Robust Design for Linear Non-Regenerative MIMO Relays With Imperfect Channel State Information

Yue Rong

Abstract—**In this correspondence, we address statistically robust multiple-input multiple-output (MIMO) relay design problems under two imperfect channel state information (CSI) scenarios: 1) a ll nodes have imperfect CSI and 2) the destination node knows the exact CSI, while the other nodes have imperfect CSI. For each scenario, we develop robust source and relay matrices by considering a broad class of frequently used objective functions in MIMO system design and the averaged transmission power constraints. Simulation results demonstrate the improved robustness of the proposed algorithms against CSI errors.**

Index Terms—**Channel state information, majorization, MIMO relay, MMSE, robustness.**

I. Introduction

Recently, there have been many research efforts on linear nonregenerative multiple-input multiple-output (MIMO) relay systems [1]–[4]. The optimal relay amplifying matrix is obtained in [1], [2] to maximize the mutual information between source and destination. In [3], optimal relay matrices are developed to minimize the mean-square error (MSE) of the signal waveform estimation at the destination. A unified framework is established for optimizing the source precoding matrix and the relay amplifying matrix of linear nonregenerative MIMO relay systems with a broad class of objective functions [4].

For MIMO relay systems, the channel state information (CSI) knowledge of all hops is required at the destination node to estimate the source signals. Moreover, in order to optimize the source and relay matrices in [1]–[4], the CSI knowledge of all hops is needed at the node which carries out the optimization procedure. However, in practical relay communication systems, the exact 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 There is always mismatch between the true and the estimated CSI due to channel noise, quantization errors and outdated channel estimates.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 in [1]–[4] will degrade due to such CSI mismatch. In [5]–[7], MMSE-based optimal relay amplifying and destination receiving matrices for a two-hop MIMO relay system have been developed taking into account the CSI mismatch. However, the source precoding matrix is not optimized in [5]–[7]. The source precoding matrix optimization under CSI mismatch is investigated in [8] and [9] using the MMSE criterion.

In this correspondence, we investigate statistically robust two-hop MIMO relay systems. In contrast to [5]–[9], we develop robust source precoding matrix and relay amplifying matrix by considering a broad class of frequently used objective functions in MIMO system design [4] (e.g., maximal mutual information, MMSE). In particular, we consider two imperfect CSI scenarios: 1) All nodes have imperfect CSI and 2) The destination node knows the exact CSI, while the source and relay nodes have imperfect CSI. The true CSI is modelled 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 [5]–[9]. We would like to point out that in [10], only the imperfect CSI case 1) is addressed.

For each of the two imperfect CSI scenarios, we show that the CSI mismatch information is intrinsically embedded in the structure of the optimal robust source and relay matrices. Moreover, the available power at the source and the relay nodes is optimally distributed among all data streams in a robust fashion against the CSI mismatch. Such robust power allocation can be implemented with the same computational complexity as the nonrobust power allocation scheme in [4]. Interestingly, when the exact CSI is available (i.e., no CSI mismatch), the robust source and relay matrices become the optimal source and relay matrices developed in [4]. Thus, this correspondence is an important generalization of [4] to the practical scenario of imperfect CSI. Simulation results demonstrate the improved robustness of the proposed approaches against the CSI mismatch.

The rest of this correspondence is organized as follows. In Section II, we introduce the model of a two-hop linear nonregenerative MIMO relay communication system. The robust source and relay matrices are developed in Sections III and IV, depending on whether the destination node has the exact CSI knowledge. In Section V, we show some numerical examples. Conclusions are drawn in Section VI.

II. System Model

We consider a three-node MIMO communication system where the source node (node 1) transmits information to the destination node (node 3) with the aid of one relay node (node 2). The ith node is equipped with $N_i,\ i=1,2,3$, antennas. We focus on the case where the direct link between the source and destination nodes is sufficiently weak to be ignored as in [1]–[4]. This scenario occurs when the direct link is blocked by an obstacle such as a mountain. Using the nonregenerative relay strategy, the received signal vector at the destination node can be written as

$$\mathbf{y} = \mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1\mathbf{s} + \mathbf{H}_2\mathbf{F}_2\mathbf{v}_2 + \mathbf{v}_3 \triangleq \mathbf{G}\mathbf{s} + \bar{\mathbf{v}} \tag{1}$$

where $\mathbf{s}$ is the $N_b \times 1$ source signal vector, $\mathbf{F}_1$ is the $N_1 \times N_b$ source precoding matrix, $\mathbf{H}_1$ is the $N_2 \times N_1$ MIMO fading channel matrix between the source and relay nodes, $\mathbf{F}_2$ is the $N_2 \times N_2$ relay amplifying matrix, $\mathbf{H}_2$ is the $N_3 \times N_2$ MIMO fading channel matrix between the relay and destination nodes, $\mathbf{v}_2$ is an $N_2 \times 1$ noise vector at the relay node and $\mathbf{v}_3$ is an $N_3 \times 1$ noise vector at the destination node. Here $\mathbf{G} \triangleq \mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1$ is the equivalent source–destination MIMO channel matrix and $\bar{\mathbf{v}} \triangleq \mathbf{H}_2\mathbf{F}_2\mathbf{v}_2 + \mathbf{v}_3$ is the equivalent noise vector.

We assume that $\mathrm{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_{N_b}$, where $\mathrm{E}[\cdot]$ stands for the statistical expectation, $(\cdot)^H$ denotes the Hermitian transpose and $\mathbf{I}_n$ is an $n \times n$ identity matrix. In order to avoid any transmission power loss at each node, there should be $N_b \leq \min(N_1, N_2, N_3)$. We also assume that all noises are independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) with zero mean and unit variance.

Using a linear receiver at the destination node, the estimated signal vector can be written as $\hat{\mathbf{s}} = \mathbf{W}^H\mathbf{y}$, where $\mathbf{W}$ is the $N_3 \times N_b$ weight matrix. The MSE matrix of the signal waveform estimation $\mathbf{E} \triangleq \mathrm{E}\left[(\hat{\mathbf{s}}-\mathbf{s})(\hat{\mathbf{s}}-\mathbf{s})^H\right]$ can be written as

$$\mathbf{E} = (\mathbf{W}^H\mathbf{G} - \mathbf{I}_{N_b})(\mathbf{W}^H\mathbf{G} - \mathbf{I}_{N_b})^H + \mathbf{W}^H\mathbf{C}_{\bar{v}}\mathbf{W} \tag{2}$$

where $\mathbf{C}_{\bar{v}} \triangleq \mathrm{E}[\bar{\mathbf{v}}\bar{\mathbf{v}}^H] = \mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^H\mathbf{H}_2^H + \mathbf{I}_{N_3}$ is the noise covariance matrix. It has been shown in [4] that a broad class of frequently used MIMO relay system design objectives such as the source–destination

Manuscript received November 15, 2010; accepted February 02, 2011. Date of publication February 10, 2011; Date of current version April 13, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Visa Koivunen. This work was supported under the Australian Research Council's Discovery Projects funding scheme (project numbers DP110100736, DP110102076). This correspondence was presented in part at the Australian Communications Theory Workshop (AusCTW), Canberra, Australia, February 2010.

