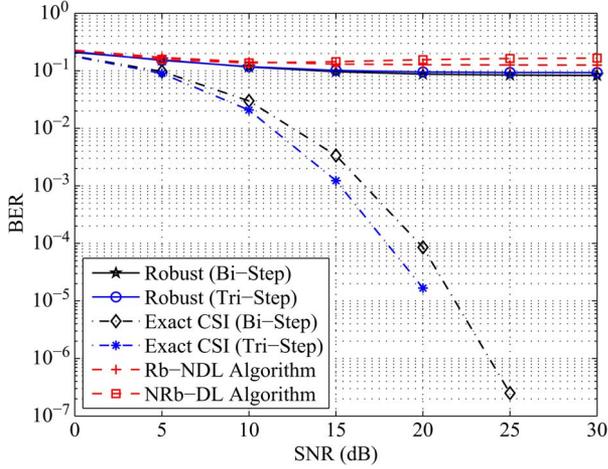
Fig. 2. Example 1: BER versus SNR, $\sigma_e^2 = 0.1$, $\Delta_{\text{SNR}} = 20$ dB.

at the receiver side is set as

$$\mathbf{\Phi}_1 = \begin{bmatrix} 1 & \phi_1 & \phi_1^2 & \phi_1^3 \\ \phi_1 & 1 & \phi_1 & \phi_1^2 \\ \phi_1^2 & \phi_1 & 1 & \phi_1 \\ \phi_1^3 & \phi_1^2 & \phi_1 & 1 \end{bmatrix},$$

$$\mathbf{\Phi}_2 = \mathbf{\Phi}_3 = \begin{bmatrix} 1 & \phi_2 & \phi_2^2 & \phi_2^3 \\ \phi_2 & 1 & \phi_2 & \phi_2^2 \\ \phi_2^2 & \phi_2 & 1 & \phi_2 \\ \phi_2^3 & \phi_2^2 & \phi_2 & 1 \end{bmatrix}.$$

In the simulations, we choose $\phi_1 = \phi_2 = 0.45$.

We compare the performance of the following six systems: (1) The nonrobust transceiver scheme with the direct link in [8] (referred to as the NRb-DL algorithm); (2) The robust transceiver design without the direct link in [10] (referred to as the Rb-NDL algorithm); (3) The proposed robust design using the Tri-Step algorithm; (4) The proposed robust design with the Bi-Step algorithm; (5) The Tri-Step algorithm with the exact CSI knowledge; (6) The Bi-Step algorithm with the exact CSI knowledge. For the robust design, the initialization of the Tri-Step and Bi-Step algorithms is listed in Tables I and II, respectively.[2] All simulation results are averaged over 1000 independent realizations of the true channel matrices. QPSK constellations are used to modulate the source symbols. We set $\text{SNR}_1 = \text{SNR}_2 = \text{SNR}$ and $\text{SNR}_3 = \text{SNR} - \Delta_{\text{SNR}}$ in the unit of dB, where Δ_{SNR} denotes the attenuation of the direct link. For both the proposed Tri-Step and Bi-Step algorithms, unless explicitly mentioned, the convergence criterion is chosen as $\varepsilon = 0.01$.

In the first simulation example, we compare the bit-error-rate (BER) performance of the six algorithms at different σ_e^2 with $\Delta_{\text{SNR}} = 20$ dB. Figs. 2–4 show the system BER versus SNR of

[2]We also tried the following two initializations: (a) $\mathbf{F}_1 = c_1\mathbf{U}$ and $\tilde{\mathbf{F}}_2 = c_2\mathbf{U}$ for the Tri-Step algorithm; and $\mathbf{F}_1 = c_1\mathbf{U}$ for the Bi-Step algorithm, where $\mathbf{U}$ is a 4×4 random unitary matrix, $c_1 = 0.5\sqrt{P_1}$, and $c_2 = \sqrt{P_2/tr(\tilde{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\tilde{\mathbf{H}}_1^H + \mathbf{I}_4)}$. (b) $\mathbf{F}_1 = c_1\mathbf{D}$ and $\tilde{\mathbf{F}}_2 = c_2\mathbf{D}$ for the Tri-Step algorithm; and $\mathbf{F}_1 = c_1\mathbf{D}$ for the Bi-Step algorithm, where $\mathbf{D}$ is a 4×4 diagonal matrix whose main diagonal entries are $[\sqrt{2}, 1, \sqrt{0.5}, \sqrt{0.5}]$. We observed that initializations (a) and (b) result in almost identical BER and MSE performance as the initialization used here.
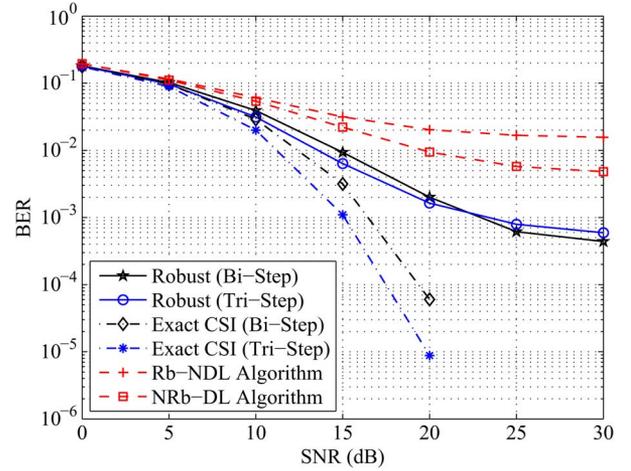
Fig. 3. Example 1: BER versus SNR, $\sigma_e^2 = 0.01$, $\Delta_{\text{SNR}} = 20$ dB.
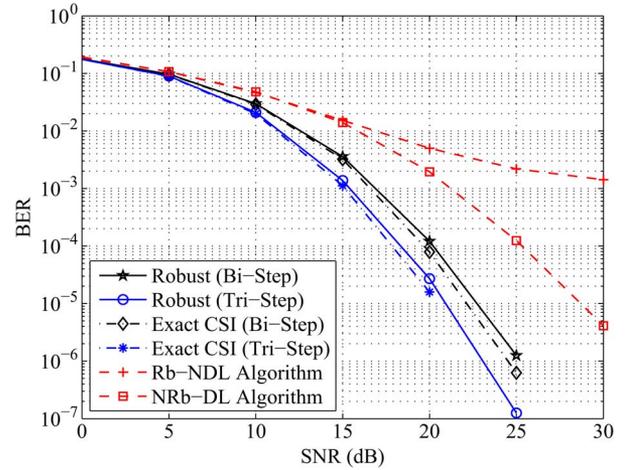
Fig. 4. Example 1: BER versus SNR, $\sigma_e^2 = 0.001$, $\Delta_{\text{SNR}} = 20$ dB.

the algorithms tested at $\sigma_e^2 = 0.1, 0.01$, and 0.001, respectively. It can be seen that

- As expected, the two algorithms with the exact CSI knowledge have the lowest BER and the performance of these two algorithms is invariant to σ_e^2.
- Since the proposed robust Bi-Step and Tri-Step algorithms consider both the CSI mismatch and the direct link, they consistently outperform the NRb-DL and Rb-NDL algorithms for all σ_e^2 tested. Moreover, their performance approaches that of the system with the exact CSI as σ_e^2 decreases.
- The BER performance of the four algorithms with CSI mismatch improves as σ_e^2 decreases. In particular, the algorithms that consider the direct link (i.e., the Bi-Step, Tri-Step, and NRb-DL algorithms) have a larger BER reduction than the Rb-NDL algorithm when σ_e^2 decreases. This verifies the importance of considering the direct link in the transceiver design.
- The Rb-NDL algorithm has a better BER performance than the NRb-DL algorithm at $\sigma_e^2 = 0.1$, indicating the gain of robust design. However, for $\sigma_e^2 = 0.01$ and 0.001, the NRb-DL algorithm performs much better than the
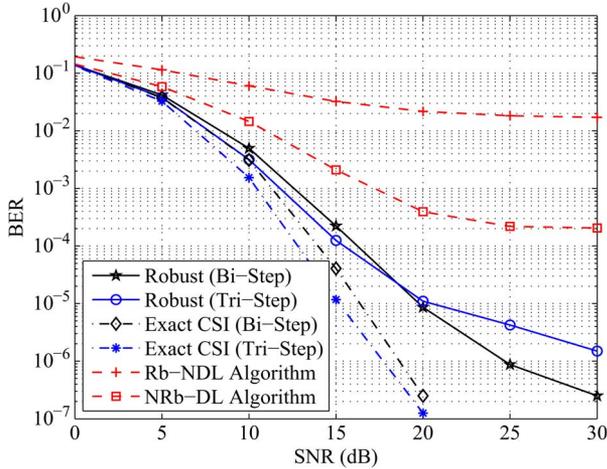
Fig. 5.   Example 2: BER versus SNR, $\sigma_e^2 = 0.01$, $\Delta_{\mathrm{SNR}} = 10$ dB.