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

Digital Object Identifier 10.1109/TSP.2011.2113376

mutual information can be written as a function of the main diagonal elements of the MSE matrix $\mathbf{E}$.

With mismatch between the true and the estimated CSI, the true channel $\mathbf{H}_i$ can be represented by the well-known Gaussian–Kronecker model [5]–[9], where $\mathbf{H}_i$ is a complex-valued Gaussian random matrix

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

Here, the mean value is the estimated channel matrix $\bar{\mathbf{H}}_i$, $\boldsymbol{\Theta}_i$ denotes the $N_i \times N_i$ covariance matrix of channel estimation error at the transmitter side, while $\boldsymbol{\Phi}_i$ is the $N_{i+1} \times N_{i+1}$ covariance matrix of channel estimation error seen from the receiver side and $\otimes$ stands for the matrix Kronecker product. In other words, we have $\mathbf{H}_i = \bar{\mathbf{H}}_i + \mathbf{A}_{\Phi_i}\mathbf{H}_{\mathrm{w}_i}\mathbf{A}_{\Theta_i}^H$, $i = 1, 2$, where $\mathbf{A}_{\Phi_i}\mathbf{A}_{\Phi_i}^H = \boldsymbol{\Phi}_i$, $\mathbf{A}_{\Theta_i}\mathbf{A}_{\Theta_i}^H = \boldsymbol{\Theta}_i^T$ and $\mathbf{H}_{\mathrm{w}_i}$ is an $N_{i+1} \times N_i$ Gaussian random matrix with i.i.d. zero mean and unit variance entries and is the unknown part in the CSI mismatch.

Lemma 1 [12]: For $\mathbf{H} \sim \mathcal{CN}(\bar{\mathbf{H}}, \boldsymbol{\Theta} \otimes \boldsymbol{\Phi})$, there is $\mathrm{E}_{\mathrm{H}}[\mathbf{H}\mathbf{A}\mathbf{H}^H] = \bar{\mathbf{H}}\mathbf{A}\bar{\mathbf{H}}^H + \mathrm{tr}(\mathbf{A}\boldsymbol{\Theta}^T)\boldsymbol{\Phi}$ and $\mathrm{E}_{\mathrm{H}}[\mathbf{H}^H\mathbf{A}\mathbf{H}] = \bar{\mathbf{H}}^H\mathbf{A}\bar{\mathbf{H}} + \mathrm{tr}(\boldsymbol{\Phi}\mathbf{A})\boldsymbol{\Theta}^T$, where $\mathrm{tr}(\cdot)$ denotes the matrix trace and $(\cdot)^T$ stands for the matrix transpose.

III. Robust MIMO Relay Design With Imperfect CSI at All Nodes

In this section, we consider the scenario where all nodes have imperfect CSI. It can be seen from (2) that if the exact $\mathbf{H}_1$ and $\mathbf{H}_2$ are unavailable at the destination node, it is impossible to design $\mathbf{W}$ that optimizes $\mathbf{E}$ in (2). If we design $\mathbf{W}$, $\mathbf{F}_1$ and $\mathbf{F}_2$ based only on $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$, there can be a great performance degradation due to the mismatch between $\mathbf{H}_i$ and $\bar{\mathbf{H}}_i$, $i = 1, 2$. Instead of optimizing $\mathbf{E}$, we design $\mathbf{W}$ to minimize $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]$, where the statistical expectation is carried out with respect to $\mathbf{H}_1$ and $\mathbf{H}_2$, with the distribution given in (3).

The statistical expectation of $\mathbf{E}$ is given by [10]

$$\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}] = \mathbf{W}^H\mathbf{A}\mathbf{W} - \mathbf{W}^H\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1 - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{W} + \mathbf{I}_{N_b} \tag{4}$$

where

$$\mathbf{A} \triangleq \bar{\mathbf{H}}_2\mathbf{F}_2\left(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\boldsymbol{\Phi}_1 + \mathbf{I}_{N_2}\right)\mathbf{F}_2^H\bar{\mathbf{H}}_2^H \alpha_2\boldsymbol{\Phi}_2 + \mathbf{I}_{N_3} \tag{5}$$

$$\alpha_1 \triangleq \mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H\boldsymbol{\Theta}_1^T) \tag{6}$$

$$\alpha_2 \triangleq \mathrm{tr}\left(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\boldsymbol{\Phi}_1 + \mathbf{I}_{N_2})\mathbf{F}_2^H\boldsymbol{\Theta}_2^T\right). \tag{7}$$

Now the weight matrix $\mathbf{W}$ which minimizes (4) is the famous Wiener filter given by

$$\mathbf{W} = \mathbf{A}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1 \tag{8}$$

where $(\cdot)^{-1}$ denotes the matrix inversion. Substituting (8) back into (4), we have

$$\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}] = \mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{A}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1. \tag{9}$$

The transmission power consumed by the relay node can be written as $p_2 = \mathrm{tr}\left(\mathbf{F}_2\left(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H + \mathbf{I}_{N_2}\right)\mathbf{F}_2^H\right)$. However, since the true $\mathbf{H}_1$ is unknown, p_2 is also unknown. In this correspondence, we consider the averaged transmission power at the relay node, which is given by

$$\begin{aligned}\mathrm{E}_{\mathrm{H}_1}[p_2] &= \mathrm{tr}\left(\mathbf{F}_2\left(\mathrm{E}[\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H] + \mathbf{I}_{N_2}\right)\mathbf{F}_2^H\right)\\ &= \mathrm{tr}\left(\mathbf{F}_2\left(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\boldsymbol{\Phi}_1 + \mathbf{I}_{N_2}\right)\mathbf{F}_2^H\right)\end{aligned} \tag{10}$$

where Lemma 1 is applied to obtain (10). As in [4], we use q to denote a unified objective function and $\mathbf{d}[\mathbf{E}]$ stands for the main diagonal elements of $\mathbf{E}$. Instead of optimizing $q(\mathbf{d}[\mathbf{E}])$ in [4], we minimize the objective function of $q(\mathbf{d}[\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]])$. Combining (9) and (10), the robust source and relay matrices optimization problem can be written as

$$\min_{\mathbf{F}_1,\mathbf{F}_2} \quad q\left(\mathbf{d}\left[\mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{A}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\right]\right) \tag{11}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}_2\left(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\boldsymbol{\Phi}_1 + \mathbf{I}_{N_2}\right)\mathbf{F}_2^H\right) \le P_2 \tag{12}$$

$$\mathrm{tr}\left(\mathbf{F}_1\mathbf{F}_1^H\right) \le P_1 \tag{13}$$

where $P_i > 0$, $i = 1, 2$, is the transmission power available at the ith node, (12) and (13) represent the transmission power constraint at the relay node and the source node, respectively. The problem (11)–(13) provides a statistically robust design of $\mathbf{F}_1$ and $\mathbf{F}_2$ when all nodes have imperfect CSI.

Let us introduce the following matrix eigenvalue decomposition (EVD) and singular value decomposition (SVD) for $i = 1, 2$

$$\boldsymbol{\Phi}_i = \mathbf{U}_{\Phi_i}\boldsymbol{\Lambda}_{\Phi_i}\mathbf{U}_{\Phi_i}^H \tag{14}$$

$$\tilde{\boldsymbol{\Lambda}}_{\Phi_i} \triangleq \alpha_i\boldsymbol{\Lambda}_{\Phi_i} + \mathbf{I}_{N_{i+1}} \tag{15}$$

$$\tilde{\mathbf{H}}_i \triangleq \tilde{\boldsymbol{\Lambda}}_{\Phi_i}^{-1/2}\mathbf{U}_{\Phi_i}^H\bar{\mathbf{H}}_i = \tilde{\mathbf{U}}_i\tilde{\boldsymbol{\Sigma}}_i\tilde{\mathbf{V}}_i^H \tag{16}$$

where $\mathbf{U}_{\Phi_i}$ and $\tilde{\mathbf{U}}_i$ are $N_{i+1} \times N_{i+1}$ unitary matrices, $\boldsymbol{\Lambda}_{\Phi_i}$ is an $N_{i+1} \times N_{i+1}$ diagonal matrix, $\tilde{\mathbf{V}}_i$ is an $N_i \times N_i$ unitary matrix and $\tilde{\boldsymbol{\Sigma}}_i$ is an $N_{i+1} \times N_i$ singular value matrix. It has been proven in [10] that for the statistically robust relay design problem (11)–(13), if q is a Schur-concave function [11] of $\mathbf{d}[\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]]$, the optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ are given by

$$\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\boldsymbol{\Lambda}_1, \quad \mathbf{F}_2 = \tilde{\mathbf{V}}_{2,1}\boldsymbol{\Lambda}_2\tilde{\mathbf{U}}_{1,1}^H\tilde{\boldsymbol{\Lambda}}_{\Phi_1}^{-1/2}\mathbf{U}_{\Phi_1}^H \tag{17}$$

where for $i = 1, 2$, $\tilde{\mathbf{V}}_{i,1}$ and $\tilde{\mathbf{U}}_{i,1}$ corresponds to N_b columns in $\tilde{\mathbf{V}}_i$ and $\tilde{\mathbf{U}}_i$ associated with the largest N_b singular values, respectively, and $\boldsymbol{\Lambda}_i$, $i = 1, 2$, are $N_b \times N_b$ diagonal matrices. If q is Schur-convex [11] with respect to $\mathbf{d}[\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]]$, the optimal $\mathbf{F}_2$ is given in (17), while the optimal $\mathbf{F}_1$ is $\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\boldsymbol{\Lambda}_1\mathbf{V}_0$, where $\mathbf{V}_0$ is an $N_b \times N_b$ unitary matrix such that $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]$ in (9) has identical main-diagonal elements.

If the exact CSI is available at all nodes, i.e., $\mathbf{H}_i = \bar{\mathbf{H}}_i$, $\boldsymbol{\Theta}_i = \mathbf{0}_{N_i}$, $\boldsymbol{\Phi}_i = \mathbf{0}_{N_{i+1}}$, $i = 1, 2$, the problem (11)–(13) becomes the MIMO relay optimization problem with the exact CSI in [4]. Therefore, the problem (11)–(13) is more general than the problem in [4]. From (6) and (7), we find that α_1 is a function of $\mathbf{F}_1$ and α_2 is a function of both $\mathbf{F}_1$ and $\mathbf{F}_2$. Consequently, it can be seen from (14)–(16) that $\tilde{\mathbf{V}}_1$ and $\tilde{\mathbf{U}}_1$ depend on $\mathbf{F}_1$ and $\tilde{\mathbf{V}}_2$ depends on both $\mathbf{F}_1$ and $\mathbf{F}_2$. Thus, from (17), we find that the explicit structure of the optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ is very difficult to find for general $\boldsymbol{\Theta}_i$ and $\boldsymbol{\Phi}_i$. In the following, we show the explicit structure of the optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ when $\boldsymbol{\Theta}_i = \tau_i\mathbf{I}_{N_i}$ and/or $\boldsymbol{\Phi}_i = \varepsilon_i\mathbf{I}_{N_{i+1}}$, $i = 1, 2$. This corresponds to the MIMO channel where the transmit and/or receiver antennas are uncorrelated as explained in detail in [7, Remark 1].

For the case of $\boldsymbol{\Phi}_i = \varepsilon_i\mathbf{I}_{N_{i+1}}$, $i = 1, 2$, the robust relay optimization problem can be written as

$$\min_{\mathbf{F}_1,\mathbf{F}_2} \quad q\left(\mathbf{d}\left[\mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{B}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\right]\right) \tag{18}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}_2\left(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \beta_1\mathbf{I}_{N_2}\right)\mathbf{F}_2^H\right) \le P_2 \tag{19}$$

$$\mathrm{tr}\left(\mathbf{F}_1\mathbf{F}_1^H\right) \le P_1 \tag{20}$$

where

$$\mathbf{B} \triangleq \bar{\mathbf{H}}_2\mathbf{F}_2\Big(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \beta_1\mathbf{I}_{N_2}\Big)\mathbf{F}_2^H\bar{\mathbf{H}}_2^H + \beta_2\mathbf{I}_{N_3} \quad (21)$$

$$\beta_1 \triangleq \varepsilon_1 \mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H\mathbf{\Theta}_1^T) + 1 \quad (22)$$

$$\beta_2 \triangleq \varepsilon_2 \mathrm{tr}(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \beta_1\mathbf{I}_{N_2})\mathbf{F}_2^H\mathbf{\Theta}_2^T) + 1. \quad (23)$$

Let us introduce the SVDs of $\bar{\mathbf{H}}_i = \mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i^H, \ i = 1,2$. It can be easily seen from (14)–(16) that for $\mathbf{\Phi}_i = \varepsilon_i\mathbf{I}_{N_{i+1}}$, we have $\tilde{\mathbf{V}}_i = \mathbf{V}_i$ and $\tilde{\mathbf{U}}_i = \mathbf{U}_i, \ i = 1,2$. Consequently, for Schur-concave q, we have

$$\mathbf{F}_1 = \mathbf{V}_{1,1}\mathbf{\Lambda}_1, \quad \mathbf{F}_2 = \mathbf{V}_{2,1}\mathbf{\Lambda}_2\mathbf{U}_{1,1}^H \quad (24)$$

where $\mathbf{V}_{i,1}$ and $\mathbf{U}_{i,1}$ corresponds to N_b columns in $\mathbf{V}_i$ and $\mathbf{U}_i$ associated with the largest N_b singular values, respectively. If q is Schur-convex, $\mathbf{F}_2$ is given in (24) and the optimal $\mathbf{F}_1$ is $\mathbf{F}_1 = \mathbf{V}_{1,1}\mathbf{\Lambda}_1\mathbf{V}_0$.