Fig. 7.   Example 3: MSE comparison of the Tri-Step and Bi-Step algorithms, $\Delta_{\mathrm{SNR}} = 20$ dB.

Fig. 6.   Example 2: BER versus SNR, $\sigma_e^2 = 0.01$, $\Delta_{\mathrm{SNR}} = 30$ dB.
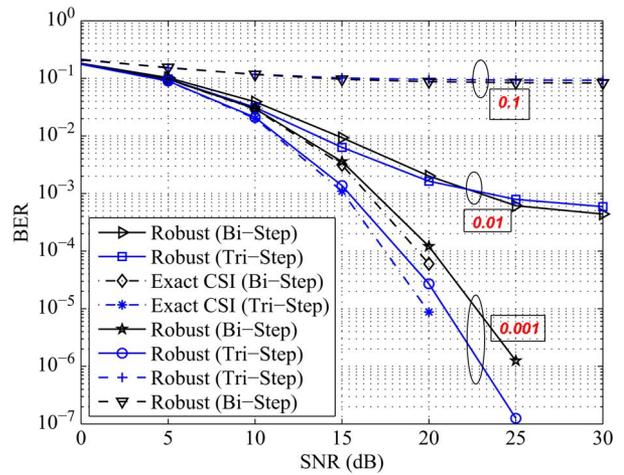
Fig. 8.   Example 3: BER comparison of the Tri-Step and Bi-Step algorithms, $\Delta_{\mathrm{SNR}} = 20$ dB.

Rb-NDL algorithm with increasing SNR, as the impact of the direct link becomes larger when SNR increases.

In the second simulation example, we compare the BER performance of the six algorithms tested at different Δ_{SNR} with $\sigma_e^2 = 0.01$. Figs. 5, 3, and 6 demonstrate the system BER versus SNR of the algorithms tested with $\Delta_{\mathrm{SNR}} = 10$ dB, 20 dB, and 30 dB, respectively. We can observe that

- The proposed robust Bi-Step and Tri-Step algorithms perform better than the NRb-DL and Rb-NDL algorithms, as the proposed algorithms consider both the CSI mismatch and the direct link.
- The BER of the Rb-NDL algorithm is invariant to the value of Δ_{SNR}.
- The BER performance of the other five algorithms (except the Rb-NDL algorithm) improves when Δ_{SNR} decreases (the strength of the direct link increases).
- The four algorithms using the mismatched CSI reach different error floors in BER at high SNR. When the direct link is relatively strong (i.e., $\Delta_{\mathrm{SNR}} = 10$ dB and 20 dB), the DR-NRb algorithm performs better than the Rb-NDL algorithm. When SNR of the direct link is low (i.e., $\Delta_{\mathrm{SNR}} = 30$ dB), the robust designs have a better BER performance.

In the third simulation example, we compare the performance of the Tri-Step and Bi-Step algorithms. Figs. 7 and 8 show the MSE and the BER of the two algorithms respectively versus SNR at $\Delta_{\mathrm{SNR}} = 20$ dB and different σ_e^2. It can be seen from Fig. 7 that the Bi-Step algorithm has better MSE performance than the Tri-Step method at different σ_e^2 when $\varepsilon = 0.01$. However, we observe from Fig. 8 that the Tri-Step algorithm has a better BER performance when σ_e^2 is small (i.e., $\sigma_e^2 = 0.001$), when σ_e^2 becomes large ($\sigma_e^2 = 0.01$ and 0.1), the Bi-Step algorithm has a better BER performance than the Tri-Step algorithm at high SNR. This can be explained as follows. The performance of both algorithms depends on several factors such as the initial point, the convergence criterion, the convergence speed, and the shape of the objective function. In the Tri-Step algorithm, both $\mathbf{F}_1$ and $\tilde{\mathbf{F}}_2$ need to be initialized, and three matrices $\mathbf{F}_1$, $\tilde{\mathbf{F}}_2$, and $\tilde{\mathbf{W}}$ are optimized in each iteration, while for the Bi-Step algorithm, only $\mathbf{F}_1$ needs to be initialized, and only two matrices $\mathbf{F}_1$ and $\tilde{\mathbf{F}}_2$ need to be updated in each iteration. Therefore, the Bi-Step algorithm converges faster than the Tri-Step algorithm and the former one has a better MSE performance when $\varepsilon = 0.01$.