Now the task is to find the $N_b \times N_b$ diagonal power loading matrices $\mathbf{\Lambda}_i, \ i = 1,2$. For Schur-concave q, substituting (24) back into (18)–(20), we obtain the following problem:

$$\min_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2} \quad q\left(\left\{\left(1 + \frac{\sigma_{1,k}^2\lambda_{1,k}^2\sigma_{2,k}^2\lambda_{2,k}^2}{\beta_1\sigma_{2,k}^2\lambda_{2,k}^2 + \beta_2}\right)^{-1}\right\}\right) \quad (25)$$

$$\text{s.t.} \quad \sum_{k=1}^{N_b}\lambda_{2,k}^2(\lambda_{1,k}^2\sigma_{1,k}^2 + \beta_1) \le P_2 \quad (26)$$

$$\sum_{k=1}^{N_b}\lambda_{1,k}^2 \le P_1 \quad (27)$$

$$\lambda_{1,k} \ge 0, \quad \lambda_{2,k} \ge 0, \quad k = 1,\ldots,N_b \quad (28)$$

where $\beta_1 \triangleq \varepsilon_1\sum_{k=1}^{N_b}\lambda_{1,k}^2[\mathbf{V}_{1,1}^H\mathbf{\Theta}_1^T\mathbf{V}_{1,1}]_{k,k} + 1$, $\beta_2 \triangleq \varepsilon_2\sum_{k=1}^{N_b}\lambda_{2,k}^2(\sigma_{1,k}^2\lambda_{1,k}^2 + \beta_1)[\mathbf{V}_{2,1}^H\mathbf{\Theta}_2^T\mathbf{V}_{2,1}]_{k,k} + 1$. Here, for $i = 1,2$, $\lambda_{i,k}$ and $\sigma_{i,k}$, $k = 1,\ldots,N_b$, are the kth largest main diagonal elements of $\mathbf{\Lambda}_i$ and $\mathbf{\Sigma}_i$, respectively, $\boldsymbol{\lambda}_i \triangleq [\lambda_{i,1},\ldots,\lambda_{i,N_b}]^T$ and for a scalar x, $\{x_k\} \triangleq [x_1,\ldots,x_{N_b}]^T$.

By introducing $a_k \triangleq \sigma_{1,k}^2$, $x_k \triangleq \lambda_{1,k}^2$, $b_k \triangleq \sigma_{2,k}^2$, $y_k \triangleq \lambda_{2,k}^2(\lambda_{1,k}^2\sigma_{1,k}^2 + \beta_1)$, $k = 1,\ldots,N_b$, the problem (25)–(28) can be simplified to

$$\min_{\mathbf{x},\mathbf{y}} \quad q\left(\left\{1 - \frac{a_kx_kb_ky_k}{(a_kx_k + \beta_1)(b_ky_k + \beta_2)}\right\}\right) \quad (29)$$

$$\text{s.t.} \quad \sum_{k=1}^{N_b}x_k \le P_1, \quad x_k \ge 0, \quad k = 1,\ldots,N_b \quad (30)$$

$$\sum_{k=1}^{N_b}y_k \le P_2, \quad y_k \ge 0, \quad k = 1,\ldots,N_b \quad (31)$$

where

$$\mathbf{x} \triangleq [x_1,\ldots,x_{N_b}]^T,$$
$$\mathbf{y} \triangleq [y_1,\ldots,y_{N_b}]^T,$$
$$\beta_1 = \varepsilon_1\sum_{k=1}^{N_b}x_k[\mathbf{V}_{1,1}^H\mathbf{\Theta}_1^T\mathbf{V}_{1,1}]_{k,k} + 1,$$
$$\beta_2 = \varepsilon_2\sum_{k=1}^{N_b}y_k\,[\mathbf{V}_{2,1}^H\mathbf{\Theta}_2^T\mathbf{V}_{2,1}]_{k,k} + 1.$$

The optimal $\mathbf{x}$ and $\mathbf{y}$ in the problem (29)–(31) can be obtained by an iterative method developed in [4]. For any Schur-convex objective function q, since the optimal $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]$ has identical main diagonal elements, it can be shown similar to [4] that the optimal power loading vectors $\mathbf{x}$ and $\mathbf{y}$ are obtained by solving the problem (29)–(31) with $q = \sum_{k=1}^{N_b}\left[1 - \frac{a_kx_kb_ky_k}{(a_kx_k+\beta_1)(b_ky_k+\beta_2)}\right]$.

For the case of $\mathbf{\Theta}_i = \tau_i\mathbf{I}_{N_i}, \ i = 1,2$, we have $\alpha_1 = \tau_1\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)$ and $\alpha_2 = \tau_2\mathrm{tr}(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2})\mathbf{F}_2^H)$. Now we show that (9) is decreasing with respect to α_1, i.e., if $\mathrm{tr}(\tilde{\mathbf{F}}_1\tilde{\mathbf{F}}_1^H) \le \mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)$, then $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\tilde{\mathbf{E}}] \succeq \mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]$, where $\succeq$ denotes matrix positive-semidefiniteness and $\tilde{\mathbf{E}}$ is obtained from (9) with $\tilde{\mathbf{F}}_1$. In fact, by introducing $\tilde{\mathbf{F}}_1 = \alpha_1^{-1/2}\mathbf{F}_1$, (9) can be written as

$$\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}] = \mathbf{I}_{N_b} - \tilde{\mathbf{F}}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\bar{\mathbf{A}}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\tilde{\mathbf{F}}_1 \quad (32)$$

where

$$\bar{\mathbf{A}} \triangleq \bar{\mathbf{H}}_2\mathbf{F}_2\Big(\bar{\mathbf{H}}_1\tilde{\mathbf{F}}_1\tilde{\mathbf{F}}_1^H\bar{\mathbf{H}}_1^H + \mathbf{\Phi}_1 + \alpha_1^{-1}\mathbf{I}_{N_2}\Big)\mathbf{F}_2^H\bar{\mathbf{H}}_2^H + \tilde{\alpha}_2\mathbf{\Phi}_2 + \alpha_1^{-1}\mathbf{I}_{N_3} \quad (33)$$

$$\tilde{\alpha}_2 \triangleq \mathrm{tr}\Big(\mathbf{F}_2\Big(\bar{\mathbf{H}}_1\tilde{\mathbf{F}}_1\tilde{\mathbf{F}}_1^H\bar{\mathbf{H}}_1^H + \mathbf{\Phi}_1 + \alpha_1^{-1}\mathbf{I}_{N_2}\Big)\mathbf{F}_2^H\Big). \quad (34)$$

It can be clearly seen from (32)–(34) that for a given $\tilde{\mathbf{F}}_1$, $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]$ is a decreasing function of α_1. It can be shown in a similar way to (32)–(34) that $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}]$ also decreases with respect to α_2. Thus, the optimal solution of $\mathbf{F}_1$ and $\mathbf{F}_2$ occurs at $\alpha_1 = \tau_1P_1$ and $\alpha_2 = \tau_2P_2$. Consequently, from (14)–(16), we find that $\tilde{\mathbf{U}}_i$ and $\tilde{\mathbf{V}}_i$ do not depend on $\mathbf{F}_1$ and $\mathbf{F}_2$.

Now the task is to find the $N_b \times N_b$ diagonal matrices $\mathbf{\Lambda}_i, \ i = 1,2$. Substituting (17) back into (11)–(13), we have

$$\min_{\mathbf{\Lambda}_1,\mathbf{\Lambda}_2} \quad q\bigg(\mathbf{d}\bigg[\Big[\mathbf{I}_{N_b} + \tilde{\mathbf{\Sigma}}_{1,1}^2\mathbf{\Lambda}_1^2\tilde{\mathbf{\Sigma}}_{2,1}^2\mathbf{\Lambda}_2^2\Big(\tilde{\mathbf{\Sigma}}_{2,1}^2\mathbf{\Lambda}_2^2 + \mathbf{I}_{N_b}\Big)^{-1}\Big]^{-1}\bigg]\bigg) \quad (35)$$

$$\text{s.t.} \quad \mathrm{tr}\Big(\mathbf{\Lambda}_2^2(\mathbf{\Lambda}_1^2\tilde{\mathbf{\Sigma}}_{1,1}^2 + \mathbf{I}_{N_b})\Big) \le P_2 \quad (36)$$

$$\mathrm{tr}(\mathbf{\Lambda}_1^2) \le P_1 \quad (37)$$

where $\tilde{\mathbf{\Sigma}}_{i,1}$ is a diagonal matrix containing the largest N_b singular values in $\tilde{\mathbf{\Sigma}}_i, \ i = 1,2$. The problem (35)–(37) can be solved by the iterative method we just developed for solving the problem (29)–(31). Before moving to the next section, we would like to mention that as can be seen from (29)–(31) and (35)–(37), the proposed robust algorithm has the same computational complexity order as the algorithm developed in [4] which requires the exact CSI at all nodes. In other words, the improved robustness in performance is achieved without increasing the computational complexity.

IV. Robust MIMO Relay Design With Exact CSI at the Destination

In some cases, channel estimation at the destination node can be accurate enough to be modelled as perfect (i.e., perfect CSI of $\mathbf{H}_2$ and $\mathbf{H}_1$), while the CSI available at the source and relay node is still imperfect due to feedback error/delay and quantization. In such case, the linear receiving matrix $\mathbf{W}$ can be designed to optimize the MSE matrix $\mathbf{E}$ in (2) as $\mathbf{W} = \Big[\mathbf{H}_2\mathbf{F}_2(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H + \mathbf{I}_{N_2})\mathbf{F}_2^H\mathbf{H}_2^H + \mathbf{I}_{N_3}\Big]^{-1}\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1$. The resulting MSE matrix, defined as $\mathbf{E}_0$, is given as

$$\mathbf{E}_0 = \Big[\mathbf{I}_{N_b} + \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{F}_2^H\mathbf{H}_2^H(\mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^H \times\mathbf{H}_2^H + \mathbf{I}_{N_3})^{-1}\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1\Big]^{-1}. \quad (38)$$

Since the exact CSI of $\mathbf{H}_1$ and $\mathbf{H}_2$ is unknown at the node performing the optimization, we consider minimizing the objective function of $q(\mathbf{d}[\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}_0]])$. However, it can be seen from (38) that it is intractable to obtain the expression of $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}_0]$. To avoid the difficulty, in the following, we derive a lower bound of $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}_0]$.

Lemma 2 [11, Ch. 16]: For a matrix function $f(\mathbf{X})$ of random matrices $\mathbf{X}$ having finite expectation $\mathrm{E}[\mathbf{X}]$, if f is matrix-convex, then $\mathrm{E}[f(\mathbf{X})] \succeq f(\mathrm{E}[\mathbf{X}])$.

It can be shown from [13] that $f(\mathbf{X}) = \mathbf{X}^{-1}$ is a matrix-convex function of $\mathbf{X}$. Using Lemma 2 and (38), we find that for a fixed $\mathbf{H}_2$,

$$\mathrm{E}_{\mathrm{H}_1}[\mathbf{E}_0] \succeq \left[\mathbf{I}_{N_b} + \mathbf{F}_1^H \mathrm{E}_{\mathrm{H}_1}[\mathbf{H}_1^H \mathbf{C}_2 \mathbf{H}_1]\mathbf{F}_1\right]^{-1} \tag{39}$$

where $\mathbf{C}_2 \triangleq \mathbf{F}_2^H\mathbf{H}_2^H(\mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^H\mathbf{H}_2^H + \mathbf{I}_{N_3})^{-1}\mathbf{H}_2\mathbf{F}_2$. Let us define $\mathbf{C}_1 \triangleq \mathbf{I}_{N_b} + \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{F}_1$ and $\mathbf{C}_3 \triangleq \mathbf{H}_1\mathbf{F}_1$, (38) can be written as

$$\begin{aligned}\mathbf{E}_0 &= \left[\mathbf{C}_1 - \mathbf{C}_3^H\left(\mathbf{F}_2^H\mathbf{H}_2^H\mathbf{H}_2\mathbf{F}_2 + \mathbf{I}_{N_2}\right)^{-1}\mathbf{C}_3\right]^{-1}\\ &= \mathbf{C}_1^{-1} + \mathbf{C}_1^{-1}\mathbf{C}_3^H\\ &\quad\times\left(\mathbf{F}_2^H\mathbf{H}_2^H\mathbf{H}_2\mathbf{F}_2 + \mathbf{I}_{N_2} - \mathbf{C}_3\mathbf{C}_1^{-1}\mathbf{C}_3^H\right)^{-1}\mathbf{C}_3\mathbf{C}_1^{-1}\end{aligned} \tag{40}$$

where the matrix inversion lemma is applied to obtain (40). Using Lemma 2, it can be seen from (40) that for a given $\mathbf{H}_1$,

$$\begin{aligned}\mathrm{E}_{\mathrm{H}_2}[\mathbf{E}_0] \succeq \mathbf{C}_1^{-1} + \mathbf{C}_1^{-1}\mathbf{C}_3^H\Big(\mathbf{F}_2^H\mathrm{E}_{\mathrm{H}_2}\Big[\mathbf{H}_2^H\mathbf{H}_2\Big]\mathbf{F}_2\\ +\mathbf{I}_{N_2} - \mathbf{C}_3\mathbf{C}_1^{-1}\mathbf{C}_3^H\Big)^{-1}\mathbf{C}_3\mathbf{C}_1^{-1}.\end{aligned} \tag{41}$$