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

SNR_1(dB)	5	10	15	20	25	30
Bi-Step (CSI mismatch)	3	3	3	4	4	4
Tri-Step (CSI mismatch)	4	6	8	10	11	11
Bi-Step (exact CSI)	3	3	4	4	4	4
Tri-Step (exact CSI)	3	5	8	12	19	25
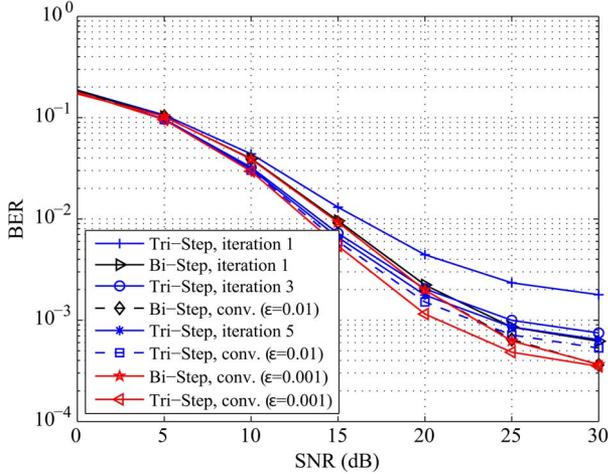
Fig. 9. Example 3: BER of the Tri-Step and Bi-Step algorithms at different number of iterations, $\sigma_e^2 = 0.01$, $\Delta_{\mathrm{SNR}} = 20$ dB.

To verify our analysis, we list the average number of iterations required by both algorithms to converge to $\varepsilon = 0.01$ at $\sigma_e^2 = 0.01$, $\Delta_{\mathrm{SNR}} = 20$ dB and varying SNR in Table III. It can be seen from Table III that the Tri-Step algorithm needs a larger number of iterations till convergence. Interestingly, the number of iterations required by the Bi-Step algorithm is almost invariant to the value of SNR, while the number of iterations needed by the Tri-Step algorithm increases with SNR. Similar results have been observed in simulation examples with different σ_e^2 and Δ_{SNR}.

The BER performance of the Tri-Step and Bi-Step algorithms at different number of iterations is shown in Fig. 9 for $\sigma_e^2 = 0.01$ and $\Delta_{\mathrm{SNR}} = 20$ dB. It can be seen that the Bi-Step algorithm converges faster than the Tri-Step algorithm, confirming the results in Table III. For $\varepsilon = 0.01$, at medium SNR range (10–20 dB), the BER reduction of the Tri-Step algorithm with iterations is larger than that of the Bi-Step algorithm. This explains the crossover of the BER curves of the two algorithms in Figs. 3, 5, and 8. Interestingly, the Bi-Step algorithm yields almost the same BER for $\varepsilon = 0.01$ and $\varepsilon = 0.001$, while the Tri-Step algorithm has a much smaller BER for $\varepsilon = 0.001$ compared with $\varepsilon = 0.01$, especially at high SNR. In fact, the Tri-Step algorithm has a better BER performance than the Bi-Step algorithm at $\varepsilon = 0.001$. This is because the Bi-Step algorithm converges after several iterations even for a smaller ε. On the other hand, the Tri-Step algorithm has a slow convergence rate, and thus, its BER performance improves when a smaller ε is chosen. We have observed that if ε is small enough, the Tri-Step algorithm always has a better BER performance than the Bi-Step algorithm in both the perfect CSI case and the CSI mismatch scenario.

Finally, we compare the computational complexity of the Tri-Step and Bi-Step algorithms. For the sake of notational simplicity, we assume $N_b = N_i = N$, $i = 1, 2, 3$. In each iteration of the Tri-Step algorithm, the complexity order of updating $\mathbf{W}$ and $\mathbf{F}_2$ is $\mathcal{O}(N^3)$, which mainly involves matrix inversion and matrix SVD. Solving the QCQP problem (38)–(40) to update $\mathbf{F}_1$ has a complexity order of $\mathcal{O}(N^6)$ [27]. Thus, the per iteration computational complexity of the Tri-Step algorithm is $\mathcal{O}(N^6)$. For each iteration of the Bi-Step algorithm, the major computation task in updating $\mathbf{F}_2$ is matrix EVD, which has a complexity order of $\mathcal{O}(N^3)$. The complexity order of updating $\mathbf{F}_1$ through solving the SDP problem (53)–(56) is $\mathcal{O}(N^7)$ [27]. Thus, the Bi-Step algorithm has a per iteration complexity order of $\mathcal{O}(N^7)$. Based on the analysis above, we can see that the Bi-Step algorithm has a higher per iteration complexity.

The overall computational complexity of the Tri-Step and Bi-Step algorithms depends also on the number of iterations required till convergence. From Table III, we find that for small $N(N \leq 3)$ and high SNR, the Bi-Step algorithm has a smaller overall complexity. On the other hand, when N is large and the SNR is low, the Tri-Step algorithm has a smaller complexity.

V. CONCLUSION

We have derived the optimal structure of the relay precoding matrix for two-hop MIMO relay systems with the direct source-destination link and imperfect CSI. We have developed two iterative algorithms to design the statistically robust source and relay matrices for the commonly used MMSE criterion. Simulation results show an improved robustness of the proposed algorithms against CSI mismatch compared with existing approaches. Besides, the iterative algorithms developed in this paper for single-user two-hop MIMO relay systems can be extended to the multiple-user case in the future.

APPENDIX A
Proof of (13)

From (7) we have

$$\bar{\mathbf{A}} = E_H[\mathbf{G}\mathbf{G}^H] + E_{H_2}[\mathbf{C}_v]. \tag{57}$$

Using (1) and the following identity from [22]

$$E_H\left[\mathbf{H}_i\mathbf{X}\mathbf{H}_i^H\right] = \bar{\mathbf{H}}_i\mathbf{X}\bar{\mathbf{H}}_i^H + tr\left(\mathbf{X}\mathbf{\Theta}_i^T\right)\mathbf{\Phi}_i \tag{58}$$

we obtain

$$E_{H_2}[\mathbf{C}_v] = \begin{bmatrix} E_{H_2}[\mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^H\mathbf{H}_2^H] + \mathbf{I}_{N_3} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_3} \end{bmatrix}$$
$$= \bar{\mathbf{C}}_v + \begin{bmatrix} tr(\mathbf{F}_2\mathbf{F}_2^H\mathbf{\Theta}_2^T)\mathbf{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{59}$$

From (8) and (58), we obtain that

$$\mathrm{E}_{H_1, H_2}\left(\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H\mathbf{F}_2^H\mathbf{H}_2^H\right)$$
$$= \mathrm{E}_{H_2}\left(\mathbf{H}_2\mathbf{F}_2\mathrm{E}_{H_1}\left(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H\right)\mathbf{F}_2^H\mathbf{H}_2^H\right)$$
$$= \mathrm{E}_{H_2}\left(\mathbf{H}_2\mathbf{F}_2\left(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1\right)\mathbf{F}_2^H\mathbf{H}_2^H\right)$$

$$= \bar{\mathbf{H}}_2 \mathbf{F}_2 \left(\bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H + \alpha_1 \mathbf{\Phi}_1 \right) \mathbf{F}_2^H \bar{\mathbf{H}}_2^H$$

$$+ tr \left(\mathbf{F}_2 \left(\bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H + \alpha_1 \mathbf{\Phi}_1 \right) \mathbf{F}_2^H \mathbf{\Theta}_2^T \right) \mathbf{\Phi}_2 \quad (60)$$

where $\alpha_1 \triangleq tr(\mathbf{F}_1 \mathbf{F}_1^H \mathbf{\Theta}_1^T)$. From (7) and (60), we obtain that

$$E_H[\mathbf{G}\mathbf{G}^H]$$

$$= E_H \left[\begin{bmatrix} \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{F}_2^H \mathbf{H}_2^H & \mathbf{H}_2 \mathbf{F}_2 \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_3^H \\ \mathbf{H}_3 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{F}_2^H \mathbf{H}_2^H & \mathbf{H}_3 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_3^H \end{bmatrix} \right]$$

$$= \bar{\mathbf{G}} \bar{\mathbf{G}}^H + \begin{bmatrix} \alpha_1 \bar{\mathbf{H}}_2 \mathbf{F}_2 \mathbf{\Phi}_1 \mathbf{F}_2^H \bar{\mathbf{H}}_2^H + \beta \mathbf{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \alpha_3 \mathbf{\Phi}_3 \end{bmatrix} \quad (61)$$

where $\beta \triangleq tr(\mathbf{F}_2(\bar{\mathbf{H}}_1 \mathbf{F}_1 \mathbf{F}_1^H \bar{\mathbf{H}}_1^H + \alpha_1 \mathbf{\Phi}_1) \mathbf{F}_2^H \mathbf{\Theta}_2^T)$ and $\alpha_3 \triangleq tr(\mathbf{F}_1 \mathbf{F}_1^H \mathbf{\Theta}_3^T)$.