From (39) and (41), we obtain $\mathbf{E}_{\mathrm{lb}}$, a lower bound of $\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}_0]$ as

$$\begin{aligned}\mathbf{E}_{\mathrm{lb}} &= \Bigg[\mathbf{I}_{N_b} + \mathbf{F}_1^H\mathrm{E}_{\mathrm{H}_1}\Bigg[\mathbf{H}_1^H\Bigg[\mathbf{I}_{N_2} - \Big(\mathbf{F}_2^H\mathrm{E}_{\mathrm{H}_2}\Big[\mathbf{H}_2^H\mathbf{H}_2\Big]\mathbf{F}_2\\ &\quad + \mathbf{I}_{N_2}\Big)^{-1}\Bigg]\mathbf{H}_1\Bigg]\mathbf{F}_1\Bigg]^{-1}\\ &= \Bigg[\mathbf{I}_{N_b} + \mathbf{F}_1^H\mathrm{E}_{\mathrm{H}_1}\Bigg[\mathbf{H}_1^H\Bigg[\mathbf{I}_{N_2} - \Big(\mathbf{F}_2^H\Big[\bar{\mathbf{H}}_2^H\bar{\mathbf{H}}_2\\ &\quad + \mathrm{tr}(\mathbf{\Phi}_2)\mathbf{\Theta}_2^T\Big]\mathbf{F}_2 + \mathbf{I}_{N_2}\Big)^{-1}\Bigg]\mathbf{H}_1\Bigg]\mathbf{F}_1\Bigg]^{-1} \qquad (42)\\ &= \Bigg[\mathbf{I}_{N_b} + \mathbf{F}_1^H\Bigg[\bar{\mathbf{H}}_1^H\Bigg[\mathbf{I}_{N_2} - \Big(\mathbf{F}_2^H\Big[\bar{\mathbf{H}}_2^H\bar{\mathbf{H}}_2\\ &\quad + \mathrm{tr}(\mathbf{\Phi}_2)\mathbf{\Theta}_2^T\Big]\mathbf{F}_2 + \mathbf{I}_{N_2}\Big)^{-1}\Bigg]\mathbf{H}_1 + \gamma\mathbf{\Theta}_1^T\Bigg]\mathbf{F}_1\Bigg]^{-1} \qquad (43)\end{aligned}$$

where Lemma 1 is used to obtain (42) and (43) and $\gamma = \mathrm{tr}\Big(\mathbf{\Phi}_1\Big[\mathbf{I}_{N_2} - \Big(\mathbf{F}_2^H\Big[\bar{\mathbf{H}}_2^H\bar{\mathbf{H}}_2 + \mathrm{tr}(\mathbf{\Phi}_2)\mathbf{\Theta}_2^T\Big]\mathbf{F}_2 + \mathbf{I}_{N_2}\Big)^{-1}\Big]\Big)$. Let us introduce $\hat{\mathbf{H}}_2^H\hat{\mathbf{H}}_2 \triangleq \bar{\mathbf{H}}_2^H\bar{\mathbf{H}}_2 + \mathrm{tr}(\mathbf{\Phi}_2)\mathbf{\Theta}_2^T$, $\tilde{\mathbf{F}}_1 \triangleq (\mathbf{\Theta}_1^T)^{1/2}\mathbf{F}_1$ and $\hat{\mathbf{H}}_1 \triangleq \bar{\mathbf{H}}_1(\mathbf{\Theta}_1^T)^{-1/2}$. Here $\mathbf{\Theta}_1^T = (\mathbf{\Theta}_1^T)^{H/2}(\mathbf{\Theta}_1^T)^{1/2}$. Then (43) can be rewritten as

$$\begin{aligned}\mathbf{E}_{\mathrm{lb}} = \Big[\mathbf{I}_{N_b} + \tilde{\mathbf{F}}_1^H\hat{\mathbf{H}}_1^H\mathbf{F}_2^H\hat{\mathbf{H}}_2^H\Big(\hat{\mathbf{H}}_2\mathbf{F}_2\mathbf{F}_2^H\hat{\mathbf{H}}_2^H + \mathbf{I}_{N_3}\Big)^{-1}\\ \times\hat{\mathbf{H}}_2\mathbf{F}_2\hat{\mathbf{H}}_1\tilde{\mathbf{F}}_1 + \gamma\tilde{\mathbf{F}}_1^H\tilde{\mathbf{F}}_1\Big]^{-1}\end{aligned} \tag{44}$$

with $\gamma = \mathrm{tr}\Big(\mathbf{\Phi}_1\mathbf{F}_2^H\hat{\mathbf{H}}_2^H\Big(\hat{\mathbf{H}}_2\mathbf{F}_2\mathbf{F}_2^H\hat{\mathbf{H}}_2^H + \mathbf{I}_{N_3}\Big)^{-1}\hat{\mathbf{H}}_2\mathbf{F}_2\Big)$.

Using (44), (12) and (13), the robust relay design problem optimizing $q(\mathbf{d}[\mathbf{E}_{\mathrm{lb}}])$ can be written as

$$\begin{aligned}\min_{\tilde{\mathbf{F}}_1,\mathbf{F}_2}\quad & q\Big(\mathbf{d}\Big[\Big[\mathbf{I}_{N_b} + \tilde{\mathbf{F}}_1^H\hat{\mathbf{H}}_1^H\mathbf{F}_2^H\hat{\mathbf{H}}_2^H\Big(\hat{\mathbf{H}}_2\mathbf{F}_2\mathbf{F}_2^H\hat{\mathbf{H}}_2^H + \mathbf{I}_{N_3}\Big)^{-1}\\ &\times\hat{\mathbf{H}}_2\mathbf{F}_2\hat{\mathbf{H}}_1\tilde{\mathbf{F}}_1 + \gamma\tilde{\mathbf{F}}_1^H\tilde{\mathbf{F}}_1\Big]^{-1}\Big]\Big) \qquad (45)\\ \text{s.t.}\quad & \mathrm{tr}\Big(\mathbf{F}_2\Big(\hat{\mathbf{H}}_1\tilde{\mathbf{F}}_1\tilde{\mathbf{F}}_1^H\hat{\mathbf{H}}_1^H + \mathrm{tr}(\tilde{\mathbf{F}}_1\tilde{\mathbf{F}}_1^H)\mathbf{\Phi}_1 + \mathbf{I}_{N_2}\Big)\mathbf{F}_2^H\Big) \le P_2 \qquad (46)\\ & \mathrm{tr}\Big(\tilde{\mathbf{F}}_1^H(\mathbf{\Theta}_1^T)^{-H/2}(\mathbf{\Theta}_1^T)^{-1/2}\tilde{\mathbf{F}}_1\Big) \le P_1. \qquad (47)\end{aligned}$$

In contrast to the case of imperfect CSI at all nodes discussed in Section III, it is very difficult to find the optimal structure of $\tilde{\mathbf{F}}_1$ and $\mathbf{F}_2$ as the solution to the problem (45)–(47). Inspired by the robust design in Section III, we adopt a (sub)optimal structure of $\tilde{\mathbf{F}}_1$ and $\mathbf{F}_2$ as $\tilde{\mathbf{F}}_1 = \hat{\mathbf{V}}_{1,1}\mathbf{\Lambda}_1$, $\mathbf{F}_2 = \hat{\mathbf{V}}_{2,1}\mathbf{\Lambda}_2\hat{\mathbf{U}}_{1,1}^H$, where $\hat{\mathbf{H}}_i = \hat{\mathbf{U}}_i\hat{\mathbf{\Sigma}}_i\hat{\mathbf{V}}_i^H$, $i = 1, 2$, is the SVD of $\hat{\mathbf{H}}_i$, $\hat{\mathbf{U}}_{i,1}$ and $\hat{\mathbf{V}}_{i,1}$, $i = 1, 2$, contain N_b columns in $\hat{\mathbf{U}}_i$ and $\hat{\mathbf{V}}_i$ associated with the largest N_b singular values, respectively. Now the robust relay design problem boils down to the optimization of the power loading matrices $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$, which is given by

$$\begin{aligned}\min_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2}\quad & q\left(\left\{\left(1 + \frac{\hat{\sigma}_{1,k}^2\lambda_{1,k}^2\hat{\sigma}_{2,k}^2\lambda_{2,k}^2}{\hat{\sigma}_{2,k}^2\lambda_{2,k}^2 + 1} + \gamma\lambda_{1,k}^2\right)^{-1}\right\}\right) \qquad (48)\\ \text{s.t.}\quad & \sum_{k=1}^{N_b}\lambda_{2,k}^2\left(\hat{\sigma}_{1,k}^2\lambda_{1,k}^2 + m_k\sum_{j=1}^{N_b}\lambda_{1,j}^2 + 1\right) \le P_2 \qquad (49)\\ & \sum_{k=1}^{N_b} d_k\lambda_{1,k}^2 \le P_1 \qquad (50)\\ & \lambda_{1,k} \ge 0,\quad \lambda_{2,k} \ge 0,\quad k = 1,\ldots,N_b \qquad (51)\end{aligned}$$

where for $i = 1, 2$, $\lambda_{i,k}$ and $\hat{\sigma}_{i,k}$, $k = 1,\ldots,N_b$, are the kth largest main diagonal elements of $\mathbf{\Lambda}_i$ and $\hat{\mathbf{\Sigma}}_i$, respectively, $\gamma = \sum_{k=1}^{N_b}\hat{\sigma}_{2,k}^2\lambda_{2,k}^2 m_k/(\hat{\sigma}_{2,k}^2\lambda_{2,k}^2 + 1)$, m_k is the kth main diagonal element of $\mathbf{M} \triangleq \hat{\mathbf{U}}_{1,1}^H\mathbf{\Phi}_1\hat{\mathbf{U}}_{1,1}$ and d_k is the kth main diagonal element of $\mathbf{D} \triangleq \hat{\mathbf{V}}_{1,1}^H(\mathbf{\Theta}_1^T)^{-H/2}(\mathbf{\Theta}_1^T)^{-1/2}\hat{\mathbf{V}}_{1,1}$.

By introducing $x_k \triangleq \lambda_{1,k}^2$, $a_k \triangleq \hat{\sigma}_{1,k}^2$, $b_k \triangleq \hat{\sigma}_{2,k}^2$, $y_k \triangleq \lambda_{2,k}^2\Big(\hat{\sigma}_{1,k}^2\lambda_{1,k}^2 + m_k\sum_{j=1}^{N_b}\lambda_{1,j}^2 + 1\Big)$, $k = 1,\ldots,N_b$, the problem (48)–(51) can be equivalently written as

$$\begin{aligned}\min_{\mathbf{x},\mathbf{y}}\quad & q\left(\left\{\left(1 + \frac{a_kb_kx_ky_k}{a_kx_k + b_ky_k + m_k\sum_{j=1}^{N_b}x_j + 1} + \gamma x_k\right)^{-1}\right\}\right) \qquad (52)\\ \text{s.t.}\quad & \sum_{k=1}^{N_b} y_k \le P_2,\quad y_k \ge 0,\quad k = 1,\ldots,N_b \qquad (53)\\ & \sum_{k=1}^{N_b} d_kx_k \le P_1,\quad x_k \ge 0,\quad k = 1,\ldots,N_b \qquad (54)\end{aligned}$$

where we have now $\gamma = \sum_{k=1}^{N_b}\frac{b_ky_km_k}{a_kx_k + b_ky_k + m_k\sum_{j=1}^{N_b}x_j + 1}$. Similar to the problem (29)–(31), the problem (52)–(54) can be efficiently solved by iteratively updating $\mathbf{x}$ and $\mathbf{y}$.

V. Numerical Examples

In this section, we study the performance of the proposed robust source and relay matrices through numerical simulations. In the simulations, the estimated channel matrices $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$ have i.i.d. complex Gaussian entries with zero-mean and variances σ_i^2/N_i for $\bar{\mathbf{H}}_i$, $i = 1, 2$. We define $\mathrm{SNR}_i = \sigma_i^2P_iN_{i+1}/N_i$ as the signal-to-noise ratio (SNR) for the ith hop, $i = 1, 2$. In all simulations, we set $N_i = 4$, $i = 1, 2, 3$, $N_b = 3$, $\mathrm{SNR}_2 = 20$ dB, and all simulation results are averaged over 1000 independent realizations of the true channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$. We will consider the following five MIMO relay algorithms.

- Non-RB (Imp. CSID): The algorithm proposed in [4] using the imperfect CSI at all nodes.
- Non-RB (Exact CSID): The algorithm developed in [4] with the exact CSI at the destination node.
- RB (Imp. CSID): The robust algorithm with imperfect CSI at the destination node developed in Section III with $q = \mathrm{tr}(\mathrm{E}_{\mathrm{H}_1,\mathrm{H}_2}[\mathbf{E}])$.
- RB (Exact CSID): The robust algorithm with the exact CSI at the destination node developed in Section IV with $q = \mathrm{tr}(\mathbf{E}_{\mathrm{lb}})$.
- Exact CSI (All nodes): The MIMO relay algorithm proposed in [4] using the exact CSI at all nodes.

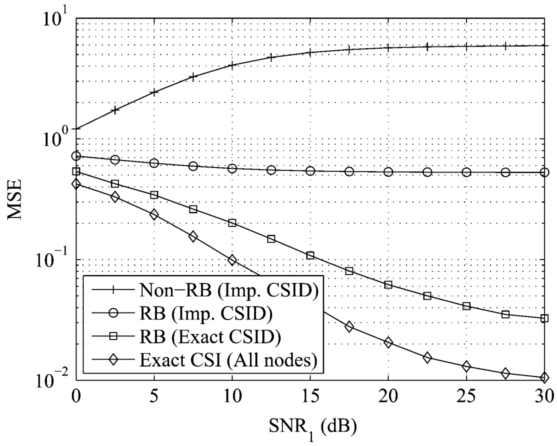

Fig. 1. Example 1: MSE versus SNR_1.