Substituting (59) and (61) back into (57), we obtain (13) with

$$\alpha_2 = \beta + tr \left(\mathbf{F}_2 \mathbf{F}_2^H \mathbf{\Theta}_2^T \right). \quad (62)$$

ACKNOWLEDGMENT

REFERENCES

[1] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A tutorial on transceiver design for amplify-and-forward MIMO relay systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1331–1346, Sep. 2012.

[2] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.

[3] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2593–2604, Jun. 2007.

[4] 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.

[5] 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, no. 12, pp. 4837–4851, Dec. 2009.

[6] Y. Rong, "Optimal joint source and relay beamforming for MIMO relays with direct link," *IEEE Commun. Lett.*, vol. 14, no. 5, pp. 390–392, May 2010.

[7] 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, no. 4, pp. 1687–1698, May 2011.

[8] F.-S. Tseng, W.-R. Wu, and J.-Y. Wu, "Joint source/relay precoder design in nonregenerative cooperative systems using an MMSE criterion," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4928–4933, Oct. 2009.

[9] C. Song, K.-J. Lee, and I. Lee, "MMSE-based MIMO cooperative relaying systems: Closed-form designs and outage behavior," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1390–1401, Sep. 2012.

[10] 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, no. 4, pp. 2273–2283, Apr. 2010.

[11] B. K. Chalise and L. Vandendorpe, "Joint linear processing for an amplify-and-forward MIMO relay channel with imperfect channel state information," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 640186-1–640186-13, Aug. 2010.

[12] C. Xing, S. Ma, Y.-C. Wu, and T.-S. Ng, "Transceiver design for dual-hop non-regenerative MIMO-OFDM relay systems under channel uncertainties," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6325–6339, Dec. 2010.

[13] 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, no. 5, pp. 1196–1209, Mar. 2013.

[14] C. Xing, M. Xia, F. Gao, and Y.-C. Wu, "Robust transceiver with Tomlinson-Harashima precoding for amplify-and-forward MIMO relaying systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1370–1382, Sep. 2012.

[15] Y. Rong, "Robust design for linear non-regenerative MIMO relays with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2455–2460, May 2011.

[16] F.-S. Tseng, M.-Y. Chang, and W.-R. Wu, "Robust Tomlinson-Harashima source and linear relay precoders design in amplify-and-forward MIMO relay systems," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1124–1137, Apr. 2012.

[17] "Relay architectures for E-UTRA (LTE-Advanced)," Sophia-Antipolis, France, TR 36.806, Mar. 2010.

[18] A. Osseiran *et al.*, "The road to IMT-Advanced communication systems: State-of-the-art and innovation areas addressed by the WINNER+ project," *IEEE Commun. Mag.*, vol. 47, no. 6, pp. 38–47, Jun. 2009.

[19] S. W. Peters and R. W. Heath, Jr., "The future of WiMAX: Multihop relaying with IEEE 802.16j," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 104–111, Jan. 2009.

[20] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4438–4446, Sep. 2011.

[21] Y. Rong, M. R. A. Khandaker, and Y. Xiang, "Channel estimation of dual-hop MIMO relay systems via parallel factor analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2224–2233, Jun. 2012.

[22] A. Gupta and D. Nagar, *Matrix Variate Distributions*. London, U. K.: Chapman & Hall, 2000.

[23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U. K.: Cambridge Univ. Press, 2004.

[25] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, no. 9, pp. 772–781, Sep. 1978.

[26] M. Grant and S. Boyd. (2010, Apr.). "CVX: Matlab software for disciplined convex programming (web page and software)." [Online]. Available: http://cvxr.com/cvx

[27] Y. Nesterov and A. Nemirovsky, "Interior-point polynomial methods in convex programming," in *Volume 13 of Studies in Applied Mathematics*. Philadelphia, PA, USA: SIAM, 1994.

Zhiqiang He (S'01–M'04) received the B.E. and Ph.D. degrees (with distinction) in signal and information processing from Beijing University of Posts and Telecommunications, China, 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.

Weipeng Jiang received the B.E. 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.

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, CA, USA, 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 an Associate 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.

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 Editor of IEEE WIRELESS COMMUNICATIONS LETTERS, 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.