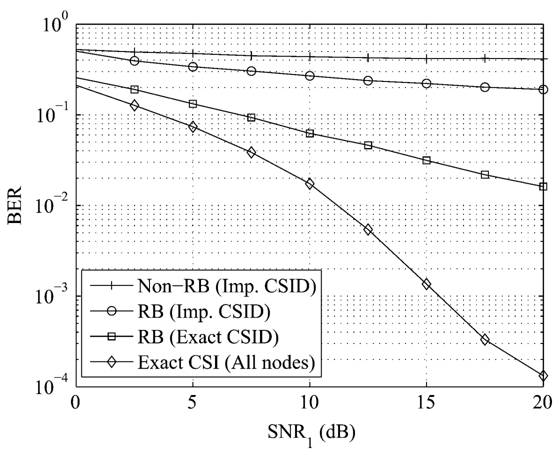

Fig. 2. Example 1: BER versus SNR_1.

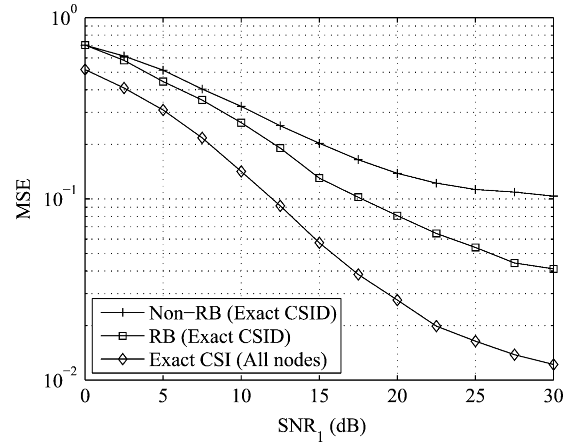

Fig. 3. Example 2: MSE versus SNR_1.

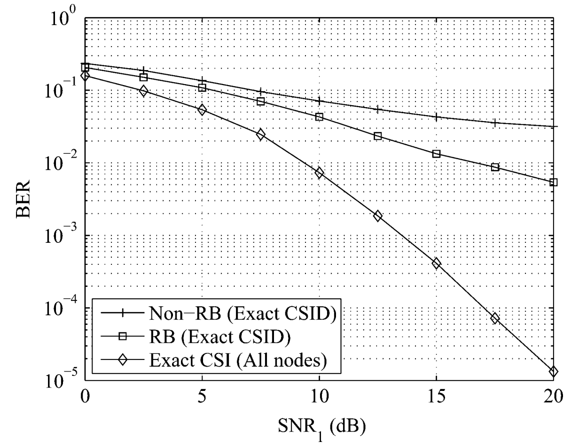

Fig. 4. Example 2: BER versus SNR_1.

In the first example, the true channel matrices are modelled as (3) with $\mathbf{\Theta}_i = \mathbf{I}_{N_i},\ i = 1, 2$ and $\mathbf{\Phi}_i,\ i = 1, 2$, are defined as Toeplitz matrices [5]–[9] with elements given by $[\mathbf{\Phi}_1]_{m,n} = 0.8^{|m-n|}$ and $[\mathbf{\Phi}_2]_{m,n} = 0.9^{|m-n|}$, respectively. Fig. 1 shows the MSE performance of four algorithms versus SNR_1. It can be seen that for a MIMO relay system with imperfect CSI at the destination node, the MSE produced by the nonrobust relay algorithm increases with SNR_1. This is due to the fact that the mismatch between $\mathbf{H}_i$ and $\bar{\mathbf{H}}_i$ is not considered by the nonrobust relay algorithm. The robust algorithm developed in Section III has a much better MSE performance compared with the nonrobust algorithm. But due to the missing of the exact CSI at the destination node, it still yields a high error-floor. With the exact CSI at the destination node, the robust relay algorithm proposed in Section IV yields a similar slope of decreasing MSE with respect to SNR_1 as the relay scheme with the exact CSI at all nodes.

Fig. 2 shows the BER performance of four relay algorithms versus SNR_1. The QPSK constellations are used to modulate the source symbols. We observe that when the exact CSI is not available at the destination node, the nonrobust algorithm has a BER close to 0.5 over the whole SNR_1 range. The robust relay algorithm has a better performance in this case. With the exact CSI at the destination node, the robust relay algorithm proposed in Section IV further improves the system BER performance. But at large SNR_1, there is still some gap between the BER of the relay system with the exact CSI at all nodes and the relay system with imperfect CSI at the source and relay nodes.

In the second example, the true channel matrices are modelled as (3) with $[\mathbf{\Theta}_1]_{m,n} = 0.7^{|m-n|}$, $[\mathbf{\Theta}_2]_{m,n} = 0.8^{|m-n|}$, $[\mathbf{\Phi}_1]_{m,n} = 0.6^{|m-n|}$, and $[\mathbf{\Phi}_2]_{m,n} = 0.7^{|m-n|}$. As mentioned in Section III, the explicit structure of the optimal robust $\mathbf{F}_1$ and $\mathbf{F}_2$ is difficult to obtain when the destination node has the imperfect CSI. Thus, we compare the performance of three relay systems with the exact CSI at the destination. Fig. 3 shows the MSE performance of three algorithms versus SNR_1, while Fig. 4 displaces the system BER produced by three algorithms versus SNR_1. Similar to Figs. 1 and 2, we observe that the robust relay algorithm developed in Section IV has improved MSE and BER performance than the nonrobust algorithm.

VI. Conclusion

We have addressed two imperfect CSI scenarios in linear nonregenerative MIMO relay communications. For each case, we have developed statistically robust source and relay matrices for most commonly used MIMO system design criteria. Simulation results show an improved robustness of the proposed algorithms against CSI errors.

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] I. Hammerström and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2798–2802, Aug. 2007.

[3] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, pp. 5062–5073, Oct. 2008.

[4] 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.

[5] C. Xing, S. Ma, and Y.-C. Wu, "Bayesian robust linear transceiver design for dual-hop amplify-and-forward MIMO relay systems," presented at the IEEE GLOBECOM, Honolulu, HI, Nov. 2009.

[6] C. Xing, S. Ma, Y.-C. Wu, T.-S. Ng, and H. V. Poor, "Linear transceiver design for amplify-and-forward MIMO relay systems under channel uncertainties," presented at the IEEE Wireless Commun. Netw. Conf. (WCNC), Sydney, Australia, Apr. 2010.

[7] 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.

[8] 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, Article ID 640186, 13 pages.

[9] 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.

[10] Y. Rong, "Robust design for linear non-regenerative MIMO relays," in *Proc. Australian Commun. Theory Workshop*, Canberra, Australia, Feb. 2–5, 2010, pp. 87–92.

[11] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.

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

[13] E. Jorswieck and H. Boche, *Majorization and Matrix-Monotone Functions in Wireless Communications*. Hanover, MA: Now Publishers, 2007.